Sampson, E. J., Fedor, J., Benkovic, P. A., and Benkovic, S. J. (1973), *J. Org. Chem. 38,* 1301.
Sillen, L., and Martell, A. E. (1964), *Chem. Soc. Spec. Publ. No. 25.*
Speakman, J. C. (1940), *J. Chem. Soc.,* 855.
Stadtman, E. R. (1957), *Methods Enzymol. 3,* 228.
Steffens, J. J., Sampson, E. J., Siewers, I. J., and Benkovic, S. J. (1973), *J. Am. Chem. Soc. 95,* 936.
Usher, D. A. (1969), *Proc. Natl. Acad. Sci. U.S.A. 62,* 661.
Usher, D. A., Richardson, D. I., and Eckstein, F. (1970), *Nature (London) 228,* 663.
Westheimer, F. H. (1968), *Acc. Chem. Res. 1,* 70.

Messenger RNA Complexity in *Drosophila melanogaster*†

Beatriz Levy W. and Brian J. McCarthy*

ABSTRACT: Complementary DNA was synthesized as a copy of polyadenylated RNA from the cytoplasm of *Drosophila* cultured cells. The kinetics of hybridization of cDNA with the RNA used as template revealed a complex distribution of frequencies in the population of polyadenylated RNA. Computer simulation suggested three frequency classes containing about 4, 190, and 6700 different RNA molecules of mean molecular weight 4×10^5. About 15% of this complementary DNA reacted with repetitive sequences of *Drosophila* DNA. The most frequent polyadenylated RNA is preferentially enriched in its content of repetitive sequences. Comparative experiments using cDNA synthesized as a complement of larval polyadenylated RNA demonstrated some stage specific changes in the populations of polyadenylated RNA.

During the recent debate concerning the number of genes constituting the eucaryotic genome (Ohno, 1972), particular attention has been directed toward *Drosophila melanogaster.* For this organism, in which the DNA content is only 5% of that of mammals, independent estimates of gene number are possible from cytological, genetic, and biochemical data (Beerman, 1972; Bishop, 1974; Lewin, 1974). The number of cytologically observable bands is between 5000 and 6000 (Bridges, 1938). Since, in limited regions of the chromosome at least, the number of chromomeres correlates well with the number of complementation groups (Judd and Young, 1973; Hochman, 1973), it is tentatively concluded that 5000–6000 is also a reasonable estimate for the total genetic potential of the organism. Other estimates of the number of genes by counting recessive lethals are in reasonable agreement (Shearn et al., 1971; Shearn and Garen, 1974). However, this line of reasoning may be criticized since mutations may not always lead to recessive lethals or to observable morphological effects.

For these reasons, it is valuable to approach the question through biochemical methodology. Molecular hybridization experiments have already established the fact that a major fraction of the *Drosophila melanogaster* genome is transcribed in larvae, pupae, and adults as well as tissue culture cells (Turner and Laird, 1973; McCarthy et al., 1973). However, these data do not bear directly on the number of genes since most of the RNA which hybridized was Hn-RNA,[1] an unknown fraction of which appears in the cytoplasm as messenger RNA. The question of the complexity of the messenger population in *Drosophila* can now be approached directly using the method of Bishop et al. (1974) in which polyadenylated messenger RNA is hybridized with complementary cDNA[1] synthesized in vitro.

Using this approach we present data concerning the complexity of mRNA in larvae, pupae, adults, and Schneider's cells, the repetitive content of this RNA, and the differences which exist among the populations of the RNA of various cells.

Experimental Section

Materials. Drosophila tissue culture cells (Schneider's cells, line 2) were grown as described elsewhere (McCarthy et al., 1973). Third instar larvae, pupae, and adult *Drosophila melanogaster* were kindly provided by Galvin Swift. The AMV DNA polymerase was kindly supplied by Dr. W. J. Rutter.

The radioactive materials were from Schwarz/Mann. Hydroxylapatite (DNA grade) was from Bio-Rad Laboratories.

Methods. PREPARATION OF SCHNEIDER CELLS CYTOPLASMIC RNA. The RNA was prepared as described by Penman (1969). For a typical preparation, 1000 ml of a suspension culture of *Drosophila* cells was collected at 2000 rpm for 20 min at 4°. The cells were washed with an isotonic saline solution and resuspended in 40 ml of 10 m*M* NaCl–10 m*M* Tris-HCl (pH 8.5)–3 m*M* $MgCl_2$. One percent diethyl pyrocarbonate was added to inhibit nucleases. NP40 (Shell Oil Co.) was added to the suspension to a concentration of 0.5% and cells were lysed in a Dounce homogenizer. Nuclei were removed by centrifugation at 3500 rpm for 5 min at 0°.

The cytoplasmic supernatant was made 0.1 *M* in NaCl, 0.01 *M* in EDTA, and 0.5% in sodium dodecyl sulfate

† From the Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143. *Received December 26, 1974.* This work was supported by National Science Foundation Grant No. GB 35704.

[1] Abbreviations used are: Hn-RNA, heterogeneous nuclear RNA; cDNA, cytoplasmic DNA; SDS, sodium dodecyl sulfate; PEB, phosphate equimolar buffer containing 1 m*M* EDTA; SSC, standard saline-citrate.

(SDS), shaken at room temperature with a mixture of phenol–chloroform (1:1), and then reextracted with chloroform containing 4% isoamyl alcohol. After five extractions, the aqueous phase layer was removed and precipitated with 2 vol of absolute alcohol overnight at −20°. The RNA was recovered by centrifugation at 6000 rpm for 30 min at 0°, resuspended in 0.1 *M* NaCl, and reprecipitated with ethanol. RNA was prepared from flies, pupae, and larvae by the same methods.

PREPARATION OF POLY(A) CONTAINING RNA. The RNA was recovered by centrifugation and resuspended in 0.1 *M* NaCl–10 m*M* Tris (pH 7.4) and applied to a column of poly(U) Sepharose, equilibrated with 0.1 *M* NaCl–10 m*M* Tris-HCl (pH 7.4)–0.5% SDS. The unadsorbed material was passed through the column three times and then flushed with 1 column volume of the same solution. The RNA retained by the column was then eluted with a solution containing 90% formamide and 10% of 10 m*M* Tris-HCl (pH 7.4)–10 m*M* EDTA. The unadsorbed material (non-poly(A)-containing) was routinely precipitated with 2 vol of ethanol and stored at −20°. The poly(A) containing RNA eluted from the column was either precipitated with ethanol, for use in cDNA synthesis, or dialyzed overnight at 0° against distilled water, for use in hybridization experiments. In each case the RNA was passed through a Sephadex SP-50 column.

PREPARATION OF cDNA. The complete reaction mixture, 100 μl, contained the following: 50 m*M* Tris-HCl (pH 8.3), 10 m*M* dithiothreitol, 6 m*M* magnesium acetate, 60 m*M* NaCl, 20 μl/ml of actinomycin D, 1 m*M* dATP, 1 m*M* dTTP, 1 m*M* dGTP, 25 μCi of [^{3}H]dCTP (20 Ci/mmol), 14.2 pmol of $(dT)_{50}$, 0.4 μg of polyadenylated RNA, and 0.2 μg of AMV DNA polymerase. After incubation at 37° for 2 hr, 5 μl of 10% SDS was added and incubated for 2 min at 37°. After addition of 25 μg of bacterial DNA as carrier, the cDNA was treated with 1 *N* NaOH for 5 min at 70°, neutralized, and applied to a Sephadex SP-50 column equilibrated with 0.3 *M* NaCl–10 m*M* sodium acetate (pH 5.0). The fractions containing Cl_3CCOOH precipitable radioactivity were pooled and dialyzed overnight at 0° against distilled water.

PREPARATION OF S1 NUCLEASE. S1 nuclease was prepared from crude (Sigma) α-amylase as described (Sutton, 1971). The enzyme preparations were routinely tested for activity against single-stranded and double-stranded DNA standards.

HYBRIDIZATION REACTIONS. Hybridization reactions were carried out in 0.24 *M* PEB (phosphate equimolar buffer containing 1 m*M* EDTA). Samples (5 μl) were sealed in capillaries, boiled for 10 min, and incubated at 70° for various time periods. At the end of each incubation period, duplicate samples were diluted with 4 ml of S1 nuclease buffer (0.3 *M* NaCl–0.03 *M* sodium acetate (pH 4.5)–0.003 *M* $ZnCl_2$). S1 nuclease was added to each sample and incubation was for 2 hr at 37° after which the samples were Cl_3CCOOH precipitated and counted.

In every experiment, controls without S1 nuclease and zero time incubations were included. The zero time control values were subtracted from each value and these corrected values expressed as percent of control without S1 nuclease.

PREPARATION OF *Drosophila* DNA. *Drosophila* DNA was extracted from Schneider's cell nuclei, prepared as described previously (McCarthy et al., 1973). The nuclear pellet was resuspended in 0.15 *M* NaCl, 0.015 *M* sodium citrate, and 1% SDS, and shaken with an equal volume of phenol–chloroform (1:1). The extraction was performed three times, after which 2 vol of cold ethanol was added to the aqueous phase. The DNA was spooled out with a glass rod and dissolved in 10 m*M* Tris–1 m*M* EDTA (pH 8.5).

For depurination and alkali cleavage of the DNA, the solution was made 0.1 *M* in acetate buffer (pH 4.2). The sample was depurinated for 30 min at 70° and then chilled to 0° (Grouse et al., 1972). For each milliliter of reaction mixture, 0.2 ml of 1 *M* NaOH was added and the mixture was incubated at 50° for 10 min. After cooling at 0°, the sample was neutralized by the addition of 0.15 ml of 1 *M* NaH_2PO_4 for each milliliter of solution. The DNA sample was dialyzed for 24 hr at 0° against 0.01 × SSC and stored at −20°.

cDNA FRACTIONATION. Aliquots of cDNA were annealed with mRNA to R_0t of 0.1 or 0.2 mol/l. sec. The cDNA in hybrid form was separated from the unreacted material by adsorption to hydroxylapatite. Reaction mixtures (1 ml), containing 50 μg of mRNA, were diluted with distilled water to a phosphate concentration of 20 m*M* and mixed with 1-ml batches of hydroxylapatite in a conical centrifuge tube. The tubes were shaken at room temperature for 15 min after which the unbound material was removed by centrifugation for 2 min at 1000 rpm. The tubes were then immersed in a water bath at 60° and washed several times with small volumes of 0.14 *M* phosphate and then 0.4 *M* phosphate. Fractions, containing single-stranded or double-stranded cDNA, were pooled and digested for 1 hr at 37° with 0.6 *N* NaOH, neutralized, and precipitated with ethanol.

The fractionated cDNA was recovered by centrifugation at 8000 rpm for 15 min and dissolved in distilled water.

THERMAL DENATURATION OF cDNA/DNA HYBRIDS. A large batch of Schneider's cell cDNA was incubated with mRNA to R_0t of 10^{-1} mol/l. sec. The hybridized and single-stranded cDNAs were separated on hydroxylapatite. Hybridized or single-stranded cDNAs were renatured with unlabeled *Drosophila* DNA in 0.24 *M* PEB at 70° in 5-μl aliquots.

After renaturation, the content of the 5-μl capillaries was diluted to a final concentration of 0.024 *M* phosphate buffer with water. Duplicate samples were incubated for 10 min at different temperatures in a thermoregulated water bath. The samples were challenged with S1 nuclease to measure the percentage of hybrid remaining at each temperature. Samples incubated at 0° were assayed with and without S1 nuclease, to determine the total amount of hybridization.

Results

Preparation of Polyadenylated RNA and Complementary DNA. Schneider's cells were labeled for 24 hr with [^{3}H]uridine and cytoplasmic polyadenylated RNA prepared by binding to poly(U)–Sepharose. The sucrose gradient profiles of the polyadenylated and nonpolyadenylated RNA are superimposed in Figure 1. The distribution of apparent molecular weight of the messenger fraction is described by a broad peak between 10 and 20 S.

The polyadenylated RNA fraction was used as a template for the preparation of cDNA with avian myeloblastosis virus reverse transcriptase. Synthesis was dependent upon the presence of oligo(dT) as primer. The cDNA was employed in various RNA–cDNA hybridization experiments to determine the complexity of Schneider's cell and other *Drosophila* RNA. Elucidation of the complexity of an RNA population is based on analysis of the kinetics of hy-

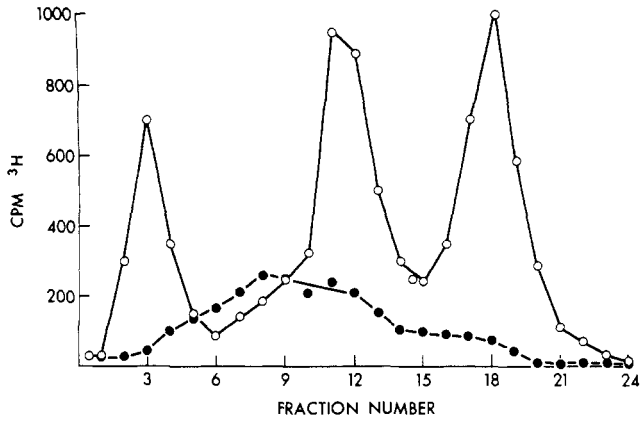

FIGURE 1: Sucrose gradient analysis of *Drosophila* RNA after poly(U)-Sepharose chromatography. Poly(A$^+$) and poly(A$^-$) RNA pellets were resuspended in 0.1 *M* NaCl-10 m*M* Tris-HCl (pH 7.5)-1 m*M* EDTA and centrifuged for 7 hr at 39,000 rpm at 24° on a 15–30% sucrose gradient in the same buffer containing 0.5% SDS. After centrifugation, 0.5-ml fractions were collected. The radioactivity of each fraction was measured by precipitation with trichloroacetic acid: (○) poly(A$^-$) RNA; (●) poly(A$^+$) RNA.

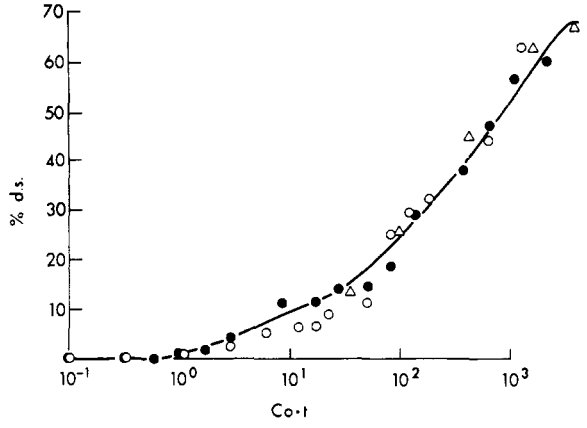

FIGURE 2: Renaturation of total cDNA of Schneider's cells with an excess of unlabeled DNA. Capillaries (5 μl), containing 500 cpm of cDNA and *Drosophila* DNA at a concentration of 1 mg/ml or calf-thymus DNA at a concentration of 4 mg/ml, were incubated in 0.24 *M* PEB[1] at 70° for different lengths of time. Samples were chilled to 0°, diluted with 4 ml of S1 buffer, and digested with S1 nuclease. Three experiments are shown performed with different cDNA and *Drosophila* preparations.

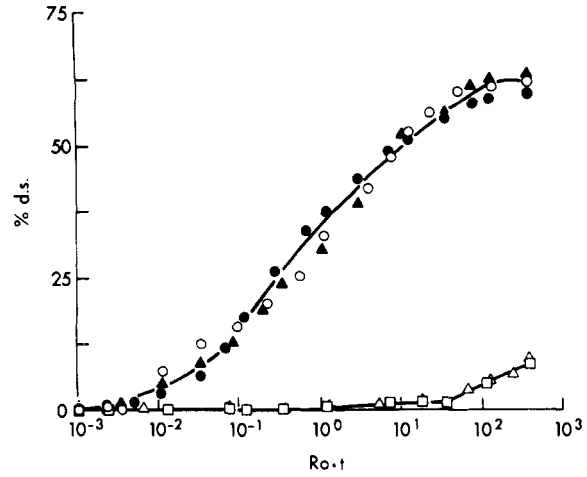

FIGURE 3: Hybridization between *Drosophila* cDNA and its mRNA template or nonpolyadenylated RNA. Two reaction mixtures were used, containing 10 and 100 μg/ml of RNA. Samples (5 μl) each containing 500 cpm of cDNA, were annealed at 70° in 0.24 *M* PEB containing 0.01% SDS and challenged with S1 nuclease. Three experiments are shown, performed with different cDNA and mRNA preparations (○, ▲, ●). Two control incubations with nonpolyadenylated RNA are shown, performed with different cDNA and RNA preparations (□,△).

bridization and comparison of rate constants with those of known standards (Bishop et al., 1974).

Bishop et al. (1974) have used a value of 6×10^{-4} for the $R_0t_{1/2}$ value for hybridization of cDNA, complementary to a mixture of globin α and β mRNA with excess RNA. Using similar hybridization conditions Gambino et al. (1974) and Imaizumi et al. (1973) report a similar value for human and duck globin cDNA. Others have obtained higher values but under somewhat different conditions of hybridization. We have obtained a value of 9×10^{-4} for cDNA made as a complement to purified ovalbumin mRNA kindly supplied by Drs. Savio Woo and Bert O'-Malley (unpublished data). Sullivan et al. (1973) reported earlier a value of about 10^{-3} for ovalbumin cDNA. Taking the combined molecular weight of Hb globin α and β mRNA as 4×10^5 and that for ovalbumin mRNA as 5.2×10^5 (Rosen et al., 1975) these two calibrations are in reasonable agreement. In order to make our calculations of complexity directly comparable with those of Bishop et al. (1974) and Ryffel and McCarthy (1975) we have used the calibration of the former group of authors.

In order to characterize the cDNA, aliquots were first renatured in the presence of excess *Drosophila* DNA. Reproducibility was checked by using three different cDNA preparations synthesized with different batches of polyadenylated RNA as template and three different *Drosophila* DNA preparations (Figure 2). All three experiments displayed essentially identical kinetics. Some 15% of the cDNA reacted rapidly with the kinetics of repetitive sequences (Laird and McCarthy, 1969; Laird, 1971), while the major reaction up to a total of more than 70% occurred with kinetics characteristic of unique sequences. No reaction occurred in the presence of calf thymus DNA. Therefore, we conclude that although most of the cDNA is representative of unique sequences, a substantial fraction seems to be complementary to repetitive DNA.

The complexity of Schneider's cell polyadenylated RNA was estimated by hybridizing it with cDNA. Again three parallel experiments were performed using different cDNA and RNA preparations. Good agreement between them was obtained (Figure 3). Two control experiments were performed using the same cDNA and the nonpolyadenylated RNA which failed to be retained by the poly(U)-Sepharose. In this case hybridization did not commence until a value of R_0t of 10^2 was reached compared to 10^{-2} for polyadenylated RNA. Therefore, the concentration of hybridizable RNA in the flow-through fraction is about 10^{-4} of that in the retained poly(A$^+$) fraction.

Clearly, the hybridization kinetics are complex, extending over 4 log units of R_0t. A computer program designed to obtain best fit to the data (G. S. Swift, manuscript in preparation) was used to obtain a description of the population distribution (Ryffel and McCarthy, 1975). The procedure is essentially analogous to that used by Silverstein et al. (1973) and Bishop et al. (1974). Data obtained as best fit imply three frequency classes with rate constants listed in Table I. Using the calibration of Bishop et al. (1974) and taking 4×10^5 as the mean molecular weight of *Drosophila*

Table I: Different Frequency Classes in *Drosophila* Polyadenylated RNA.[a]

Component	P	$R_0t_{1/2}$ Obsd	$R_0t_{1/2}$ If Pure	No. of 400,000 Dalton Sequences	Molecules of Each Sequence per Cell
1	0.041	0.064	0.0026	4	1×10^4
2	0.380	0.29	0.113	190	230
3	0.220	18.8	4.13	6700	1

[a] The data were obtained from Figure 2. P denotes the fraction of hybridizable cDNA (Bishop et al., 1974). The solution represents best fit as determined by a computer program, designed to solve for three components (G. S. Swift, in preparation).

polyadenylated RNA, approximate frequencies were calculated. The experimental data can be fit to three frequency classes containing approximately 4, 190, and 6700 different sequences. Obviously more complex solutions may be obtained but this triphasic solution represents a useful approximation.

Fractionation of Complementary DNA. In order to verify that individual cDNAs are represented in widely disparate proportions in the total population, we attempted to fractionate cDNA through incomplete hybridization with RNA. cDNA was incubated with RNA to R_0t of 0.2 where approximately 25% reaction had occurred. Hybridized and unreacted moieties were recovered by hydroxylapatite chromatography and again incubated with Schneider's cell polyadenylated RNA. The kinetics of reaction were markedly distinct. Reaction of previously hybridized cDNA was essentially complete by R_0t = 1, while that for unreacted cDNA was displaced by approximately 1.5 log units (Figure 4). The difference between the two reactions is even more apparent when the data are plotted in a form linear in R_0t (Figure 5). As demonstrated for HeLa RNA by Bishop et al. (1974), this form of presentation reveals the three transitions and values of $R_0t_{1/2}$ apparent from the computer simulation (Table I).

Reaction of cDNA with cellular DNA suggested that some cDNA was complementary to repetitive DNA. Therefore, it is reasonable to inquire whether this repetitive DNA is preferentially responsible for the most abundant molecules of polyadenylated RNA. If this is the case, cDNA fractionated to obtain complements of the most abundant RNA would react preferentially with repetitive sequences in *Drosophila* DNA. This possibility was tested by using aliquots of fractionated cDNA used in the experiments illustrated in Figures 4 and 5. In effect the prediction was verified: more than 30% of the previously hybridized cDNA reacted by C_0t = 10 (Figure 6). Thus, by selecting cDNA for a rapid reaction with polyadenylated RNA, one obtains a population enriched in repetitive sequences. However, not all the rapidly hybridizing cDNA is repetitive; about 50% renatures in the range expected for unique sequences.

Nature of the Repetitive Sequences in cDNA. The kinetics of renaturation of cDNA with DNA suggest that sequences coding for some polyadenylated RNA are repetitive. This would imply the existence of classes of repetitive genes other than those for ribosomal RNA, 5S RNA, tRNA, and histone genes. However, other possible interpretations must be considered. It is possible that the DNA sequences with which the cDNA reacts are related in sequence but not identical, perhaps representing a family of

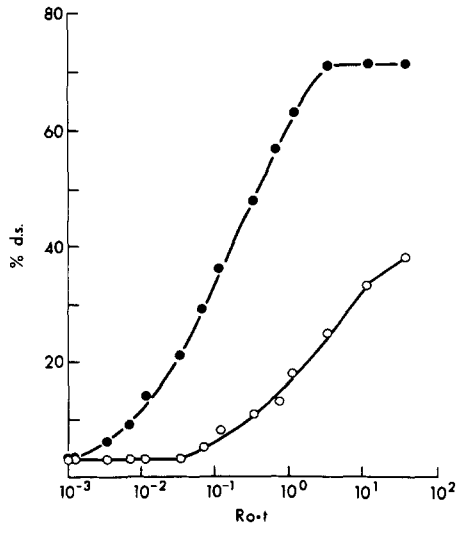

FIGURE 4: Hybridization of Schneider's cell cDNA fractionated on hydroxylapatite with its mRNA template. cDNA (5×10^5 cpm) was annealed with mRNA to an R_0t of 0.2 and fractionated on hydroxylapatite. Aliquots with 500 cpm of previously reacted cDNA (●) or unreacted cDNA (○) were annealed with mRNA. Two reaction mixtures were used containing 10 or 100 μg/ml of RNA. The data shown represent the arithmetic mean of three experiments.

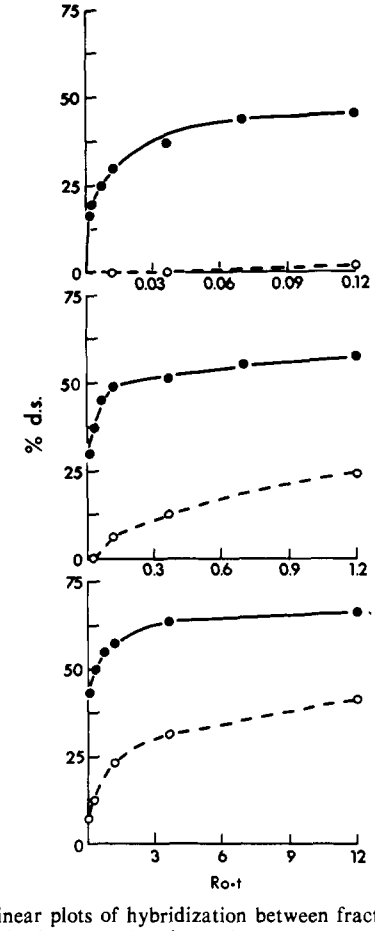

FIGURE 5: Linear plots of hybridization between fractionated cDNA and mRNA. The data are taken from Figure 4.

related proteins. Alternatively they may represent insect proteins containing internal homologies of amino acid sequence (Suzuki et al., 1972; Lambert, 1973). In either of these cases, base pairing between cDNA and repetitive DNA would be imperfect. This possibility was tested by establishing the thermal dissociation profile of duplexes

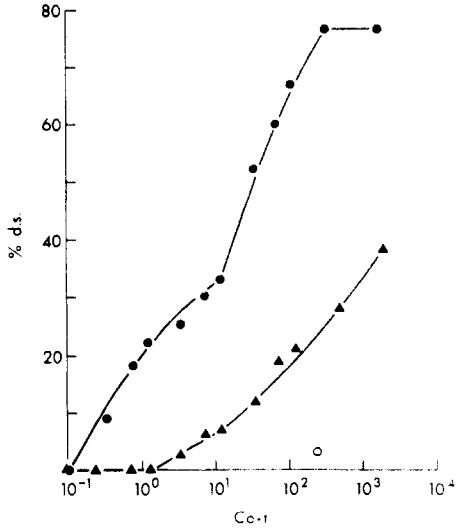

FIGURE 6: Renaturation of fractionated cDNAs with an excess of unlabeled *Drosophila* DNA. cDNA synthesized with Schneider's cell polyadenylated RNA as template was annealed to the mRNA to an R_0t of 0.1 and fractionated on hydroxylapatite as described in the Experimental Section. Aliquots containing 500 cpm of previously hybridized cDNA (●) or unreacted cDNA (▲) were annealed with unlabeled *Drosophila* DNA. Two experiments were performed, using different cDNA preparations. A control is included in which previously reacted cDNA was annealed with bacterial DNA to a C_0t value of 250 (○).

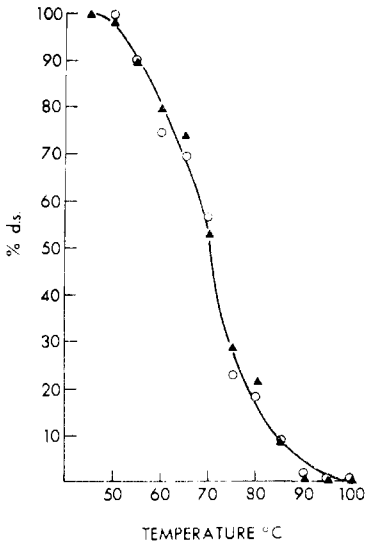

FIGURE 7: Thermal stability of abundant cDNA-DNA and infrequent cDNA-DNA duplexes. DNA-cDNA duplexes were formed by renaturation of abundant and infrequent cDNA with *Drosophila* DNA to C_0t values of 1.0 and 240, respectively. After hybridization, the samples were diluted and the thermal dissociation profile determined as described in the Experimental Section: (○) repetitive cDNA-DNA duplexes; (▲) infrequent cDNA-DNA duplexes.

formed between the two fractions of cDNA and cellular DNA. These were in fact indistinguishable (Figure 7) implying that no gross difference in the fidelity of base pairing exists. The T_m value, 70°, in 0.024 *M* phosphate buffer, is consistent with that expected for well-paired duplexes (Schildkraut and Lifson, 1965; Rosbash et al., 1974).

Comparison of Transcription Patterns in Different Stages of Development. Synthesis of cDNA using preparations from various stages of development of *Drosophila* makes it possible to compare the complexity of various populations of RNA and the degree of overlap among them. Several ex-

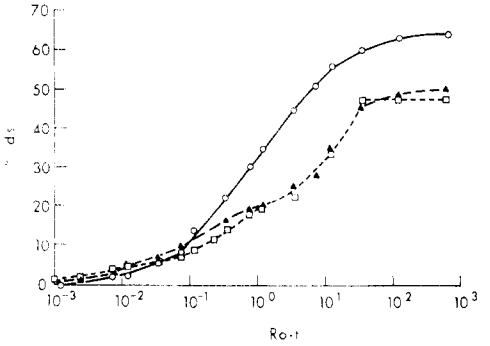

FIGURE 8: Hybridization between larval cDNA and mRNA from different tissues. Larval cDNA was made as described in the Experimental Section for the Schneider's cell cDNA but using larval mRNA as a template. Samples containing 500 cpm of larval cDNA were annealed as described before with mRNA from (○) larvae; (▲) Schneider's cells; and (□) adult flies.

periments of this type have been performed and examples will be described. cDNA was prepared with polyadenylated RNA of third instar larvae as template. The hybridization kinetics of this cDNA with its template (Figure 8) are similar to those obtained for Schneider's cell cDNA and RNA (Figure 3). Calculation of the complexity and distribution among the three frequency classes gives data essentially the same as that in Table I. However, when the same cDNA was hybridized with adult or Schneider's cell RNA, clear differences are apparent. In both cases most of the curve was displaced to higher values of R_0t by a factor of about five. Furthermore, reactions reached a plateau at high R_0t values implying that some 10–15% of the cDNA was either not represented or poorly represented by complementary RNA in the population from adults or Schneider's cells. One further feature of the data merits comment. The three curves are almost coincident for the first 10% of the hybridization reaction implying a high degree of homology among those polyadenylated RNAs most abundant in the three populations. Such RNAs may represent functions common to all cells involving abundant structural proteins.

The degree of overlap among the most frequent RNAs present at various developmental stages was investigated directly by employing cDNA selected to be representative of abundant RNA. When this fractionated cDNA was hybridized with polyadenylated RNA of Schneider's cells, larvae, pupae, and adults, a high efficiency of reaction was obtained in each case (Figure 9). Displacements of the curve to higher values of R_0t are clear especially for pupal RNA. However, it is evident that most species of abundant RNA are present in each case at concentrations which do not differ substantially.

Discussion

The enumeration of the number of genes in *Drosophila* has been a matter of intense activity over the past several years (Lewin, 1974). About 5000 to 6000 bands are observable in polytene chromosomes (Bridges, 1938) and recent fine-structure mapping suggests that each band is a single complementation group perhaps responsible for the specification of a single protein product (Judd and Young, 1973: Hochman, 1973). More recently the problem has been approachable through biochemical analysis. In principle, kinetic analysis of the hybridization reaction of synthetic DNA, made as a complement to polyadenylated RNA, with its template RNA, yields data concerning the complexity of

the RNA population (Bishop et al., 1974). The theoretical basis of this approach is analogous to methods for estimating complexity of DNA through renaturation kinetics (Britten and Kohne, 1968), or of viral messenger by hybridization with viral DNA (Frenkel and Roizmann, 1972). In a footnote to a recent paper, Bishop et al. (1974) state the conclusion from experiments of this type with *Drosophila* RNA, namely that some 4000 genes are active.

Before discussing the present data, it is appropriate to consider the validity of the approach. The overall kinetic approach has been validated through many experiments since the pioneering work of Britten and Kohne (1968) and Wetmur and Davidson (1968). In the case of reaction of cDNA with polyadenylated RNA, the implicit assumption is that each RNA molecule is randomly copied by the reverse transcriptase so that the relative frequency of that cDNA is proportional to the frequency of the template RNA. Although this is probably not true in detail (Duesberg et al., 1971; Taylor et al., 1974), the assumption remains valid, provided that the efficiency of copying is not directly related to the abundance of a particular RNA. For so long as one is concerned with approximations of the frequency distribution, inaccuracies of this type will not be severe (Bishop et al., 1974). Calculation of the number of genes active in a given cell depends upon accurate determination of the $R_0t_{1/2}$ value for the most slowly reacting component. The problem is entirely analogous to cases where the analytical complexity of a genome is estimated from the rate constant for renaturation of unique sequence DNA. When the kinetics of renaturation are such that repetitive and unique sequences do not overlap, this measurement can be quite precise (Britten and Kohne, 1968; Laird, 1971). However, in the case of plant DNA where no sharp transitions are observed it is very difficult to obtain values of $C_0t_{1/2}$ for unique sequences (Bendich and McCarthy, 1970). Similar considerations apply to cDNA–RNA kinetics. In fact, clear transitions do not occur and reactions of abundant and rare RNA species tend to overlap in time. In order to surmount this problem the data have been plotted on a scale linear in R_0t and analyzed for best fit with a computer program. Using this approach Bishop et al. (1974) calculate about 40,000 genes active in HeLa cells and 4000 in *Drosophila* and Ryffel and McCarthy (1975) calculate 8000 genes in mouse L-cells and 20,000 in mouse brain. From the present data we estimate a total of 6900 for *Drosophila* tissue culture cells. However, the error in this calculation could be at least a factor of two due to the difficulty of estimating the value of $R_0t_{1/2}$ for the slowest component. It should be noted that the calculations depend upon an assumption for the mean size of polyadenylated RNA. We have taken a value of 4×10^5. If Bishop et al. assumed a higher figure, e.g., 6×10^5, as they did for HeLa RNA, then the two estimates of complexity for *Drosophila* mRNA are in close agreement.

Although the data are fragmentary, it is possible to interpret them in several ways. For example, if gene number is correlated with DNA content, then 40,000 for HeLa and 7000 for *Drosophila* are reasonable. However, there may be no such correlation, for in amphibia, at least, the complexity of oocyte mRNA is independent of genome size (Rosbash et al., 1974). Furthermore, closer agreement would be expected for HeLa cells and L-cells, two established mammalian tissue culture lines. In fact, the estimates are very similar for L-cells and Schneider's cells, two lines of disparate origin, but whose conditions of exponential growth in

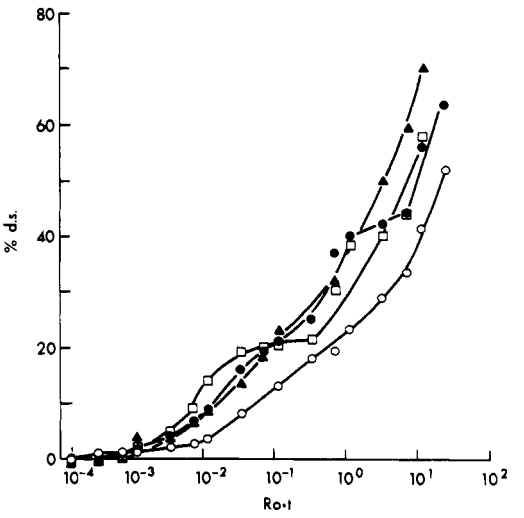

FIGURE 9: Hybridization of cDNA representing frequent messages of *Drosophila* Schneider's cells with mRNA from different stages of *Drosophila melanogaster* development. cDNA (10^6 cpm) from Schneider's cell RNA was annealed with homologous mRNA to an R_0t of 0.1 and the hybridized cDNA was separated. Samples containing 500 cpm of this cDNA were annealed with polyadenylated mRNA from (▲) Schneider's cells; (□) larvae; (○) pupae; (●) adult flies. Two reaction mixtures were used in each experiment, containing 1 and 100 μg/ml of RNA.

vitro are superficially similar. For one can argue that the number of genes necessary for growth in culture should be very similar. The gene numbers obtained using cDNA as a probe should be regarded as a minimum because poly(A^-) mRNAs other than histone mRNA may also be present in the mRNA population as is the case in HeLa cells (Milcarek et al., 1974), and they would not contribute to the measured complexity.

These experiments also demonstrate that cDNA can be fractionated by limited hybridization to obtain a probe for the most abundant RNA in the cell. Furthermore, these cDNA sequences derive preferentially from repetitive sequences in the cell. Assuming that these sequences are repetitive genes, the result suggests that production of abundant RNA species is accomplished partially through gene dosage. However, the equation between repetitive sequences and abundance of the transcript in the cell is far from complete. A major fraction of the cDNA corresponding to abundant RNA renatures with unique sequences. This result is consistent with other cases where single gene copies are responsible for large amounts of mRNA (Suzuki et al., 1972; Sullivan et al., 1973; Stavnezer et al., 1974). The possibility should also be considered that the repetitive DNA with which some cDNA reacts does not, in fact, represent nuclear genes. Schneider's cells and insects in general are known to be infected by viruses. Large numbers of virus genomes may exist contaminating the host genome and contributing to the poly(A) containing RNA fraction. Moreover, the nuclear DNA preparation cannot be guaranteed free of mitochondrial DNA. At present the only counter argument to this interpretation is that the RNA complementary to the abundant cDNA is present in all stages of development. Although this may seem to be most consistent with messengers for common structural proteins, it could still be attributable to contaminating genomes.

Apart from the most frequent polyadenylated RNA which appear to be common to cells and developmental stages examined, the data are consistent with some degree

of transcriptional regulation. For example, the reaction of adult or Schneider's cell RNA with larval cDNA is incomplete. Although this is not surprising, we are not yet in a position to evaluate the magnitude of this transcriptional control. This must await the results of experiments in progress, using fractionated stage specific cDNA probes.

Acknowledgment

We are indebted to Galvin Swift for computer analysis of the data and to Dr. Gerhart Ryffel for valuable advice.

References

Beerman, W. (1972), in Developmental Studies on Giant Chromosomes, Beerman, W., Ed., New York, N.Y., Springer-Verlag New York, p 1.

Bendich, A. J., and McCarthy, B. J. (1970), *Genetics 65,* 545.

Bishop, J. O. (1974), *Cell 2,* 81.

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

Bridges, C. B. (1938), *J. Hered. 29,* 11.

Britten, R. J., and Kohne, D. E. (1968), *Science 161,* 529.

Doenecke, D., and McCarthy, B. J. (1975), *Biochemistry 14,* 1373.

Duesberg, P., Helm, K. V. D., and Canaani, E. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 2505.

Frenkel, N., and Roizman, B. (1972), *Proc. Natl. Acad. Sci. U.S.A. 69,* 2654.

Gambino, R., Kacian, D., O'Donnell, J., Ramirez, F., Marks, P. A., and Bank, A. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 3966.

Grouse, L., Chilton, M. D., and McCarthy, B. J. (1972), *Biochemistry 11,* 798.

Hochman, B. (1973), *Cold Spring Harbor Symp. Quant. Biol. 38,* 581.

Imaizumi, T., Diggelman, H., and Scherrer, K. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 1122.

Judd, B. H., and Young, M. W. (1973), *Cold Spring Harbor Symp. Quant. Biol. 38,* 573.

Laird, C. D. (1971), *Chromosoma 32,* 378.

Laird, C. D., and McCarthy, B. J. (1969), *Genetics 63,* 865.

Lambert, B. (1973), *Cold Spring Harbor Symp. Quant. Biol. 38,* 637.

Lewin, B. (1974), *Nature (London) 251,* 373.

McCarthy, B. J., Nishiura, J. T., Doenecke, D., Nasser, D. S., and Johnson, C. B. (1973), *Cold Spring Harbor Symp. Quant. Biol. 38,* 763.

Milcarek, C., Price, R., and Penman, S. (1974), *Cell 3,* 1.

Penman, S. (1969), *Fundam. Tech. Virol. 1969,* 35.

Rosbash, M., Ford, P. J., and Bishop, J. O. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 3746.

Rosen, J. M., Woo, S. L. C., Holder, J. W., Means, A. R., and O'Malley, B. W. (1975), *Biochemistry 14,* 69.

Ryffel, G. U., and McCarthy, B. J. (1975), *Biochemistry 14,* 1385.

Schildkraut, C., and Lifson, S. (1965), *Biopolymers 3,* 195.

Shearn, A., and Garen, A. (1974), *Proc. Natl. Acad. Sci. U.S.A. 71,* 1393.

Shearn, A., Rice, T., Garen, A., and Gehring, W. (1971), *Proc. Natl. Acad. Sci. U.S.A. 68,* 2594.

Silverstein, S., Bachenheimer, S. L., Frenkel, N., and Roizman, B. (1973), *Proc. Natl. Acad. Sci. U.S.A. 70,* 2101.

Stavnezer, J., Huang, R. C. C., Stavnezer, E., and Bishop, J. M. (1974), *J. Mol. Biol. 88,* 43.

Sullivan, D., Palacios, R., Stavnezer, J., Taylor, J. M., Faras, A. J., Kiely, M. L., Summers, N. M., Bishop, J. M., and Schimke, R. T. (1973), *J. Biol. Chem. 248,* 7530.

Sutton, W. D. (1971), *Biochim. Biophys. Acta 240,* 522.

Suzuki, Y., Gage, L. P., and Brown, D. D. (1972), *J. Mol. Biol. 70,* 637.

Taylor, J., Varmus, H., Faras, A., Levinson, W., and Bishop, J. M. (1974), *J. Mol. Biol. 84,* 217.

Wetmur, J. G., and Davidson, N. (1968), *J. Mol. Biol. 31,* 349.