Nucleotide sequence of a mutant eukaryotic gene: The yeast tyrosine-inserting ochre suppressor *SUP4-o*

(transfer RNA/eukaryotic DNA/intervening sequences)

HOWARD M. GOODMAN*, MAYNARD V. OLSON, AND BENJAMIN D. HALL

Genetics Department, SK-50, University of Washington, Seattle, Washington 98195

ABSTRACT   One of the eight endonuclease *Eco*RI fragments of yeast DNA that hybridize to yeast tRNA^{Tyr} has been identified with the genetically defined nonsense-suppressor locus *SUP4*. This identification was achieved by analyzing the meiotic linkage between the genetic determinant for the *SUP4* phenotype and that for an electrophoretic variant of the *Eco*RI fragment. The *SUP4* gene was then cloned from an ochre-suppressing yeast strain and analyzed by DNA sequencing. A wild-type *SUP4* gene and two other genetically unidentified tRNA^{Tyr} genes were also sequenced. The sequence of the ochre suppressor differs from that of the wild-type genes by virtue of a G·C → T·A transversion in the base pair that codes for the wobble position base of the tRNA^{Tyr} anticodon. All four genes contain, immediately to the 3' side of the anticodon triplet, a 14 base pair tract that is not present in mature tRNA^{Tyr}. Although the four genes, which represent three unlinked chromosomal loci, all encode the same mature tRNA sequence, there is virtually no observable sequence homology between the three loci in the region preceding the 5' end of the mature tRNA^{Tyr} sequences.

The yeast tyrosine transfer RNA genes provide a eukaryotic gene system that is usually amenable to both genetic and biochemical analysis. There are eight unlinked genetic loci, each of which can mutate to a tyrosine-inserting nonsense suppressor (1–3); correspondingly, eight different endonuclease *Eco*RI fragments of yeast DNA hybridize to tRNA^{Tyr} (4). Because only one tRNA^{Tyr} sequence has been detected in wild-type yeast cells (5), the eight genes are presumed to be identical in the region that encodes the mature tRNA sequence.

We have now identified one of the eight *Eco*RI fragments with the particular tyrosine-inserting suppressor locus *SUP4*, and this fragment has been cloned from both *sup4*$^+$ and *SUP4-o* yeast strains. The ochre-suppressing mutation involves a G·C → T·A transversion in the base-pair coding for the first base in the anticodon triplet. Our DNA sequencing data also reveal an unexpected feature of the yeast tRNA^{Tyr} genes—a tract of 14 base pairs, adjacent to the 3' end of the anticodon triplet, which does not appear in mature tRNA^{Tyr}. The identical sequence occurs in both the wild-type and mutant *SUP4* genes, and, with the exception of one altered base pair, in two other genetically unidentified tRNA^{Tyr} genes that were also sequenced.

MATERIALS AND METHODS

Yeast Strains. The *SUP4* strains used in this work (J15-13C and J15-8D) were provided by J. Kurjan. They are haploid segregants from the diploid strain 164RW87, which is heterozygous for a spontaneously derived *SUP4-o* mutation; 164RW87 was isolated by Rothstein *et al.* (6). D311-3A was obtained from Sherman *et al.* (7). The bacteriophage λ clones containing wild-type tRNA^{Tyr} genes were derived from the

yeast strains B596 [which is in the same genetic background as D311-3A, and was also obtained from Sherman *et al.* (7)] and S288C [obtained from G. Fink, originally isolated by R. K. Mortimer (8)]; the isolation of these recombinant λ will be described elsewhere.†

Cloning. The λ cloning was carried out essentially as described (9, 10). The transfection plates, containing 10^3 plaques per plate, were screened directly by the method of Benton and Davis (11). Nick-translation of plasmid DNA for hybridization to these filters was done using Boehringer DNA polymerase (grade I) without any DNase pre-treatment of the plasmid DNA (12). Two early batches of nick-translated hybridization probe were kindly prepared for us by M. Smith. Hybridization was carried out for 12 hr at 65° in 6 × SCP (20 × SCP is 2.0 M NaCl/0.6 M Na$_2$HPO$_4$/0.02 M EDTA, adjusted to pH 6.2 with HCl), 1% sodium dodecyl sulfate; filters were wetted with a minimal volume of solution containing 10^6 cpm/ml of nick-translated probe with a specific activity of 5 × 10^6 cpm/μg. Positive plaques were picked from these plates, and the λ strains were re-tested and plaque-purified before their DNA was characterized.

The yeast *Eco*RI fragments cloned in λgt were cleaved with restriction nuclease *Hin*dIII and the sub-fragments containing tRNA^{Tyr} genes were recloned in plasmid pBR322 (13). Plasmid DNA was digested with either *Hin*dIII or *Hin*dIII + *Eco*RI endonucleases and mixed with a corresponding digest of λ DNA. When fragments were being inserted into the *Hin*dIII site of pBR322, the linear plasmid DNA was treated with alkaline phosphatase to prevent ligation of the vector DNA in the absence of an insert (14). Ligations were carried out overnight at 15° using T4 DNA ligase; typical reaction mixtures contained 100 ng of plasmid DNA and 500–1000 ng of λ DNA in 50 μl of 66 mM Tris·HCl, pH 7.6/6.6 mM MgCl$_2$/10 mM dithiothreitol/1 mM ATP.

Colony screening was performed using filters prepared as described by Beckmann *et al.* (15) from ampicillin plates containing grids of transformants set up with toothpicks. The filters were hybridized to ^{125}I-labeled tRNA^{Tyr}. Positive colonies were replated and single colonies were retested before plasmid DNA was prepared.

Biological experiments involving recombinant DNA were carried out under P2-EK1 containment conditions.

DNA Sequencing. DNA sequencing was carried out by the method of Maxam and Gilbert (16) on genes that had been recloned from λ into pBR322. All the plasmids that were used in sequencing contained a single *Xma* I site, which was located, as predicted from the known yeast tRNA^{Tyr} sequence, near the

Abbreviations: *SUP4-o*, ochre-suppressing allele of *SUP4*; kb, kilobase (1000 base pairs).
* Permanent address: Department of Biochemistry and Biophysics, University of California, San Francisco, CA 94143.
† M. V. Olson, J. R. Cameron, R. W. Davis, and B. D. Hall, unpublished data.
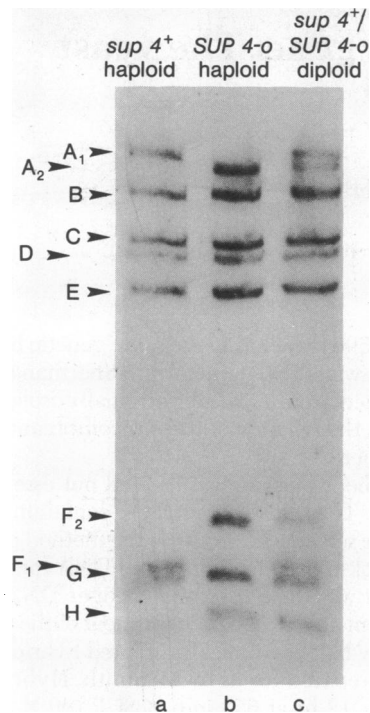
FIG. 1.   The pattern of bands that results when the yeast *Eco*RI fragments are fractionated by size and hybridized to tRNA^{Tyr}. Total yeast DNA was cleaved with *Eco*RI and fractionated on an agarose slab gel; smaller fragments are at the bottom of the figure. The DNA was transferred to a nitrocellulose sheet as described by Southern (17) and hybridized to ^{125}I-labeled tRNA^{Tyr}. Details of the procedure are being published separately;[†] see also ref. 4. The yeast strains are: lane a, D311-3A; b, J15-13C; c, the diploid D311-3A × J15-13C.

3' end of the cloned tRNA^{Tyr} gene. The *Xma* I was the generous gift of Louis DeGennaro and David Gelfand. After *Xma* I cleavage and treatment with alkaline phosphatase, the linear DNA was end-labeled by incubation with polynucleotide kinase and [γ-^{32}P]ATP. The labeled DNA was then digested with an additional restriction enzyme that was chosen to give two labeled fragments that could be readily separated on a 6% polyacrylamide gel. The separated, end-labeled fragments were electroeluted from the sliced gel, and sequenced by the published method (16).

RESULTS

Identification of a Specific tRNA^{Tyr}-Hybridizing *Eco*RI Fragment with the *SUP4* Locus. To identify a particular yeast *Eco*RI fragment with the specific tyrosine-inserting suppressor locus *SUP4*, we exploited the natural strain-dependent variation that occurs in the sizes of some of the hybridizing fragments. Labeled yeast tRNA^{Tyr} hybridizes to *Eco*RI fragments of eight different molecular weights in each of two haploid yeast strains, one of which carries the *SUP4-o* mutation (Fig. 1, lanes a and b). Six of the fragments have the same size in both strains, while two show strain-dependent variations in molecular weight. When the two haploids are crossed, the resulting diploid is heterozygous for the two variable fragments and gives a 10-band pattern (Fig. 1, lane c). Hybridization analysis of a large number of yeast strains has clearly demonstrated that A$_1$ and A$_2$ arise from corresponding chromosomal segments, as do F$_1$ and F$_2$ (M. V. Olson, K. Loughney, and B. D. Hall, unpublished data). The genotype of the diploid, therefore, can be represented as

$$\frac{A_1\ F_1\ sup^+}{A_2\ F_2\ SUP4}.$$

Our tentative identification of the *SUP4* locus with the A band was based on the observation that A$_2$ and *SUP4* cosegregated when this diploid was sporulated. Altogether, eight haploid segregants were analyzed, including one complete tetrad. All three segregating markers A$_1$/A$_2$, F$_1$/F$_2$, and *sup$^+$*/*SUP4* segregated 2:2 in the tetrad, and all eight of the analyzed segregants displayed one of the parental combinations A$_1$*sup$^+$* or A$_2$*SUP4* (data not shown). These results provided a strong inference that the *SUP4* gene was on the A$_2$ fragment. Direct biochemical proof was then sought by cloning the A$_2$ fragment from a *SUP4-o* strain and analyzing the nucleotide sequence of its *tRNA^{Tyr}* gene.

Cloning of the *SUP4-o* Gene in a λ Vector. To clone the A fragment from a *SUP4* strain, we constructed a λ-yeast phage pool by joining *Eco*RI-cleaved DNA from the vector λgt·λB to *Eco*RI-digested yeast DNA from the *SUP4-o* strain J15-8D. Plaques from this pool were screened by hybridization to ^{32}P-labeled DNA from a recombinant plasmid (pYT-A) that was made by transferring a 6.4-kilobase (kb) segment of A$_1$ from an already available λ strain into the *Hin*dIII site of the plasmid pBR322 (13). This probe crossreacted weakly with some of the other *Eco*RI fragments that contain tRNA^{Tyr} genes and with an unidentified fragment of molecular weight, 3.5 × 10^6, but it hybridized strongly only to the A$_2$ fragment of J15-8D. In the first screening experiment, filters were made from six plates containing a total of 6000 plaques. When these filters were hybridized to ^{32}P-labeled pYT-A DNA, one very strong positive was obtained along with several weaker ones. The strong positive arose from a phage strain that had a single *Eco*RI insert that was identical in size to A$_2$ and hybridized to ^{125}I-labeled tRNA^{Tyr} as well as to the nick-translated probe.

DNA Sequencing of the *SUP4-o* and Three Wild-Type tRNA^{Tyr} Genes. To carry out DNA sequencing on this potentially mutant tRNA^{Tyr} gene, it was necessary to locate a restriction site in or near the gene where we could end-label the DNA and initiate a sequencing run. Ideally, the restriction enzyme used for this purpose should make a minimum number of cuts in the neighboring yeast and vector DNA. Inspection of the known sequence of tRNA^{Tyr} (5) allowed the prediction that each tRNA^{Tyr} gene should have an *Xma* I site,

$$\overset{\displaystyle\downarrow}{\underset{\displaystyle\uparrow}{\begin{array}{l}\text{C-C-C-G-G-G}\\ \text{G-G-G-C-C-C}\end{array}}},\tag{18}$$
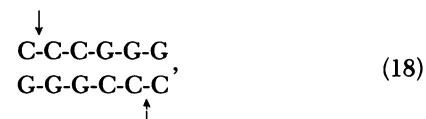
near the 3' end of the gene. Such sites are expected to occur infrequently in yeast DNA, which is (A+T)-rich, and therfore to provide good starting points for DNA sequencing. We were able to optimize the utility of this site by re-cloning the tRNA^{Tyr} genes in pBR322 which, unlike λ, contains no *Xma* I sites.

Each of the plasmids that we used in the sequencing experiments proved to have only the single *Xma* I site that was predicted from the tRNA^{Tyr} sequence. We initiated sequencing runs in both directions from this starting point, which in the transcribed strand is four nucleotides from the start of the C-C-A terminus of the tRNA sequence. The sequence of the transcribed strand was determined, starting two nucleotides from the *Xma* I cut and extending through the gene to a point ~20 base pairs in front of the 5' end of the tRNA sequence. In the other direction, we sequenced the opposite strand from near the *Xma* I cut to a point some 40–60 nucleotides beyond the 3' end of the tRNA sequence. We have inferred the sequence, therefore, of some 150 base pairs for each gene. The results for any particular region depend on the sequencing of a single strand and have not yet been cross-checked by sequencing the second strand. The results for all four of the sequenced genes
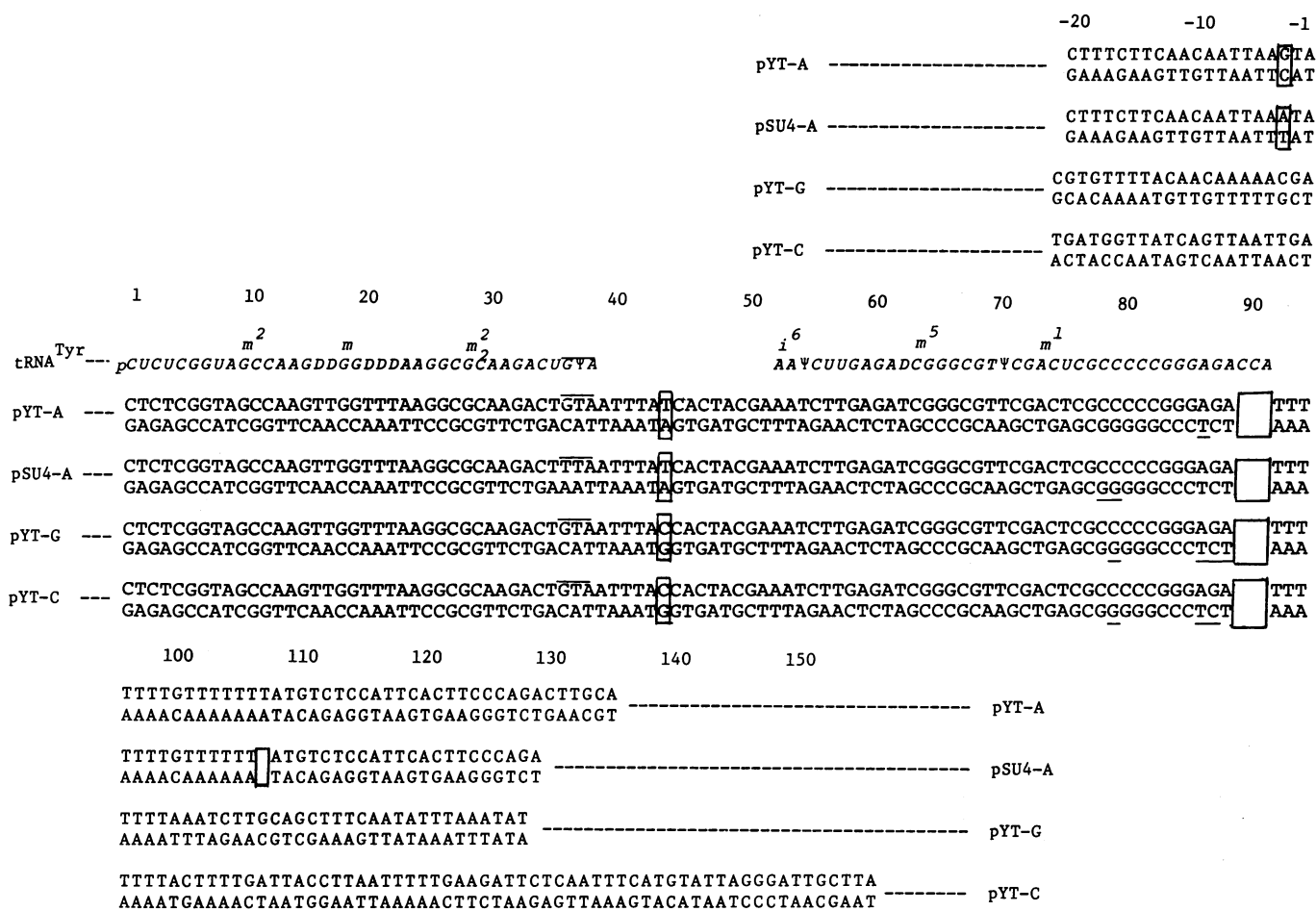
Biochemistry: Goodman *et al.*

Proc. Natl. Acad. Sci. USA 74 (1977)   5455

```
                                                                    -20        -10       -1
                                                         CTTTCTTCAACAATTAACTA
pYT-A    ---------------------                           GAAAGAAGTTGTTAATTCAT

                                                         CTTTCTTCAACAATTAAATA
pSU4-A   ---------------------                           GAAAGAAGTTGTTAATTTAT

                                                         CGTGTTTTACAACAAAAACGA
pYT-G    ---------------------                           GCACAAAATGTTGTTTTGCT

                                                         TGATGGTTATCAGTTAATTGA
pYT-C    ---------------------                           ACTACCAATAGTCAATTAACT
```

```
          1      10      20      30      40        50       60       70       80       90
             m2      m        2                        i6       m5       m1
tRNA^Tyr---  pCUCUCGGUAGCCAAGDDGGDDDAAGGCGCAAGACUGYA       AAYCUUGAGADCGGGCGTYCGACUCGCCCCCGGGAGACCA
```

```
pYT-A  --- CTCTCGGTAGCCAAGTTGGTTTAAGGCGCAAGACTGTAATTTATCACTACGAAATCTTGAGATCGGGCGTTCGACTCGCCCCCGGGAGA[ ]TTT
           GAGAGCCATCGGTTCAACCAAATTCCGCGTTCTGACATTAAATAGTGATGCTTTAGAACTCTAGCCCGCAAGCTGAGCGGGGGCCCTCT[ ]AAA

pSU4-A --- CTCTCGGTAGCCAAGTTGGTTTAAGGCGCAAGACTTTAATTTATCACTACGAAATCTTGAGATCGGGCGTTCGACTCGCCCCCGGGAGA[ ]TTT
           GAGAGCCATCGGTTCAACCAAATTCCGCGTTCTGAAATTAAATAGTGATGCTTTAGAACTCTAGCCCGCAAGCTGAGCGGGGGCCCTCT[ ]AAA

pYT-G  --- CTCTCGGTAGCCAAGTTGGTTTAAGGCGCAAGACTGTAATTTACCACTACGAAATCTTGAGATCGGGCGTTCGACTCGCCCCCGGGAGA[ ]TTT
           GAGAGCCATCGGTTCAACCAAATTCCGCGTTCTGACATTAAATGGTGATGCTTTAGAACTCTAGCCCGCAAGCTGAGCGGGGGCCCTCT[ ]AAA

pYT-C  --- CTCTCGGTAGCCAAGTTGGTTTAAGGCGCAAGACTGTAATTTACCACTACGAAATCTTGAGATCGGGCGTTCGACTCGCCCCCGGGAGA[ ]TTT
           GAGAGCCATCGGTTCAACCAAATTCCGCGTTCTGACATTAAATGGTGATGCTTTAGAACTCTAGCCCGCAAGCTGAGCGGGGGCCCTCT[ ]AAA
```

```
          100       110       120       130       140       150
```

```
TTTTGTTTTTTTATGTCTCCATTCACTTCCCAGACTTGCA
AAAACAAAAAAATACAGAGGTAAGTGAAGGGTCTGAACGT    ------------------------------------    pYT-A

TTTTGTTTTTT[ ]ATGTCTCCATTCACTTCCCAGA
AAAACAAAAAA[ ]TACAGAGGTAAGTGAAGGGTCT        ------------------------------------    pSU4-A

TTTTAAATCTTGCAGCTTTCAATATTTAAATAT
AAAATTTAGAACGTCGAAAGTTATAAATTTATA           ------------------------------------    pYT-G

TTTTACTTTTGATTACCTTAATTTTTGAAGATTCTCAATTTCATGTATTAGGGATTGCTTA
AAAATGAAAACTAATGGAATTAAAAACTTCTAAGAGTTAAAGTACATAATCCCTAACGAAT    --------    pYT-C
```

FIG. 2. The nucleotide sequence of yeast tRNA^Tyr and four tRNA^Tyr genes. The anticodon and its coding triplet are overlined in both the tRNA and the gene sequences. The empty boxes in all four genes at positions 90–92 indicate the absence of a coding triplet for the terminal C-C-A. Also enclosed in boxes are the variable base pair within the 14 base pair "insert" and two points of divergence between the genes on pYT-A and pSU4-A. The position of the insert, as displayed here, is arbitrary in that it could equally well be displaced one base pair in either direction. Because the *Xma* I site (C-C-C-G-G-G at positions 81–87) was the starting point for the sequencing runs, the sequence in this region was inferred, in part, from the known sequence of the *Xma* I recognition site (18). In some cases, there are one or two nucleotides outside the recognition site that have been included here for clarity but were not actually sequenced; these nucleotides are underlined. The four plasmids have the following structures: pYT-A contains a 6.4-kb segment of the A₁ fragment from yeast strain B596 inserted into the *Hind*III site of pBR322; pSU4-A contains a 4.1-kb segment of the A₂ fragment from J15-8D inserted between the *Eco*RI and *Hind*III sites of pBR322; pYT-G contains the entire 1.2-kb G fragment from S288C inserted into the *Eco*RI site of pBR322; pYT-C contains a 3-kb segment of the C fragment from S288C inserted between the *Eco*RI and *Hind*III sites of pBR322. In the sequencing experiments all four plasmids were digested with *Xma* I and the 5' ends were labeled. The two labeled *Xma* I ends were then separated on a 6% acrylamide gel after secondary cleavages with the following enzymes: pYT-A and pSU4-A, *Hpa* II; pYT-G, *Eco*RI; pYT-C, *Alu* I. The larger of the two labeled fragments contained sequences from the 3' end of the tRNA^Tyr genes of pYT-A and pSU4-A and the 5' end of the tRNA^Tyr genes on pYT-G and pYT-C.

are displayed in Fig. 2 along with the sequence of yeast tRNA^Tyr.

The first gene that we sequenced, present in recombinant plasmid pYT-A, was from the *Eco*RI fragment A₁ of a wild-type yeast strain. The DNA sequence confirmed all of Madison and Kung's tRNA^Tyr sequence (5), but contained in addition a tract of 14 base pairs in the middle of the coding region that has no counterpart in mature tRNA^Tyr. This "insertion" occurs immediately to the 3' side of the anticodon triplet. The tRNA sequence in this region is UGΨAA^i⁶ AΨ (anticodon overlined, A^i⁶ is N⁶-isopentenyladenosine); the predicted coding sequence is

$$\begin{array}{ccc} & \downarrow \downarrow \downarrow \\ \text{T-G-T-A-A-A-T} \\ \text{A-C-A-T-T-T-A} \end{array}$$
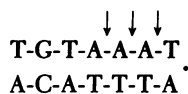
The precise positioning of the extra 14 base pairs within this sequence is arbitrary in that the actual DNA sequence can be derived from the predicted coding sequence by inserting 14 base pair tract at any one of the three positions marked with arrows.

The tRNA sequence contains nine different modified bases whose unmodified precursors can be inferred from the DNA sequence. The 3' terminal C-C-A sequence of the mature tRNA (residues 90–93) is not encoded in the gene. Immediately following the A·T base pair whose transcript would precede the C-C-A in mature tRNA^Tyr, the transcribed strand contains the tract A₇-C-A₇.

The sequence of the tRNA^Tyr gene on the *Eco*RI fragment A₂ from the *SUP4-o* strain confirmed our supposition that the tRNA^Tyr gene on the A fragment is *SUP4*. Fig. 3 shows autoradiograms of sequencing gels for the mutant and wild-type genes. The site of the mutation is the base pair at the 5' end of the anticodon triplet. The wild-type triplet G-T-A has mutated to T-T-A. The mutant gene contains the same 14 base pair "insertion" as the wild-type gene.

Two other tRNA^{Tyr} genes were also sequenced. It was hoped that a comparison of these sequences with those of the genes from the *SUP4* locus would provide further insight into the significance of the *SUP4* insertions, as well as revealing the extent of homology between the various tRNA^{Tyr} loci. Genes were sequenced from the wild-type *Eco*RI fragments C and G in plasmids pYT-C and pYT-G, respectively, neither of which has been identified with a particular suppressor locus. Both genes proved to contain identical 14 base pair insertions whose sequences differed at one base pair from the *SUP4* sequence. The regions of the genes that correspond to the mature tRNA^{Tyr} sequence are identical. Directly adjacent to the 3' end of the tRNA structural sequence, the transcribed strands of both genes, as in the case of pYT-A and pSU4-A, contain tracts of seven As. Following the A$_7$ tracts, there are generally more As, but the detailed sequences differ from case to case. Beyond this short region the C and G genes show no detectable homology either with each other or with the *SUP4* gene.

At their other ends, the three genes have in common a single base pair next to the 5' end of the mature tRNA sequence (Fig. 2, residue −1); beyond that point there is no recognizable homology. Even the sequences from pYT-A and pSU4-A differ in at least three positions in this region (see Fig. 2 and the legend to Fig. 3).

DISCUSSION

The DNA sequencing data on one mutant and three wild-type genes confirm that *Eco*RI fragments A, C, and G contain tRNA^{Tyr} genes and that the gene on the A fragment is *SUP4*. However, all four of these DNA sequences differ from that of yeast tRNA^{Tyr} by virtue of a 14 base pair tract,

A-T-T-T-A-Y-C-A-C-T-A-C-G-A

T-A-A-A-T-R-G-T-G-A-T-G-C-T

(Y = pyrimidine, R = purine), which occurs just to the 3' side of the anticodon triplet. The existence of this intervening DNA sequence‡ within the tRNA^{Tyr} genes is reminiscent of those that have been found recently in several other eukaryotic genes. For some of the 28S rRNA cistrons of *Drosophila melanogaster* (19–21), for mouse β-globin genes (S. M. Tilghman, D. C. Tiemeier, and P. Leder, personal communication), and for the late region of the adenovirus chromosome (22, 23), there are DNA sequences of highly variable length that separate sequences that are contiguous in the corresponding RNA.

At present, it is difficult to muster any real evidence on how such genes or gene systems might work. Even the basic question of whether or not the "extra" interior DNA sequences are transcribed must await further work in each case. For the yeast tyrosine tRNA genes, the position of the extra DNA sequence tempts us to hypothesize that it is transcribed and that its function may be to prevent interaction of the immature tRNA with actively translating mRNA–ribosome complexes while still allowing regions outside the anticodon loop to assume the basic three-dimensional conformation of mature tRNA^{Tyr}.

In a less speculative vein, we can at least rule out some "pseudogene" hypotheses concerning the role of the extra interior DNA sequences. These hypotheses would hold that only a subset of the eight chromosomal DNA sequences that hybridize to tRNA^{Tyr} represents the functional genes, while the remaining sequences have been inactivated by their insertions. Our results indicate that the insert is present in at least one

‡ We use this term, recently proposed by Philip Leder, to describe DNA sequences that intervene between nucleotide sequences that adjoin each other in the corresponding RNA.
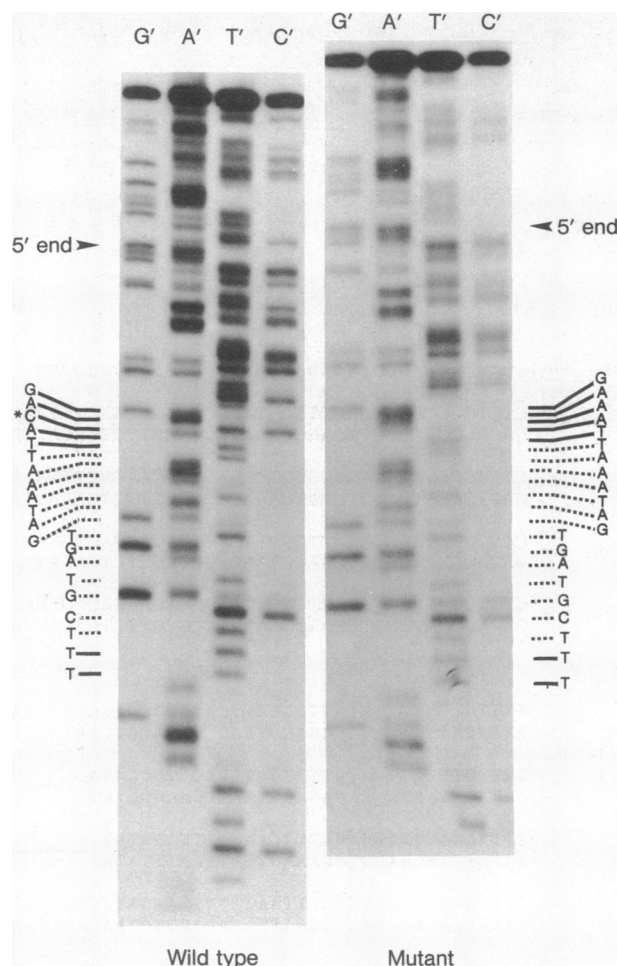

FIG. 3. Autoradiogram of a sequencing gel of the transcribed strand of the wild-type and mutant *SUP4* genes. Each well was loaded with one of the four reaction mixtures G', A', T', C'. These reactions cause cleavages after particular nucleotides with the following specificities: G', G; A', A ≅ G; T', T ≅ C; C', C. The 5' end-labeled cleavage products were fractionated on an acrylamide gel; the smaller fragments are at the bottom of the figure. Details of the procedure are in *Materials and Methods,* the caption to Fig. 2, and ref. 16. The nucleotide sequence through the 14 base pair insert is shown. Dotted arrows indicate nucleotides within the insert. The asterisk indicates the site of the ochre-suppressing mutation. The position marked "5' end", shows the G coding for the 5'-terminal C of mature tRNA^{Tyr}. Comparison of the C' wells of the two autoradiograms immediately above this point clearly reveals the C in the wild-type strain that is absent in the mutant (where it is, in fact, a T; see Fig. 2). Comparison of the G' wells in this region shows that there are at least two further sites of divergence between the two strains beyond the point at which the complete sequence can be read. In the A' well of the wild-type strain the next-to-the-smallest band was interpreted as artifactual on the basis of its anomalous spacing relative to the neighboring bands. It was present in the identical position in sequencing gels from all four genes; if it is read as an A it would imply the presence of a U between positions 62 and 63 in the tRNA sequence (Fig. 2). In order to test the correctness of our interpretation of this band, it will be necessary to sequence the other strand.

demonstrably active tRNA^{Tyr} gene, namely the mutant *SUP4-o* gene. This gene must be active, because the mutation confers a dominant phenotype on *SUP4-o* strains. The only potential escape from the conclusion that the presence of the insert must be compatible with gene function rests on the possibility, which we consider unlikely, that the 14 base pairs are not present in yeast DNA at all but were inserted during the cloning process. A cloning artifact of this type would require a sequence-spe-

Biochemistry: Goodman et al.

Proc. Natl. Acad. Sci. USA 74 (1977)   5457

cific, deterministic insertion system in *Escherichia coli* that has no known analogue.

Because the sequenced tRNA^{Tyr} genes come from a variety of chromosomal locations, the results offer an intriguing glimpse at the nature of variation between redundant genetic loci that are too widely spaced to retain their sequence homogeneity by conventional correction mechanisms (24). Our most striking observation is that perfect sequence homogeneity exists within the DNA sequences coding for mature tRNA^{Tyr} while divergence is apparent in all other regions of the three genes. From a point two base pairs before the 5' end of the tRNA^{Tyr} sequence to the limits of our data in that direction, the sequences of the three genes are almost totally dissimilar. If the transcription units of these genes resemble those for prokaryotic tRNA genes (25), this region would be expected to be transcribed into precursor tRNA. Our results indicate that if such precursors exist, the processing system places remarkably weak constraints on the sequence of the 5' leader regions. Even the *SUP4-o* and *sup4*$^+$ genes, which were isolated from different genetic backgrounds, show some divergence in this region.

In the region corresponding to the 3' end of mature tRNA^{Tyr}, all four sequences lack a coding triplet for the terminal C-C-A. The structures of known prokaryotic tRNA precursors suggest that in bacteria there is no uniform pattern of presence or absence of encoded C-C-A in tRNA genes (25), and the same may prove true in eukaryotes. The A-rich tracts in the transcribed strand beyond the 3' end of the region encoding the mature tRNA sequence are reminiscent of those observed in the yeast (26, 27) and *Xenopus* (28) 5S genes. Beyond the initial A$_7$ sequence the detailed structure of these tracts is variable in the three tRNA^{Tyr} genes that were sequenced, and even the A$_7$ unit is not fully shared with the 5S gene sequences.

The general pattern of total sequence conservation within regions encoding the mature tRNA sequence and slight to extreme divergence outside these regions suggests selection for tRNA function as a likely mechanism for the maintenance of homogeneity between the different yeast tRNA^{Tyr} loci. A corollary to this interpretation is that *all* regions of the mature tRNA sequence may be functionally significant in their interactions with processing enzymes, aminoacyl tRNA synthetases, ribosomes, or other cellular components.

Finally, our results provide clear evidence on the nature of the mutational change that gives rise to tyrosine-inserting ochre suppressors in yeast. This question has been the subject of considerable speculation, largely because these suppressors are UAA-specific, unlike the analogous *E. coli* ochre suppressors, which respond to UAA and UAG codons with roughly comparable efficiencies (29). To explain the UAA specificity of tyrosine-inserting yeast ochre suppressors, two different anticodons have been proposed for suppressor tRNA (1, 30). One proposal is that, like their *E. coli* analogues (31), the yeast ochre suppressors result from the anticodon change GΨA → [U]ΨA, but with a different wobble position U modification than occurs in *E. coli*. The other proposal is that the yeast ochre suppressors arise from the change GΨA → IΨA. The IΨA anticodon is predicted to respond to UAA, UAC, and UAU codons, the latter two being the normal tyrosine triplets.

The spontaneously derived ochre suppressor that we sequenced is clearly transcribed into a tRNA with a [U]ΨA anticodon, because the mutation changes the anticodon triplet from GTA to TTA. Inosine is derived by deamination of A, so an IΨA anticodon would require a GTA to ATA transition. It remains to be seen whether all yeast tyrosine-inserting ochre suppressors result from the same mutational change, particularly because Hawthorne and Leupold (1) have proposed on the

basis of mutagenesis studies that both G·C → T·A and G·C → A·T changes can give rise to tyrosine-inserting ochre suppressors with indistinguishable phenotypes.

Note Added in Proof. Results similar to ours have been obtained with yeast tRNA^{Phe} by Valenzuela *et al.* (32).

1. Hawthorne, D. C. & Leupold, U. (1974) *Curr. Top. Microbiol. Immunol.* **64**, 1–47.
2. Hawthorne, D. C. (1976) *Biochimie* **58**, 179–182.
3. Sherman, F., Liebman, S. W., Stewart, J. W. & Jackson, M. (1973) *J. Mol. Biol.* **78**, 157–168.
4. Olson, M. V., Montgomery, D. L., Hopper, A. K., Page, G. S., Horodyski, F. & Hall, B. D. (1977) *Nature* **267**, 639–641.
5. Madison, J. T. & Kung, H. K. (1967) *J. Biol. Chem.* **242**, 1324–1330.
6. Rothstein, R. J., Esposito, R. E. & Esposito, M. S. (1977) *Genetics* **85**, 35–54.
7. Sherman, F., Stewart, J. W., Jackson, M., Gilmore, R. A. & Parker, J. H. (1974) *Genetics* **77**, 255–284.
8. Fink, G. R. & Styles, C. A. (1974) *Genetics* **77**, 231–244.
9. Thomas, M., Cameron, J. R. & Davis, R. W. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 4579.
10. Cameron, J. R., Panasenko, S. M., Lehman, I. R. & Davis, R. W. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 3416–3420.
11. Benton, W. D. & Davis, R. W. (1977) *Science* **196**, 180–182.
12. Maniatis, T., Jeffrey, A. & Kleid, D. G. (1975) *Proc. Natl. Acad. Sci. USA* **72**, 1184–1188.
13. Bolivar, F., Rodriguez, R. L., Greene, P. J., Betlach, M. C., Heyneker, H. L., Boyer, H. W., Crosa, J. H. & Falkow, S. (1977) *Gene*, in press.
14. Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W. J. & Goodman, H. M. (1977) *Science* **196**, 1313–1319.
15. Beckmann, J. S., Johnson, P. F. & Abelson, J. (1977) *Science* **196**, 205–208.
16. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–564.
17. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
18. Endow, S. A. & Roberts, R. J. (1977) *J. Mol. Biol.* **112**, 521–529.
19. White, R. L. & Hogness, D. S. (1977) *Cell* **10**, 177–192.
20. Wellauer, P. K. & Dawid, I. B. (1977) *Cell* **10**, 193–212.
21. Pellegrini, M., Manning, J. & Davidson, N. (1977) *Cell* **10**, 213–224.
22. Chow, L. T., Gelinas, R. E., Broker, T. R. & Roberts, R. J. (1977) *Cell* **12**, 1–8.
23. Berget, S. M., Moore, C. & Sharp, P. A. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3171–3175.
24. Smith, G. (1973) *Cold Spring Harbor Symp. Quant. Biol.* **38**, 507–513.
25. Smith, J. D. (1976) *Progr. Nucleic Acid Res. Mol. Biol.* **16**, 25–73.
26. Valenzuela, P., Bell, G. I., Masiarz, F. R., DeGennaro, L. J. & Rutter, W. J. (1977) *Nature* **267**, 641–643.
27. Maxam, A. M., Tizard, R., Skryabin, K. G. & Gilbert, W. (1977) *Nature* **267**, 643–645.
28. Brown, R. D. & Brown, D. D. (1976) *J. Mol. Biol.* **102**, 1–14.
29. Feinstein, S. I. & Altman, S. (1977) *J. Mol. Biol.* **112**, 453–470.
30. Piper, P. W., Wasserstein, M., Engbaek, F., Kaltoft, K., Celis, J. E., Zeuthen, J., Liebman, S. & Sherman, F. (1976) *Nature* **262**, 757–761.
31. Altman, S. (1976) *Nucleic Acid Res.* **3**, 441–448.
32. Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R. & Rutter, W. J. (1978) *Proc. Natl. Acad. Sci. USA* **75**, in press.