THESIS

Identification of a *HeT-A* Transposable Element in *Drosophila teissieri*

Submitted by

Christopher T. Ostler

In partial fulfillment of the requirements for the
Thesis in Zoology
Weber State University
Ogden, Utah
April 2001

THESIS

Identification of a *HeT-A* Transposable Element in *Drosophila teissieri*

Submitted by

Christopher T. Ostler

In partial fulfillment of the requirements for the
Thesis in Zoology
Weber State University
Ogden, Utah
April 2001

Weber State University

We hereby recommend that the thesis prepared under our supervision by Christopher T. Ostler entitled "Identification of a *HeT-A* Transposable Element in *Drosophila teissieri*" be accepted as fulfilling partial requirements for the thesis in Zoology.

Thesis committee

Gloria Z. Wurst

Roy Meyer

James B. Clark

Advisor

Samuel I. Zeveloff

Department Chair

Abstract

A telomere is a repetitive segment of DNA located at chromosome ends in eukaryotic cells. Telomere maintenance is vital to the cell's ability to divide and is believed to be important in both cancer and aging. Most eukaryotic cells utilize telomerase to rebuild telomeres after each cell division. However, in *Drosophila melanogaster* telomere maintenance is performed by *HeT-A*, a transposable element. In spite of the detailed information about *HeT-A* in *D. melanogaster,* little is known about the distribution of *HeT-A* outside of this species. The main objective of this research project was to determine whether other species of *Drosophila* also possess *HeT-A* transposable elements within their genomes.

Using the polymerase chain reaction (PCR), the *HeT-A* transposable element is amplified from the genome of *D. teissieri*, a close relative of *D. melanogaster*. This represents the first time that a *HeT-A* sequence has been identified using PCR. Comparison of the *HeT-A* sequence from *D. teissieri* to a *HeT-A* sequence from *D. melanogaster* reveals a 60.6% nucleotide identity. In contrast, an 83.0% sequence identity is reported for comparison of the *D. teissieri HeT-A* sequence to the *HeT-A* sequence from another species, *D. yakuba*. Identification of *HeT-A* transposable element in other species of *Drosophila* is important for future research to elucidate whether or not the *HeT-A* transposable element performs the same telomere maintenance function in other *Drosophila* species as it does in *Drosophila melanogaster*.

Introduction

Telomeres are repetitive DNA segments located at the ends of eukaryotic chromosomes. Most eukaryotic cells utilize the enzyme telomerase to maintain telomeres after DNA replication (Meyerson et al. 1997). However, in *Drosophila melanogaster* telomere maintenance is performed by *HeT-A*, a transposable element (Biessmann et al. 1992). Transposable elements are genes that are capable of excising themselves from one location in the genome and reinserting into another. In spite of the detailed information available about *HeT-A* in *D. melanogaster,* little is known about the distribution of *HeT-A* in other species of *Drosophila*. To date, a *HeT-A* sequence has been reported from only one other species, *D. yakuba* (Danilevskaya et al. 1998a).

The objective of this research project is to determine whether other species of *Drosophila* also possess the *HeT-A* transposable element in their genomes. Accomplishing this goal will require (1) Genomic DNA extraction from selected species of *Drosophila*, (2) Optimization of the Polymerase Chain Reaction used to amplify the *HeT-A* sequences, (3) Isolation and Cloning of amplified *HeT-A* genes from *Drosophila* and (4) Sequencing *HeT-A* gene fragments for comparison to the known sequences of the *HeT-A* elements of *D. melanogaster* and *D. yakuba*.

Materials and Methods

Flies were obtained as live cultures from the National Drosophila Species Resource Center, Bowling Green, OH. Genomic DNA was isolated by homogenizing approximately one hundred adult flies in an extraction buffer (0.1M NaCl, 0.2M sucrose, 0.01M EDTA, .03M Tris/ pH 8.0). Cells of resulting homogenized tissue were lysed using 220µl of a lysis buffer (20% SDS, 0.25M EDTA, 0.1M Tris/pH 8.0), which rendered the membrane components soluble. The cellular debris was then pelleted in a centrifuge (10,000 xg) leaving the nucleic acids and some proteins in the supernatant. Two Phenol-ether extractions followed which removed the proteins from the nucleic acids. 20µl of RNase was then added to digest the RNA, leaving only DNA. The DNA was precipitated with excess ethanol and the pellet suspended in TE Buffer (0.01M Tris and 0.001M EDTA/ pH 8.0).

The concentration and purity of DNA was determined by UV spectroscopy. A 100-fold dilution of the DNA sample was made in diH_2O and the UV absorbance was measured at wavelengths between 220nm and 320nm. The concentration was then calculated using the extinction coefficient of DNA and the purity was determined by comparing the absorbance of the DNA sample to the absorbance of known impurities such as protein and phenol (Figure 1). Once the purity and concentrations were determined, the volume of the sample was adjusted to achieve a working concentration of 400ng DNA/µl of solution.

The Polymerase Chain Reaction (PCR) is a DNA amplification procedure that mimics the normal process of DNA replication. In PCR, genomic DNA is added to a reaction tube with all of the ingredients necessary for DNA replication. General PCR conditions for a single reaction are described in Table 1. Specific primers, listed in Table 2, are used to amplify only a relatively small segment of DNA (<10,000 base pairs) that corresponds to the gene of interest. The primers used here target the more highly conserved regions of the *HeT-A* gene sequence. In addition, two sets of rRNA primers were used for positive controls. The reaction tube was continually cycled over three different temperatures within a 36 block Perk and Elmer Thermocycler: 94°C for 1 min to denature the DNA strands, 37°-50°C for 1 min for annealing of the primer, and 72°C for 1 min for polymerization of the new strands of DNA. Just as DNA synthesis prior to cell division doubles the entire genome of a cell, each cycle of PCR doubles the amount of DNA of the gene being amplified in a reaction tube. After a sufficient number of cycles, the PCR product was visualized by 1.0% agarose gel electrophoresis to reveal the size and amount of the amplified product.

Table 1. Reagents Used For the Polymerase Chain Reaction

	H2O	dNTPs	PRIMER 1	PRIMER 2	Buffer	MgCl2	Taq	Template
Final Concentration	(---)	200 μM	250 nM	250 nM	1X	5.0 mM	0.125U	5-100 ng/μl
Volume/0.025 ml RXN	15.625μl	2μl	1.25μl	1.25μl	2.5μl	1.25μl	0.125μl	1μl

Table 2. Oligonucleotide Primers Used to Amplify the HeT-A and Ribosomal mRNA Genes		
Primer	Gene Target	Sequence 5`-3`
5'UTR-1F	HeT-A	AAAGCCAGGGTATTTATACC
ORF-2F	HeT-A	ATGAGYTTYTCNGARGTGGTNGC
ORF-1R	HeT-A	TCCCGYTTNGCNGCAGGNCGG
ORF-3F	HeT-A	CARGCNAARGAYATGACNGC
ORF-3R	HeT-A	GNGTCATRTCYTTNGCYTG
ORF-5R	HeT-A	CGCCAYTTRTTNAGNCGYTTYTG
ITS-1 7246	rRNA	GCTGCGTTCTTCATCGAC
ITS-1 7247	rRNA	CGTAACAAGGTTTCCGTAGG
18S- 937	rRNA	GTGCTTCATACGGGTAG
18S- 938	rRNA	CAGCACCATAATCCTG

Following electrophoresis, bands containing amplified DNA fragments were excised from the gel, eluted and then purified using a silica-based DNA purification method (Gene Clean, Bio 101). Purified DNA was ligated into the cloning vector PCR 2.1-TOPO (Invitrogen) at a 1:3 (vector: insert) ratio to create recombinant plasmids. The plasmids were then used to transform competent *E. coli* cells. Blue/White screening was used to identify those bacteria containing plasmids with an insert of the PCR product. LB broth was inoculated with white colonies and the cultures were grown at 37°C overnight to amplify the plasmid DNA.

Plasmid DNA was isolated using the Qiagen plasmid isolation kit. To verify that the plasmids did indeed carry the *HeT-A* insert, a sample of the purified circular plasmid was digested with the restriction enzyme *Eco* R I. The digest was examined on an agarose gel that included a 1kb size standard. This allowed an estimate to be made of the sizes of both the plasmid and the *HeT-A*

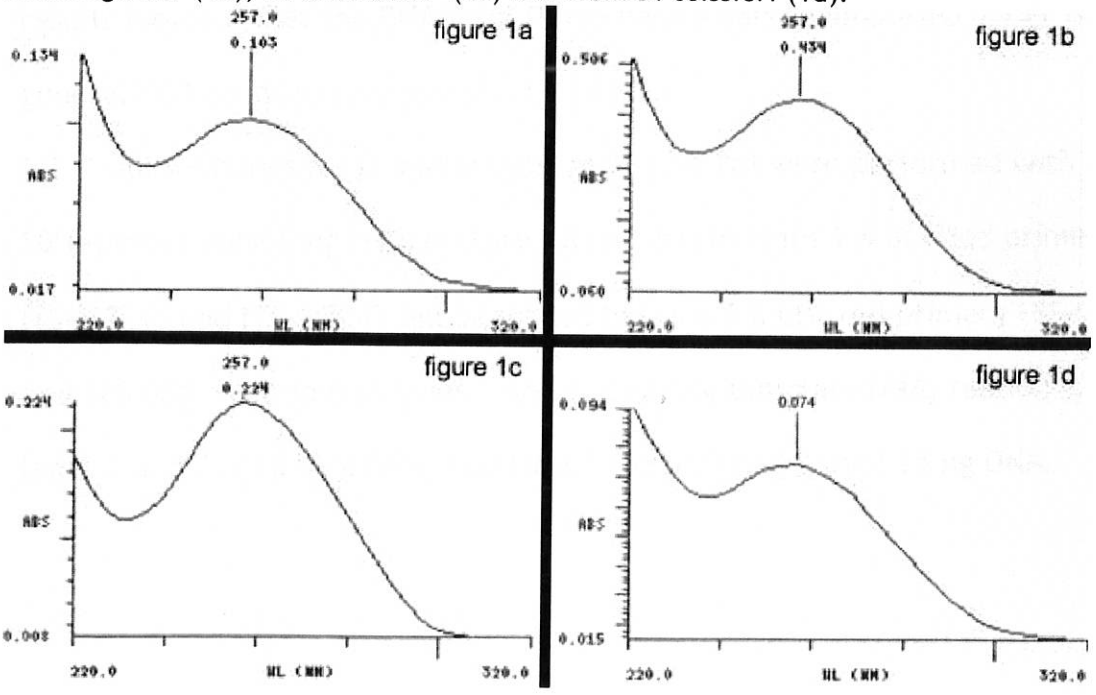
PCR fragment. Once verified, plasmids were sent to the University of Utah for automated DNA sequencing. The DNA sequence was then compared to the *HeT-A* sequence from *D. melanogaster* and *D. yakuba* using the method of Person et al. (1997), available at http://www2.igh.chrs.fr/bin/align-guess.cgi.

Results

DNA was isolated from four different species of flies. These species were chosen because they all belong to the *melanogaster* species group of the genus *Drosophila.* Figure 1 shows spectrophotometric data of DNA isolated from *D. melanogaster, D. ananassae, D. takahashii and* from *D. teissieri.* Using the values of absorbance at 257 nm, the DNA concentration of each sample was calculated using the following equation:

(A_{257})(Dilution Factor)(Extinction Coefficient) = DNA Concentration (ng/µl)

Figure 1. Spectrophotometric data of DNA isolated from *D. ananassae* (1a), *D. melanogaster* (1b), *D. takahashii* (1c) and from D. *teissieri* (1d).



To verify that the DNA was of sufficient purity required for PCR, PCR was performed using one of two primer sets, specific for ribosomal RNA gene sequences (Table 2). The primer set ITS-1 7246 and ITS-1 7247 targets the internal transcribe spacer separating the 18s rRNA gene and the 5.8 s rRNA gene. The 18S-937 and 18S-938 primer set targets a portion of the 18s rRNA gene. Since all flies have rRNA genes, amplification with these primers is expected to be successful. These constitute positive control reactions.

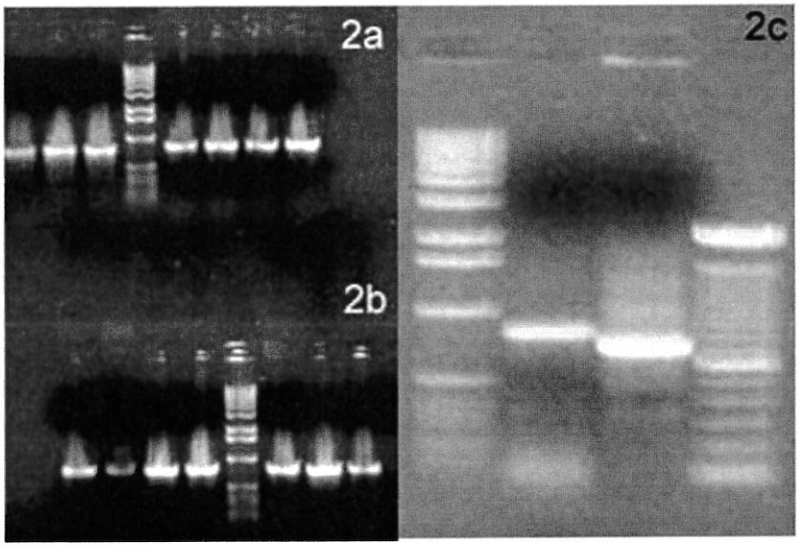
Different concentrations of template DNA were used to determine if the DNA concentration influences the success of the PCR. All reactions for *D. ananassae* (Figure 2a) used a 50°C primer annealing temperature. Reactions in lanes 1-3 utilized primers ITS-1 7246 and ITS-1 7247, and reactions in lanes 5-8 utilized primers 18S-937 and 18S-938. Reactions in lanes 1 and 6 used 5ng template DNA. Reactions in lanes 2 and 7 used 25ng DNA; reactions in lanes 3 and 8 used 100ng DNA. The reaction in lane 5 used 1ng DNA. These positive results indicate that the DNA from *D. ananassae* can be amplified under the general PCR conditions described in Table 1.

All reactions for *D. melanogaster* (Figure 2b) were performed with a 50°C primer annealing temperature. Reactions in lanes 1-4 utilized primers ITS-1 7246 and ITS-1 7247 and reactions in lanes 6-8 utilized primers 18S-937 and 18S-938; reactions in lanes 1 and 6 used 1ng template DNA; reactions in lanes 2 and 7 used 5 ng DNA; reactions in lanes 3 and 8 used 25 ng DNA.

The reaction in lane 4 used 100ng DNA. These positive results indicate that the genomic DNA from *D. melanogaster* can be amplified under general PCR conditions (Table 1).

The reactions for *D. takahashii* (Figure 2c) all used a 50°C primer annealing temperature. The reactions in lane 2 utilized primers ITS-1 7246 and ITS-1 7247, and the reaction in lane 3 utilized primers 18S-937 and 18S-938. A 100 bp size standard (Life Technologies) is shown in lane 4.

Figure 2. Results of PCR Using rRNA specific primers. Template DNA was from *D. ananassae* (2a), *D. melanogaster* (2b), and *D. takahashii* (2c). The 1 kb size standard (Life Technologies) is located in Figure 2a lane 4; 2b, lane 5; and 2c, lane 1. The 1 kb size standard has bands of 12.216 kb; 11.198 kb; 10.180 kb; 9.162 kb; 8.144 kb; 7.126 kb; 6.108 kb; 5.090 kb; 4.072 kb; 3.054 kb; 2.036 kb; 1.636 kb; and 1.018 kb, top to bottom.



After verifying that the various DNA samples were pure, it was necessary to search for the optimal conditions for using the *HeT-A* primers. PCR is a method in which DNA is replicated *in vitro*. Among the variables that affect the success of PCR are DNA template concentration, primer sequences, $MgCl_2$

concentration, dNTP concentration, and cycling temperatures (Innis and Gelfand 1990). The primer annealing temperature is particularly important because PCR will be successful only if the primers anneal to the target sequence. The *HeT-A* primers used for PCR here were designed from the *HeT-A* sequences that have been reported from *D. melanogaster* (Danilevskaya et al. 1998a) The predicted DNA fragment sizes for the primer sets used are listed in Table 3.

Table 3. Primer Combinations

Pimer 1	Primer 2	Expected Fragment Size
5`-UTR-1F	ORF-1R	1040 bp
ORF-2F	ORF-3R	630 bp
ORF-2F	ORF-5R	1400 bp
ORF-3F	ORF-5R	785 bp
5`-UTR-1F	ORF-3R	1730 bp

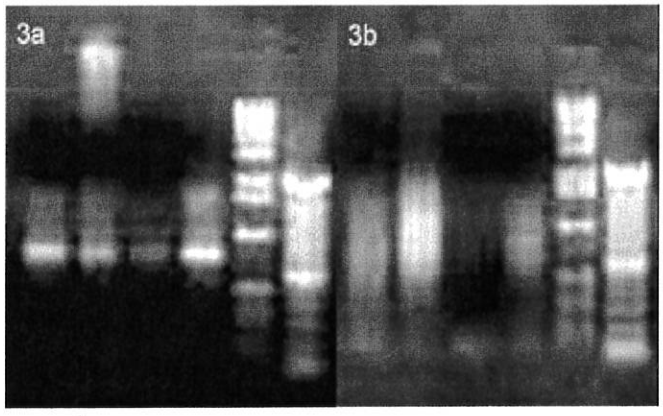


Figure 3. PCR results from primers ORF-3F and ORF-5R. *D. melanogaster*, Figure 3a and *D. ananassae*, Figure 3b. Lanes 5 and 6, Figures 3a and 3b, represent 1kb and 100bp size standards, respectively.

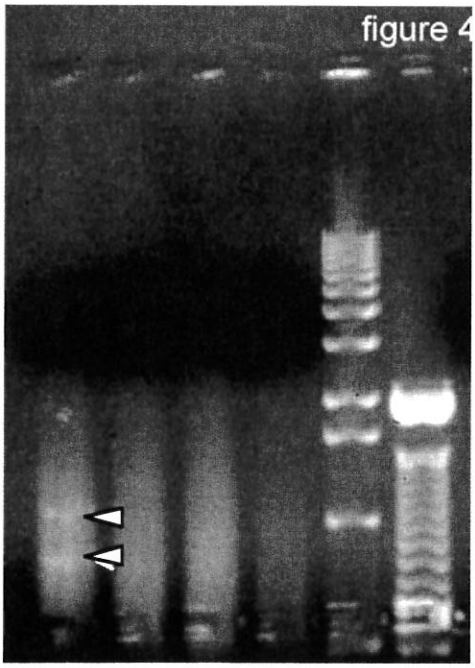


Figure 4. Results of PCR using variable amounts of $MgCl_2$ and primers with DNA from *D. ananassae*. Lanes 5 and 6 represent 1 kb and 100 bp size standards, respectively. The bottom arrowhead shows a faint band of the expected size (785 bp) and the top arrowhead shows a band of anomalous size.

The *HeT-A* primers ORF-3F and ORF-5R were first tried with DNA from *D. ananassae*. The reaction conditions included annealing temperatures that began at 37°C and rose slowly over a two minute interval to 72°C. It was reasoned that relatively low annealing temperatures would be forgiving if the primers did not match the target sequence exactly. The results for *D. melanogaster* DNA are shown in Figure 3a, and for *D. ananassae* in Figure 3b. All Reactions used 10mM $MgCl_2$ concentration. For both template DNAs, reactions in lanes 1 and 3 used 10ng of template DNA, and the reactions in lanes 2 and 4 used 100ng of template DNA. The reactions using DNA from *D. ananassae* are characterized by the lack of discrete bands and excess smearing (Figure 3b). In contrast, the reactions from *D. melanogaster* show bands of the correct size and little smearing (Figure 3a).

Since $MgCl_2$ concentration and primer concentration can both affect primer binding specificity (Innis and Gelfand 1990), reactions were tried with a range of concentrations of each. These results are shown in Figure 4. Reactions in lanes 1 and 2 used a $MgCl_2$ concentration of 10 mM, and reactions in lanes 3 and 4 used a $MgCl_2$ concentration of 15 mM. Reactions in lanes 2 and 4 used a primer concentration of 1μM, and reactions in lanes 1 and 3 used 500nM primers. All reactions were run with a primer annealing temperature of 50°C. Although a faint band of the expected size is seen in lane 1, there is an anomalous band present as well. None of the other reaction conditions were successful.

These results suggest that the primers designed from *D. melanogaster* do not work well with DNA from *D. ananassae*. Additional experiments, utilizing different primers, were performed on DNA from *D. takahashii and D. ananassae* (Figure 5). Reactions in lanes 1 and 2 of Figure 5 used primers UTR-1F and ORF-3R and 20ng *D. ananassae* DNA with a 37°C primer annealing temperature. These same reactions also used a dNTP concentration of 400μM and a $MgCl_2$ concentration of 15mM. Reactions represented in lanes 3 and 4 used the primers 5′-UTR-1F and ORF-3R, the reactions in lanes 5 and 6 used the primers ORF-3F and ORF-5R.

The results in Figure 5 show that the primer combinations 5′-UTR-1F and ORF-3R (lanes 1-4) and ORF-3F and ORF-5R (lanes 5 and 6) amplify DNA fragments similar to the sizes expected. However, due to the less stringent reaction conditions used here, numerous additional bands are also seen.

Thus, these reactions are not appropriate for cloning the *HeT-A* fragments. Since much time was spent exploring various PCR conditions with no success, it was decided to try DNA from a third species, *D. teissieri*. Among these three species, *D. teissieri* is more closely related to *D. melanogaster* than either *D. ananassae* or *D. takahashii* (Bock 1980). Thus, we would expect the *HeT-A* genes from this species to be more similar to those from *D. melanogaster*.

After trying a number of conditions (data not shown), amplification with *D. teissieri* DNA was found to be optimal with a $MgCl_2$ concentration of 15mM, 20 ng template DNA, the primer combination 5′-UTR-1F and ORF-3R, and a 37°C primer annealing temperature. As seen in Figure 6, lane 1, a single intense band of approximately 1600 bp is produced under these conditions. This fragment is only slightly smaller than the 1730 bp DNA fragment predicted from the *HeT-A* sequence from *D. melanogaster*.

The fact that the PCR fragment amplified from *D. teissieri* is only 1600 bp suggests that some of the nucleotides have been deleted in the *HeT-A* gene from this species. To verify this and to examine the molecular differentiation between the *HeT-A* genes from *D. teissieri* and *D. melanogaster*, this fragment was subsequently isolated from the gel and cloned into the vector PCR-TOPO 2.1. Following transformation, six white colonies were used to inoculate LB broth and cultures were grown overnight. Following DNA isolation, a plasmid carrying an insert of the correct size (1600 bp) was prepared for automated DNA sequencing.

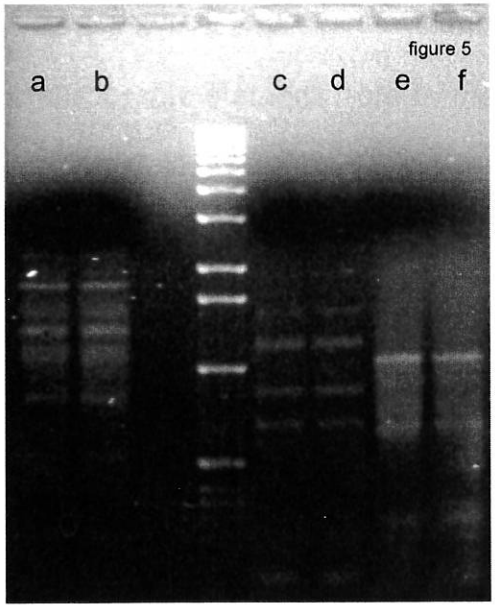


Figure 5. PCR results for *D. ananassae* (lanes a and b) and *D. takahashii* (lanes c-f). The primer combination 5′-UTR-1F and ORF-3R was used for reactions in lanes a-d. The primers ORF-3F and ORF-5R were used in lanes e and f. The 1 kb size standard is shown in lane 4.

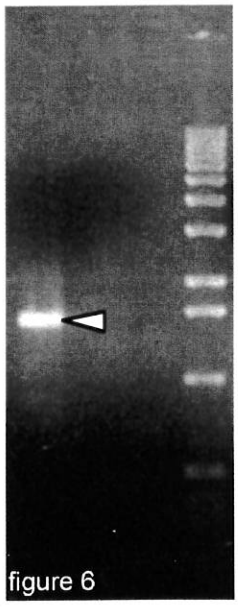
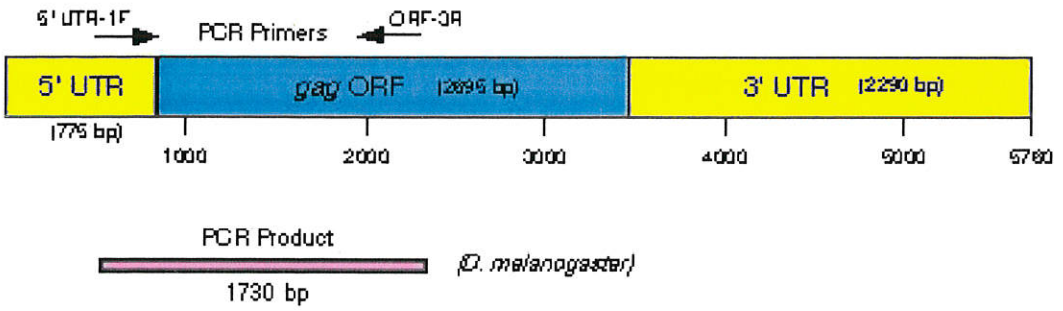


Figure 6. Results of PCR from *D. teissieri*. The arrowhead shows an intense discrete band of 1600 bp in lane 1. The 1kb size standard is shown at the far right.

Figure 7 is a schematic representation of a coding region of the *HeT-A* transposable element from *D. melanogaster*. Relative locations of primers 5′-UTR-1F and ORF-3R are shown in relation to the open reading frame (ORF) of the gag protein. *Gag* genes compose a significant amount of the coding region of retrotransposons and retroviruses and are necessary for active transposition of most replication-competent retroelements and retroviruses (Pardue et al. 1996). The *HeT-A gag* sequence is expected to be somewhat conserved among different species. Although the precise role of the 5′ untranslated region (UTR) is not known, portions of this sequence are also conserved among different species (Danilevskaya et al. 1998a).

Figure 7. Schematic representation of the *HeT-A* element from 5′ *D. melanogaster*. Depicted are the 5′ - untranslated region (UTR) and 3′ untranslated region (3′ UTR) and the open reading frame (ORF) for the gag protein. Also shown is the 1730 bp fragment expected from amplification with primers 5′-UTR-1F and ORF-3R.



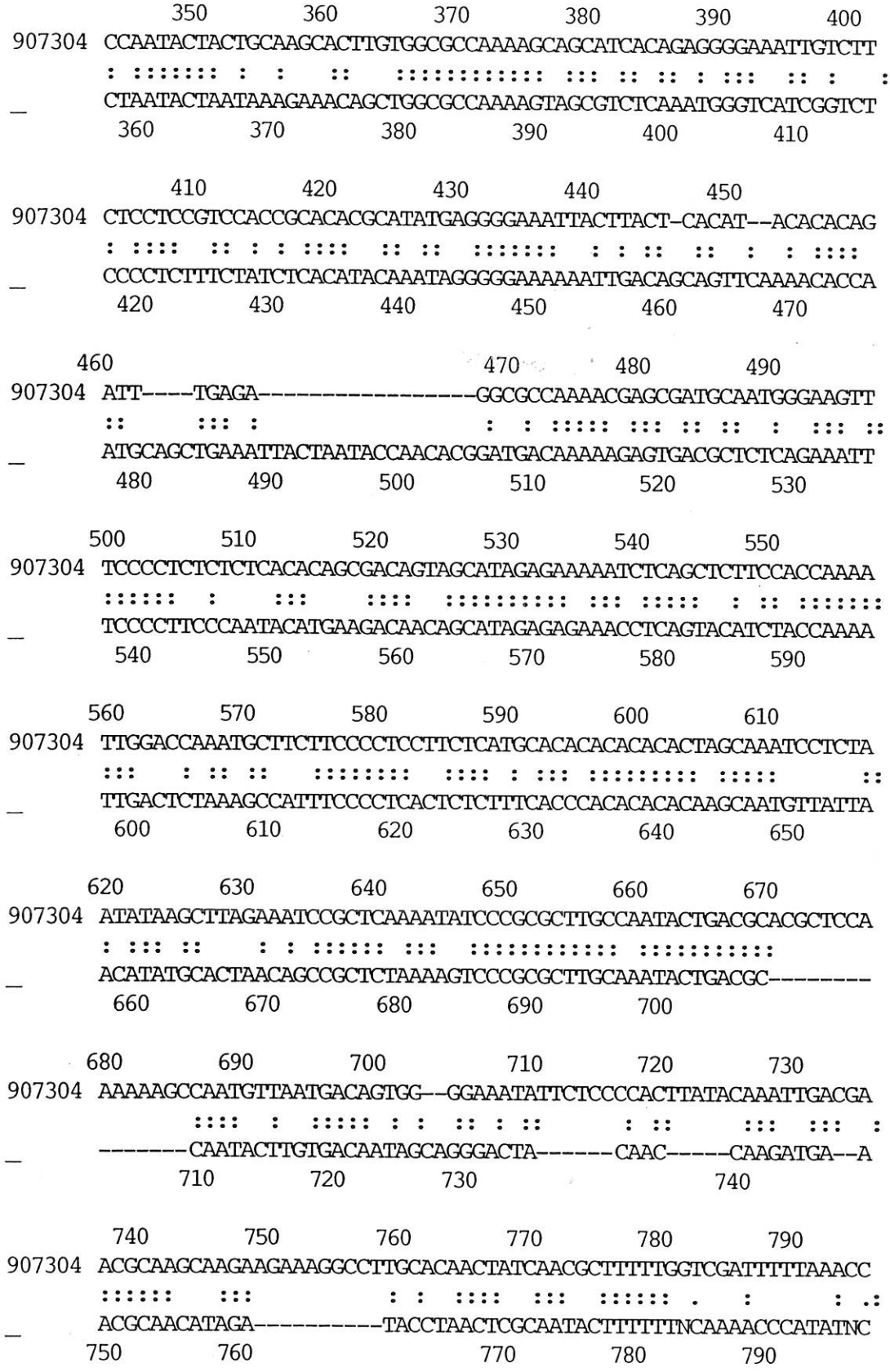
The DNA sequence of the PCR fragment from *D. teissieri* was compared to a *HeT-A* sequence from *D. melanogaster* and revealed a 60.6% nucleotide identity (Figure 8a). The sequence from *D. teissieri* was also compared to a *HeT-A* sequence from *D. yakuba* and was found to be 83.0% identical.(Figure 8b)

Figure 8a. Sequence alignment of *HeT-A* elements of *D. melanogaster* and *D. teissieri*. Sequence from *D. melanogaster* is shown on the top lines and sequence from *D. teissieri* is shown on the bottom lines. Dots between these lines indicate identical sequence between the two species. Dashes within the sequences were introduced by the program to maximize the alignment.

GeneStream align Home Page http://www2.igh.cnrs.fr/bin/align-guess.cgi

Please site: Person, W.R., Wood, T., Zhang, Z., and Miller, W. (1997)
Comparision of DNA sequences with protein sequences, Genomics 46: 24-36

```
>_ D. melanogaster/9D4a                                  1589 nt vs.
>_ D. teissieri/Dt-1                                     1647 nt
scoring matrix: , gap penalties: -12/-2
60.6% identity;          Global alignment score: 1995

                10         20          30          40
907304 AA-GCCAGGGTATTTATACTTTACAA--GTATCGTT--TTAACATA-----------AT
       :: ::::::::::::::::   : ::  :::::::  :: :::::            ::
_      AAAGCCAGGGTATTTATACCAGAAAAAAGTATCGTTCCTTTACATATAGTCAGCAGATAT
               10         20          30          40          50          60

            50          60          70          80          90          100
907304 AAAAACCAACATGTCCATGTCCGACAACCTATTTTCTGACGATGAGGTACTTTCAATTTC
        :   :::: ::      : :: :::   :: : :::  :::: :: :  ::  :  : :
_      TATGGCCAATATAATACTCTCTGAC--GCT-TCTTCAAACGACGAAGCGCTCACCCTATT
               70          80          90          100         110

           110         120         130         140         150         160
907304 CTCAAGCCCAGAACAGCGATCTTCCCCTTTCTACTTAAATATATCGCCCATGTCCAACGG
       ::: :::::::::    : :  : :::::::::::: :  : ::::::::::::::: : :
_      CTCGAGCCCAGAAAGACAAAATACCCCTTTCTACCTGGAAATATCGCCCATGTCCCATGA
       120         130         140         150         160         170

           170         180         190         200         210         220
907304 ATCAGATAATTCTCAGATTAATACTATAATTTTTAATCAGAAGAAATTGCCCTCAAATCA
        ::  : :: :: ::: : ::::  :  ::  ::::   ::: ::: :: : :::::  :
_      TTCTAACAACTCCCAGTTCAATATCAACATAATTAACATGAAAAAACTGTCTTCAAACCC
       180         190         200         210         220         230

           230         240         250         260         270         280
907304 AGCAAACATAAGTCTAAAAAACTCTTCTGGGGCTGCTATAAAAATTGTTAGTTCC-CTTT
        : ::  :  ::::::::::::: :::: : :::::::::::: :: :::  ::: :::
_      TGAAATTAACAGTCTAAAAAACCCTTCCGAGGCTGCTATAAAA-TTCTTAAATCCACTTA
       240         250         260         270         280         290

           290         300         310         320         330         340
907304 CACATAAGAAGAAAGAGAACTCAAACACTAATAATGCCCAAAAAGACCCCCTCTCACTCA
       :: :::::  :::::::::     :: :   : ::::::::::::: :::: ::::: :: :
_      CATATAAGGGGAAAGAGAATAGGAAAAACGAAAATGCCCAAAAAAACCCGCTCTCCCTTA
        300         310         320         330         340         350
```


350 360 370 380 390 400
907304 CCAATACTACTGCAAGCACTTGTGGCGCCAAAAGCAGCATCACAGAGGGGAAATTGTCTT
CTAATACTAATAAAGAAACAGCTGGCGCCAAAAGTAGCGTCTCAAATGGGTCATCGGTCT
360 370 380 390 400 410
410 420 430 440 450
907304 CTCCTCCGTCCACCGCACACGCATATGAGGGGAAATTACTTACT-CACAT--ACACACAG
CCCCTCTTTCTATCTCACATACAAATAGGGGGAAAAAATTGACAGCAGTTCAAAACACCA
420 430 440 450 460 470
460 470 480 490
907304 ATT----TGAGA----------GGCGCCAAAACGAGCGATGCAATGGGAAGTT
ATGCAGCTGAAATTACTAATACCAACACGGATGACAAAAAGAGTGACGCTCTCAGAAATT
480 490 500 510 520 530
500 510 520 530 540 550
907304 TCCCCTCTCTCTCACACAGCGACAGTAGCATAGAGAAAAATCTCAGCTCTTCCACCAAAA
TCCCCTTCCCAATACATGAAGACAACAGCATAGAGAGAAACCTCAGTACATCTACCAAAA
540 550 560 570 580 590
560 570 580 590 600 610
907304 TTGGACCAAATGCTTCTTCCCCTCCTTCTCATGCACACACACACACTAGCAAATCCTCTA
TTGACTCTAAAGCCATTTCCCCTCACTCTCTTTCACCCACACACACAAGCAATGTTATTA
600 610 620 630 640 650
620 630 640 650 660 670
907304 ATATAAGCTTAGAAATCCGCTCAAAATATCCCGCGCTTGCCAATACTGACGCACGCTCCA
ACATATGCACTAACAGCCGCTCTAAAAGTCCCGCGCTTGCAAATACTGACGC--------
660 670 680 690 700
680 690 700 710 720 730
907304 AAAAAGCCAATGTTAATGACAGTGG--GGAAATATTCTCCCCACTTATACAAATTGACGA
-------CAATACTTGTGACAATAGCAGGGACTA------CAAC-----CAAGATGA--A
710 720 730 740
740 750 760 770 780 790
907304 ACGCAAGCAAGAAGAAAGGCCTTGCACAACTATCAACGCTTTTTGGTCGATTTTTAAACC
ACGCAACATAGA----------TACCTAACTCGCAATACTTTTTTNCAAAACCCATATNC
750 760 770 780 790

```
          800         810         820         830          840         850
907304 CAAGCCGGACGTTACTAAACCAAGCCTAAATAGGAAACCCA--CCAATCCCACTAGAAAT
       ::: :: ::: :  . .::  :::  :::: : ::::::    ::: :: : :  :::::
_      CAAACCAGACATCTNCNAATTAAGTTTAAAAACGAAACCTTACCCACTCTCGCG-GAAAT
      800         810         820         830         840         850

           860         870         880         890         900         910
907304 ACTGGGAAAA--AATGCATCTCCCCTCACAAAAAGAGCGCTTCTTTATGCCCTCCCGCTC
        :    ::::  :: :: :::::::.:::    ::  ::: :::: :  : :: ::::::::
_      CCCCAAAAAAGAAAAGCGTCTCCCNTCAACTGAAAGGCGTTTCTATGAGTCCCCCCGCTC
       860         870         880         890         900         910

               920         930         940         950         960
907304 AGGATGATTTAAATTTAAATTTAAACCCCAAATCTAGCGCCAAGCCCACTGTGG---TGA
       :::      ::::::: ::::::::       :::: : : ::: ::  ::: :::   : :
_      AGG-----TTAAATT-AAATTTAA------AATCCACCACCACGCATACTATGGGAATCA
       920                 930                 940         950         960

     970         980         990        1000        1010        1020
907304 ATTTACCAGCTGCCCGCATCCTAAGCCGGCCTGCAGCCAAGCGGGATTTATTTAAATCAT
        ::   :: ::    : : ::: : ::::::::::::::::::::  : :::::::  :
_      CTTCTACATCTA---GAACCCTCAACCGGCCTGCAGCCAAGCGGGACCT-TTTAAAT--T
         970             980         990        1000        1010

    1030        1040        1050        1060        1070        1080
907304 CATCCTCACGAAGCCCAGACGAGCAGCCTATGAGTTTTTCGGAAGTGGTTGCTGGCACGG
       :: :: :    ::: : : ::  :::::::::: :::::::::::::: :: :: :: :
_      CACCCCCCAATAGCACGGGCGCACAGCCTATGAGATTTTCGGAAGTGGTGGCCGGAACCG
    1020        1030        1040        1050        1060        1070

    1090        1100        1110        1120        1130        1140
907304 GTTCAATTTTTGCGGCACCTTGTGCCCCGGCACCTCTAACGAAAACTCCAGGCAAGCGGA
       :: ::    : : :::: : : ::  ::::::::      ::::::: :: :::::::: :
_      GTACAGGACTAGTGGCATCCTCTGATCCGGCACCAACTTCGAAAACCCCGGGCAAGCGCA
    1080        1090        1100        1110        1120        1130

    1150             1160        1170             1180        1190
907304 CAA------ACGACGATCTTGACTGCT----CCAA--CTTTAAGACGCCCAATAAAAAAA
        ::       ::   ::  : ::  :::    ::::  ::  :: ::::: ::::::::
_      AAATTTCCAACATGGACATGGATAGCTTCAGCCAAAGCTACAAAACGCCTAATAAAAAGG
    1140        1150        1160        1170        1180        1190

      1200        1210        1220        1230        1240        1250
907304 TTATGCGCGACTTCCAACTTTGTAACTCCCAGCATTTTTCCGCCGCTCATCACTCCCGTT
        ::::: : : ::::  ::::  :: :: : :  : : :: :: ::    :: ::: :
_      C-ATGCGTGCCCTCCATTTTTGCGACCCCAAACCCTCTCCCCCCCCCTAGCTACCCCCATC
    1200        1210        1220        1230        1240        1250
```

```
       1260        1270        1280        1290        1300        1310
907304 TTCAAGAGCAAGGCAGCTCAATCTGTTTACGAGGAATCCAAAGCCAGAAATGGACCCCCC
       :: :: ::::: :: ::  ::: : ::  ::::::: :::::::              ::
_      TTTAAAAGCAAAGCGGCCAAATGTATTGCCGAGGAAATCAAAGCC-------------CC
     1260        1270        1280        1290        1300

       1320        1330        1340        1350        1360        1370
907304 CCGCCGGCCACCGCCTGCAGCATCAATGCCTCTGCTCGCAGCGCAGCG-CGCCACCCGGA
       ::::   :::  : :   ::  : :: ::: :: : ::::::::: :: ::: : : :
_      CCGCTACCCA--GTCGAGAGTGTAAAGGCCACTACACGCAGCGCACCGGCGCAAACTGAT
        1310        1320        1330        1340        1350        1360

       1380        1390        1400        1410        1420        1430
907304 ATTGCCCCCCTACCCCCTCAAAACACAGATACAGAGCTGCCTCCATGGAAAATCGTGCCC
       ::::::::::   :: :: ::::  : : ::: :: ::::: :: ::: :::: :::::
_      ATTGCCCCCC---CCTCTAAAAATTCGGCTACTGAACTGCCCCCGTGGCAAATAGTGCCA
          1370        1380        1390        1400        1410        1420

       1440        1450        1460        1470        1480        1490
907304 CAGAGCCGTAGAGCACCCCCTATACTCGTCAATGATGTAAAGGAAATTGTACCTCTACTG
       :::::::: :: :: :::::::: :   :  :: :::: : ::::::::: :: ::: ::
_      CAGAGCCGCAGGGCCCCCCCTATTCATATAGATAATGTTAGGGAAATTGTGCCACTATTG
             1430        1440        1450        1460        1470        1480

       1500        1510        1520        1530        1540        1550
907304 GAAAAGCTGAACTACACAGCAGGAGTCTCCAGCTACACTACCAGG---GCTATAGAAGGA
       ::::: :: :: ::: ::::::: ::   ::: : ::: :: :     :::::::   :
_      GAAAAACTAAATTACTCAGCAGGGGTGGACAGTTTCACAACAAAAACAGCTATAG---GT
             1490        1500        1510        1520        1530

          1560        1570        1580
907304 AACGGGGTCAGGATACAGGCAAAGGACATGACCGCC
       :::::  : :  :: ::::: :: :: ::::: ::
_      AACGGCATAACTATTCAGGCCAAAGATATGACAGCA
```

Figure 8b. Sequence alignment of *HeT-A* elements of *D. yakuba* and *D. teissieri*. Sequence from *D. yakuba* is shown on the top lines and sequence from *D. teissieri* is shown on the bottom lines. Dots between these lines indicate identical sequence between the two species. Dashes within the sequences were introduced by the program to maximize the alignment.

GeneStream align Home Page http://www2.igh.cnrs.fr/bin/align-guess.cgi

Please site: Person, W.R., Wood, T., Zhang, Z., and Miller, W. (1997)
Comparision of DNA sequences with protein sequences, Genomics 46: 24-36

>_ D. yakuba 1581 nt vs.
>_ D. teissieri/Dt-1 1647 nt
scoring matrix: , gap penalties: -12/-2
83.0% identity; Global alignment score: 4726

```
                10        20        30        40        50
586874 AAAGCCAGGGTATTTATACCACAAAAA-GTATCGTTCCTTTACATATAGTCAGCAAATAT
       :::::::::::::::::::: ::::: :::::::::::::::::::::::::::: ::::
_      AAAGCCAGGGTATTTATACCAGAAAAAAGTATCGTTCCTTTACATATAGTCAGCAGATAT
                10        20        30        40        50        60

      60        70        80        90       100       110
586874 TATGGCCCAAGTAATTCTCTCTGACGACTCTTCAAATGATGAGGTGCTCTCTCTTTTCTC
       ::::::: : :::: :::::::::: :::::::: :: :: : :::: : :: :::::
_      TATGGCCAATATAATACTCTCTGACGCTTCTTCAAACGACGAAGCGCTCACCCTATTCTC
                70        80        90       100       110       120

     120       130       140       150       160       170
586874 GAGCCCAGAAAGTCAAAATACCCCTTTCTACCTGGAAATCTCGCCCATGTCCCAAAATTC
       :::::::::::: :::::::::::::::::::::::::: ::::::::::::::  ::::
_      GAGCCCAGAAAGACAAAATACCCCTTTCTACCTGGAAATATCGCCCATGTCCCATGATTC
               130       140       150       160       170       180

     180       190       200       210       220       230
586874 TAACAACTCTCAGTTTAATATCAGCATAATAAACATGAAAAAATCGTCTTCCAACTCTGC
       ::::::::: ::::: ::::::: :::::: :::::::::::  :::::: ::: :::
_      TAACAACTCCCAGTTCAATATCAACATAATTAACATGAAAAAACTGTCTTCAAACCCTGA
               190       200       210       220       230       240

     240       250       260       270       280       290
586874 AATTAACAGTCTAAAAAACCCTTCCGAGGCTGCTATAAAAATTATAAATTCACTTACATA
       :::::::::::::::::::::::::::::::::::::::: : ::::: ::::::::::
_      AATTAACAGTCTAAAAAACCCTTCCGAGGCTGCTATAAAATTCTTAAATCCACTTACATA
               250       260       270       280       290       300

     300       310       320       330       340       350
586874 TAAGGGGAAAGAGAATAGAGAAAACAAAAATGCCCAAAAAGACCCGCTCTTCCTTACTAA
       :::::::::::::::::  ::::: :::::::::::::: ::::::::: :::::::::
_      TAAGGGGAAAGAGAATAGGAAAAACGAAAATGCCCAAAAAAACCCGCTCTCCCTTACTAA
               310       320       330       340       350       360
```

```
      360         370         380         390         400         410
586874 TACCAATAAAGAAACAGCTGGCGCCAAAAGTAGCGTCTCAAATGGGCCAGTGGTTTCCCT
       ::: :::::::::::::::::::::::::::::::::::::::::::  ::  ::: ::::
_      TACTAATAAAGAAACAGCTGGCGCCAAAAGTAGCGTCTCAAATGGGTCATCGGTCTCCCC
                370         380         390         400         410         420

      420         430         440         450         460         470
586874 TCTTTCTGTCACACATACAAATAGGGGGAAAAAATTGACAACAGCTCACAACACCAATGC
       ::::::: :: ::::::::::::::::::::::::::::::: ::: ::: ::::::::::
_      TCTTTCTATCTCACATACAAATAGGGGGAAAAAATTGACAGCAGTTCAAAACACCAATGC
                430         440         450         460         470         480

      480         490         500         510         520         530
586874 AGCTGAAACCACTAATACCAACATGGATGACAAAAAGAGCGGCGCTCTTAGAAATTTCCC
       ::::::::  :::::::::::::: ::::::::::::::: : :::::: ::::::::::
_      AGCTGAAATTACTAATACCAACACGGATGACAAAAAGAGTGACGCTCTCAGAAATTTCCC
                490         500         510         520         530         540

      540         550         560         570         580         590
586874 TTTCCCCACACATGAAGACAACAGCATGGAGAGAAACCTCAGCTCATCTACAAAAATTGG
       :::::  : ::::::::::::::::: :::::::::::::  ::::::: :::::::
_      CTTCCCAATACATGAAGACAACAGCATAGAGAGAAACCTCAGTACATCTACCAAAATTGA
                550         560         570         580         590         600

      600         610         620         630         640         650
586874 CTCTAAAAGCATTTCCCCTCACTCTCTCTCACCTACACACACAAGCAAGGTTATTAACAT
       :::::::  ::::::::::::::::: ::::: :::::::::::::  ::::::::::
_      CTCTAAAGCCATTTCCCCTCACTCTCTTTCACCCACACACACAAGCAATGTTATTAACAT
                610         620         630         640         650         660

      660         670         680         690         700         710
586874 AAGCACAAACAGCCGCTCAAAAAGTCCCGCGCTTGCAAATGCTGACACACTACATAAACT
       : :::: :::::::::::  ::::::::::::::::::::::  :::::
_      ATGCACTAACAGCCGCTCTAAAAGTCCCGCGCTTGCAAATACTGAC--------------
                670         680         690         700

      720         730         740         750         760         770
586874 AGCCAATATATATGACAATAGCAGGGACCACAACCAAGGTGAAACACAACATAAATTTAT
       :::::::  : ::::::::::::::: ::::::::: :::::: ::::::: ::    :
_      -GCCAATACTTGTGACAATAGCAGGGACTACAACCAAGATGAAACGCAACATAGATACCT
         710         720         730         740         750         760

      780         790         800         810         820         830
586874 AACTCGCAATACTTTTTTCCAAAACTTGTATCCTAAACCTGACATTTCCAAACTAAGCTT
       ::::::::::::::::::.::::::    :::.: ::::: ::::: :.:.:: :::: ::
_      AACTCGCAATACTTTTTTNCAAAACCCATATNCCAAACCAGACATCTNCNAATTAAGTTT
         770         780         790         800         810         820
```

```
        840         850         860         870         880         890
586874  AAAAAATAAGGCCA--CCATTCTCGCGAAAA-CTACAAAAAATAAATGCATCTCCCCCCA
        :::::  ::  :    ::: ::::::: ::: : ::::::: ::: :: ::::::. ::
_       AAAAACGAAACCTTACCCACTCTCGCGGAAATCCCCAAAAAAGAAAAGCGTCTCCCNTCA
          830         840         850         860         870         880

          900         910         920         930         940         950
586874  ACTGAAAGGCGCTTCTTTGTGTTCCCCTGTTCAGCCTAATTTAAATTTCAAAGTCACCAC
        ::::::::::: :::: :: :: :::: : ::::  ::: :::::::: ::: ::::::
_       ACTGAAAGGCGTTTCTATGAGTCCCCCCCGCTCAGGTTAAATTAAATTTAAAATCCACCAC
          890         900         910         920         930         940

          960         970         980         990        1000        1010
586874  TACACACTCTATGGGTAACACATCTGCATCCAGAACCCTAAACCGGCCTGCAGCCAAGCG
        :: ::  ::::::: : ::: ::: :::: :::::::: ::::::::::::::::::::
_       CACGCATACTATGGGAATCACTTCTACATCTAGAACCCTCAACCGGCCTGCAGCCAAGCG
          950         960         970         980         990        1000

        1020        1030        1040        1050        1060        1070
586874  GGACCTTTTTAATTCACCCCCCAATAGCACGAACGCACTGCCTATGAGTTTTTCGGAAGT
        ::::::::: :::::::::::::::::::::  ::::: ::::::::: ::::::::::
_       GGACCTTTTAAATTCACCCCCCAATAGCACGGGCGCACAGCCTATGAGATTTTCGGAAGT
          1010        1020        1030        1040        1050        1060

        1080        1090        1100        1110        1120        1130
586874  GGTGGCCGGAACCGGTCCAGGAATTGTGGCATCCTCTGATCCGGCACCAATTACGAAAAC
        ::::::::::::::: ::::: : ::::::::::::::::::::::::: : :::::::
_       GGTGGCCGGAACCGGTACAGGACTAGTGGCATCCTCTGATCCGGCACCAACTTCGAAAAC
          1070        1080        1090        1100        1110        1120

        1140        1150        1160        1170        1180              1190
586874  CCCGGGCAAGCGCACAAATACCAACATGGACATGGATAGCTTTAGC------TACAAAAC
        :::::::::::::: :: : :::::::::::::::::::::: :::       ::::::::
_       CCCGGGCAAGCGCAAAATTTCCAACATGGACATGGATAGCTTCAGCCAAAGCTACAAAAC
          1130        1140        1150        1160        1170        1180

              1200        1210        1220        1230        1240        1250
586874  GCCCAATAAAAAAGCATGTGTGCCCACCAATTTTGCGACCCCAAACCCTTTCCCCCCCCCT
        ::: :::::::: ::::: :::::: ::: :::::::::::::::::: ::::::::::
_       GCCTAATAAAAAGGCATGCGTGCCCTCCATTTTTGCGACCCCAAACCCTCTCCCCCCCCT
          1190        1200        1210        1220        1230        1240

              1260        1270        1280        1290        1300        1310
586874  AGCCACCCCCCATCTTTAAAAGTAAGGCGGCCAAAAGTATTGTCGAGGAAATAAAAGCCCC
        ::: :::::::::::::::: :: ::::::::: :::::: ::::::::: ::::::::
_       AGCTACCCCCCATCTTTAAAAGCAAAGCGGCCAAATGTATTGCCGAGGAAATCAAAGCCCC
          1250        1260        1270        1280        1290        1300
```

```
                1320        1330        1340        1350        1360        1370
586874  CCGCCACCCAGTCGAGAGTGAAAAGGCCTCTGCACGCAGCGTACCGGCGCAAACTGAAAT
        :::: ::::::::::::::: ::::::: :: ::::::::: :::::::::::::::: ::
_       CCGCTACCCAGTCGAGAGTGTAAAGGCCACTACACGCAGCGCACCGGCGCAAACTGATAT
         1310        1320        1330        1340        1350        1360

                1380        1390        1400        1410        1420        1430
586874  TGCCCCCCCCTCCCCCTAAAAATTCGGCTACCGAACTGCCCCCGTGGCAAATAGTTCCACA
        ::::::::: :  ::::::::::::::: :::::::::::::::::::: :::::
_       TGCCCCCCC---CTCTAAAAATTCGGCTACTGAACTGCCCCCGTGGCAAATAGTGCCACA
         1370        1380        1390        1400        1410        1420

                1440        1450        1460        1470        1480        1490
586874  GAGCCGCAGGGCCCCCCCTATTCATATAAAAAATGTTAGGGAAATCGTGCCACTATTGGA
        :::::::::::::::::::::::::: : :::::::::::::: ::::::::::::::
_       GAGCCGCAGGGCCCCCCCTATTCATATAGATAATGTTAGGGAAATTGTGCCACTATTGGA
             1430        1440        1450        1460        1470        1480

                1500        1510        1520        1530        1540        1550
586874  AAAGCTAAATTACTCAGCAGGGGTAGACAGTTTTACAACAAAAACGTCTATAGGAAACGG
        ::: :::::::::::::::::: :::::::: :::::::::::  ::::::: :::::
_       AAAACTAAATTACTCAGCAGGGGTGGACAGTTTCACAACAAAAACAGCTATAGGTAACGG
             1490        1500        1510        1520        1530        1540

                1560        1570        1580
586874  TGTAACTATACAGGCTAAAGACTTGACTGCA
          ::::::: ::::: ::::: :::: :::
_       CATAACTATTCAGGCCAAAGATATGACAGCA
             1550        1560        1570
```

Discussion

HeT-A is a member of the non-LTR family of retrotransposons; DNA sequences that excise themselves and reinsert in the genome utilizing an RNA intermediate (Pardue et al. 1996). These mobile DNA sequences are unique, however, in that they insert exclusively at the ends of the *Drosophila* chromosomes. The central open reading frame of the *HeT-A* sequence codes for the Gag protein, which is similar in sequence to the retroviral Gag (group-specific-antigen) proteins (see Figure 7). The *gag* gene encodes several polypeptides that form the capsid of the virus. Since retrotransposons are not infectious, they do not possess capsid proteins. Here the Gag proteins are thought to be involved in forming a transposition complex that is essential for the mobility of the retrotransposon (Pardue et al. 1996). Sequence comparisons indicate that the *gag* genes of both retroviruses and retrotransposons are evolving relatively rapidly. In the case of retroviruses, this provides an evolutionary advantage as the viruses try to evade the host's immune system. It is unclear why the *gag* gene of retrotransposons also evolves rapidly, since they are not targets of the host's immune system. However, this may be related to their mechanism of transposition (Pardue et al. 1996).

Telomeres are repeated segments of DNA at the chromosome ends of eukaryotic organisms. They serve as an important buffer for the internal portions of chromosomes, which codes for vital genetic information. Some of the telomere is lost during each cell division, and it must be repaired if the cell

is to continue to divide. Most eukaryotic organisms utilize telomerase to perform this function. Telomerase adds a short repetitive DNA sequence to the ends of each chromosome. By repeatedly transposing itself within the telomeric region of the chromosome, *HeT-A* performs this important function for *D. melanogaster* (Biessmann et al. 1992). This latter mechanism of telomere maintenance is unique among eukaryotic organisms.

Prior to this study, *HeT-A* DNA sequences were available from only two species, *D. melanogaster* (Danilevskaya et al. 1994) and *D. yakuba* (Danilevskaya et al. 1998b). Like *D. teissieri*, these two species belong to the *melanogaster* species group of the genus *Drosophila*. This study demonstrates that the *HeT-A* element from *D. teissieri* is clearly similar to the sequence of the *HeT-A* elements from *D. melanogaster* and *D. yakuba*. Thus, it is now known that *D. teissieri*, a third species of *Drosophila*, contains the *HeT-A* gene in its genome. This study represents the first time that the *HeT-A* sequence has been isolated using PCR. In the case of *D. melanogaster* and *D. yakuba*, the *HeT-A* sequences were obtained by screening a genomic library with a *HeT-A* probe, a time-consuming task. Thus, the innovation described here may enhance the search for *HeT-A* in other species of *Drosophila*.

Polymerase Chain Reaction

PCR, while fairly forgiving in some aspects, such as the amount of template DNA, is a fastidious technique. A slight alteration of even a single variable in the reaction can lead to a dramatic alteration in the binding

specificity of the primers and thus can affect the success of the reaction. Many of the figures shown above have multiple DNA bands in a single lane, which is the result of non-specific binding of the primers. Primers are designed to have a specific melting temperature (Tm) which is an important consideration for the heat cycling of PCR. During PCR, an annealing temperature that is less than or equal to the Tm of the primers must be used. It is only after the primers have annealed to the denatured template DNA that polymerization of the DNA between the primers can be initiated, producing a PCR fragment.

$MgCl_2$ seemed to be the variable with the most pronounced effect on the PCR conditions described here. In general, $MgCl_2$ concentration is inversely proportional to primer annealing specificity (Innis and Gelfand 1990). Examples of this are illustrated in many of the experiments described here in which multiple DNA bands appear under sub-optimal $MgCl_2$ concentrations. It was determined here that the optimal $MgCl_2$ concentration for amplification of the *HeT-A* gene sequence from *D. teissieri* was 15 mM, which is considerably higher than the 5mM usually used (see Table 1).

Sequence differentiation of HeT-A elements in Drosophila

Prior to the onset of this project, there was a great deal of uncertainty associated with amplifying the *HeT-A* element from *D. teissieri*. Transposable elements evolve rapidly because there are so many copies in a single genome and because, if active, they tend to accumulate mutations in their sequence during transposition (Xiong and Eickbush 1990). Non-mobile genes evolve much

more slowly. For example, it was relatively straightforward to amplify the rRNA genes from all species of *Drosophila* examined here (Figure 2). This illustrates how conserved these genes are among different species.

The primers for the *HeT-A* element were designed based on the sequences of *HeT-A* elements from the genomes of *D. melanogaster* and *D. yakuba,* sequences that differ from each other by about 45% at the nucleotide level (Danilevskaya et al. 1998b). Because of this extreme sequence divergence, it is not possible to predict if the primers designed will recognize the *HeT-A* element in other species. If the *HeT-A* elements in another species was sufficiently diverged from the *HeT-A* elements in *D. melanogaster* and *D. yakuba,* then the primers would not amplify the *HeT-A* element in that species. When this is coupled with the variables associated with PCR conditions, finding *HeT-A* in other species becomes a difficult task. In this study, *HeT-A* was detected in only one species, *D. teissieri*. PCR was not successful with *D. ananassae* or *D. takahashii* using the same primer combinations. Thus it will be necessary to design new primers for detecting *HeT-A* in these other species.

There are hundreds of species of flies in the genus *Drosophila* and, due to their importance to the study of genetics, evolutionary relationships among the species have been studied carefully. Flies are placed into species groups, which consist of species that are obviously related based on morphology and biogeography. Within each species group, flies are further classified based on morphology and studies of reproductive isolation and are placed into species

subgroups. *D. melanogaster, D. yakuba* and *D. teissieri* belong to the melanogaster subgroup (Bock 1980). *D. ananassae* belongs to the *melanogaster* species group and *ananassae* subgroup, and *D. takahashii* belongs to the *melanogaster* species group and the *takahashii* subgroup (Jeffs et al. 1994). Based on the results described here, the *HeT-A* primers used may only work for species in the *melanogaster* subgroup.

The sequence comparison of the *HeT-A* elements from *D. melanogaster, D. teissieri and D. yakuba* yielded interesting results. Phylogenetically, we expect little or no difference among the *HeT-A* elements from these three species because they all belong to the same subgroup. However, the results presented here show that the *HeT-A* sequence from *D. yakuba* is more closely related to that from *D. teissieri* than either is to *D. melanogaster*. Such unexpected findings are common when examining the phylogeny of transposable elements. These sequences are typically present in hundreds of copies in the genome, and different copies may have different evolutionary histories (Clark and Kidwell 1997). When PCR is used to amplify a particular *HeT-A* gene in a species, that sequence represents only one of many sequences of the genome. In order to obtain a more comprehensive picture of *HeT-A* evolution, it will be necessary to amplify and sequence multiple copies of the *HeT-A* gene from different species.

Future Research

In order to examine the evolutionary history of *HeT-A* in *Drosophila*, multiple clones of the *HeT-A* element must be obtained from different species. More *D. teissieri* sequences may be obtained using the same PCR conditions outlined here. In addition, having the *HeT-A* sequence from a third species of *Drosophila* will help in designing new primers for use in PCR with other species. PCR conditions for those new primers will then need to be optimized so they can be used to amplify *HeT-A* in these species. Once this is done, the PCR fragments can be cloned and the resulting sequences compared to one another. In this way, it will be possible to sample a large number of *HeT-A* sequences in order to examine their evolutionary history. Ultimately, it may be possible to determine how this unusual telomere function arose in *Drosophila*.

Further research may also elucidate whether or not the *HeT-A* transposable element performs the same telomere maintenance function in other *Drosophila* species as it does in *D. melanogaster*. This can be accomplished by *in situ* hybridization to polytene chromosomes. In *D. melanogaster,* such experiments show that *HeT-A* is confined to the ends of the chromosomes. This type of experimentation should be extended to other species, like *D. teissieri*.

This study is currently unable to draw conclusions regarding the diversity *HeT-A* elements within the genome of a particular species. Recall that there are multiple copies of these elements within a single genome. Transposable elements can be seen as populations of genes within a species. Transposable

elements tend to move actively through a species initially but, over time, may become inactive, leaving a veritable graveyard of DNA sequences (Pennisi,1998). When PCR is used to amplify these segments, it is detecting not only the transposable elements that are active, but also those segments that are inactive. The human genome project has revealed that only a small fraction, about 5%, of the genome codes for proteins (Lander et al. 2001). This value is similar to that of other eukaryotes. One explanation for such a large percentage of non-coding DNA in many eukaryotic genomes is the ebb and flow of transposable elements (Pennisi,1998). Though the evolutionary constraints of transposable elements are still unknown, it is clear that they have been coexisting with eukaryotic organisms for millions of years.

Acknowledgements

Dr. Jonathan B. Clark, Primary Investigator

For his exemplary instruction, insight, expertise, technical advice and patience in all aspects of this research endeavor.

Thesis committee members and advisors

Thanks to Dr. Ron A. Meyers, Dr. Gloria Z. Wurst and Dr. Samuel I. Zeveloff for all the support and much enjoyed intellectual interaction.

Weber State University and the Department of Zoology

Thank you for a fantastic undergraduate experience.

References

Barinaga, M. 1997. The Telomerase Picture Fills In. Science 276:528-529.

Biessmann, H., Valgeirsdottir K., Lofsky A., Chin C., Ginther B., Levis R.W., Pardue M.L. 1992. HeT-A, a transposable element specifically Involved in "healing" broken chromosome ends in *Drosophila melanogaster*. Molec. Cell Biol 12:3910-3918.

Bock, I. 1980. The current status of the *Drosophila melanogaster* species group. Syst. Entomol. 5:341-356.

Clark, J.B. and M.G. Kidwell. 1997. A phylogenetic perspective on P transposable element evolution in *Drosophila*. Proc. Natl. Acad. Sci. USA. 94:11428-11433.

Danilevskaya, O.N., D.A. Petrov, M.N. Pavlova, A. Koga, E.V. Kurenova, D.L. Hartl. 1992. A repetitive DNA element, associated with telomeric sequences in *Drosophila melanogaster*, contains open reading frames. Chromosoma 102:32-40

Danilevskaya, O.N., F. Slot, M. Pavlova, and M. L. Pardue. 1994. Structure and function of the *Drosophila HeT-A* transposon: A retrotransposon like element forming telomeres. Chromosoma 103:215-224

Danilevskaya, O.N., K. Lowch haupt, and M.L. Pardue. 1998a. Conserved subfamilies of the Drosophila HeT-A telomeric-specific retrotransposon. Genetics 148:233-242.

Danilevskaya, O.N., C. J. Wung, M. Aibhai, and M.L. Pardue. 1998b. Unusual features of the *Drosophila melanogaster* telomere transposable element *HeT-A* are conserved in *D. yakuba* telomere elements. Proc. Natl. Acad. Sci. USA 95: 3770-3775.

Innis, M.A. and D.H. Gelfand. 1990. Optimization of PCR. pp. 3-12 in: *PCR Protocols.* M.A. Innis et al., eds. Academic Press, San Diego, CA.

Jeffs, P.S., E.C. Holmes, and M. Ashburner. 1994. The molecular evolution of the alcohol dehydrogenase gene in the *Drosophila melanogaster* species group. Molec. Biol. Evol. 11:287-304.

Lander, E.S. et al. 2001. Initial sequencing and analysis of the human genome. Nature 409:860-921.

Meyerson, M., C.M. Counter, E.N. Eaton, L.W. Ellison, P. Steiner, S.D. Caddle, L. Ziaugra, R.L. Beijersbergen, M.J. Davidoff, Q.Y. Liu, S. Baccetti, D.A. Haber, R.A. Weinberg. 1997. hEST2, the putative human telomerase subunit gene, is up-regulated in tumor cells and during immortalization. Cell, 90:785-95.

Pardue, M-L, O.N Danilevskaya, K. Lowenhaupt, J. Wong, K. Erby. 1996. The *gag* coding region of the *Drosophila* telomeric retrotransposon, *HeT-A*, has an internal frame shift and a length polymorphic region. J Mol Evol 43:572-583

Pennisi, E. 1998. How the genome readies itself for evolution. Science 281:1131-1134.

Person, W.R., T. Wood, Z. Zhang, and W. Miller. 1997. Comparison of DNA sequences with protein sequences. Genomics 46:24-36

Xiong, Y. and Eickbush, T.H. 1990. Origin and evolution of retroelements based upon their reverse transcriptase sequences. EMBO J. 9:3353-3362.