Volume 14 Number 23 1986 Nucleic Acids Research

Nucleotide sequence of a retrotransposon *297* isolated from *Drosophila simulans*

Yoshinori Matsuo[1], Wataru Kugimiya[2], Yoichi Kadokami[3] and Kaoru Saigo*

Department of Biochemistry, Kyushu University 60 School of Medicine, Fukuoka 812, Japan

Submitted 5 November 1986

The genome of D. melanogaster has been shown to conatin a large group of retrotransposons, which are related in evolution to retroviruses and suggested to be reverse transcribed on translocation(1). The evolutionary rate of any RNA genomes including retrovirus is suggested to be more than a million times greater than that of DNA genomes(2). Thus,if retrotransposons frequently move on the chromosome, one may expect the evolutionary rate of retrotransposons to be much higher than that of normal genes. To examine this possibility, the nucleotide sequences of 297 isolated from D.simulans (pWK1020 & pMS248) were determined (Figure) and compared with that of the authentic 297 of D.melano-

```
┌►pWK1020 *           *           *            *            *            *            *       +  *           *            *
tttccgaccatatatgtgaagtgtatatAGTGACGTATTTGGGTGGTCCAAACCAGCCACTTCTATTATTTCAAAGGAATCAGTAATGCACTCTAGTAAT
                          target └──►5'LTR
TTTCCATAAC-GTATCCCAGCTGCGCAGACTCGTTTATCTTTGGCAGCGCAGCGTTCTT-GTAAACATCCTAAAG-CCTGACCTAAGCAGATTTGACTGC
CCTCTTTCAACGCTACCTAATCTTAAGAACCCAAGAGCGAGGCTCTCCCGAAATACAAATAATGTTCAAATACTGAGGCTTCTCCTCAATCCAATTTGCA
TTTGATTTTTAGTCTTAAGCTGAGATCCAAAGAATAAAGTCGTGAAACTATTTCTCCTAAAAACTATTTTTTATTTCTTGGCGTTGTCCTTAGTCAACTG
    RNA start
ACAGGACATTAGTTCGACTCAAAAAAATAAAACAACAATTTTACTGGCGCAGTCGGTGGGATACAAAAGTATCCGAAAAAAAAGAACCTTCGAGTGGAAA
                                          5'LTR◄┘           pbs
ATAAGTTAAATTTTATAGTCCAGTGCTCGAAACATCTCCCAAAATAAATTCGTGAAAACTCTTCAACTTGAATTATAATTCCAATTCGGTTATCCAATAA
TAAGTGGAAGTGAAATACGAAACAAAAATATTAAGTCCAAAGGCAACTAAGTTTTAAAACCAACATATAAAAATAAAAAATAAAAACAATATAGAATTTT
                                                                                    pMS248      ORF2(pol)
AATAATACAACACAAAAA-------AAAAAAAAAAAAAACAAGTGAAACTAGAAAGCTT(728).......(4265)ATCATACTTCAATCACATGAAAA
ACTTTTACACCCTGGTATACAGAAAATGACAAAATTATTTAAAGAAAATCACTTATTTCCAAATAGCCAACTATTAACTCAGAATTTAATAAACGAATGC
AACATATGCAATTTGGCCAAAACAGAACATAGAAACACCAAAATGCCTTTAAAAATCACACCCAACCCGGAACATTG(4454).......(4815)GAAA
CAATCCTCTACACATACAACCATAAAATTAAACATGACACTACTGGACAGACACCTGCTCAAATTTTCTTATACGCTGGGCATCCCATATTAGACACTCA
AAAAATTAAAGAGAAGAAAATAGACAAAATAAATGAAGACAGACAGGAATTTAATATTGACACTAATTACAGAAAAGGTCCACTACAGAAAGGCAAATTA
GAAAACCCATTTAAACCAACCAAAAATGTAGAACAGACAGACCCTGACCATTACAAAATCACTAATAGAAATAGAGTTACGCACTACTACAAAACACAAT
                ORF2(pol)     ┌──────────►             ◄───┐           ORF3
TCAAAAAACAAAAGAAAAATAATAAACTCTCAATTTCACAGGCACCTGGTACCC(5173--deletion--5863)TACCCATTTACCTTTCACTAATAA
GCACAATTAAAATAATTCCTTACCCAGACTCCAACGGCTATCAGCTAGATTACACAGACACACAATCATATTTTGAAAAAGAAAATAAAGTTTATATTAC
CGAAAATAAAGAAGTAAAAAATGAATGTGTCACCAATATTATTAAACACTTAAATCCAATTTGTAATTTTAAGCCAGTACACACGAACGAAATAATAAAA
TACATAGAACCAAACACAATTGTAACCTGGAACTTAACCCAAACAATTCTTAACCAAAATTGTCAAAATTCAATTAATAAAATAAAAATAGAAGGAAACA
AAATGATAAGAGTAA(6201).......(6413)TTTAAGATATGTATCATTTAAACCATTTATGATGCTGTATGCAAAACTTAAAATAAGAAAAAATCA
AAATCAAAACACACCACAACAAACAGAAATAGAAGAAACTCCATTTCCCACACTATATCCATCAATCCCAGCCCAAGTATAGGCTTCTCTTTAAGGGAAG
 ┌────►        3'LTR                                                                      ORF3◄┘       polypurine
GGGAGTGACGTATTTGGGTGGTCCAAACCAGCCACTTCTATTATTTCAAAGAAATCAGTAATGCACTCTAGTAATTTTCCATAA-TGTATCCCAGCTGCG
  ├──►pWK1020(only difference in sequence is shown)
CAGACTCGTTTATCTTTGGCAGCGCAGCGTTCTT-GTAAACATCCTAAAG-CCTGACCTAAGCAGATTTGACTGCCCTCTTTCAACACTTCCTAATCTTA
                                                                                         g  a
AGAACCCAAGAGCTAAGCTCTGCCGAAATACAAATATTGTTCAAATACTGAGGCTTCTCCTCAATCCAATTTGCATTTGATTTTTAGTCTTAAGCTGAGA
              g g     c              a
TCCAAAGAATAAAGTCGTGAAACTATTTCTCCTAAAAACTATTTTTTATTTCTTGGCGTTGTCCTTAGTCAACTGACGGGACATTAGTTCGACTCAAAAA
 polyA add'n 3'LTR◄┐ target                                                                  a
AATAAAACAACAATTTTACTatatatatatatatacg (pMS248)
                    atatattcttaatc    (pWK1020)
```

gaster (3). Differences in sequence are shown by underlines. We estimated the rate of nucleotide substitution of *297* to be about 5x10^{-9} per year per site throughout the genome. This value is essentially identical to that for normal genes (2), supporting the notion that *297* have not moved so frequently in evolution.

Present addresses: [1]Department of Psychiatry, Washington University School of Medicine, Audubon, St. Louis, MO 63110, USA, [2]Fuji Oil Co. Ltd., Izumi-Sano, Osaka 593 and [3]Life Science Institute, Advance R & D Co. Ltd., Chofu, Tokyo 182, Japan

*To whom correspondence should be addressed

References (1)Saigo,K. *et al*.(1984) Nature 312,658-661.(2)Gojobori,T. & Yokoyama,S.(1985) Proc.Natl.Acad.Sci.USA 82,4198-4201.(3)Inouye,S. *et al*.(1986)Eur.J.Biochem.154,417-425.