Dual tagging gene-trap of novel genes in *Drosophila melanogaster*.

Tamas Lukacsovich, Zoltan Asztalos, ~~Wakae Awano, Kotaro Baba and Daisuke Yamamoto.~~

Personal communication to FlyBase

July 15th, 1999.

FB-t 0108534

Results

The construction of the vectors for the gene-trap system

The gene-trap vector pGT1 used in this work was created by the following consecutive steps.

1) A synthetic, double-stranded oligonucleotide, flanked by *Hin*dIII and *Not*I sticky ends, was inserted into the polycloning site of the pCasper3 vector (Pirrotta 1988).
This oligonucleotide contained an artificial consensus splicing acceptor site, followed by *Eco*RI and *Bam*HI restriction sites. The insertion destroyed the flanking *Hin*dIII and *Not*I sites.

2) A 3.5-kb BamHI-SpeI fragment from plasmid pGaTB (Brand and Perrimon 1993) containing the Gal4 gene and SV40 transcriptional terminator sequence, was inserted into the BamHI-XbaI sites of the modified polylinker.

3) The small BamHI-SphI fragment of the 5' part of the Gal4 gene (including a short 5'UTR sequence and its first few amino acid codons) was replaced with a double-stranded, synthetic oligonucleotide devoid of the 5' UTR but containing sequence identical to the first few codons of Gal4 that were removed.

4) Another double-stranded, synthetic oligonucleotide, containing so-called "stop-start" signal, was inserted into the EcoRI-BamHI sites, just upstream from the Gal4 coding region. This sequence has consecutive overlapping transcriptional stop and start codons with which the otherwise variable reading frames can be set into the proper reading frame for the Gal4 coding region. This idea of artificial frameshifting with overlapping stop and start codons was "stolen" from the *Drosophila* gypsy transposon, which produces a fusion mRNA with a similar frameshift sequence that enables the host cell to translate the mRNA in the correct frame (Lee et al., 1991). In practice, we could not confirm whether the stop-start sequence is effective for adjusting the reading frame, because none of the trapped genes isolated so far contained any coding sequence upstream of the insertion site.

The frameshift sequence was also designed to be similar to the functional *kozak* site (Vavener 1987), enabling it to initiate translation from the Gal4 start codon if there was only an untranslated exon spliced to the Gal4 mRNA (and this actually happened in the majority of the gene-trap events).

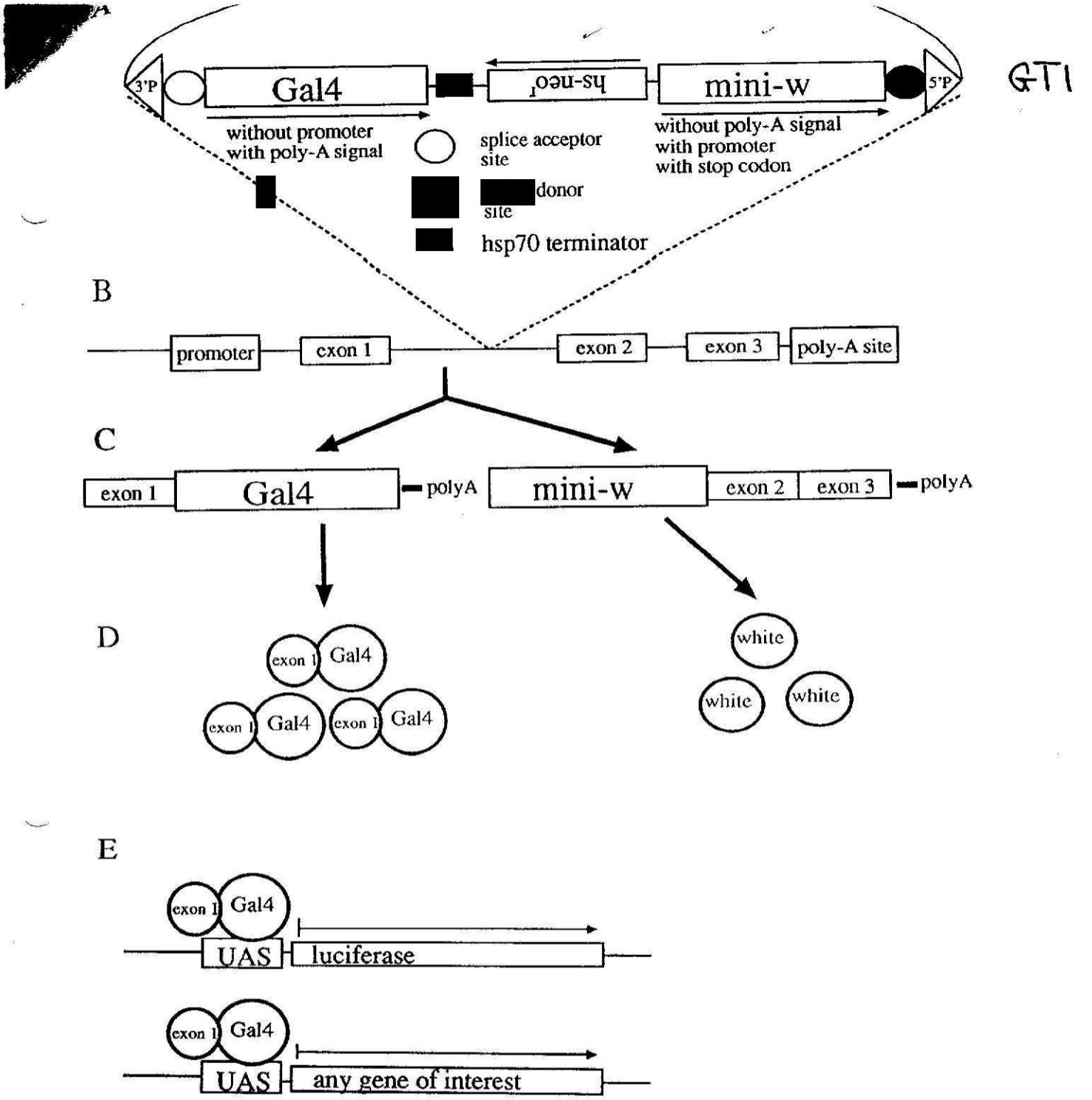
5) The 3' UTR sequence of the mini-w gene was removed and replaced with a synthetic, artificial consensus splicing donor site. To achieve this, the synthetic double-stranded oligonucleotide, which is followed by a 140-bp long NotI-HindIII fragment of plasmid pGawB (Brand and Perrimon 1993), replaced the affected part of the vector, which lies between an ApaI restriction site 17 bp upstream of the transcriptional stop codon and a HindIII site in the 5' end of the P-element sequence. The synthetic oligonucleotide, having ApaI and NotI sticky ends, completed the mini-w coding region down to

the stop codon, which was followed by the consensus splicing donor site. The 140-bp long NotI-HindIII fragment was used to complete the removed part of the 5' P-element sequence, required for mobilization of the vector.

6) In the absence of a successful gene-trapping event, the truncated mini-w gene used here would not confer any eye color on the fly. Therefore we inserted the *neomycin-phosphotransferase* gene driven by a heatshock promoter (hs-neo) in the opposite orientation into the StuI site (between the Gal4 and mini-w cassettes) for help in the selection of primary transformants based on their existence to the antibiotic. This gene is transcribed from the strand complementary to that which harbors the Gal4 and mini-w genes.

After introducing the construct, it was found that the truncated mini-w gene supports very weak pigmentation of the compound eye, leading to a pale yellow color which was distinguishable from the w^- background, and therefore selection with G418 was not necessary in establishing transformant lines. The nucleotide sequence of the entire vector has been submitted to the DDBJ database under the accession number: AB028139

The schematic map of the construct and the principle of its operation in a gene-trap system is illustrated in Fig. 1.


A
3'P
Gal4
without promoter
with poly-A signal
splice acceptor site
donor site
hsp70 terminator
hs-neo^r
mini-w
without poly-A signal
with promoter
with stop codon
5'P
GT1
B
promoter
exon 1
exon 2
exon 3
poly-A site
C
exon 1
Gal4
polyA
mini-w
exon 2
exon 3
polyA
D
exon 1
Gal4
white
E
UAS
luciferase
any gene of interest