THE COMPLEXITY OF POLY(A)⁺ RNA IN THE ALGA CHLORELLA FUSCA

N. K. McBRIDE* AND C. F. THURSTON†
Microbiology Department, Queen Elizabeth College, Campden Hill Road, London W8 7AH, U.K.

SUMMARY
RNA excess hybridization with radioactively labelled complementary DNA (cDNA) was used to reveal the complexity of poly(A)⁺ RNA from Chlorella fusca var. vacuolata. RNA was tested from cells during photosynthetic exponential growth and during adaptation to heterotrophic growth in the dark on acetate. Both RNA populations were resolved into abundant, intermediate and rare sequence classes. Abundant sequences (200–400 copies per cell) constituted a significantly larger proportion of total poly(A)⁺ RNA in acetate-adapting cells than in exponentially growing autotrophic cells. Both types of RNA contained a rare sequence class of complexity consistent with a composition of about 20,000 different sequences. This indicated a substantially greater complexity of genes expressed in this alga and Euglena than in fungi and slime moulds. Heterologous hybridization, and hybridization with fractionated cDNA, showed that the majority of differences between RNA populations from exponentially growing and adapting cells were changes in relative abundance of groups of sequences, rather than presence of different sets of sequences in the two populations.

INTRODUCTION
Hybridization of excess messenger RNA with labelled complementary DNA previously prepared from it by reverse transcription may be used to resolve the complexity of mRNA populations. Since this was first done for HeLa cell mRNA (Bishop, Morton, Rosbash & Richardson, 1974) a wide range of eukaryotic organisms have been found to generate cytoplasmic poly(A)⁺ RNA populations broadly consisting of message sequences at three distinct relative frequencies: abundant, intermediate and rare sequence classes (reviewed by Lewin, 1980). In this paper we have made such an analysis of poly(A)⁺ RNA from the unicellular green alga Chlorella fusca var. vacuolata.

MATERIALS AND METHODS
Organism and conditions of culture
Chlorella fusca var. vacuolata, Cambridge Culture Collection of Algae and Protozoa strain 211/8p, was grown autotrophically at 25 °C in a mineral medium at pH 6.7 with conditions of aeration and continuous illumination that have all been fully described elsewhere (John, Thurston & Syrett, 1970). Liquid cultures of up to 21 in volume grew exponentially until the cell density was about 10⁸ cells ml⁻¹, which was measured with a haemocytometer. RNA was extracted from autotrophic

* Present address: Courtauld Institute of Biochemistry, Middlesex Hospital Medical School, Cleveland Street, London W1, U.K.
† Author to whom correspondence should be addressed.
N. K. McBride and C. F. Thurston

cultures at 4 to 5 \((10^6)\) cells ml\(^{-1}\) (exponential-autotrophic cells) and at 2 to 3 \((10^7)\) cells ml\(^{-1}\) (light-limited cells). Cultures at about 10\(^7\) cells ml\(^{-1}\) were transferred to the dark and supplied with 25 mM-sodium acetate for 2h prior to RNA extraction (acetate-adapting cells; Dunham & Thurston, 1978).

**Preparation of poly(A)\(^+\) RNA**

Cells broken by grinding under liquid N\(_2\) in a detergent medium gave total RNA by precipitation with ethanol after deproteinization with phenol/cresol. Poly(A)\(^+\) RNA was obtained by affinity chromatography on oligo(dT)-cellulose. These methods have been described previously (Leonard, Dunham & Thurston, 1981).

**Synthesis of cDNA**

This was accomplished essentially as described by Craig *et al.* (1979). In a total volume of 100\(\mu\)l, 4\(\mu\)g of poly(A)\(^+\) RNA were incubated for 5 min on ice with 0.5 \(\mu\)g oligo(dT)\(_{12-18}\), 0.5 mm-dATP, 0.5 mm-dGTP, 0.5 mm-dTTP, 5\(\mu\)g actinomycin D and 100\(\mu\)Ci \([\text{U}^3\text{H}]\)dCTP (19 Ci mmol\(^{-1}\)); 20 units of avian myoblastosis virus reverse transcriptase were added and after 20 min at 46°C the reaction was stopped by addition of 10\(\mu\)l 0.2 m-EDTA and 5 \(\mu\)l 20 \%(w/v) sodium dodecyl sulphate. The mixture was deproteinized with phenol/chloroform (1:1, v/v) and nucleic acid was recovered by precipitation with 3 vol. ethanol after addition of 50 \(\mu\)g *Escherichia coli* transfer RNA and sodium acetate to a concentration of 0.3 M. RNA was degraded by incubation of the dissolved pellet at 70°C for 20 min and cDNA was separated from ribo- and deoxyribonucleotides (and small oligo-deoxyribonucleotides) by chromatography on Sephadex G-50 in which the excluded fraction was collected. cDNA preparations were precipitated with ethanol and stored as solutions in sterile distilled water at -70°C. cDNA specific activity was about 10\(^6\) c.p.m. ng\(^{-1}\).

**RNA-excess hybridization**

The hybridization mixture contained 1–3000 c.p.m. of cDNA in 3\(\mu\)l vol. of 10 mm-Hepes-NaOH, 600 mm-NaCl, 1 mm-EDTA (pH 7.4) sealed in siliconized glass capillaries with RNA at 0.1–1000 \(\mu\)g ml\(^{-1}\). The capillaries were transferred from boiling water to a water bath at 70°C for 30 s to 44 h and washed into 1 ml of 100 mm-sodium acetate, 50 mm-NaCl, 1 mm-ZnSO\(_4\) (pH 4.6) containing 20 \(\mu\)g calf thymus DNA after rapid cooling in an ice/salt bath. Single-stranded nucleic acid was digested by incubation at 37°C for 2 h with 50 \(\mu\)g *Aspergillus oryzae* nuclease S\(_1\) (20 units, Sigma). For each sample 450 \(\mu\)l were incubated with nuclease, and an equal volume without. Undigested cDNA was precipitated in the presence of 50 \(\mu\)g bovine serum albumin and 10 \%(w/v) trichloroacetic acid for 30 min at 0°C. Radioactivity was counted with precipitates collected on 0.45 \(\mu\)m pore size filters (Millipore). Percentage of cDNA in hybrids with RNA was calculated from the ratio of radioactivity precipitated with and without nuclease digestion and was corrected for the small fraction (3–8 \%) of cDNA that resisted nuclease digestion in the absence of RNA.

**Preparation of abundant fraction cDNA by selective hybridization**

cDNA (0.26 \(\mu\)g) and poly(A)\(^+\) RNA (12 \(\mu\)g) were incubated in 0.2 ml of the hybridization buffer mixture described above, at 70°C for 26 min to allow hybridization to an \(R_0\) value of 0.30 mol \(1^{\text{st}}\). The reaction mixture was treated with S\(_1\) nuclease as described above and deproteinized by extraction with phenol. cDNA hybridized to RNA was precipitated with 2 vol. ethanol at -20°C in the presence of 50 \(\mu\)g ml\(^{-1}\) *E. coli* tRNA. The precipitate was dissolved in 0.1 m-NaOH, 1 mm-EDTA and incubated at 70°C for 20 min to hydrolyse RNA, releasing single-stranded cDNA, which was recovered by precipitation with ethanol.

**Numerical analysis of hybridization data**

Two computer programs were used (Pearson, Davidson & Britton, 1977; B. D. Young, personal communication). Both programs employ a least-squares algorithm to fit data to a pseudo-first-order rate equation with the minimum number of components. The kinetic components so derived (see Table 1) were essentially similar whichever program was used, and if the initial parameter estimates were varied the same results were still obtained.
RESULTS

Homologous hybridization analysis of total poly(A)⁺ RNA

Homologous hybridization of cDNA to excess of the RNA population from which it was copied was followed for poly(A)⁺ RNA prepared from cultures under three different physiological conditions (Fig. 1). With preparations from both 'exponential-autotrophic' and 'acetate-adapting' cultures about 70% of cDNA was driven into DNA/RNA hybrids at log \( R_{t} = 2.5 \). Computer-generated least-squares analysis of the data gave best fit with three-component curves in both cases. Table 1 shows the analysis of nucleotide complexity classes so resolved, which indicated a broadly similar distribution. The hybridizations of exponential-autotrophic and acetate-adapting
Table 1. Complexity and abundance classes of C. fusca poly(A) + RNA

<table>
<thead>
<tr>
<th>Physiological condition</th>
<th>Frequency class</th>
<th>Fraction of cDNA hybridizing</th>
<th>$R_{DA}$ (mol l⁻¹ s⁻¹)</th>
<th>Nucleotide complexity (mol)</th>
<th>Number of sequences</th>
<th>Average no. of copies of each sequence (per cell)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponential-autotrophic</td>
<td>Abundant</td>
<td>0.117</td>
<td>0.174</td>
<td>0.0163</td>
<td>3.13 x 10⁴</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>0.464</td>
<td>1.41</td>
<td>0.523</td>
<td>1.00 x 10⁵</td>
<td>667</td>
</tr>
<tr>
<td></td>
<td>Rare</td>
<td>0.418</td>
<td>47.6</td>
<td>15.9</td>
<td>3.05 x 10⁷</td>
<td>20300</td>
</tr>
<tr>
<td>Light-limited</td>
<td>Abundant</td>
<td>0.313</td>
<td>0.149</td>
<td>0.0373</td>
<td>7.16 x 10⁴</td>
<td>48</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>0.406</td>
<td>3.28</td>
<td>1.06</td>
<td>2.03 x 10⁶</td>
<td>1353</td>
</tr>
<tr>
<td></td>
<td>Rare</td>
<td>(0.281)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acetate-adapting</td>
<td>Abundant</td>
<td>0.340</td>
<td>0.296</td>
<td>0.0805</td>
<td>1.54 x 10⁵</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td>0.406</td>
<td>1.94</td>
<td>0.630</td>
<td>1.21 x 10⁶</td>
<td>807</td>
</tr>
<tr>
<td></td>
<td>Rare</td>
<td>0.254</td>
<td>112.6</td>
<td>22.9</td>
<td>4.39 x 10⁷</td>
<td>23900</td>
</tr>
<tr>
<td>Acetate-adapting (abundant)³</td>
<td>(1)</td>
<td>0.185 (0.544)</td>
<td>0.00670</td>
<td>0.291</td>
<td>1.28 x 10⁴</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>(2)</td>
<td>0.155 (0.456)</td>
<td>0.150</td>
<td>0.0547</td>
<td>1.05 x 10³</td>
<td>70</td>
</tr>
</tbody>
</table>

* Physiological conditions of C. fusca cultures from which poly(A)⁺ RNA was isolated are described in Materials and Methods.

b Fraction of cDNA hybridizing was calculated assuming hybridization had been driven to completion at $R_{DA} = 250$ for exponential-autotrophic and acetate-adapting RNA, where about 70% of cDNA radioactivity was $S₁$ nuclease-resistant. As insufficient RNA was available to measure hybridization of light-limited cDNA to completion, it was assumed that the rare sequence class had an intermediate abundance between those measured for the other two conditions, in order to derive values for the fraction of cDNA hybridizing in the abundant and intermediate classes for this condition.

c $R_{DA}$ was corrected for the proportion of total cDNA in each complexity class taken as a single component and for the contamination of driving poly(A)⁺ RNA with ribosomal RNA, which was taken to be 20%. After one passage through oligo(dT)-cellulose, poly(A)⁺ RNA was found to contain 15–20% ribosomal RNA (data not shown) as has been found elsewhere (Hynes et al. 1977).

d Nucleotide complexity as moles of nucleotide residue (average $M_r = 330$) was derived from the $R_{DA}$ for pure globin mRNA–cDNA hybridization (6.88 x 10⁻⁴ mol l⁻¹ s⁻¹) using the data shown in Fig. 1, and 1320 nucleotides as the chain length of this RNA. Average mRNA chain length for C. fusca was taken as 1500 nucleotides based on the data of Leonard et al. (1981), in order to calculate the number of sequences in each complexity class.

*The average number of copies of each sequence =

\[
\text{weight of poly(A)⁺ RNA per cell} \times \frac{\text{fraction hybridizing}}{6 \times 10^{12}} \times 330 \times \text{complexity}
\]

The weight of poly(A)⁺ RNA per cell was taken as 0.055 pg (Leonard et al. 1981).

1 After selection by limited hybridization as described in Materials and Methods, the abundant class from acetate-adapting cells could be resolved further into more (1) and less abundant (2) sub-fractions (see Fig. 4).
poly(A)$^+$ RNA to cDNA were not however identical, as hybridization of acetate-adapting RNA was uniformly greater than for exponential-autotrophic RNA over the range of $R_J$ values tested. This was largely due to the substantially greater proportion of abundant sequences in acetate-adapting RNA (Table 1). Intermediate abundance sequences, in contrast, were very similar and constituted the largest class in all preparations. The relative size of the rare sequence class was larger in the exponential-autotrophic RNA than in the acetate-adapting RNA, but the proportion of poly(A)$^+$ RNA in this class was not obtained with high precision, as insufficient RNA was obtained to drive the hybridization to completion (100% of cDNA radioactivity in hybrids), or to a clear plateau value. Plateau values at 70–80% hybridization have, however, been found commonly in comparable procedures (reviewed by Lewin, 1980) and, notably, even if our estimates of the magnitude of the rare class are too low the number of sequences in the rare class was large when compared with other eukaryotic microorganisms (Hereford & Rosbash, 1977; Rozek, Orr & Timberlake, 1978) with the exception of Euglena, which is described below. The difference in number of sequences in the rare class for exponential-autotrophic and acetate-adapting RNA shown in Table 1 may also be statistically insignificant as the estimate for acetate-adapting RNA was derived from the too small a number of data points.

Although data for hybridization of 'light-limited' RNA are incomplete it is apparent from the $R_J$ curves (Fig. 1) that light-limitation resulted in an increase in abundant

![Graph](image_url)

**Fig. 2.** Heterologous hybridization of acetate-adapting cDNA to exponential-autotrophic poly(A)$^+$ RNA. RNA concentrations used were 60, 300 and 1000 μg ml$^{-1}$. The broken line shows the homologous acetate-adapting cDNA/RNA hybridization from Fig. 1.
sequences when compared with non-light-limited (exponentially growing) cells. The apparent increase in complexity of the intermediate class (Table 1) is likely to be overestimated by the least-squares routines in the absence of data for larger $R_{dG}$ values.

**Heterologous hybridization analysis comparing exponential-autotrophic and acetate-adapting RNA**

In order to test the similarity of composition of poly(A)$^+$ RNA populations from exponential-autotrophic and acetate-adapting cells poly(A)$^+$ RNA from each condition was hybridized to cDNA from the other. When acetate-adapting cDNA was hybridized to excess exponential-autotrophic RNA, the curve was shifted to higher $R_{dG}$ values (Fig. 2) than for the homologous hybridization of this cDNA. The heterologous hybridization showed very similar kinetics to the exponential-autotrophic homologous hybridization. This result could arise if the acetate-adapting sequences were the same as those present in exponential-autotrophic cells but were generally more abundant, or if acetate-adapting cells contained significant quantities of sequences that were not represented in exponential-autotrophic poly(A)$^+$ RNA. These alternatives are not mutually exclusive.

In the reverse combination where exponential-autotrophic cDNA was hybridized to acetate-adapting RNA (Fig. 3), the hybridization was greater than the homologous
hybridization of this cDNA at low $R_d$ values ($<0.3$) and less at high $R_d$ values ($>0.3$). This showed that some abundant-class sequences in exponential-autotrophic RNA were significantly more abundant in acetate-adapting RNA and that many intermediate-class sequences were correspondingly very much less abundant (or absent) in the latter RNA population.

**Hybridization of acetate-adapting abundant-class cDNA**

The abundant sequence class from acetate-adapting cDNA was isolated as described in Materials and Methods. Kinetics of hybridization of this cDNA to excess total poly(A)$^+$ RNA from both acetate-adapting (= homologous) and exponential-autotrophic cells is shown in Fig. 4. This homologous hybridization was resolved by least-squares analysis (Table 1) into two components: a super-abundant sequence class of only about nine different sequences present as about 1400 molecules per cell...
on average, and a less abundant class of about 70 sequences present on average as about 150 copies per cell. In both homologous and heterologous hybridization of this cDNA it was possible to drive >90% of cDNA into hybrids. In consequence the proportions of super-abundant and less-abundant sequences were more precisely derived than the major classes obtained from analysis of total cDNA hybridization. Also, it is evident that essentially all abundant-class sequences in acetate-adapting RNA were present in exponential-autotrophic RNA, but at lower abundance.

**DISCUSSION**

**Complexity of Chlorella poly(A) + RNA in relation to other organisms**

The resolution of cytoplasmic mRNA populations into three complexity classes has been described for a wide range of eukaryotic cells including several eukaryotic microorganisms (reviewed by Lewin, 1980). Our data show that *Chlorella* expresses a greater diversity of sequence than fungi such as yeast (Hereford & Rosbash, 1977) and *Achlya* (Rozek *et al.* 1978) or the cellular slime mould *Dictostelium* (Blumberg & Lodish, 1980), in which the mRNA populations contain 3–4000 sequences. With a total sequence complexity of about 20 000 sequences in poly(A) + RNA, *Chlorella* is more comparable to higher plants such as tobacco (Goldberg, Hoschek, Kamalay & Timberlake, 1978) and *Euglena* (Verdier, 1979). This suggests a much greater evolutionary gap in the development of complexity of gene expression between the fungi and the algae than between the algae and higher plants (or higher animals).

**Poly(A)-containing RNA composition under different physiological conditions**

*Chlorella* suspensions growing exponentially in the light are making the full complement of proteins required to make new biomass from CO₂ (and other inorganic nutrients). As the growth of such cells slows down due to light limitation (as the cultures become sufficiently dense), or when they are forced to adapt to heterotrophic growth in the dark (on acetate), there are changes in the relative rate of synthesis of many proteins and in the spectrum of proteins made *in vitro* when heterologous cell-free translation assays are directed by poly(A)-containing RNA extracted from cells in the three different conditions (John *et al.* 1982; McBride, 1982). The hybridization data presented here show changes in both abundant and intermediate mRNA sequence classes that are largely consistent with such changes in the relative rates of synthesis of many different proteins. A striking biochemical feature of this strain of *Chlorella* is the enormous amount of isocitrate lyase protein made upon adaptation to growth on acetate in the dark, when this enzyme accumulates to constitute about 8% of total soluble protein (John & Syrett, 1968). The mRNA for this protein is very probably one component of the enlarged abundant sequence class found in acetate-adapting cells. It is less clear what other sequences become very abundant in either light-limited or acetate-adapting mRNA populations.

We thank Dr S. Malcolm for much helpful advice, Dr J. W. Beard for a gift of AMV reverse transcriptase, Dr G. Cowling for a gift of purified rabbit globin mRNA and SERC for a studentship for N. K. McB.
REFERENCES


(Received 22 February 1983 – Accepted 12 April 1983)