

# Comparing protein and mRNA expression level on a genome scale: posttranscriptional regulation in yeast *Saccharomyces cerevisiae*

Roberto Olivares, Renata Usaite, Kiran R. Patil, Jens Nielsen<sup>1</sup>

Center for Microbial Biotechnology, BioCentrum-DTU, Building 223, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark. <sup>1</sup>email: jn@biocentrum.dtu.dk

In this study we explore posttranscriptional regulation using high-quality global gene and protein expression level from *Saccharomyces cerevisiae*. Datasets were generated identifying levels for approximately 1800 different proteins and for mRNA it was possible to measure levels for 5900 genes. Protein and transcript levels were measured for three mutants with deletions in genes SNF1, SNF4 and double deletion SNF1-SNF4 and wild type strain. Performing a Student t-test on differential data, between wild type and each mutant, the dataset was reduced to approximately 250 ORFs for each dataset, see **Table 1**. These reduced datasets have the characteristic of contains genes that have significant changes in mRNA and protein levels. In this way it was possible to reduce the presence of noise in the final measurements.

The reduced dataset confers reliability allowing a deep inspection of the features derived from gene and protein sequence and its respective fold change and deviation from the correlation mRNA-proteins. Only 45 ORFs were common for the three data sets with GO term associated to eight different processes which makes an independent dataset that measure the response to different stress conditions.

A first inspection of the data was to find a correlation between fold changes in mRNA and fold changes in protein. The data presents a weakly positive Pearson and Spearman rank correlation that can be significant but insufficient to predict protein levels from mRNA (data not show).

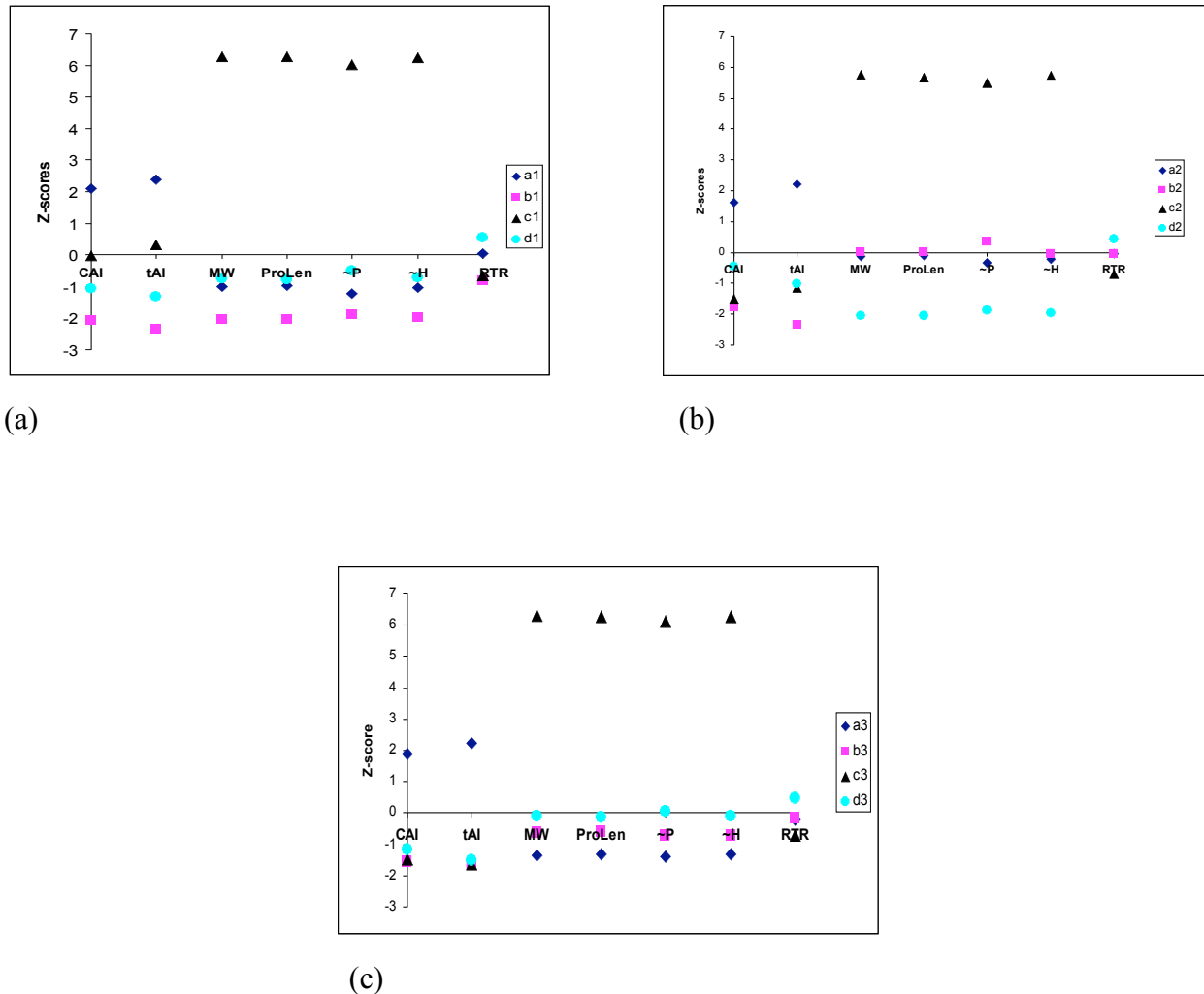
**Table 1.** Groups with different fold changes in protein and mRNA levels. For each mutant we identified different number of ORFs with significant changes,  $\Delta snf1 = 238$ ,  $\Delta snf4 = 237$  and  $\Delta snf1snf4 = 303$ .

		<i>Fold change up in mRNA level</i>			<i>Fold change down, in mRNA level</i>			
<i>Fold change up</i>								
<i>in protein level:</i>								
		x=1.2 <sup>a</sup>	1.3	1.4		1.2	1.3	1.4
$\Delta snf1$	a <sub>1</sub> . <sup>b</sup>	63	38	20	b <sub>1</sub> .	12	7	3
$\Delta snf4$	a <sub>2</sub> .	77	53	36	b <sub>2</sub> .	8	5	2
$\Delta snf4snf1$	a <sub>3</sub> .	122	95	77	b <sub>3</sub> .	11	5	1
<i>Fold change down</i>								
<i>in protein level:</i>								
		1.2	1.3	1.4		1.2	1.3	1.4
$\Delta snf1$	c <sub>1</sub> .	12	5	2	d <sub>1</sub> .	79	62	50
$\Delta snf4$	c <sub>2</sub> .	11	7	4	d <sub>2</sub> .	76	66	54
$\Delta snf4snf1$	c <sub>3</sub> .	8	3	0	d <sub>3</sub> .	95	81	71

<sup>a</sup> x represent the fold change between mutant and WT. The fold changes are 1.2, 1.3 and 1.4.

<sup>b</sup> For instance,  $\Delta snf1$  groups a<sub>1</sub>, b<sub>1</sub>, c<sub>1</sub> and d<sub>1</sub> covers all the directions in fold change.

The first calculations were a size-independent Z-score was calculated for each group included in the **Table 1**. With Z-score is possible to make comparable the descriptors for each group independently of the size. We can compare features as protein length, molecular weight, codon adaptation index (CAI) translational efficiency (tAI), relative translation rate and metabolic energy cost ( $\sim P$  and  $\sim H$ ) for each amino acid present the expressed protein. High values of Z-score represent high significance in the respective features among the groups.



**Figure 1** . Z-scores for each group with fold change  $x=1.2$ . (a)  $\Delta snf1$ , (b)  $\Delta snf4$ , (c)  $\Delta snf1\Delta snf4$ .

For the three mutants, **Figure 1(a)**, a special trend in the Z-score is observed when the protein fold change goes down, groups c1, c2 and c3, meaning that the cell decreases the amount of protein for these genes due the high cost, and this in accordance with the low translational efficiency, low values of CAI, tAI and RTR. When the amount of protein goes up, group b for each mutant, these proteins have a low energy cost with low translational efficiency concluding that this group has a significant biological importance for the cell. To express genes in groups a1, a2 and a3, the cell does not required much energy and these genes have a considerable translational efficiency.

1. Greenbaum D, Colangelo C, Williams K, Gerstein M (2003) Comparing protein abundance and mRNA expression levels on a genomic scale. *Genome Biol* 4: 117
2. Carbone A, Zinovyev A and Kepes F (2003) Codon adaptation index as a measure of dominating codon bias. *Bioinformatics* **19**:2005-15
3. Man O and Pilpel Y,(2007) Differential translation efficiency of orthologous genes is involved in phenotypic divergence of yeast species. *Nature Genetics* **39**:415-421
4. dos Reis M, Savva R and Wernisch L (2004) Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Res* **32**:5036-44
5. Gygy SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between Protein and mRNA abundance in yeast. *Mol. & Cel. Rese.* **19**(3):1720-30
6. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R and Hood L (2001) Integrated genomic and proteomic analysess of a systematically perturbed metabolic network. *Science* **292**:929-34
7. Akashi H and Gojobori T (2002) Metabolic efficiency and amino acid composition in the proteomes of E. coli and Bacilus subtilis. *Proce. Nat. Acad. Scien.*, **99**(6):3695-3700