Cell, Vol. 14, 123–131, May 1978, Copyright © 1978 by MIT

Sequence Complexity of Nuclear and Polysomal RNA in Leaves of The Tobacco Plant

Robert B. Goldberg,* Gisela Hoschek and Joseph C. Kamalay
Department of Biology
University of California
Los Angeles, California 90024
William E. Timberlake
Department of Biology
Wayne State University
Detroit, Michigan 48202

Summary

The first measurements are reported of the sequence complexity of nuclear and polysomal RNA contained within the cells of a higher plant. Polysomal RNA from tobacco leaves was prepared by a procedure which minimized contamination with nuclear RNA. Hybridization of ^{3}H-cDNA complementary to polysomal poly(A) RNA with an excess of tobacco DNA indicated that >95% of the poly(A) mRNA was transcribed from single-copy sequences. RNA excess hybridization reactions with polysomal poly(A) RNA and ^{3}H-cDNA revealed the presence of three abundance classes in the poly(A) mRNA. The best least-squares solution indicated that these classes comprise 9, 52, and 39% of the poly(A) mRNA and contain sequences present an average of 4500, 340 and 17 times per cell. Hybridization reactions containing an excess of nuclear or total polysomal RNA and ^{3}H-single-copy DNA indicated the complexity of these RNA populations to be 1.19×10^8 nucleotides (nuclear) and 3.33×10^7 nucleotides (polysomal). Thus only 28% of the nuclear RNA sequence diversity (27,000 average-sized mRNA sequences) is represented in leaf polysomes. These results suggest that there is a general similarity in the basic transcriptional processes of metaphytan and metazoan cells.

Introduction

In all animal cells thus far investigated, the sequence complexity of hnRNA is greater than that of cytoplasmic or mRNA active in translation by a factor of 4–10 (Getz et al., 1975; Hough et al., 1975; Liarakos et al., 1975; Ryffel, 1976; Bantle and Hahn, 1976; Levy, Johnson and McCarthy, 1976; Kleiman et al., 1977). The function of hnRNA is not presently known, but has been implicated in mRNA biosynthesis as well as gene regulation (Darnell, Jelinek and Molloy, 1974; Herman, Williams and Penman, 1976; Davidson, Klein and Britten, 1977). On the other hand, the cells of simpler eucaryotes—for example, protistans and fungi—do not have a separate class of complex hnRNA molecules (Firtel and Lodish, 1973; Hereford and Rosbash, 1977; Timberlake, Shumard and Goldberg, 1977; Rozek, Orr and Timberlake, 1978). In these eucaryotes, the sequence complexities of polysomal, nuclear and total cellular RNA are the same (Hereford and Rosbash, 1977; Timberlake et al., 1977; Rozek et al., 1978). These results suggest that whatever function is finally attributed to hnRNA will be one which is unique to higher eucaryotes (that is, animals and plants).

Little quantitative information exists regarding the nature of transcriptional processes in higher plants other than the biosynthesis of rRNA (Leaver and Key, 1970; Rogers, Loening and Fraser, 1970). For example, the presence of a class of hnRNA molecules which might serve as mRNA precursors has not been established. It is also not known to what extent nuclear DNA sequences are transcribed, what number of structural genes is required to maintain a given cell type or tissue, or what DNA sequence components are represented in the structural gene set. Although cellular and developmental processes in higher plants differ significantly from those of animals, higher plant cells do undergo intricate programs of differentiation which require complex regulatory mechanisms (Raghavan, 1976). The fact that higher plant genomes are similar in size, complexity and sequence organization to those of animal cells (Walbot and Dure, 1976; Zimmerman and Goldberg, 1977; Goldberg, 1978) is consistent with this requirement. Plant genomes, however, generally contain a higher proportion (>60%) of repeated DNA than do those of animal cells (Flavell et al., 1974).

Here we report the first sequence complexity measurements of the nuclear and polysomal RNA populations present in the cells of a higher plant. These studies reveal that there are approximately 27,000 structural genes expressed in the leaf, and that >95% of these genes are represented once per haploid genome. As is the case with animals, the sequence complexity of nuclear RNA exceeds that of mRNA by a factor of 4.

Results

Preparation and Size of Polysomal Poly(A) RNA

Tobacco leaves were chosen for this study for several reasons. First, quantitative information exists regarding the DNA sequence organization of the tobacco genome (Zimmerman and Goldberg, 1977). Second, procedures have been developed for the isolation of intact tobacco polysomes (Jackson and Larkins, 1976). Finally, leaves can be grown in large numbers under environmental conditions which favor developmental uniformity. Pol-

* To whom requests for reprints should be addressed.

ysomal RNA was isolated from leaves using an EDTA release procedure which minimizes contamination with sequences of nuclear origin (Goldberg et al., 1973; Galau, Britten and Davidson, 1974; Galau et al., 1976; Timberlake et al., 1977). Sucrose density gradients of polysomes concentrated from the postmitochondrial supernate are shown in Figures 1A (analytical) and 1B (preparative). Greater than 75% of the cellular RNA was contained in the polysomal pellet. Approximately 60% of the ribonucleoprotein in the polysomal pellet was present in polysomes sedimenting at >100S (fraction I, Figure 1B). Although the total mRNA content in the >100S polysomes could not be measured due to rRNA synthesis, we determined that this fraction contained >90% of the poly(A) RNA in the polysomal pellet. A sucrose density gradient of EDTA-treated >100S polysomes is shown in Figure 1C. After EDTA treatment, all of the ribosomal ribonucleoprotein and >95% of the polysomal poly(A) RNA sedimented at <80S (fraction II, Figure 1C). Despite minor losses at each fractionation step, we estimate that fraction II contains most of the messenger ribonucleoprotein since >85% of the polysomal poly(A) RNA was present.

The size of leaf polysomal poly(A) RNA was measured by hybridization of an excess of ^{3}H-poly(U) across a sucrose density gradient containing RNA isolated from fraction II ribonucleoprotein (Bishop, Rosbash and Evans, 1974a). The results of this experiment are shown in Figure 1D. Assuming that the poly(A) size is not a function of poly(A) RNA length, we calculate that the leaf polysomal poly(A) RNA has a number average size of 1340 ± 85 nucleotides (six determinations). This value is in close agreement with the size of polysomal poly(A) RNA in other higher plants (Key and Silflow, 1975; Gray and Cashmore, 1976; Walbot and Dure, 1976) and eucaryotes in general (Davidson and Britten, 1973).

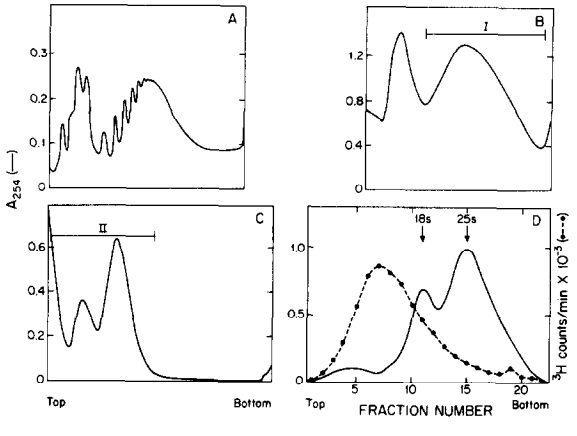

Figure 1. Preparation of Polysomes and Size Distribution of Polysomal Poly(A) RNA

Details of protocols are presented in Experimental Procedures. (A) Analytical sedimentation of leaf polysomes; (B) preparative sedimentation of leaf polysomes. This sucrose gradient contained approximately 100 times the polysomal material as that shown in (A). Polysomes sedimenting at >100S (fraction I) were collected and treated with EDTA. (C) Sedimentation of EDTA-treated polysomes; represents sedimentation of fraction I after EDTA treatment. Ribonucleoprotein sedimenting at <80S (fraction II) was collected and precipitated with EtOH. (D) Size distribution of polysomal poly(A) RNA. RNA was extracted from fraction II, and a portion was heated for 3 min at 60°C, cooled to 0°C and sedimented in a sucrose density gradient. The gradient was fractionated, and each fraction was processed for hybridization to ^{3}H-poly(U). (——) sedimentation of polysomal RNA; (- -●- -) sedimentation of polysomal poly(A) RNA.

DNA Sequence Representation in Polysomal Poly(A) RNA

The spectrum of DNA sequence classes represented in a RNA population can be estimated either by hybridization of an excess of DNA with RNA of high specific radioactivity (Gelderman, Rake and Britten, 1971; Melli et al., 1971) or with labeled cDNA prepared from poly(A) RNA (Bishop et al., 1974b); Rosbash, Ford and Bishop, 1974). We adopted the second approach because procedures for labeling leaf mRNA to high specific activity in vivo are currently unavailable. The results of hybridizing ^{3}H-cDNA (640 nucleotides) to an excess of unlabeled tobacco DNA are presented in Figure 2 (open circles). For comparison, the reassociation kinetics of ^{3}H-single-copy DNA (300 nucleotides) are also shown (closed circles).

The ^{3}H-cDNA reassociated with the driver DNA as a pure second-order reaction, having a K of 1.5 × 10^{-3} M^{-1} sec^{-1}. This rate constant is approximately 3 times higher than that observed for the ^{3}H-single-copy DNA (4.8 × 10^{-4} M^{-1} sec^{-1}). After appropriate fragment length corrections are made (Davidson et al., 1973; Smith, Britten and Davidson, 1975); the latter value is in excellent agreement with our previous measurements of the single-copy reassociation rate in total short tobacco DNA (Zimmerman and Goldberg, 1977). Since the ratio of the ^{3}H-cDNA to ^{3}H-single-copy DNA fragment lengths is 2.13 (640/300), this result indicates that sequences in the ^{3}H-cDNA population are homologous to sequences present only once or twice per haploid genome.

Despite the fact that >75% of the driver DNA fragments contain at least one repeated sequence (dashed curve, Figure 2; Zimmerman and Goldberg, 1977), we were unable to detect any repetitive DNA in the ^{3}H-cDNA population. In this type of experiment, however, a small fraction (<5%) of repeated DNA would go undetected. If we assume that sequences at the 5′ end of polysomal poly(A) RNA have the same reiteration frequency as those at the 3′ end, and that reverse transcriptase copies repeated and single-copy DNA transcripts with equal efficiency, the results presented in Figure 2

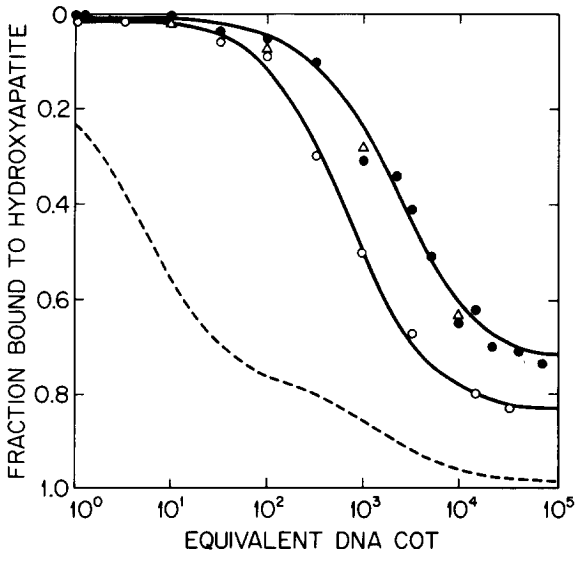

Figure 2. Reassociation of ^{3}H-cDNA and ^{3}H-Single-Copy DNA with Total Leaf DNA

Trace amounts of 640 nucleotide ^{3}H-cDNA (○) or 300 nucleotide ^{3}H-single-copy DNA (●) were reassociated in the presence of an excess of 390 nucleotide total leaf DNA (mass ratio 20,000/1 and 100,000/1, respectively) as described in Experimental Procedures. The extent of reassociation was measured by hydroxyapatite chromatography. The solid curves through the data points represent the best least-squares solution for one second-order component. Attempts to fit these data to more than one component were unsuccessful. The kinetic parameters of these curves were: ^{3}H-cDNA, K of 1.5×10^{-3} M^{-1} sec^{-1}, RMS error 1.7%; ^{3}H-single-copy, K of 4.8×10^{-4} M^{-1} sec^{-1}, RMS error 2.6%. Included in the ^{3}H-single-copy curve are data points obtained from the reassociation of ^{3}H-single-copy DNA isolated from polysomal RNA-^{3}H-DNA hybrids (△). The dashed curve represents the reassociation of the driver DNA, which has been quantitatively described elsewhere (Zimmerman and Goldberg, 1977).

indicate that at least 95% and probably all of the leaf polysomal poly(A) RNA is transcribed from single-copy DNA.

Sequence Complexity and Abundance Classes in Polysomal Poly(A) RNA

To ascertain the range of abundance classes in higher plant mRNA and provide and estimate of their sequence diversity, a large excess of leaf polysomal poly(A) RNA was hybridized to complementary ^{3}H-cDNA, and the extent of hybridization was assayed by resistance to S1 nuclease (Bishop et al., 1974b). The results of this experiment are presented in Figure 3 and summarized in Table 1.

Figure 3 demonstrates that the hybridization reaction of ^{3}H-cDNA with polysomal poly(A) RNA occurs over a large range of RNA Cot values (5 log units). Since a pseudo-first-order hybridization reaction with a homogeneous population of RNA and tracer DNA extends for only 1.5 log RNA Cot units (Galau, Britten and Davidson, 1977), polysomal poly(A) RNA sequences must vary widely in their cellular concentration. Our best approximation of

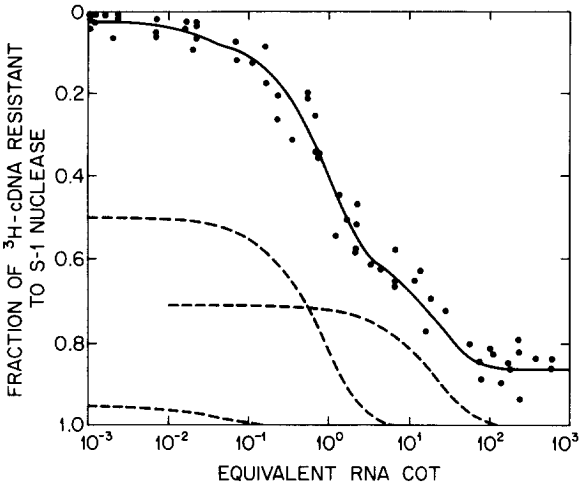

Figure 3. Hybridization of ^{3}H-cDNA to Polysomal Poly(A) RNA

Trace amounts of ^{3}H-cDNA were hybridized to a 1000 fold excess of polysomal poly(A) RNA as described in Experimental Procedures. The extent of hybridization was measured by resistance to S1 nuclease. The solid curve through the data points represents the best least-squares solution for three pseudo-first-order components. The dashed curves represent elements of the overall solution. The kinetic parameters and an analysis of this hybridization curve are presented in Table 1. The RMS error of this curve is 5%. Other solutions to these data, resulting in <15% increase in RMS error, are possible.

the data in Figure 3 is that they are fit by three pseudo-first-order components, representing RNA classes differing in abundancy by several orders of magnitude (Table 1). These classes represent 9, 52 and 39% of the RNA mass and contain an average of 10, 770 and 11,300 diverse polysomal poly(A) RNA sequences each.

We must emphasize that this description of the abundance classes in leaf polysomal poly(A) RNA should not be considered unique. Other solutions to the data in Figure 3 are possible with little change in RMS error. For example, we were able to obtain a range of 6,000–24,000 diverse RNA sequences in class 3 with only a 15% increase in RMS error (see Table 1). Whatever the exact solution, however, these results indicate that leaf polysomal poly(A) RNA contains a large number of different RNA sequences which vary in cellular concentration, and that >90% of the sequence diversity is represented by molecules present in only a few copies per cell.

Sequence Complexity of Total Polysomal mRNA

The hybridization reaction presented in Figure 3 provided an estimate of the poly(A) mRNA sequence diversity. In plant cells, however, only 35–50% of the mRNA is polyadenylated (Key and Sillow, 1975; Gray and Cashmore, 1976; Ragg, Schroder and Hahlbrock, 1977). Our own measurements indirectly support this observation, since we estimate from the data presented in Figure 1 that

Table 1. Abundance Classes of Polysomal Poly(A) RNA

Class	Fraction of ^{3}H-cDNA	Fraction of Polysomal Poly(A) RNA Mass[a]	K (M^{-1} sec^{-1})	K_{pure} (M^{-1} sec^{-1})[b]	Complexity (Nucleotides)[c]	Number of 1240 Nucleotide Sequences[d]	Number of Molecules per Cell per Sequence[e]
1	0.07	0.09	14	156	1.2×10^4	10	4500
2	0.44	0.52	0.983	1.89	9.5×10^5	770	340
3	0.33	0.39	0.049	0.126	1.4×10^7	11,300	17

[a] Normalized to 100% ^{3}H-cDNA reactability.

[b] $K_{pure} = \frac{K}{\text{fraction of polysomal poly(A) RNA}}$.

[c] $C_{\text{abundance class}} = \frac{K_{\text{globin}} \times C_{\text{globin}}}{K_{pure}}$

The pseudo-first-order rate constant of 600 nucleotide rabbit globin ^{3}H-cDNA hybridized to its template RNA was determined to be 1375 M^{-1} sec^{-1} using our hybridization conditions (data of Larry Lasky). This rate constant is in close agreement to that of 1170 M^{-1} sec^{-1} calculated from the data of Galau et al. (1977) and Chamberlin et al. (1978) for an RNA with a complexity of 1300 nucleotides (C_{globin}).

[d] The number average size of leaf polysomal poly(A) RNA is 1340 nucleotides (Figure 1D). We have determined the number average size of poly(A) on this RNA to be 100 nucleotides (data not shown). The average number of RNA sequences in each abundance class was computed from the relationship: average number of sequences = $C_{\text{abundance class}}/1240$. This estimate of the sequence diversity in polysomal poly(A) RNA should not be considered unique. Other least-squares solutions to the hybridization data presented in Figure 3 are possible. The range of RNA sequences which can be calculated for each abundance class by varying the RMS error of the least-squares solution by <15% are: class 1, 7–10; class 2, 400–1500; class 3, 6000–24,000. A value of 24,000 diverse sequences of rare polysomal poly(A) RNA agrees with the results obtained from the hybridization of ^{3}H-single-copy DNA to total polysomal RNA (Figure 4). In this solution, class 3 represents 23% of the polysomal poly(A) RNA, and each sequence is represented by 5 molecules per cell.

[e] The ratio of RNA to DNA in 1–3 cm tobacco leaves is 14. Since each leaf cell contains 6.6 pg of DNA (Zimmerman and Goldberg, 1977), there are 92 pg of RNA per cell. Of this amount, 40% is cytoplasmic polysomal RNA, the remainder being nuclear, chloroplast, transfer and nonpolysomal RNA. Hence there are 38 pg of polysomal RNA per cell. Approximately 1% of this RNA is polysomal poly(A) RNA, averaging 1340 nucleotides in length, or a total of 5×10^5 polysomal poly(A) molecules per cell. Using this value, the average number of molecules per cell in each abundance class was calculated from the relationship:

$$\text{Number of molecules} = \frac{\text{fraction of polysomal Poly(A) RNA} \times 5 \times 10^5}{\text{number of 1240 nucleotide sequences}}.$$

2.8% of the polysomal RNA mass should be mRNA (see Table 2, note[c]), but actually isolate only 1% poly(A) RNA. Because the sequence relationship between plant poly(A) and poly(A-) mRNA is not yet known, a significant fraction of the mRNA complexity could be contained in the nonpolyadenylated class. To obtain a more precise estimate of the leaf mRNA sequence complexity, an excess of total polysomal RNA was hybridized to the ^{3}H-single-copy DNA preparation described in Figure 2. The amount of ^{3}H-single-copy DNA which hybridized to the RNA at saturation then provided a direct estimate of the mRNA complexity (Galau, Britten and Davidson, 1974). The results of this experiment are presented in Figure 4 (closed circles) and are summarized in Table 2.

At saturation (RNA Cot >40,000), 1.86% of the ^{3}H-single-copy DNA hybridized to the polysomal RNA (Figure 4A). To use this value to estimate the mRNA complexity, it was important to establish that all of the hybridization was to single-copy sequences in the tracer, and not to a minor repetitive DNA contaminant (although none was revealed by the ^{3}H-single-copy reassociation kinetics presented in Figure 2). Polysomal RNA was hybridized to the ^{3}H-single-copy DNA (RNA Cot of 30,000) and the ^{3}H-DNA in DNA/RNA hybrids was isolated as described in Experimental Procedures. The tracer DNA which hybridized was then annealed with an excess of total tobacco driver DNA. The results of this experiment are presented in Figure 2 (open triangles). There is no evidence of any repetitive sequence contamination since the hybridized ^{3}H-DNA reassociated with kinetics identical to those of the parental tracer (closed circles). Hence all of the polysomal RNA hybridization shown in Figure 4 was with single-copy DNA sequences.

After correcting the amount of ^{3}H-single-copy DNA which hybridized to the polysomal RNA for asymmetric transcription and tracer reactability, we calculate that 5.2% of the single-copy DNA in the tobacco genome is represented in leaf polysomes (Table 2). This corresponds to 3.33×10^7 nucleotides of mRNA transcripts or approximately 27,000 diverse average-sized sequences. This value is roughly 2.5 times higher than our best estimate of the poly(A) mRNA sequence diversity (Table 1), but within the range of values permitted by the polysomal poly(A)-^{3}H-cDNA hybridization kinetics (Figure 3). We cannot state with certainty, therefore, what proportion of the total mRNA sequence diversity is contributed by the poly(A) or poly(A-)

Table 2. Sequence Complexity of Polysomal and Nuclear RNA by Hybridization to ^{3}H-Single-Copy DNA

RNA	Saturation Value[a]	Corrected Saturation Value[b]	Complexity (Nucleotides)[c]	K_{obs} (M^{-1} sec^{-1})	K_{exp} (M^{-1} sec^{-1})[d]	F[e]	Number of Molecules per Cell per Sequence[f]	RMS Error
Polysomal	1.86	5.20	3.33×10^7	8.19×10^{-5}	1.4×10^{-2}	5.9×10^{-3}	12	3.8%
Nuclear	8.02	18.65	1.19×10^8	1.46×10^{-4}	3.1×10^{-3}	4.7×10^{-2}	2	4.4%

[a] Represents the terminal values obtained from the least-squares solution of the hybridization data presented in Figure 4. The standard deviation, calculated from the plateau hybridization values (RNA Cot > 40,000) presented in Figure 4, was 1.86 ± 0.12% for polysomal RNA and 8.02 ± 0.61% for nuclear RNA.

[b] Corrected saturation value = $\frac{\text{saturation value} \times 2}{\text{reactability of } ^3\text{H-single-copy DNA}}$.

This value corrects for asymmetric transcription and reactability of the ^{3}H-single-copy DNA preparation used (see Experimental Procedures). The ^{3}H-single-copy DNA reactability was 86% and 71.5% for the nuclear and polysomal hybridization curves, respectively.

[c] C_{RNA} = corrected saturation value × 6.4×10^8.
The single-copy complexity of the tobacco genome is 6.4×10^8 nucleotide pairs (Zimmerman and Goldberg, 1977).

[d] Pseudo-first-order rate constant expected for an RNA population of known complexity (Galau et al., 1977; Chamberlin et al., 1978). Calculated from the relationship:

$$K_{exp} = \left(\frac{5374 \times 200}{C_{RNA}}\right) \times \left(\frac{300}{L}\right)^{1/2}.$$

5374 nucleotides is the complexity of a population of φX174 RNA, and 200 M^{-1} sec^{-1} is the pseudo-first-order rate constant for a hybridization reaction with excess φX174 RNA and 300 nucleotide tracer RF DNA (Galau et al., 1977). L is the mass average size of the tobacco dirver RNA, and 300 is the DNA tracer length (Chamberlin et al., 1978). Values of 1700 and 2600 nucleotides were used for the polysomal and nuclear RNA, respectively.

[e] Fraction of nuclear or polysomal RNA represented in this complexity class. Calculated from the ratio of K_{obs} to K_{exp} (Galau et al., 1974; Hough et al., 1975). The average leaf polysome contains 8–10 ribosomes (Figure 1A). The molecular weight of 18S and 25S rRNA is 2×10^6 daltons (Leaver and Key, 1970). The mass average weight of poly(A) polysomal RNA is 575,000 daltons (1700 nucleotides; data not shown). Hence 2.8% of the polysomal RNA mass is mRNA. The fraction of polysomal mRNA which is driving the ^{3}H-single-copy DNA, therefore, is 21%.

[f] Number of molecules = $\frac{F \times \text{pg of RNA} \times 6 \times 10^{23} \text{ nucleotides per mole}}{C_{RNA} \times 339 \text{ g/mole nucleotide}}$.

Approximately 4% of the cellular RNA is in the nucleus. Thus leaf nuclei contain 3.68 pg of RNA (0.04 × 92 pg). There are approximately 38 pg of polysomal RNA per leaf cell (Table 1).

mRNA populations.

The ^{3}H-single-copy DNA hybridized to the polysomal RNA as a pure pseudo-first-order reaction (Figure 4B), suggesting that the reacting mRNA sequences are present in leaf cells in about the same concentration (that is, they consist of an abundance class). To estimate the fraction of mRNA represented in this class, we calculated the rate constant expected for an RNA population having a complexity of 3.33×10^7 nucleotides (K_{exp}). Using this value and that of the observed rate constant (K_{obs}), we computed the fraction of polysomal RNA driving the ^{3}H-single-copy DNA into DNA/RNA hybrids (Galau et al., 1974). These calculations, presented in Table 2, indicate that 0.59% of the polysomal RNA drives the hybridization reaction. Since we estimated that only 2.8% of the polysomal RNA is mRNA, however, 21% $\left(\frac{0.59}{2.8} \times 100\right)$ of the mRNA is contained in this abundance class. Although these calculations should not be considered absolute, they serve to illustrate that the vast majority of the leaf mRNA diversity is contained in a small fraction of the mRNA mass.

Sequence Complexity of Nuclear RNA

Up to this point, we have considered only the polysomal mRNA of leaf cells. The possibility existed, however, that a more complex class of RNA sequences was contained in leaf nuclei (that is, hnRNA). To demonstrate whether higher plants contain hnRNA, an excess of leaf nuclear RNA was hybridized to ^{3}H-single-copy DNA. The results of this experiment are presented in Figure 4 (open circles) and an analysis is given in Table 2.

In sharp contrast to the results with polysomal RNA, 8% of the ^{3}H-single-copy DNA hybridized to the nuclear RNA at RNA Cot values >40,000 (Figure 4A). Since there was no further hybridization beyond RNA Cot 40,000, and since the ^{3}H-single-copy DNA contained no repetitive contamination (Figure 2), this value can be used to estimate the nuclear RNA sequence complexity. After correcting for asymmetric transcription and tracer reactability (Table 2), we calculate that almost 19% of the tobacco single-copy DNA is represented in the leaf nuclear RNA population. This value is 3.6 times higher than that obtained with the polysomal RNA and corresponds to 1.19×10^8 nucleotides of diverse nuclear RNA transcripts.

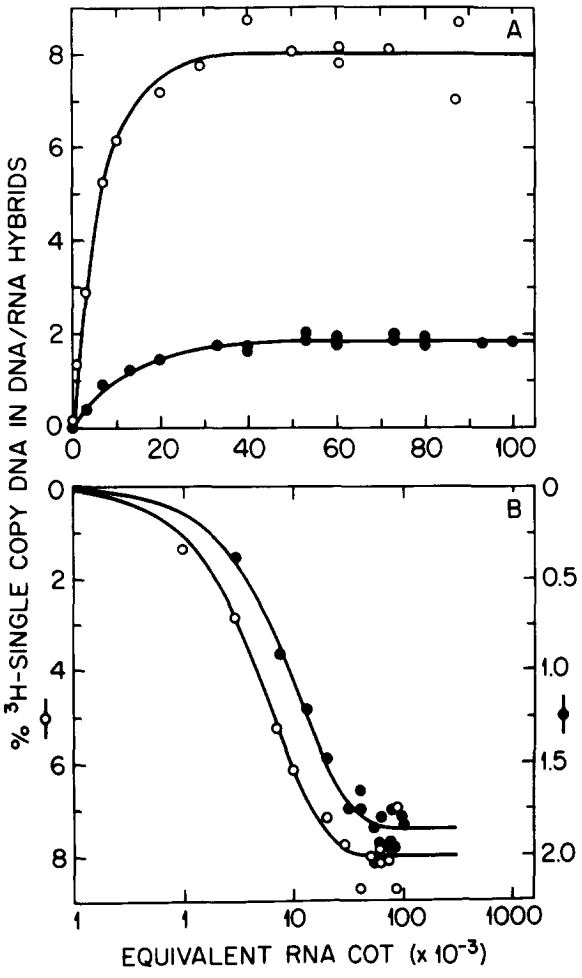

Figure 4. Hybridization of ³H-Single-Copy DNA to Polysomal and Nuclear RNA

Trace amounts of ³H-single-copy DNA were hybridized to a 2200 fold excess of polysomal RNA (●) or a 4400 fold excess of nuclear RNA (○), and analyzed by hydroxyapatite chromatography as described in Experimental Procedures. The solid curves through the data points represent the best least-squares solution for one pseudo-first-order component. The kinetic parameters and an analysis of these hybridization curves are presented in Table 2. (A) Linear plot; (B) semi-logarithmic plot.

The nuclear RNA hybridization reaction exhibits pure pseudo-first-order kinetics (Figure 4B). Each nuclear sequence in the complex class is therefore present in the nucleus in about the same concentration. We calculate that 4.7% of the nuclear RNA drives the ³H-single-copy hybridization reaction (Table 2). This corresponds to roughly two molecules per cell of each nuclear RNA sequence in this abundance class.

Discussion

DNA Sequences Represented in the Structural Gene Set

Two experiments presented in Figure 2 demonstrate that the vast majority of the diverse structural genes expressed in the leaf are contained within the single-copy DNA class. First, ³H-cDNA transcribed from poly(A) mRNA was shown to be homologous to single-copy sequences in the tobacco genome. Second, ³H-DNA isolated from total mRNA/DNA hybrids reassociated with a rate expected for tobacco single-copy DNA. In both of these experiments, up to 5% of the tracer DNA could consist of repeated sequences. We must emphasize, however, that there is no proof for the presence of repetitive DNA in either the ³H-cDNA or the selected ³H-DNA population.

Our previous studies on the DNA sequence organization in tobacco (Zimmerman and Goldberg, 1977), revealed that >60% of the chromosomal DNA is reiterated and has a relatively high sequence complexity (3×10^6 nucleotide pairs). Although some of the repeated DNA consists of short interspersed sequences, the majority are long and nondivergent and therefore good candidates for clustered repetitive genes. The striking feature of the results presented here is that few, if any, of these sequences are contained in the structural gene set active in the leaf.

Number of Structural Genes Expressed in the Leaf

The hybridization reaction between polysomal RNA and ³H-single-copy DNA (Figure 4) revealed that the complexity of leaf mRNA is 3.33×10^7 nucleotides. This corresponds to about 27,000 diverse structural gene transcripts actively being translated on leaf polysomes. Due to uncertainties in the analysis of the polysomal poly(A) RNA-³H-cDNA hybridization experiment presented in Figure 3, we are unable to estimate with precision what fraction of the structural genes code for poly(A) mRNA. Certainly, however, this value is not <25% (Table 1).

The leaf is one of the three major organ systems in a higher plant. As such, it is composed of a number of differentiated cell types, each performing a specialized function. Because of this cellular diversity, we cannot state the number of structural genes which are expressed in any individual cell type. The preponderance of leaf cells (>90%), however, are parenchyma engaged in the major metabolic processes of the leaf. Since mRNA transcripts from the vast majority of structural genes are represented about 12 times per leaf cell (Table 2), our measurements probably reflect primarily the structural gene activity in these cells.

Abundance of Leaf Polysomal mRNA

Our best approximation of the hybridization data presented in Figure 3 is that there are three abundance classes in the poly(A) mRNA population (Table 1). The existence of a unique low abundance

class is implicit in the pseudo-first-order hybridization kinetics of the ^{3}H-single-copy experiment presented in Figure 4. Although we have no direct evidence for the existence of two discrete abundant poly(A) mRNA classes, such classes have been demonstrated in other higher eucaryotic cells (Hastie and Bishop, 1976; Axel, Feigelson and Schutz, 1976).

Abundant mRNA molecules in the leaf comprise 60–75% of the poly(A) mRNA (Table 1) and 75% of the total mRNA (Table 2). We estimate from the ^{3}H-cDNA hybridization reaction presented in Figure 3 that 400–1500 diverse structural gene transcripts are represented in the abundant poly(A) mRNA population (Table 1). If we assume that this estimate reflects that of the entire abundant mRNA population, only 2–6% of the active structural genes in the leaf contribute to the bulk of the mRNA mass.

Quantitative levels of specific mRNA species correlate well with the cellular concentration of proteins they code for (Axel et al., 1976; Paterson and Bishop, 1977; Hynes et al., 1977). A major function of the leaf is to manufacture food for the plant. Due to this function, a large fraction of leaf proteins are associated with photosynthetic activities. It might be expected, therefore, that the abundant mRNA population performs a central role in this process.

Complexity of Leaf Nuclear RNA

The nuclear RNA hybridization experiment presented in Figure 4 indicates that leaf nuclear RNA has a complexity of 1.19×10^8 nucleotides, or roughly 4 times that of polysomal mRNA (Table 2). This result is strikingly similar to that obtained with metazoan cells (Getz et al., 1975; Hough et al., 1975; Liarakos et al., 1975; Ryffel, 1976; Bantle and Hahn, 1976; Levy et al., 1976; Kleiman et al., 1977), but very different from the situation in simple eucaryotes where the ratio of nuclear to polysomal mRNA complexity is 1 (Hereford and Rosbash, 1977; Timberlake et al., 1977). At present, we do not know the size distribution of hnRNA in leaf nuclei. Preliminary experiments suggest that the mass average size of nuclear poly(A) RNA is 21S or 2600 nucleotides. Due to potential degradation during nuclear isolation, however, this must be regarded as a minimum estimate. Thus leaf poly(A) nuclear RNA is larger than poly(A) mRNA, but precisely how much larger we cannot state with certainty.

Transcription in Plant and Animal Cells

The biology of higher plants is vastly different from that of animals. Plants synthesize their own food, have an alternation of spore- and gamete-producing generations, respond developmentally to environmental fluctuations, have cells which are totipotent and possess numerous other specific biological characteristics. Despite the enormous evolutionary distance separating these two kingdoms, a major conclusion of the results presented here is that there is a striking resemblance in the transcriptional processes and levels of gene activity in metaphytan and metazoan cells. Coupled with the general similarity in DNA sequence organization, these results imply that processes which regulate gene expression at the transcriptional level are the same in both kingdoms of higher eucaryotes. What these processes are, and in particular, what unique biological events interface with these processes in higher plants, remain to be learned.

Experimental Procedures

Growth of Plants

Nicotiana tabacum L. cv "Samsun" was grown under standard greenhouse conditions. Developmental uniformity was assured by harvesting 1–3 cm leaves, surrounding the apical meristem, from 15–20 cm plants. All harvests were at noon.

Preparation of Polysomal RNA

Leaf polysomes were prepared according to the procedure of Jackson and Larkins (1976), except that 2% Triton X-100 was incorporated into the extraction buffer. To remove any contaminating nuclear ribonucleoprotein, the procedure of Goldberg et al. (1973) was used. Preparation of purified polysomal RNA was according to previously published procedures (Timberlake et al., 1977; Rozek et al., 1978). Poly(A) RNA prepared according to these procedures was determined to be >95% poly(A) RNA by hybridization to ^{3}H-poly(U) (Bishop et al., 1974a).

Preparation of Nuclear RNA

Leaf nuclei were prepared according to the following procedure. Deribbed leaves were suspended in 5 vol of buffer H [500 mM sucrose, 10 mM Tris-HCl (pH 7.6), 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 0.1% diethylpyrocarbonate] and homogenized at maximum speed in the Omni-Mixer (Sorvall) for 30 sec at 0°C (Hamilton, Kunsch and Temperli, 1972). The homogenate was filtered through two layers of boiled Miracloth, and the nuclei were pelleted for 10 min at 1000 × g. The pellet (containing nuclei and chloroplasts) was resuspended in buffer H + 0.5% Triton X-100, and the nuclei were again pelleted and resuspended in the same buffer. The suspension was then layered over a 20 ml cushion of 1.8 M sucrose containing buffer H + 0.5% Triton X-100, and the nuclei were pelleted for 1 hr at 25,000 × g in the Sorvall HB-4 rotor. The final pellet contained nuclei and starch granules, and was judged free of visible cytoplasmic contamination by phase-contrast microscopy. The nuclei contained approximately 4% of the total cellular RNA.

Nuclear RNA was extracted according to a modification of the procedure of Penman (1966). The final nuclear pellet was resuspended in HSB [500 mM NaCl, 50 mM $MgCl_2$, 10 mM Tris-HCl (pH 7.6), 500 μg/ml heparin], and the DNA was digested with 50 μg/ml of DNAase for 3 min at 0°C. The solution was then adjusted to 1% SDS and deproteinized with 500 μg/ml Proteinase K for 3 hr at room temperature, and the nucleic acids were precipitated with EtOH. The precipitate (containing RNA and low molecular weight DNA) was collected by centrifugation and resuspended in LSB (HSB with 100 mM NaCl), and the DNA was further digested with 50 μg/ml of DNAase at room temperature for 30 min. DNAase was removed by treatment with Proteinase K and SDS, the mixture was deproteinized twice with chloroform and the nucleic acids were precipitated with sodium acetate and 100% EtOH. RNA was separated from deoxyoligonucleotides by Sephadex G-100 chro-

matography. Greater than 80% of the nuclear RNA was in the exclusion peak.

At this stage, the RNA was free of DNA and proteins, but was contaminated with a small amount of pigment. To remove this contaminant, purification was continued using a modification of the CsCl procedure of Glisin, Crkvenjakov and Byus (1972). Precipitated RNA was pelleted, dried and dissolved in 10 mM Tris-HCl (pH 7.6), 0.1 mM EDTA, 0.1% sodium dodecyl sarcosinate (TESar buffer). Solid CsCl (Optipur, EM Labs) was added (0.5 g/ml), and the solution was layered over a 1.2 ml cushion of 5.7 M CsCl (ρ = 1.701 g/ml) containing TESar buffer. The solutions were centrifuged for 45 hr at 35,000 rpm in the Sorvall AH650 rotor at 25°C. Using this procedure, RNA molecules (ρ = 1.9 g/ml) with sedimentation coefficients >6S will pellet, while polysaccharides (ρ = 1.66 g/ml) will not. The unpigmented pellet containing >95% of the input RNA was resuspended in TESar buffer, and residual CsCl was removed by Sephadex G-100 chromatography. The excluded RNA was precipitated with sodium acetate and EtOH.

Hybridization to ^{3}H-Poly(U)

Polysomal RNA was hybridized to ^{3}H-poly(U) (Miles; 10 μCi/μmole P) according to the procedure of Bishop et al. (1974a). The number average size of polysomal poly(A) RNA was computed as previously described (Timberlake et al., 1977; Rozek et al., 1978).

Preparation of ^{3}H-cDNA

^{3}H-cDNA was synthesized from polysomal poly(A) RNA according to the procedure of Rozek et al. (1978). The ^{3}H-cDNA had a modal size in alkaline sucrose of 640 nucleotides and a specific activity of 4×10^6 cpm/μg. Approximately 0.14% of the ^{3}H-cDNA bound to hydroxyapatite at 60°C in 0.12 M PB [PB is an equimolar mixture of NaH_2PO_4 and Na_2HPO_4 (pH 6.8)] at a Cot (mole nucleotide L^{-1} sec) of $<10^{-5}$.

Preparation of ^{3}H-Single-Copy DNA

Total leaf DNA was prepared according to the procedure of Zimmerman and Goldberg (1977) and sheared to a modal single-stranded length of 390 nucleotides in the Virtis 60 homogenizer (Britten, Graham and Neufeld, 1974). This DNA was used for the isolation of single-copy DNA as well as driver in all DNA excess reassociation reactions.

Unlabeled single-copy DNA was prepared by several cycles of renaturation and hydroxyapatite fractionation (Zimmerman and Goldberg, 1977). After the final cycle of annealing and hydroxyapatite fractionation, the single-stranded DNA was reassociated to Cot 15,000 and labeled in vitro with ^{3}H-TTP (New England Nuclear; 55 Ci/mmole), using E. coli DNA polymerase I (Boehringer-Mannheim grade I) and a protocol developed by Galau et al. (1976). Foldback regions were removed by two cycles of annealing to Cot $<10^{-5}$ and hydroxyapatite fractionation at 50°C (Galau et al., 1976). ^{3}H-single-copy DNA had a single-strand fragment length of 300 nucleotides in alkaline sucrose and a specific activity of 1×10^7 cpm/μg. Before using the ^{3}H-single-copy DNA in a hybridization reaction with RNA, its reactability was assayed by annealing to Cot 50,000 in the presence of an excess of unlabeled driver DNA. At the time of labeling, approximately 65% of the ^{3}H-single-copy DNA bound to hydroxyapatite at this Cot.

DNA/DNA Reassociation

The specific procedures we use for DNA reassociation experiments have been described (Zimmerman and Goldberg, 1977; Goldberg, 1978).

DNA/RNA Hybridization with ^{3}H-cDNA

^{3}H-cDNA was mixed with a 1000 fold excess of polysomal poly(A) RNA, denatured at 100°C and incubated in 0.3 M NaCl-0.01 M PIPES (pH 6.7), 0.1 mM EDTA, 0.1% SDS at 64°C to various RNA Cot values. All RNA Cot values have been corrected to equivalent RNA Cot by multiplying by a factor of 2.32 (Britten et al., 1974). Reactions were terminated by freezing in liquid N_2. The mixture was thawed rapidly at 60°C and diluted into a 2.2 ml solution of 0.3 M NaCl, 0.03 M sodium acetate (pH 4.5), 1.8 mM $ZnSO_4$ containing 10 μg/ml of 300 nucleotide single-stranded DNA (Leong et al., 1972). This mixture was then split into two 1 ml aliquots, 400 units of S1 nuclease (Sigma) were added to one and both were incubated for 45 min at 40°C. The reactions were terminated by chilling to 4°C; 20 μg of yeast tRNA were added, and both samples were then analyzed for cold TCA-precipitable radioactivity. The fraction of ^{3}H-cDNA which hybridized to the poly(A) RNA was calculated by taking the ratio of TCA-precipitable cpm in the aliquot with S1 nuclease to the aliquot without S1 nuclease.

DNA/RNA Hybridization with ^{3}H-Single-Copy DNA

^{3}H-single-copy DNA was mixed with a >2000 fold excess of polysomal or nuclear RNA, denatured at 107°C and annealed in 0.5 M PB, 0.1 mM EDTA, 0.2% SDS at 68°C to various RNA Cot values. The Cot values were corrected to equivalent RNA Cot by multiplying by a factor of 5.82 (Britten and Smith, 1970). Generally, each reaction mixture contained approximately 50,000 cpm of ^{3}H-DNA and 20 μg of RNA, adjusted to an RNA concentration of 10–15 mg/ml. Hybridization reactions were carried out in sealed, silicladed capillary tubes and terminated by freezing in liquid N_2.

The fraction of ^{3}H-DNA in DNA/RNA hybrids was determined by a modified "two-column" procedure of Galau et al. (1976). This protocol is tedious, but consistently yields more precise and reproducible data than others which we have tried. Frozen hybridization mixtures were thawed rapidly at 60°C, diluted into 2 ml of 0.05 M PB and dispersed at 37°C. The dilution was split into two 0.9 ml aliquots, one of which was used to assay ^{3}H-DNA in both DNA/DNA duplexes and DNA/RNA hybrids (sample T), and the other to assay ^{3}H-DNA in DNA/DNA duplexes only (sample D). Sample T was adjusted to 0.25 M PB by adding 0.9 ml of 0.45 M PB; RNAase A was added to 10 μg/ml, unhybridized RNA was hydrolyzed for 1 hr at room temperature and the digestion was terminated by freezing in liquid N_2. Unhybridized RNA and RNA in DNA/RNA hybrids were hydrolyzed in sample D by incubation with 10 μg/ml of RNAase A for 15 hr at 37°C. Sample D was then processed analogous to T and frozen in liquid N_2. Both samples were thawed at 37°C, diluted to 0.12 M PB by the addition of 0.01 M PB and deproteinized with chloroform. The organic and aqueous phases were separated by centrifugation, and the aqueous phase was brought to 0.2% SDS. Each sample was then passed over a 1 ml hydroxyapatite column at 60°C, and the ^{3}H-DNA which bound to hydroxyapatite was determined by elution at 98°C. The amount of ^{3}H-DNA in DNA/RNA hybrids was calculated by subtracting the fraction of ^{3}H-DNA binding to hydroxyapatite in sample D from that binding in sample T.

^{3}H-DNA from polysomal RNA/DNA hybrids was isolated according to the procedure of Galau et al. (1974) incorporating the modifications described above.

Computer Analysis

The kinetic parameters describing the DNA/DNA reassociation and DNA/RNA hybridization curves were obtained with the aid of a computer program which uses a least-squares procedure designed for the analysis of both second-order and pseudo-first-order reactions (Britten et al., 1974; Galau et al., 1976; Pearson, Davidson and Britten, 1977).

Acknowledgments

We are indebted to Mr. Larry Lasky for performing the hybridization experiment with rabbit globin mRNA and for many invaluable suggestions. We thank Dr. Glenn Galau for advice on the single-copy hybridization experiments. This research was supported by a National Science Foundation Grant to R.B.G.

The costs of publication of this article were defrayed in part by

the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received January 7, 1978; revised February 24, 1978

References

Axel, R., Feigelson, P. and Schutz, G. (1976). Cell *7*, 247–254.

Bantle, J. A. and Hahn, W. E. (1976). Cell *8*, 139–150.

Bishop, J. O., Rosbash, M. and Evans, D. (1974a). J. Mol. Biol. *85*, 75–86.

Bishop, J. O., Morton, J. G., Rosbash, M. and Richardson, M. (1974b). Nature *250*, 199–204.

Britten, R. J. and Smith, J. (1970). Carnegie Institution Washington Yearbook *68*, 378–386.

Britten, R. J., Graham, D. E. and Neufeld, B. R. (1974). Methods in Enzymology, *29E*, S. P. Colowick and N. O. Kaplan, eds. (New York: Academic Press), pp. 363–418.

Chamberlin, M. E., Galau, G. A., Britten, R. J. and Davidson, E. H. (1978). Nucl. Acids Res., in press.

Darnell, J. E., Jelinek, W. R. and Molloy, G. R. (1974). Science *181*, 1215–1221.

Davidson, E. H. and Britten, R. J. (1973). Quart. Rev. Biol. *48*, 565–613.

Davidson, E. H., Hough, B. R., Amenson, C. S. and Britten, R. J. (1973). J. Mol. Biol. *77*, 1–23.

Davidson, E. H., Klein, W. H. and Britten, R. J. (1977). Dev. Biol. *55*, 69–84.

Firtel, R. A. and Lodish, H. F. (1973). J. Mol. Biol. *79*, 295–314.

Flavell, R. B., Bennett, M. O., Smith, J. B. and Smith, D. B. (1974). Biochem. Genet. *4*, 257–269.

Galau, G. A., Britten, R. J. and Davidson, E. H. (1974). Cell *2*, 9–21.

Galau, G. A., Britten, R. J. and Davidson, E. H. (1977). Proc. Nat. Acad. Sci. USA *74*, 1020–1023.

Galau, G. A., Klein, W. H., Davis, M. M., Wold, B. J., Britten, R. J. and Davidson, E. H. (1976). Cell *7*, 487–505.

Gelderman, A. H., Rake, A. V and Britten, R. J. (1971). Proc. Nat. Acad. Sci. USA *68*, 172–176.

Getz, M. J., Birnie, G. D., Young, B. D., MacPhail, E. and Paul, J. (1975). Cell *4*, 121–129.

Glisin, V., Crkvenjakov, R. and Byus, C. (1974). Biochemistry *13*, 2633–2637.

Goldberg, R. B. (1978). Biochem. Genet. *16*, 45–68.

Goldberg, R. B., Galau, G. A., Britten, R. J. and Davidson, E. H. (1973). Proc. Nat. Acad. Sci. USA *70*, 3516–3520.

Gray, R. E. and Cashmore, A. R. (1976). J. Mol. Biol. *108*, 595–608.

Hamilton, R. H., Kunsch, U. and Temperli, A. (1972). Anal. Biochem. *49*, 48–57.

Hastie, N. D. and Bishop, J. O. (1976). Cell *9*, 761–774.

Hereford, L. M. and Rosbash, M. (1977). Cell *10*, 453–462.

Herman, R. C., Williams, J. G. and Penman, S. (1976). Cell *7*, 429–437.

Hough, B. R., Smith, M. J., Britten, R. J. and Davidson, E. H. (1975). Cell *5*, 291–299.

Hynes, N. E., Groner, B., Sippel, A. E., Nguyen-Huu, M. C. and Schutz, G. (1977). Cell *11*, 923–932.

Jackson, A. O. and Larkins, B. A. (1976). Plant Physiol. *57*, 5–10.

Key, J. L. and Silflow, C. (1975). Plant Physiol. *56*, 364–369.

Kleiman, L., Birnie, G. D., Young, B. D. and Paul, J. (1977). Biochemistry *16*, 1218–1223.

Leaver, C. J. and Key, J. O. (1970). J. Mol. Biol. *49*, 671–680.

Leong, J., Garapin, A., Jackson, N., Fanshier, L., Levinson, W. and Bishop, J. M. (1972). J. Virol. *9*, 891–902.

Levy, B., Johnson, C. B. and McCarthy, B. J. (1976). Nucl. Acids Res. *3*, 1777–1789.

Liarakos, C. B., Rosen, J. M. and O'Malley, B. W. (1973). Biochemistry *12*, 2809–2816.

Melli, M., Whitfield, C., Rao, K. V., Richardson, M. and Bishop, J. O. (1971). Nature New Biol. *231*, 8–12.

Paterson, B. M. and Bishop, J. O. (1977). Cell *12*, 751–765.

Pearson, W. R., Davidson, E. H. and Britten, R. J. (1977). Nucl. Acids Res. *4*, 1727–1737.

Penman, S. (1966). J. Mol. Biol. *17*, 117–130.

Ragg, H., Schroder, J. and Hahlbrock, K. (1977). Biochim. Biophys. Acta *474*, 226–233.

Raghavan, V. (1976). Experimental Embryogenesis in Vascular Plants (New York: Academic Press).

Rogers, M. E., Loening, U. E. and Fraser, R. S. S. (1970). J. Mol. Biol. *49*, 681–692.

Rosbash, M., Ford, P. J. and Bishop, J. O. (1974). Proc. Nat. Acad. Sci. USA *71*, 3746–3750.

Rozek, C. E., Orr, W. C. and Timberlake, W. E. (1978). Biochemistry *17*, 716–722.

Ryffel, G. U. (1976). Eur. J. Biochem. *62*, 417–423.

Smith, M. J., Britten, R. J. and Davidson, E. H. (1975). Proc. Nat. Acad. Sci. USA *72*, 4805–4809.

Timberlake, W. E., Shumard, D. S. and Goldberg, R. B. (1977). Cell *10*, 623–632.

Walbot, V. and Dure, L. S. (1976). J. Mol. Biol. *101*, 503–536.

Zimmerman, J. L. and Goldberg, R. B. (1977). Chromosoma *59*, 227–252.