Volume 17 Number 7 1989 Nucleic Acids Research

Three mitochondrial tRNA genes from *Arabidopsis thaliana*: evidence for the conversion of a $tRNA^{Phe}$ gene into a $tRNA^{Tyr}$ gene

Hsu-Ching Chen[1,2§], Henri Wintz[1+], Jacques-Henry Weil[2] and Datta T.N.Pillay[1*]

[1]Department of Biological Sciences, University of Windsor, Windsor, Ontario N9B 3P4, Canada and [2]Institut de Biologie Moléculaire des Plantes, 12 rue du Général Zimmer, 67000 Strasbourg, France

Received January 20, 1989; Accepted Febuary 27, 1989 EMBL accession no. X14110

ABSTRACT

Three tRNA genes have been isolated from a genomic library of *Arabidopsis thaliana*: a $tRNA^{Ser}$ (GCU), a $tRNA^{Tyr}$ (GUA) and a $tRNA^{Glu}$ (UUC) genes. These genes are located closely on the same DNA fragment. The $tRNA^{Ser}$ and the $tRNA^{Glu}$ genes have both 99% sequence similarity with their mitochondrial counterparts from higher plants indicating that these three tRNA genes are mitochondrial. The $tRNA^{Tyr}$ gene shows a particular high sequence similarity with the mitochondrial $tRNA^{Phe}$ pseudogene from maize, and both genes are flanked by a $tRNA^{Ser}$ gene in the upstream region. Extensive sequence comparisons of the *Arabidopsis thaliana* mitochondrial sequence containing the three tRNA genes and the corresponding region from maize and soybean mitochondria have shown evidence that the $tRNA^{Tyr}$ gene has been generated from a mitochondrial $tRNA^{Phe}$ gene. The conversion was accomplished by three genetic events: a 4 base-pair deletion, a mutation and a recombination, which led to the transformation of the acceptor stem and the anticodon.

INTRODUCTION

It is believed, based on sequence homology, that most (if not all) tRNA genes derive from the same ancestor gene. High level of sequence similarities have been observed not only between the same species of tRNA genes from different organisms, but also between different species of tRNA genes. For example, in higher plant chloroplast the $tRNA^{Met}$ elongator gene has 68% sequence similarity with the $tRNA^{Thr}$ gene (1). The nuclear $tRNA^{Phe}$ gene has 70% sequence similarity with the nuclear $tRNA^{Tyr}$ gene from *Xenopus laevis* (2). However, besides the sequence similarity data, there is no direct evidence that one tRNA gene has derived from another tRNA gene.

Recently, it has been shown that the identity of a tRNA can be changed *in vitro* by altering only a few nucleotides. The nucleotides which determine the identitiy of a tRNA reside along and around the inside of the L-shaped tRNA structure including the acceptor stem (3,4,5), the dihydrouridine stem (5), and the anticodon (6,7). However, the nucleotides which reside in a patch on the surface of the tRNA molecule where the two loops (D and T loop) interact, termed variable pocket, can also play an important role on the identity of tRNAs (8). These data demonstrate that one tRNA can be converted *in vitro* into another tRNA by changing only a

limited number of nucleotides, whereas the determinant nucleotides vary from one to another tRNA.

We have characterized a cluster of three mitochondrial (mt) tRNA genes from *Arabidopsis thaliana* coding for a $tRNA^{Ser}$, a $tRNA^{Tyr}$ and a $tRNA^{Glu}$. We present here the first evidence of the conversion of a tRNA gene *in vivo*. The data presented here suggest strongly that the $tRNA^{Tyr}$ gene has been generated from a $tRNA^{Phe}$ gene in *A.thaliana* mitochondrion.

MATERIALS & METHODS

A lambda genomic library of *A.thaliana* constructed in vector EMBL4 was kindly provided by E.M. Meyerowitz (9). Recombinant lambda DNA was prepared from 5 ml overnight liquid cultures and purified according to Helms *et al.* (10).

DNA restriction fragments were transferred from agarose gels onto nylon membranes (Hybond, Amersham) (11). Hybridizations of nick-translated DNA probes (12) to filter-bound DNA were carried out as described (13).

DNA sequencing was performed by using the dideoxynucleotide chain termination method (14) after subcloning DNA fragments into M13mp19 vector. DNA fragments for sequencing were generated by DNaseI partial hydrolysis (15) or exonuclease III digestion. For the exonuclease III digestion, a series of four reactions were performed: each reaction contained one µg of purified DNA restriction fragment and 0.5 µl, 1 µl, 2 µl, or 3 µl of enzyme (65 U/µl, BRL) respectively in a total volume of 40 µl. Ten µl aliquots were removed from each reaction at 10 min intervals and transferred into a tube containing 5 µl of 0.5 M EDTA in ice. After the completion of the reaction, the reaction mixture was extracted with phenol/chloroform, chloroform and then precipitated with ethanol in the presence of 2 M of ammonium acetate at -70 °C for 30 min. The DNA was recovered by centrifugation and subjected to S1 nuclease digestion (12). The generated DNA fragments were end-filled by using the klenow fragment of DNA polymerase I (12) and cloned into M13mp19 for sequencing.

RESULTS

Isolation and identification of three tRNA genes

We have screened five equivalent genomes of the *A.thaliana* genomic library by using an *A.thaliana* mitochondrial $tRNA^{Tyr}$ ($tRNA^{Tyr-1}$) pseudogene, previously characterized (Chen et al., in preparation), as a probe. One lambda clone, hybridizing with the probe, has been subcloned into pUC19 after EcoRI digestion. An EcoRI fragment of 6 kb, which hybridized with the $tRNA^{Tyr-1}$ gene probe has been subcloned into phage M13mp19 and sequenced partially. A nucleotide sequence of 880 bp has been determined. Comparison of the sequence with the sequence of tRNAs and tRNA genes known so far (1) has revealed the presence of three tRNA

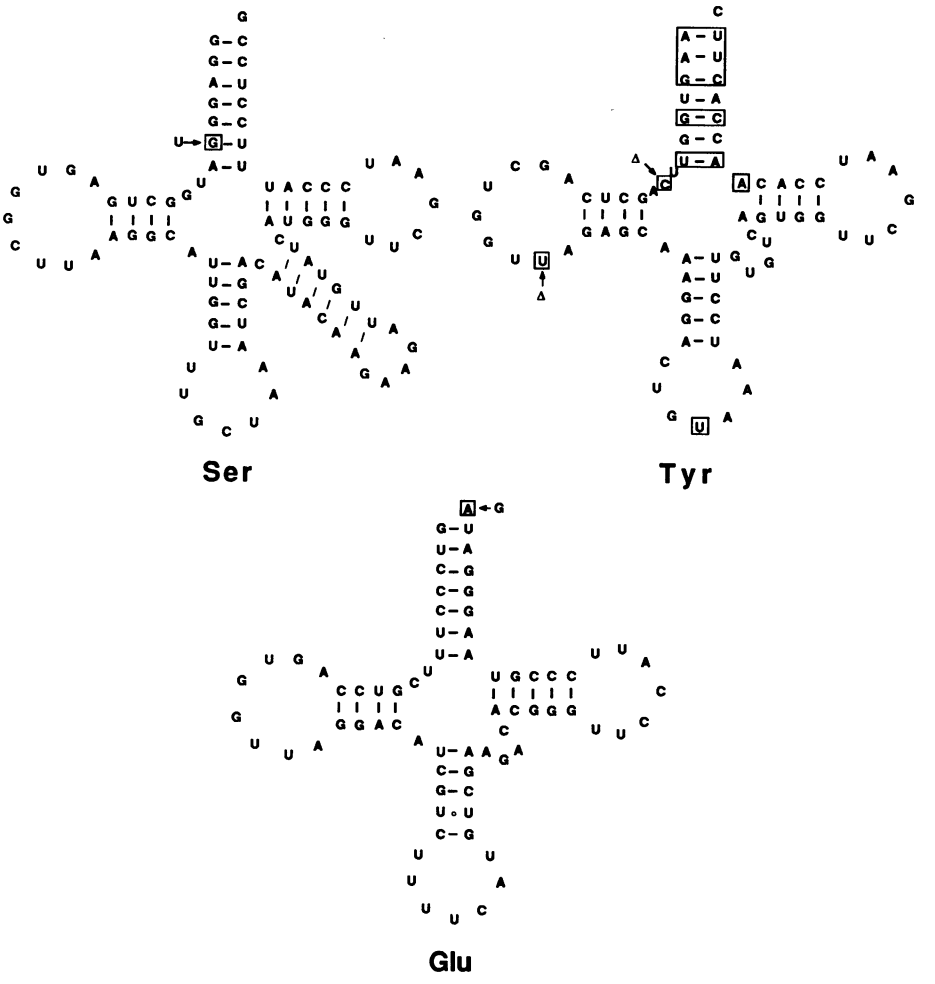

Figure 1. Secondary structure of three *Arabidopsis thaliana* mitochondrial tRNAs deduced from the corresponding gene sequence. The difference between the *A. thaliana* mt $tRNA^{Ser}$ gene and the maize (16) or wheat (17) mt $tRNA^{Ser}$ gene is boxed, and the corresponding nucleotide in maize and wheat genes is indicated. The differences between the *A. thaliana* mt $tRNA^{Tyr}$ gene and the bean mt $tRNA^{Phe}$ (20) are boxed. The nucleotides which are not present in the bean mitochondrial $tRNA^{Phe}$ sequence are indicated by Δ. The difference between the *A. thaliana* mt $tRNA^{Glu}$ gene and the soybean mt $tRNA^{Glu}$ gene (18) is boxed, and the corresponding nucleotide in the soybean gene is illustrated.

genes: a $tRNA^{Ser}$ (GCU), a $tRNA^{Tyr(Tyr-2)}$ (GCA) and a $tRNA^{Glu}$ (UUC) genes. The $tRNA^{Ser}$ is located 304 bp upstream and the $tRNA^{Glu}$ gene is located 38 bp downstream of the $tRNA^{Tyr}$ gene.

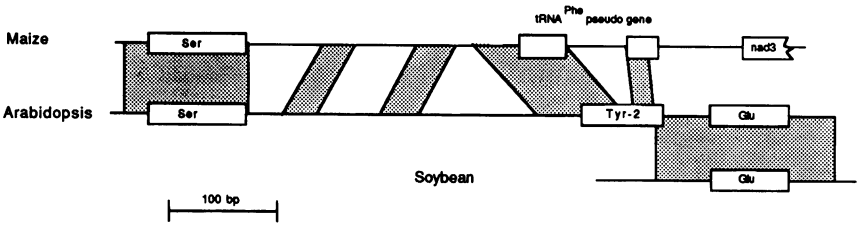

Figure 2. Comparison of the gene organization in the mitochondrial genomes of *A.thaliana*, maize and soybean. Genes are represented by boxes (*nad* 3 = gene coding for subunit III of NADH dehydogenase). The shaded areas represent the regions which are homologous (degree of similarity greater than 85%).

Analysis of the tRNA genes

The secondary structures of three tRNAs deduced from the gene sequences are presented in Fig.1. The $tRNA^{Ser}$ gene differs from the mt $tRNA^{Ser}$ gene from maize (16) and wheat (17), by only one nucleotide corresponding to nucleotide position 6 of the tRNA (G instead of U) which changes the mispaired U-U into a paired G-U in the *A.thaliana* $tRNA^{Ser}$. The $tRNA^{Glu}$ gene differs from the soybean mt $tRNA^{Glu}$ gene (18) by its 3' terminal nucleotide only (A instead of G).

The $tRNA^{Tyr}$ gene does not show sequence similarity with the $tRNA^{Tyr}$ and $tRNA^{Tyr}$ gene from bean (19) and wheat (17) mitochondria respectively. But it shows a high sequence similarity with the $tRNA^{Phe}$ from bean mitochondria (20) and the $tRNA^{Phe}$ pseudogene from maize (16) and wheat (21) mitochondria. The $tRNA^{Tyr}$ differs from the mt $tRNA^{Phe}$ pseudogene from maize by the anticodon and the acceptor stem (Fig. 1), if we do not take into account the insertion present in the variable loop of maize mt $tRNA^{Phe}$ pseudogene (16). The differences between the $tRNA^{Tyr}$ gene and the bean mitochondrial $tRNA^{Phe}$ reside also in the acceptor stem and the anticodon, plus one extra U in the D stem (position 20A; numbering of tRNA nucleotide is according to ref: 1) and one nucleotide substitution in the T stem (position 49). These two positions are conserved in the maize mitochondrial $tRNA^{Phe}$ pseudogene (16). Interestingly, both $tRNA^{Tyr}$ gene and maize mt $tRNA^{Phe}$ pseudogene are flanked by a $tRNA^{Ser}$ gene in the upstream region.

We have not been able to purify the mitochondrial DNA from *Arabidopsis thaliana*, however we can conclude that the three tRNA genes that we have sequenced are mitochondrial genes because of their high sequence similarity with known plant mt tRNA genes.

Analysis of the sequence containing the three putative mitochondrial tRNA genes

We have compared the 880 bp sequence containing three putative tRNA genes with the nucleotide sequence of the region from maize mitochondrial genome containing the $tRNA^{Ser}$

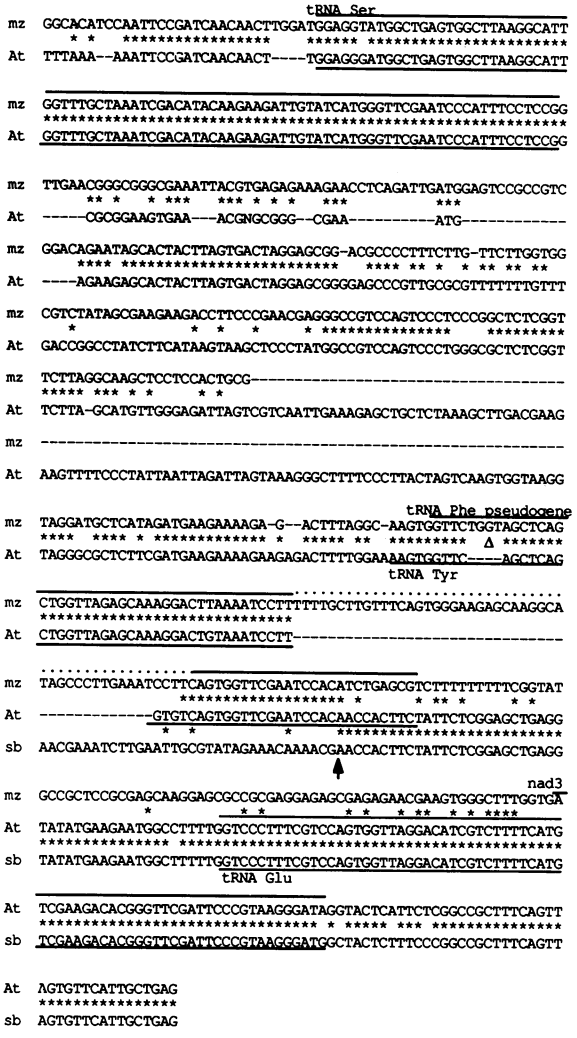

Figure 3. Nucleotide sequence comparison of the mt tRNA^{Ser}, tRNA^{Tyr} and tRNA^{Glu} genes from *A.thaliana* with the nucleotide sequences of the tRNA^{Ser} and the tRNA^{Phe} pseudogene from maize mitochondria (16) and the mt tRNA^{Glu} gene from soybean (18). The genes are underlined or overlined. The sequence inserted in the maize pseudogene (16) is marked by the dotted line. The 4 bp deletion is indicated by Δ and the point of recombination is indicated by a vertical arrow. Dashes in the sequence represent gaps introduced to get the best sequence alignment.

gene and the tRNA^{Phe} pseudogene (16) on one hand, and with the soybean mt tRNA^{Glu} gene and its flanking regions on the other hand. The comparison of the gene organization and sequence

homology of these three sequences is shown in the Fig. 2. A detailed nucleotide sequence comparison is illustrated in Fig. 3.

The sequence similarity between the mitochondrial $tRNA^{Ser}$ genes from maize and *A.thaliana* starts 22 bp upstream of the genes with a 4 bp deletion in the *A.thaliana* sequence relative to the maize sequence, and ends one nucleotide proceeding the 3' end of the genes. There is no similarity immediately downstream of the genes. However, two regions of 30 and 37 bp located 30 and 118 bp respectively downstream of the *A.thaliana* $tRNA^{Ser}$ gene (65 and 151 bp downstream of the maize gene) are highly conserved. Downstream of those regions the similarity is interrupted by an extra sequence of 94 bp in the *A.thaliana* sequence.

The sequence similarity between $tRNA^{Tyr}$ gene and $tRNA^{Phe}$ pseudogene starts 42 bp upstream of the $tRNA^{Tyr}$ gene, and ends abruptly at the position corresponding to the first nucleotide of 3' strand of the acceptor stem of the two tRNAs. No futher sequence similarity can be observed downstream the genes, since the two sequences diverge in their gene content. A mt $tRNA^{Glu}$ gene is present 38 bp downstream of the mt $tRNA^{Tyr}$ gene in *A.thaliana* whereas a gene coding for subunit III of NADH dehydrogenase (*nad* 3) is present 73 bp downstream of the $tRNA^{Phe}$ pseudogene in maize (Fig. 3).

The comparison also reveals that the sequence corresponding to the 5' strand of the acceptor stem of the $tRNA^{Tyr}$, is also present in the maize sequence (Fig. 3). The sequence of the 5' strand of the acceptor stem of $tRNA^{Tyr}$ gene can be derived, in that region, from the $tRNA^{Phe}$ pseudogene sequence by a 4 bp deletion (TGGT).

We have compared the *A.thaliana* mt $tRNA^{Glu}$ gene, located 38 bp downstream of the mt $tRNA^{Tyr}$ gene, and its flanking regions with its counterpart from soybean (18) (Fig. 2,3). It shows that the sequence of the two mt $tRNA^{Glu}$ genes is highly conserved not noly in the gene coding sequence, but also in the flanking regions. The similarity starts 47 bp upstream of the genes and extends to at least 43 bp downstream of the genes (the sequence of *A. thaliana* has not been determined futher). There is only one difference in the upstream region and three differences in the downstream region between these two sequences. The similarity between the two mt $tRNA^{Glu}$ genes starts exactly where the similarity between the $tRNA^{Tyr}$ gene and the maize $tRNA^{Phe}$ pseudogene ends, i.e. the first nucleotide of the 3' strand of acceptor stem of the $tRNA^{Tyr}$ (Fig. 3). It seems that the last eight nucleotides of the $tRNA^{Tyr}$ gene constituting the 3' strand of the acceptor stem, have been brought by recombination from another part of the mitochondrial genome of *A.thaliana*.

The recombination event explains the difference in the gene organization between *A.thaliana* and maize mitochondrial DNA downstream the mt $tRNA^{Tyr}$ gene and $tRNA^{Phe}$ pseudogene.

DISCUSSION

We have characterized three closely linked putative tRNA genes from *A.thaliana* coding for a $tRNA^{Ser}$ (GCU), a $tRNA^{Tyr(Tyr-2)}$ (GCA) and a $tRNA^{Glu}$ (UUC). The genes have been identified as being mitochondrial tRNA genes by their homology with known plant mitochondrial tRNA genes. A high degree of sequence similarity can be observed in the coding sequences and the flanking regions.

The notable feature of the identified $tRNA^{Tyr}$ gene is its low sequence similarity with mitochondrial $tRNA^{Tyr}$ from bean (19) and $tRNA^{Tyr}$ genes from wheat (17). This is unusual to what has been observed between other plant mitochondrial genes. However there is evidence for evolutionnary homology with mitochondrial $tRNA^{Phe}$ from bean (20). The $tRNA^{Tyr}$ gene shows a high level of similarity with bean mitochondrial $tRNA^{Phe}$ and maize mitochondrial $tRNA^{Phe}$ pseudogene. Both $tRNA^{Tyr}$ gene and $tRNA^{Phe}$ pseudogene are located on a sequence with similar gene organization where a $tRNA^{Ser}$ gene is located upstream of these genes. The major difference between $tRNA^{Tyr}$ gene and the $tRNA^{Phe}$ pseudogene is in the acceptor stem and the anticodon, if we do not take into account an extra sequence of 49 nucleotide present in the maize pseudogene.

A detailed sequence comparison by alignment of the $tRNA^{Tyr}$ gene with the maize $tRNA^{Phe}$ pseudogene and the soybean mt $tRNA^{Glu}$ gene has revealed that the $tRNA^{Tyr}$ gene and the $tRNA^{Phe}$ pseudogene derive from a common sequence. It seems that the $tRNA^{Tyr}$ gene has derived from a $tRNA^{Phe}$ gene (or $tRNA^{Phe}$ pseudogene) by three genetic events which led to the remodeling of the acceptor stem and the conversion of the anticodon.

The sequence comparison data suggest that the 5' end of the $tRNA^{Phe}$ gene underwent a 4 bp deletion whereas the 3' end underwent a recombination restoring the proper base-pairing of the tRNA acceptor stem. The 4 bp deletion and the recombination events have remodelled the acceptor stem and resulted in the presence of three nucleotides between the acceptor and the D stems of $tRNA^{Tyr}$.

One mutation, affecting the second position of the anticodon, has converted the original Phe GAA anticodon into a Tyr GUA anticodon. In wheat mitochondrial genome, a $tRNA^{Phe}$ pseudogene has also been identified (21). The wheat and maize $tRNA^{Phe}$ pseudogenes share a high degree of sequence similarity and are both located on a sequence with the same gene organization, where a $tRNA^{Ser}$ gene is located on the upstream region and a *nad3* gene is located on the downstream region. If we exclude the extra 68 bp insertion present in the D loop of the wheat pseudogene there are only 4 nucleotide differences between the wheat and the maize pseudogenes. Two differences occur in the 48 bp insertion in the variable loop and two other differences are located in the anticodon. The anticodon in the maize $tRNA^{Phe}$ pseudogene is UAA specific for leucine (21), whereas in wheat it is GCA, specific for serine. This shows that the

anticodon in the $tRNA^{Phe}$ pseudogene undergoes frequent nucleotide substitutions without the rest of the sequence beeing affected.

The mt $tRNA^{Tyr}$ gene probably codes for a functional tRNA since the cloverleaf structure of the tRNA deduced from the gene sequence is conserved with all the invariant and semi-invariant nucleotides of functional tRNAs. In addition, there are two other copies of the mt $tRNA^{Tyr}$ gene in the *A.thaliana* mitochondrial genome (Chen et al., in preparation). These two copies have probably been generated by duplication of the gene presented here, but they have an insertion and a deletion respectively relative to the $tRNA^{Tyr}$. They are probably pseudogenes. It appears that the structure of the $tRNA^{Tyr}$ gene presented here has been maintained by selective constraints and the redundant copies of the gene which were not under the same selective pressure have evolved rapidly into nonfunctional structures.

Recent studies have shown that the acceptor stem and the anticodon can determine the identity of a tRNA (3,4,5). Earlier studies have shown that the acceptor stem plays an important role in determining the identity of $tRNA^{Tyr}$ from *E. coli* (22). These two regions have been the major sites of transformation for the conversion of $tRNA^{Phe}$ gene to $tRNA^{Tyr}$ gene. Therefore the tRNA encoded by the $tRNA^{Tyr}$ gene is probably recognized by the tyrosyl-tRNA synthetase in mitochondrial protein synthesis. Moreover, it is interesting to note that the structure of the $tRNA^{Tyr}$ gene is not alien to a known $tRNA^{Tyr}$ structure as it presents a high degree of similarity (70%) with the higher plant nuclear $tRNA^{Tyr}$ (5,6,7).

One $tRNA^{Phe}$ has been purified from bean mitochondria but the corresponding gene has not been sequenced and may not be present in the mitochondrial genome, since by using the purified bean mt $tRNA^{Phe}$ as a probe to hybridize the mitochondrial genomes of maize and wheat, only one homologous fragment has been identified which carries the mt $tRNA^{Phe}$ pseudogene (16). Thus it is possible that in maize and wheat mitochondria the $tRNA^{Phe}$ gene has evolved into a pseudogene. The $tRNA^{Phe}$ which is used in the elongation step of mitochondrial protein synthesis might be coded for by the nuclear genome as is the case for bean mt $tRNA^{Leu}$ (23), or coded for by a mitochondrial gene with a different nucleotide sequence.

The results presented here show the first evidence that the identity of a tRNA gene has been changed in the course of evolution by a limited number of substitutions. This supports the idea that most of the tRNA genes (if not all) may have evolved from a single ancestor tRNA gene sequence. Interestingly, the nuclear $tRNA^{Tyr}$ and $tRNA^{Phe}$ genes from *Xenopus laevis* also display a high degree of similarity (2), suggesting a similar relationship between nuclear $tRNA^{Tyr}$ and $tRNA^{Phe}$ genes.

ACKNOWLEDGEMENTS

This work was supported by a grant from Natural Sciences and Engineering Research Council of Canada to D.T.N.P.

*To whom correspondence should be addressed

Present addresses: [+]Department of Genetics and Development, Cornell University, Ithaca, NY 14853 and [§]Boyce Thompson Institute, Cornell University, Tower Road, Ithaca, NY 14853, USA

REFERENCES

1. Sprinzl,M., Hartmann,T., Maissner,F., Moll,J. and Vorderwulbecke,T. (1987) Nucleic Acids Res. 15suppl, r15-r188.
2. Muller,F.,Clarkson,S.G. (1980) Cell 19,345-353.
3. Hou,Y.-M.,Schimmel,P. (1988) Nature 333,140-145.
4. McClain,W.H. and Foss,K. (1988) Science 240, 793-796.
5. Normanly,J., Ogden,R.C., Horvath,S.J. and Abelson,J. (1986) Nature 321, 213-219.
6. Muramatsu,T., Nishikawa,K., Nemoto,F., Kushino,Y., Nishimura,S., Miyazawa,T. and Yokoyama,S. (1988) Nature 336, 179-181.
7. Schulman,L.H. and Pelka,H. (1988) Science 242, 765-768.
8. McClain,W.H. and Foss,K. (1988) Science 241, 1804-1807.
9. Chang,C. and Meyerowitz,E.M. (1986) Proc.Natl.Acad.Sci.USA 83, 1408-1412.
10. Helms,C., Graham,M.Y., Dutchik,J.E. and Olson,M.V. (1985) DNA 4, 39-50.
11. Southern,E.M. (1975) J. Mol. Biol. 98, 503-518.
12. Maniatis,T., Fritsch,E. and Sambrook,J. (1982) Cold Spring Harbor Laboratory press, Cold Spring Harbor, N.Y.
13. Wintz,H., Chen,H.-C. and Pillay,D.T.N. (1988) Curr. Genet. 13, 255-260.
14. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
15. Anderson,S. (1981) Nucleic Acids Res. 9, 3015-3027.
16. Wintz,H., Grienenberger,J.M., Weil,J.H. and Lonsdale,D.M. (1988) Curr. Genet. 13, 247-254.
17. Joyce,P.B.M.,Spencer,D.F.,Bonen,L.,Gray,M.W. (1988) Plant Mol. Biol. 10,251-262.
18. Wintz,H., Chen,H.-C. and Pillay,D.T.N. (1987) Nucleic Acids Res. 15, 10588.
19. Marechal,L., Guillemaut,P. and Weil,J.H. (1985) Plant Mol. Biol. 5, 347-351.
20. Marechal,L., Guillemaut,P., Grienenberger,J.M., Jeannin,G. and Weil,J.H. (1985) FEBS Lett. 184, 289-293.
21. Gualberto,J.M., Wintz,H., Weil,J.H. and Grienenberger,J.M. (1988) Mol. Gen. Genet. in press.
22. Ghysen,A. and Celis,J.E. (1974) J.Mol.Biol. 83, 333-351.
23. Marechal-Drouard,L., Weil,J.H. and Guillemaut,P. (1988) Nucleic Acids Res. 16, 4777-4788.