Proc. Natl. Acad. Sci. USA
Vol. 76, No. 5, pp. 2209–2212, May 1979
Biochemistry

Distinct single-copy sequence sets in sea urchin nuclear RNAs

(complexity/embryo nuclear RNA)

SUSAN G. ERNST*, ROY J. BRITTEN*†, AND ERIC H. DAVIDSON*

*Division of Biology, California Institute of Technology, Pasadena, California 91125; and †Kerckhoff Marine Laboratory, California Institute of Technology, Corona Del Mar, California 92625

Contributed by Roy J. Britten, February 21, 1979

ABSTRACT **The purpose of this study was to determine whether nuclear RNAs (nRNA) of sea urchin embryos and adult tissues contain identical or partially distinct sets of single-copy sequence transcripts. A DNA tracer was prepared consisting mainly of sequences absent from gastrula nRNA; 3.6% of this tracer reacted with adult intestine nRNA but not with gastrula nRNA. The existence of a differentially transcribed DNA fraction was verified by its partial purification and rehybridization to intestine and gastrula nRNAs. About one-third of the genomic single-copy sequence is represented in both nRNAs, or about 2 $\times 10^8$ nucleotides. The differentially transcribed portion of the single-copy genome identified in this work includes about 3.5 $\times 10^7$ nucleotides.**

Recent measurements indicate that the nuclear RNAs (nRNAs) of several sea urchin embryo stages and adult tissues are of approximately equal sequence complexity. About a third of the single-copy DNA sequence is represented in the nRNAs of blastula (1), gastrula (2, 3), and pluteus (1) embryos and of previtellogenic ovarian oocytes (4) and adult intestine (5). Kleene and Humphreys (1) showed in addition that the single-copy sequence sets of blastula and pluteus nRNAs are largely overlapping. They concluded that with respect to single-copy sequence content, these nRNAs are in fact identical.

Complete overlap in the single-copy sequence sets represented in diverse nRNAs would be paradoxical. The implication is that in the sea urchin differentiation need not be accompanied by alteration in the regions of the genome transcribed into nRNAs of typical prevalence. There are several reasons to question the conclusion that sea urchin nRNAs are identical. Scheller *et al.* (6) found recently that gastrula and adult intestine nRNAs differ strikingly in the representation of a number of repetitive sequence families. In the genome, these repetitive sequences are in general interspersed with single-copy sequence [ref. 7, and unpublished data regarding the same cloned repetitive sequence families as studied by Scheller *et al.* (6)], and the nRNA transcripts studied by Scheller *et al.* are long enough to contain both single-copy and repetitive sequences. Thus, a likely though not the only possible implication of the results of Scheller *et al.* is that qualitative differences in the single-copy sequence transcripts should exist between gastrula and intestine nRNAs and, by implication, between the nRNAs of other cells and tissues. Furthermore, several studies have indicated clear distinctions in the single-copy sequence sets of nRNAs extracted from various mammalian tissues (see *Discussion* for refs.).

The complexity of sea urchin nRNAs is so large—about 2×10^8 nucleotides—that minor differences that are not easily detected could be of great biological significance. With this in mind, we have carried out a detailed comparison of the sequence sets found in adult intestine and gastrula embryo nRNAs. The approach we have used is that first applied by Galau *et al.* (8) in a comparison of sea urchin mRNA sequence sets. A single-copy DNA tracer consisting mainly of sequences absent from gastrula nRNA was prepared, and this tracer was then reacted with intestine nRNA and ovary nRNA. As Hough-Evans *et al.* (4) have shown, there is little difference between gastrula and ovary nRNA single-copy sequence sets. However, the present experiments demonstrate a significant set of single-copy sequences in intestine nRNA that is not found in gastrula nRNA.

MATERIALS AND METHODS

Preparation of Nuclear RNAs. *Gastrula. Strongylocentrotus purpuratus* were spawned, the eggs were fertilized, and the embryos were grown for 36 hr to the gastrula stage. nRNA was isolated as described earlier (6).

Intestine. The intestine was removed from adult sea urchins, rinsed in seawater/EGTA [ethylene glycol-bis(β-aminoethyl ether)-*N,N,N′,N′*-tetraacetic acid] and lysed on ice in 50 mM Tris (pH 8.0), 5 mM magnesium chloride, 50 mM ammonium chloride, 5 mM EGTA, and 1% Triton X-100. Nuclei were pelleted through a 0.5 M sucrose cushion and the RNA was extracted as described earlier (5).

Ovary. Nuclei were isolated from ovaries that contained previtellogenic oocytes and a high vitellogenic oocyte content, and the RNA was extracted by the methods of Hough-Evans *et al.* (4).

Preparation of Labeled Single-Copy DNA. Single-copy DNA was renatured to 10–20 times the $C_0t_{1/2}$ (moles of nucleotide liter^{-1}·sec) to form hyperpolymers (7, 8). The DNA was labeled *in vitro* by the "gap translation" method with *Escherichia coli* polymerase I (8, 9). The single-copy tracer had a specific activity of about 10^7 cpm/μg and a mode fragment size of about 200–500 nucleotides.

Preparation of Single-Copy Tracer Depleted of Gastrula nRNA Sequences. The single-copy tracer was reacted with gastrula nRNA at a mass ratio of 1800 to an RNA C_0t of 78,000. Single-stranded fragments were separated by hydroxyapatite (HAP) chromatography, concentrated, and again reacted with gastrula nRNA at a mass ratio of 1:1400 to an RNA C_0t of 90,000, and the single-stranded fraction was collected as above. This tracer fraction is termed "gastrula null nDNA." When driven to kinetic termination with total sea urchin DNA, 70% of the null nDNA tracer bound to HAP. The final yield of the null nDNA tracer was 8.6% of the starting single-copy tracer. This yield would result from an average 88.5% recovery of fragments at each of the 20 individual steps in the procedure, including dialyses, precipitations, and reconcentrations not described here. These procedures are detailed in previous publications (5, 8, 9).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: nRNA, nuclear RNA; HAP, hydroxyapatite; C_0t, moles of nucleotides per liter × sec.

Hybridization Conditions. Hybridization was carried out in 0.41 M phosphate buffer, 1–5 mM EDTA, and 0.5% sodium dodecyl sulfate at 60°C. Hybridization was assayed by binding to HAP. The content of DNA duplex was assayed by prior treatment with RNase A under low salt conditions, followed by HAP binding. These procedures have been described in detail earlier (5, 8, 9). All C_0t values are given as *equivalent* C_0t values—i.e., the C_0t has been adjusted for acceleration in hybridization rate due to Na^+ concentration >0.18 M (10). Hybridization data were reduced by a nonlinear least-squares analysis assuming pseudo-first-order kinetics (11).

RESULTS

Sequence Complexity of Gastrula, Intestine, and Ovary nRNAs. The major object of the experiments described in this paper is to compare the single-copy sequences represented in gastrula and adult intestine nRNAs. The nRNA preparations were initially reacted with single-copy [^{3}H]DNA in order to determine their sequence complexities. In Table 1 the results of these measurements are shown to compare closely with previous data (2, 5). Thus, the complexity of the gastrula nRNA was found to be about 1.8×10^8 nucleotides, or 30% of the total single-copy sequence in the sea urchin genome. The kinetics of this reaction are shown in Fig. 1 *upper*. The complexity of the intestine nRNA is the same or perhaps very slightly higher. The two separate measurements cited in Table 1 for this nRNA indicate complexities of 1.9×10^8 and 2.2×10^8 nucleotides or, respectively, 31% and 36% of the genomic single-copy sequences. Table 1 also lists the complexity of nRNA from ovaries containing both previtellogenic and vitellogenic oocytes. This measurement was published earlier by Hough-Evans *et al.* (4), who used the same nRNA preparation as used in the present study.

Reactions of a Single-Copy [^{3}H]DNA Fraction Depleted of Gastrula nRNA Sequences. To prepare a sensitive probe for intestine nRNA sequences *absent* from the gastrula nRNA, we twice reacted single-copy tracer DNA with gastrula nRNA, each time recovering the single-stranded portion (see *Materials and Methods* for procedures). The resulting single-copy [^{3}H]DNA fraction is termed the "gastrula null nDNA" tracer.

Table 1. Sequence complexity of nRNAs

nRNA	Complexity* and rate constant†	
	This work	Previous measurements
Gastrula	1.7×10^8 nucleotides	1.7×10^8 nucleotides
	7.1×10^{-5} M^{-1} sec^{-1}	1.1×10^{-4} M^{-1} sec^{-1} (ref. 2)
Adult intestine	1.9×10^8 nucleotides	2.2×10^8 nucleotides
	1.1×10^{-4} M^{-1} sec^{-1}	2.5×10^{-4} M^{-1} sec^{-1} (ref. 5)
Ovary		1.7×10^8 nucleotides
		2.3×10^{-5} M^{-1} sec^{-1} (ref. 4)

* Complexity, N, is calculated [assuming asymmetric transcription of single-copy sequences as shown by Hough *et al.* (2)] as follows, where F is the fraction of reactable single-copy DNA tracer recovered as RNA·DNA hybrid at termination of the reaction and S is the single-copy sequence length of the genome, here taken as 6.1×10^8 nucleotides (7); $N = 2FS$.

† Pseudo-first-order rate constant for nRNA excess reactions with single-copy DNA tracer.

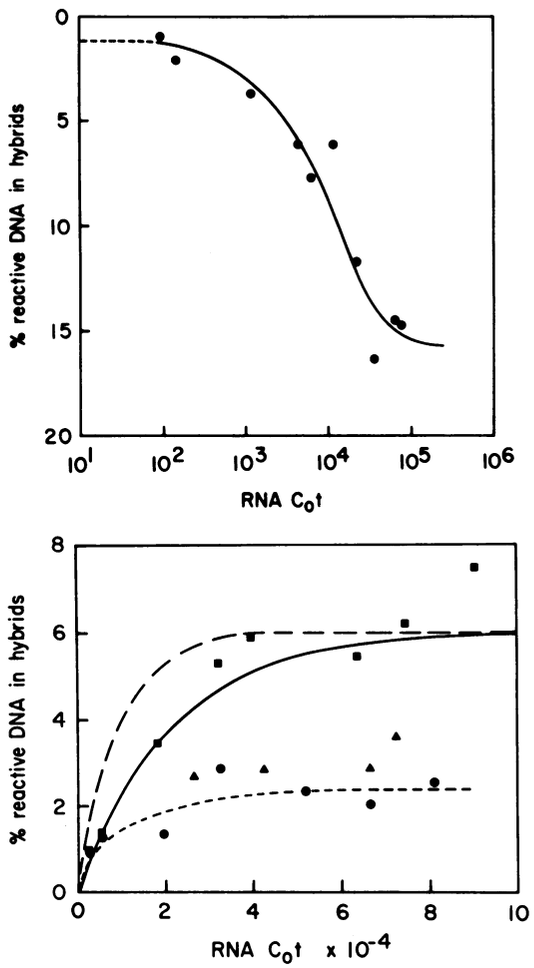

FIG. 1. Reaction of nRNAs with single-copy DNA and gastrula null nDNA. (*Upper*) Reaction of gastrula nRNA with single-copy [^{3}H]DNA. Data have been corrected for a tracer reactivity of 85%. The rate constant for the reaction is 7.1×10^{-5} M^{-1} sec^{-1}. The fraction of reactive [^{3}H]DNA tracer included in nRNA·DNA hybrid at termination of the reaction is 14.6%. (*Lower*) Reactions of gastrula null nDNA tracer with gastrula nRNA (●), intestine nRNA (■), and ovary nRNA (▲). Dashed curves portrayed for these reactions were generated by fixing the terminal values as the mean of the data obtained at RNA $C_0t \geq 20{,}000$ and imposing the rate constants determined for the reaction of these nRNAs with total single-copy [^{3}H]DNA (Table 1). Solid curve for the intestine nRNA reaction shows the least-squares solution for the kinetics of this reaction with the terminal value fixed as above. The rate constant obtained is 5×10^{-5} M^{-1} sec^{-1}, which is about a factor of 2 less than that of the dashed curve. Data have been corrected for the 70% reactivity of the null nDNA tracer.

It reacted with gastrula nRNA to about 2.4%, compared with the 14.6% hybridization of single-copy [^{3}H]DNA observed initially. This result is illustrated in Fig. 1 *lower*. The kinetics of the residual reaction of the null nDNA tracer with gastrula nRNA are consistent with the kinetics of the reaction of total single-copy DNA with this nRNA. Therefore, those fragments of the null nDNA that react with gastrula nRNA are not a specific sequence subset that is represented only by very rare nRNA transcripts. Instead, the null nDNA probably contains *all* of the gastrula nRNA sequences but at a concentration reduced to about one-sixth the concentration of all other single-copy sequences. The presence of the small amount of residual "gastrula sequences" in the null nDNA tracer is due to failure to trap 100% of the nRNA·DNA hybrids during the preparative HAP fractionations. The persistence of these sequences does not interfere with the use of the null nDNA tracer for the detection of "nongastrula sequences," as shown below.

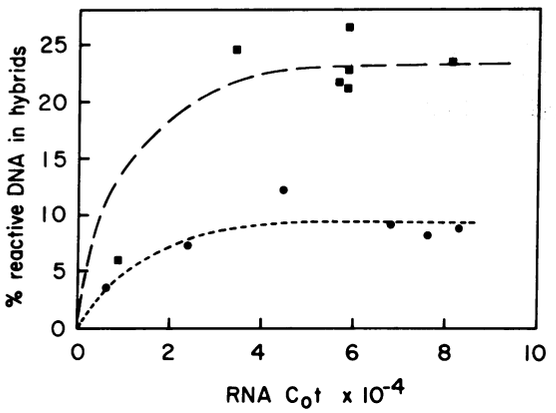

FIG. 2. Reaction with gastrula and intestine nRNAs of gastrula null nDNA tracer enriched for intestine nRNA sequences. A gastrula null nDNA tracer fraction was selected to contain sequences represented in intestine nRNA. This tracer fraction was reacted with intestine (■) and gastrula (●) nRNAs. Data have been corrected for the 70% reactivity of the tracer. The kinetic curves shown were generated as described in the legend to Fig. 1 *lower*. The terminations are at 23.4% (intestine) and 9.5% (gastrula).

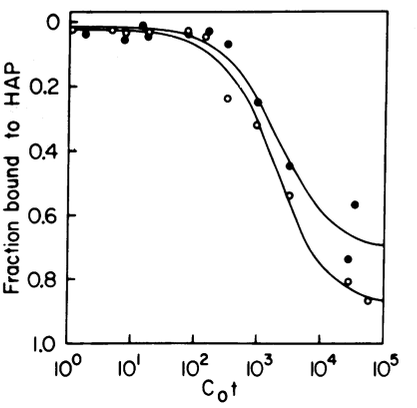

FIG. 3. Reactions of the selected tracer fraction and total single-copy [^{3}H]DNA with sheared sea urchin DNA driver. A large excess of sea urchin DNA sheared to about 300 nucleotides was renatured with total single-copy [^{3}H]DNA (○) and with the gastrula null nDNA tracer fraction enriched for intestine nRNA sequences (●). The second-order rate constant for the reaction of the null nDNA with the driver DNA was 5.0×10^{-4} M^{-1} sec^{-1}, compared to 5.6×10^{-4} M^{-1} sec^{-1} for the reaction of the total single-copy tracer with the driver DNA.

Reactions between the gastrula null nDNA tracer and the intestine and ovary nRNAs are also shown in Fig. 1 *lower*. The ovary nRNA reacted to a slightly higher level (3%) than did gastrula nRNA (2.4%). However, this difference may not be significant. On the other hand, the sequence content of the intestine nRNA is clearly distinguishable from that of gastrula nRNA. The intestine nRNA reaction terminates at 6%, 2.5 times higher than the gastrula nRNA reaction with the null nDNA tracer. The difference between the two reactions amounts to 3.6% of the hybridizable single-copy sequence in the null nDNA tracer. The experiment suggests that intestine nRNA includes a significant set of single-copy transcripts that do not exist at detectable levels in gastrula nRNA. Furthermore, it can be seen from the reaction kinetics in Fig. 1 *lower* that the concentration of this specific set of transcripts is within a small factor of that of most of the intestine nRNA species that react with unfractionated single-copy DNA tracer.

Enrichment for Intestine nRNA Sequences Absent from Gastrula nRNA. The gastrula null nDNA was reacted on a preparative scale with intestine nRNA in order to partially purify the specific intestine nRNA sequence set. To ensure termination, the reaction was carried out to an RNA C_0t of 9×10^4, and the intestine nRNA/[^{3}H]DNA mass ratio was 1800. On passage over a HAP column, 10.9% of the tracer bound. Of this, about 6.4% is attributed to renatured tracer (a [^{3}H]DNA C_0t of 50 M·sec was generated during the reaction, and the tracer was 70% reactive), leaving 4.5% nRNA·DNA hybrid (the expected hybrid yield was $0.7 \times 6\%$ or 4.2%). To save procedural steps, the entire bound fraction was collected and hybridized with both gastrula and intestine nRNAs, with the results shown in Fig. 2. Here it can be seen that the ratio of the intestine and gastrula nRNA reactions remains about 2.5, just as in the experiment of Fig. 1 *lower*. However, the absolute amount of the tracer reacting with the intestine nRNA has increased from 6% to 23.4%. The fraction of the tracer representing intestine nRNA sequences absent from gastrula nRNA rose from 3.6% to 14%. About 62% of the bound [^{3}H]DNA fraction is expected to be random sequence included initially as renatured [^{3}H]DNA tracer. In the 38% of the bound tracer that reacted with the intestine nRNA, the sequence set absent from gastrula was purified about 8.6-fold as a result of the additional reaction (close to 45% of the maximum possible purification). This is a minimum estimate of the extent of purification since with each successive reaction, the fraction of tracer capable of forming recoverable RNA·DNA hybrid structures decreases.

The result shown in Fig. 2 demonstrates that the difference in the reaction of the null nDNA tracer with intestine and gastrula nRNAs is not artifactual. Had the 3.6% of the null nDNA tracer reacting specifically with intestine nRNA and not with gastrula nRNA (Fig. 1 *lower*) consisted of a random set of single-copy sequences, it could not have been significantly purified by a further additional hybridization reaction with intestine nRNA. Therefore, a specific set of sequences complementary to about 3.6% of the mass of the gastrula null nDNA tracer exists in intestine nRNA.

Since in this study we are dealing with small portions of the genome, it is important to show that the relevant tracer fractions are indeed single-copy. A sample of the enriched tracer used for the experiment of Fig. 2 was reassociated with excess sea urchin DNA. The kinetics of this reaction are shown in Fig. 3. The rate constant for the reaction with driver DNA of the unfractionated single-copy tracer used in this work was 5.6×10^{-4} M^{-1} sec^{-1}. Except for the slightly decreased reactivity, the kinetics of the reassociation of the enriched tracer fraction with driver DNA are nearly identical. The rate constant for this reaction was 5.0×10^{-4} M^{-1} sec^{-1}. An almost complete absence of reaction at driver DNA $C_0t < 100$ M·sec is observed. Thus, there is no evidence for any repetitive sequence component. Even though only a minor fraction ($\geq$14%) of the reactable selected tracer consists of sequences represented in intestine nRNA and absent from gastrula nRNA, it is clear that most of these sequences are single-copy, or near single-copy, in genomic frequency.

DISCUSSION

These experiments provide a measure of the complexity of the single-copy sequence represented in one nRNA and absent from another in the sea urchin. Quantitatively different results would no doubt have been obtained had other nRNAs been compared. The sequence distinction between gastrula and ovary nRNA appears smaller, as shown in Fig. 1 *lower* and ref. 4, and there appears to be less difference between blastula and pluteus nRNA (1) than between the nRNAs compared in this work. However, sea urchin oocytes and embryos also display larger

overlaps in their mRNA sequence sets relative to intestine and embryo (5, 8, 9). On the other hand, there is no reason to assume that intestine and gastrula contain the most disparate nRNAs in the organism. The main qualitative conclusion of this measurement does not depend on the arbitrary choice of tissues or the exact size of the nonshared nRNA sequence set. The results we report exclude the proposition that nRNAs of diverse sea urchin tissues all contain identical single-copy nRNA sequence sets.

Considering the very high total complexities of these nRNAs (Table 1), the amount of sequence present in intestine nRNA and absent from gastrula nRNA is not insignificant. Fig. 1 *lower* shows that 3.6% of the reactable gastrula null nDNA tracer hybridizes with intestine nRNA, but not with gastrula nRNA. The complexity of this intestine nRNA sequence subset can be calculated as follows. We assume, as shown earlier (2), that transcription of the relevant nRNA sequences is asymmetric. The null nDNA tracer lacks about 12.2% of its original mass—i.e., the fraction physically removed by hybridization with gastrula nDNA (14.6% − 2.4%). The maximum portion that *could* have hybridized with intestine nRNA is then half of 75.6% of the original tracer, or 37.8/87.8 of the null nDNA tracer. The length of single-copy sequence in the genome of *S. purpuratus* is 6.1×10^8 nucleotides, and the complexity of the "nongastrula" single-copy sequences is 70.8% of this value or 4.3×10^8 nucleotides. Thus, the complexity of the intestine nRNA sequence subset absent from gastrula nRNA is (0.036) (87.8/37.8) (4.3×10^8) or about 3.6×10^7 nucleotides. This is approximately 20% of the total intestine nRNA complexity. This amount of single-copy sequence is several times larger than the cytoplasmic mRNA complexity of intestine cells (5, 8) and, in fact, exceeds the complexity of any of the sea urchin mRNA populations so far measured (8, 9, 12).

The recent studies of Wold *et al.* (5) show that essentially all mRNA sequences that are expressed in embryos and absent from intestine polysomes are nonetheless represented at typical concentrations in intestine nRNA. In the sea urchin, most or all single-copy structural gene sequences thus appear to be represented in an ubiquitous fashion in nRNAs, irrespective of whether the genes are producing cytoplasmic messages in given cell types. It follows that the 3.6×10^7-nucleotide intestine nRNA sequence subset lacking from gastrula nRNA is probably not to be interpreted as a set of structural gene transcripts (including possible intervening sequences) destined to become specific intestine messages absent in gastrula. The distinctions among nRNA single-copy sequence sets from diverse mammalian tissues appear greater than those among sea urchin tissues, though here also there is a large "core" of shared sequences (e.g., see refs. 13 and 14; earlier data reviewed in ref. 15). Although the conclusion is far from secure, studies with specific genes suggest that in mammals and birds, nRNA transcripts bearing mRNA sequences may also be found in cell types not expressing these genes. Globin gene transcripts have been claimed to be present in nonerythropoietic tissues (16–19), though a conflicting report also exists (20). Similarly, ovalbumin mRNA sequences are reported in RNAs from spleen, liver, and other tissues (21, 22), though again these sequences are supposed to be absent from fibroblast RNA (20, 23). It may well be that, as in the sea urchin, the large tissue-specific single-copy sequence sets of mammalian nRNAs include transcripts of sequences that are not structural genes.

Scheller *et al.* (6) recently discovered that intestine and gastrula nRNAs differ sharply in the concentration of particular repetitive sequence transcripts. Some repeat families are more highly represented in intestine nRNA relative to gastrula while the reverse is true for others. As pointed out in the introductory portion of this paper, differences in the extent to which specific repeat families are represented in various nRNAs may well imply differences in the nRNA transcripts deriving from *single-copy* regions flanking the expressed repeats. Assuming average interspersion distances, the 3.7×10^7-nucleotide single-copy sequence subset of intestine nRNA could include some $2–3 \times 10^4$ interspersed repetitive sequence transcripts per genome which would be present in intestine nRNA but absent from gastrula nRNA. This is sufficient to account for the amount of increase in repeat sequence representation implied by the measurements of Scheller *et al.* (6). Thus, single-copy sequence subsets specific to given nRNAs may exist as interspersed transcripts bearing specific populations of nRNA repeats. The function of such transcripts is unknown and could be related to that of the repetitive sequence elements which they include.

We thank Mr. Clifford Beall for technical assistance. This research was supported by grants from the National Institutes of Health, HD-05753 and RR-07003 (awarded by the Biomedical Research Support Grant Program). S.G.E. was supported by National Institutes of Health Fellowship GM-05273.

1. Kleene, K. C. & Humphreys, T. (1977) *Cell* **12,** 143–155.
2. Hough, B. R., Smith, M. J., Britten, R. J. & Davidson, E. H. (1975) *Cell* **5,** 291–299.
3. Rodgers, W. H. & Gross, P. R. (1978) *Cell* **14,** 279–288.
4. Hough-Evans, B. R., Ernst, S. G., Britten, R. J. & Davidson, E. H. (1979) *Dev. Biol.* **69,** 225–236.
5. Wold, B. J., Klein, W. H., Hough-Evans, B. R., Britten, R. J. & Davidson, E. H. (1978) *Cell* **14,** 941–950.
6. Scheller, R. H., Costantini, F. D., Kozlowski, M. R., Britten, R. J. & Davidson, E. H. (1978) *Cell* **15,** 189–203.
7. Graham, D. E., Neufeld, B. R., Davidson, E. H. & Britten, R. J. (1974) *Cell* **1,** 127–137.
8. Galau, G. A., Klein, W. H., Davis, M. M., Wold, B. J., Britten, R. J. & Davidson, E. H. (1976) *Cell* **7,** 487–505.
9. Hough-Evans, B. R., Wold, B. J., Ernst, S. G., Britten, R. J. & Davidson, E. H. (1977) *Dev. Biol.* **60,** 258–277.
10. Britten, R. J., Graham, D. E. & Neufeld, B. R. (1974) *Methods Enzymol.* **29E,** 363–418.
11. Pearson, W. R., Davidson, E. H. & Britten, R. J. (1977) *Nucleic Acids Res.* **4,** 1727–1737.
12. Galau, G. A., Britten, R. J. & Davidson, E. H. (1974) *Cell* **2,** 9–21.
13. Chikaraishi, D. M., Deeb, S. S. & Sueoka, N. (1978) *Cell* **13,** 111–120.
14. Hahn, W. E., Van Ness, J. & Maxwell, I. H. (1979) *Proc. Natl. Acad. Sci. USA* **75,** 5544–5547.
15. Davidson, E. H. (1976) *Gene Activity in Early Development* (Academic, New York).
16. Gilmore, R. S., Harrison, P. R., Windass, J. D., Affara, N. A. & Paul, J. (1974) *Cell Differ.* **3,** 9–22.
17. Humphries, S., Windass, J. & Williamson, R. (1976) *Cell* **7,** 267–277.
18. Gottesfeld, J. M. & Partington, G. A. (1977) *Cell* **12,** 953–962.
19. Therwath, A. & Scherrer, K. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 3776–3780.
20. Groudine, M. & Weintraub, H. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4464–4468.
21. Axel, R., Feigelson, P. & Schutz, G. (1976) *Cell* **7,** 247–254.
22. Roop, D. R., Nordstrom, J. L., Tsai, S. Y., Tsai, M.-J. & O'Malley, B. W. (1978) *Cell* **15,** 671–685.
23. Spector, D. H., Smith, K., Padgett, T., McCombe, P., Roulland-Dussoix, D., Moscorici, C., Varmus, H. E. & Bishop, J. M. (1978) *Cell* **13,** 371–379.