03/06/2009

Annotation update for pHyg3

1. Labeling of the rbcS2 intron in pHyg3

The labeling of the *rbcS2* intron in our respective publication (Berthold et al. 2002. Protist, Vol. 153, 401–412) is wrong. The four bases in the red box in the following figure taken from our publication in fact **do** belong to the *rbcS2* intron.

The box for the intron was inadvertently misplaced in the published figure (using the wrong GT as the intron 5' border). For annotation of the sequence of the complete pHyg3 plasmid on our homepage the sequence in the paper was used as a reference for the marker cassette within the whole plasmid sequence and the wrong intron 5' end was used for counting.

2. rbcS2 3' untranslated region

In our publication we didn't resolve the individual sequences but stated: "The 3' end contains 230 bp (in fact there are 233 bp) of the 3' untranslated region of the *C. reinhardtii rbcS2* gene ....".

In our "lab terminology" all sequence downstream of the *aph7* stop codon in pHyg3 is 3'-UTR. This is in fact not correct.

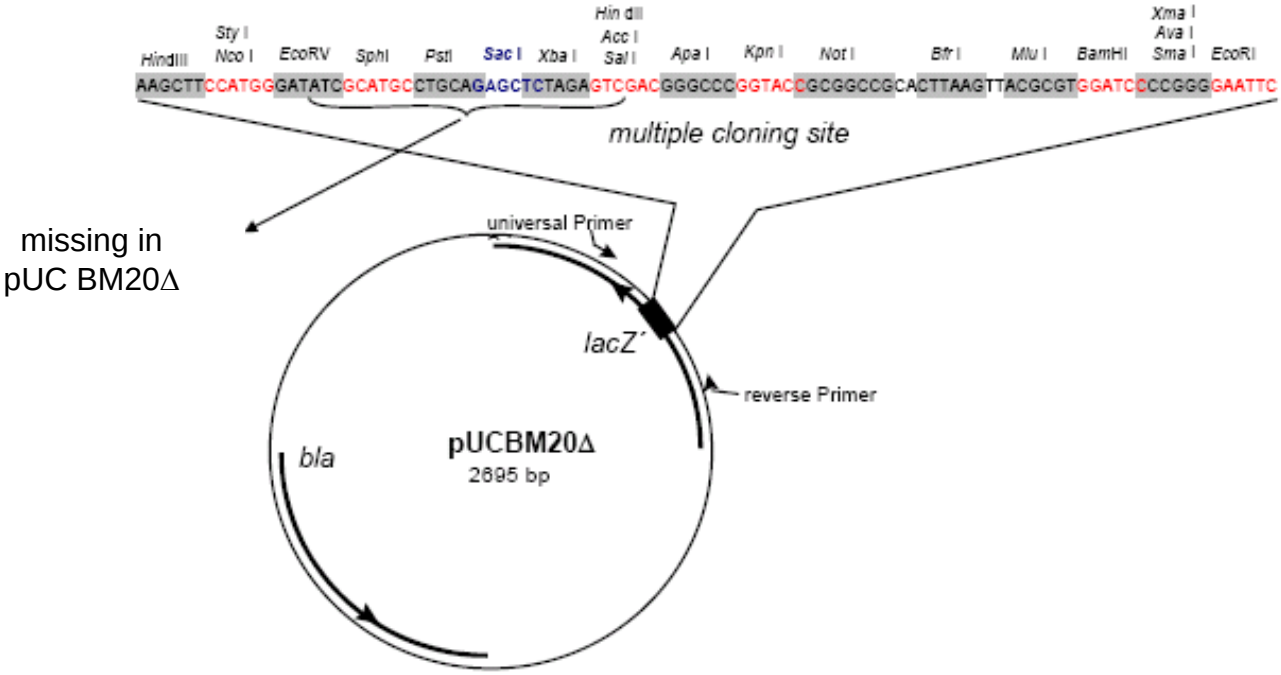
Altogether the sequences are as follows (numbering refers to the complete sequence of pHyg3):

1703	Last base of stop codon of *S. hygroscopicus aph7* gene
1704-1706	additional stop codon introduced purposely with the amplification primer (see below)
1707-1712	*Bam*HI site used for cloning *aph7* with additional stop codon into pUC BM20Δ already containing 3' region from synthetic ble-GFP construct.
1713	one additional base (**C**) found in the synthetic ble construct of Fuhrmann's (see below)
1714-1946	actual *rbcS2* 3' UTR (including the last 3 codons and the TAA stop codon of *rbcS2*)
1947-1952	*Kpn*I site from synthetic ble-GFP construct (or pUC BM20Δ).
1953-1976	linker from pUC BM20Δ (see figure of pUC BM20Δ)
1971-1976	*Hin*dIII-site from pUC BM20Δ

The following data show in more detail where the short base sequences came from. In order to understand it, it may be necessary to know some cloning details first.

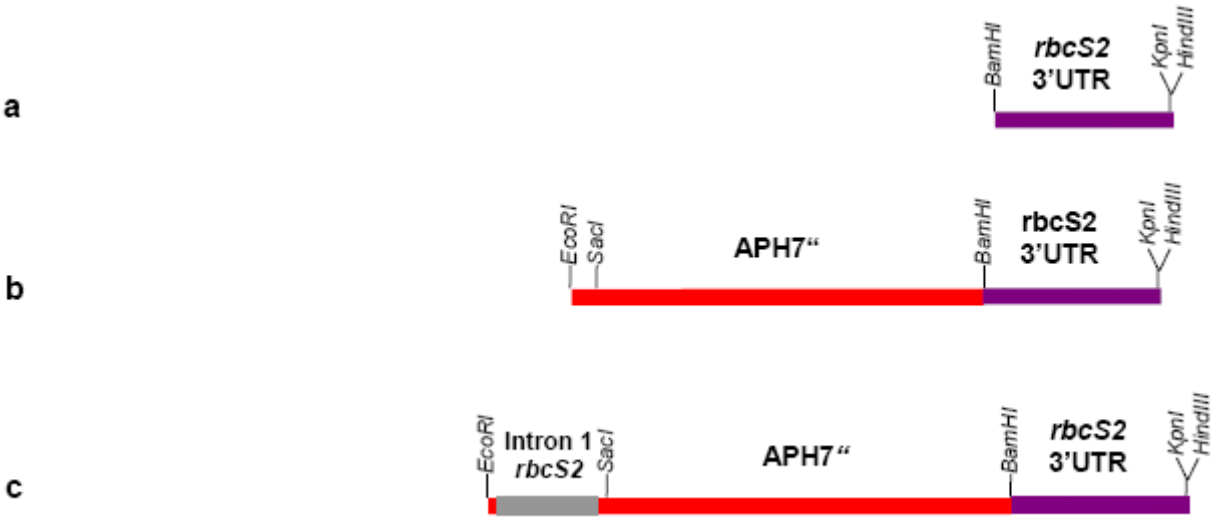
The construction of pHyg3 was done in pUC BM20 (a pUC-derivative made by former company Boehringer Mannheim which now belongs to Roche). For pUC BM20 check this: http://seq.yeastgenome.org/vectordb/vector_descrip/COMPLETE/PUCBM20.SEQ.html

We deleted restriction sites between and including EcoRV and SalI (HindII) by cutting pUC BM20 with EcoRV and Hind II and religating it to give pUC BM20Δ:



The following figure shows the cloning strategy for cloning the fragment containing the 3' region (**a**), part of the *aph7* gene (**b**) and insertion of the *rbcS2* intron 1 (**c**). I will not describe cloning and fusion of the *aph7* 5' part with the *Chlamydomonas* tubulin promoter at this point, because it is not in question.

Part of the cloning strategy for pHyg3



a

In step **a** the region containing the *rbcS2* 3' UTR was cloned via *Bam*HI and *Kpn*I into pUC BM20Δ. We used a fragment from a plasmid (pMF59) with a synthetic ble-GFP-construct that we had obtained from Markus Fuhrmann in the Hegemann lab.

The following BLAST result shows that the fragment from Fuhrmann's construct is identical with the 3' part of pHyg3 from the *Bam*HI site starting at position 1707 in pHyg3.

The *Bam*HI and *Kpn*I sites used for cloning the 3' region into pUC BM20Δ are shown in green.

An additional nucleotide (**C**) contained in "Fuhrmann's" fragment is shown in pink. **The actual rbcS2 3' UTR is shown in blue** !

```
Synthetic construct fusion Ble/GFP/delAPHVIII protein (ble::gfp::del_aphVIII)
gene, complete cds
Length=2791

 Score =  488 bits (246),  Expect = 6e-135
 Identities = 246/246 (100%), Gaps = 0/246 (0%)
 Strand=Plus/Plus
```

1707 in pHyg3

```
Query  4     GGATCCCCGCTCCGTGTAAATGGAGGCGCTCGTTGATCTGAGCCTTGCCCCCTGACGAAC  63
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2546  GGATCCCCGCTCCGTGTAAATGGAGGCGCTCGTTGATCTGAGCCTTGCCCCCTGACGAAC  2605

Query  64    GGCGGTGGATGGAAGATACTGCTCTCAAGTGCTGAAGCGGTAGCTTAGCTCCCCGTTTCG  123
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2606  GGCGGTGGATGGAAGATACTGCTCTCAAGTGCTGAAGCGGTAGCTTAGCTCCCCGTTTCG  2665

Query  124   TGCTGATCAGTCTTTTTCAACACGTAAAAAGCGGAGGAGTTTTGCAATTTTGTTGGTTGT  183
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2666  TGCTGATCAGTCTTTTTCAACACGTAAAAAGCGGAGGAGTTTTGCAATTTTGTTGGTTGT  2725
```

```
Query  184   AACGATCCTCCGTTGATTTTGGCCTCTTTCTCCATGGGCGGGCTGGGCGTATTTGAAGCG  243
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2726  AACGATCCTCCGTTGATTTTGGCCTCTTTCTCCATGGGCGGGCTGGGCGTATTTGAAGCG  2785

Query  244   GGTACC  249
             ||||||
Sbjct  2786  GGTACC  2791
```

b

Next, part of the *S. hygroscopicus aph7* gene was amplified from an *aph7*-containing plasmid that we had obtained from Dr. A. Jimenez (see: Zalacain et al. 1986. Nucl. Acids Res. 14:1565-1581). At the 3' end we used a primer that contained (the complement of) an additional TAA stop codon immediately after the TGA stop codon of the *aph7* gene. We introduced this second stop codon in order to make sure that translation of *aph7* would really stop at this point in transgenic *Chlamydomonas* cells.

The primer also contains a *Bam*HI site for cloning the *aph7* fragment upstream of the *Bam*HI/*Kpn*I-fragment that contains the 3' region in pUC BM20Δ (see step **a** in cloning strategy)

The sequence of pHyg3 at the end of the *aph7* gene and the downstream primer we used for amplification of the *aph7* gene are shown in the following (numbering is from pHyg3 as on our web page):

orange	3' end of *aph7*
red	additional stop codon introduced purposely
green	*Bam*HI site introduced for cloning
pink	additional base in 3' region from Fuhrmann's ble-GFP-aphVIII construct
blue	*rbcS2* 3' end