

# Yeast U3 localization and correct sequence (*snR17a*) and promotor activity (*snR17b*) identified by homology search

Sir,

*snR17a* and *snR17b* from yeast, the two copy gene encoding the yeast homologue of vertebrate small nuclear RNA U3 (Hughes *et al.*, 1987), were recently discovered to contain similar introns (Myslinski *et al.*, 1990). Three groups (Lagosky *et al.*, 1987; Fling *et al.*, 1988; Barclay *et al.*, 1988) have independently described the cloning of the dihydrofolate reductase gene (*DFR1*) from yeast and another (Gilbert, 1989) looked for transcription start sites on the opposite

strand to the *DFR1* locus. None of these groups seems to have realized from the accumulating data that they all had sequenced the same DNA region. However, a computer search for sequences homologous to the intron reveals that the 3' end of *snR17a* is located exactly 311 bp upstream from the start of the open reading frame of the *DFR1* gene. A contiguous sequence of 3234 bp can be formed from all the sequences (fig. 1).

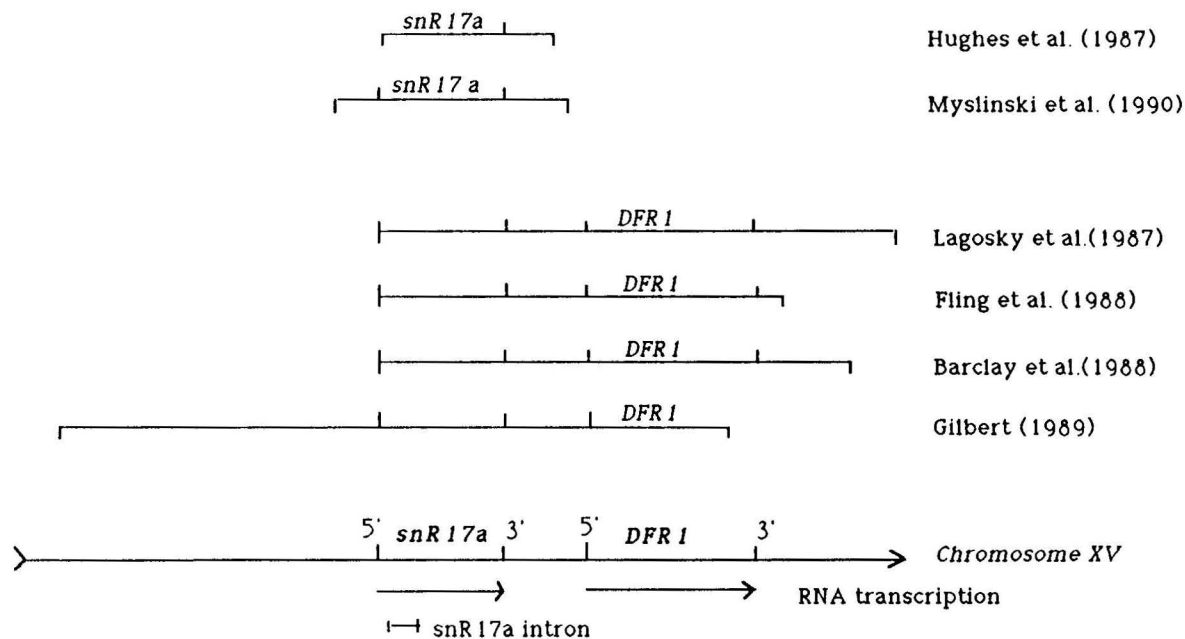


Figure 1. DNA region containing *snR17a* and *DFR1*. The regions sequenced by each of the groups are shown as lines. Start- and endpoints of the two genes on the sequenced stretches are marked. The genes already recognized by the different groups are labeled. The contiguous sequence on chromosome XV is sketched schematically at the bottom, direction and length of transcription of the respective RNAs and the position of the intron in the *snR17a* gene are indicated.

**Table 1** List of the sequences containing yeast U3 found by the search

Acc. No. <sup>a</sup>	Reference	described gene			additional gene identified			total length
		from	to	from	to	from	to	
X05499	Hughes et al. (1987)	<i>snR17a</i>	> 1	455	—	—	—	654 bp
X05498	Hughes et al. (1987)	<i>snR17b</i>	194	655	—	—	—	900 bp
M26648	Myslinski et al. (1990)	<i>snR17a</i>	182	671	—	—	—	724 bp
Y00887	Lagosky et al. (1987)	<i>DFR1</i>	800	1433	<i>snR17a</i>	1	490	1957 bp
M26667	Fling et al. (1988)	<i>DFR1</i>	799	1432	<i>snR17a</i>	1	489	1540 bp
M26668	Barclay et al. (1988)	<i>DFR1</i>	799	1432	<i>snR17a</i>	> 1	489	1782 bp
X16541	Gilbert (1989)	<i>DFR1</i>	524	>1(C)	<i>snR17a</i>	1324	836(C)	2601 bp
M23394	Santangelo et al. (1988)	random	1	559	<i>snR17b</i>	411	551	559 bp

<sup>a</sup> Respective EMBL accession numbers are given in this column. The feature tables have now been updated accordingly.

The termination signals for *snR17a* partly overlap with the transcription start signals of the *DFR1* gene. *snR17a* can now also be mapped to chromosome XV by the map position of *DFR1* (Barclay et al., 1988). The four differences noted by Myslinski et al. (1990) in their *snR17a* gene sequence in comparison to the data of Hughes et al. (1987) are all confirmed in the data from each of the four groups (Lagosky et al., 1987; Fling et al., 1988; Barclay et al., 1988; Gilbert, 1989) investigating the *DFR1* gene. Thus the sequence of Myslinski et al. (1990) is probably the correct sequence and the folding model of the RNA encoded by the *snR17a* gene (Hughes et al., 1987) has an increased hairpin 4. Apart from this, the upstream putative general amino acid control motif for *DFR1* suggested by Lagosky et al. (1987) and Barclay et al. (1988) becomes less convincing as it lies within the *snR17a* RNA coding region.

A similar search for the intron from *snR17b* finds a sequence reported to be a random promotor from yeast (Santangelo et al., 1988) which turns out to be the upstream half of *snR17b* (up to six nucleotides before the 3' end of the intron) plus 410 upstream nucleotides. The S1 mapping data of this group suggests that the transcription initiation site of their fusion construct lies indeed close to the revised start of the *snR17b* coding region as inferred by Myslinski et al. (1990) by the comparison of the *snR17b* gene to its RNA sequence. The related sequences are summarized in the table.

It is striking that in this case two well known genes from yeast could be sequenced repeatedly without noticing their close neighborhood and moreover, that

DNA from another copy of the first gene was considered to be random. With growing sequencing data, sequence comparisons will become more important to prevent ignoring of genes already characterized, but this will also yield valuable additional data like chromosomal location or direct evidence for promotor activity as in this case.

Thomas Dandekar  
European Molecular Biology Laboratory  
Postfach 10.22.09; D-69 Heidelberg  
Germany

## REFERENCES

- Barclay, B.J., Huang, T., Nagel, M.G., Misener, V.L., Game, J.C. and Wahl, G.M. (1988) Mapping and sequencing of the dihydrofolate reductase gene (*DFR1*) of *Saccharomyces cerevisiae*. *Gene* 63, 175–185.
- Fling, M.E., Kopf, J. and Richards, C.A. (1988) Nucleotide sequence of the dihydrofolate reductase gene of *Saccharomyces cerevisiae*. *Gene* 63, 165–174.
- Gilbert, S.C. (1989) Isolation of sequences upstream of the *Saccharomyces cerevisiae* *DFR1* gene containing a new promotor. Submitted by tape to EMBL data library, ID-code SCDFR1G.
- Hughes, J.M.X., Konings, D.A.M. and Cesarini, G. (1987) The yeast homologue of U3 snRNA. *EMBO* 6, 2145–2155.
- Lagosky, P.A., Taylor, G.R. and Haynes, R.H. (1987) Molecular characterization of the *Saccharomyces cerevisiae* dihydrofolate reductase gene (*DFR1*). *Nucl. Acids Res.* <15, 10355–10371.
- Myslinski, E., Ségault, V. and Branlant, C. (1990) An Intron in the genes for U3 small nucleolar RNAs of the yeast *Saccharomyces cerevisiae*. *Science* 247, 1213–1216.
- Santangelo, G.M., Tornow, J., McLaughlin, C.S. and Moldave, K., (1988) Properties of promoters cloned randomly from the *Saccharomyces cerevisiae* genome. *Molecular and cellular Biology* 8, 4217–4224.