Proceedings of the National Academy of Sciences
Vol. 68, No. 1, pp. 172–176, January 1971

Transcription of Nonrepeated DNA in Neonatal and Fetal Mice

A. H. GELDERMAN*, A. V. RAKE†, AND R. J. BRITTEN

Department of Terrestrial Magnetism, Carnegie Institution of Washington, Washington, D.C. 20015

Communicated by R. B. Roberts, October 28, 1970

ABSTRACT **The transcription of nonrepeated DNA sequences was measured by hybridization of RNA from neonatal and fetal mice to mouse DNA using three different techniques. The measurements indicate that a large part (about 70%) of the rapidly-labeled fetal RNA is transcribed from nonrepeated DNA sequences. It appears that more than 12% of the single-copy DNA sequences are represented in the RNA of newborn mice.**

The transcription of genetic information from DNA into complementary RNA has been studied fruitfully in many laboratories by measuring the hybridization of the RNA to DNA. The recognition that the DNA of higher organisms contains repeated sequences strongly affects the interpretation of these studies (1). Much of the repeated DNA occurs in sets of 10,000–1,000,000 similar (but not identical) sequences of nucleotides. Taken together, these sets make up a large fraction of the DNA (about $^1/_3$ in the case of the mouse when the fraction is measured in the customary way). In hybridization experiments, RNA fragments form double-stranded structures with complementary or partially complementary sequences in the DNA. The rate of the hybridization of the RNA depends on the effective concentration of the complementary DNA sequences. If the RNA has been transcribed from sets of repeated DNA sequences, large numbers of sequences complementary to the RNA will, of course, be present in the DNA for the hybridization reaction. As a result, the hybridization of RNA to repeated DNA sequences proceeds at a much faster rate than that of RNA for which only one complementary DNA sequence is present in the genome. Because of the very large genome of most higher cells, the collision of RNA fragments with unique or nonrepeated DNA sequences is quite rare and a very long time or high concentration is required to hybridize RNA to nonrepeated DNA sequences. Thus, most measurements of the hybridization RNA to DNA have actually been based on the hybridization of RNA to repeated DNA.

The measurements described here, previously briefly reported (2), represent the first case for which evidence has been presented that hybridization of RNA with nonrepeated DNA sequences is observable, and thus that nonrepeated sequences are transcribed in higher cells. Reports of related observations have recently been published (3).

Estimates are made both of the fraction of rapidly labeled RNA that is complementary to nonrepeated sequences and the fraction of nonrepeated DNA that is transcribed, under specific circumstances. Neonatal or fetal mice were chosen for this study since the wide variety of rapidly growing tissues would be expected to supply a mixed population of RNA suitable for initial exploratory work.

METHODS

The distinction between repeated and nonrepeated DNA sequences

The boundary between repeated and nonrepeated DNA sequences is, in a sense, arbitrary. In fact, a very wide range occurs in what may be described as the precision or nearness of relationship of such sets. It is possible to discriminate effectively against less similar members, or perhaps whole sets, by varying the conditions of hybridization. Under the conditions of incubation and assay used in this work, about $^2/_3$ of the mouse genome appears to be made up of DNA sequences that occur in only one copy; the remaining third shows a great degree of repetition, ranging from a million copies down to a few thousand. There is, as yet, no evidence for the existence of repeated DNA with a small (2–10) number of copies (4).

Calculation of hybridization rates

Several parameters determine the rate of reassociation. Temperature, salt concentration, and fragment size are important but can be standardized and need little discussion. If these are maintained constant, then only the concentration of the complementary nucleotide sequences need be considered. In the simplest case, the DNA concentration is much greater than that of the RNA for all individual sequences and controls the rate of the reaction. The Cot* (calculated from the DNA concentration) required for half completion of hybridization with nonrepeated DNA is simply proportional to the genome size. The mammalian genome contains about 3×10^9 nucleotide pairs of DNA. This large quantity dilutes greatly the individual sequences. As a result, the rate of reassociation or hybridization of the nonrepeated DNA is low; high concentrations and long incubation times are required to get practical degrees of hybridization. For mouse DNA under our standard conditions (60°C, 0.12 M phosphate buffer pH 7), about 500

Abbreviation: SSC, saline–sodium citrate (0.15 M NaCl–0.015 M sodium citrate).

* Present address: U.S. Department of Health, Education, and Welfare, National Institutes of Health, National Institute of Allergy and Infectious Diseases, Laboratory of Parasitic Diseases, Bethesda, Md. 20014.

† Present address: 618 Life Sciences Building, Biophysics Department, Pennsylvania State University, University Park, Pa. 16802.

* Cot = the product of nucleic acid concentration and time expressed in mol/liter × sec.

nucleotide DNA fragments see below) the Cot for half-reassociation of nonrepeated DNA is about 4000 or 10 mg/ml for 40 hr.† If the salt concentration is raised to about 1 molar, the rate is increased by a factor of 8 and the Cot ($^1/_2$) becomes 500. Thus, at high salt concentrations, 10 mg/ml for 40 hr would bring the DNA reassociation reaction to about 90% completion. The available data is inadequate to establish the relative rates of RNA–DNA hybrid formation and DNA reassociation for comparable populations of complementary sequences. Measurements of ribosomal-RNA hybridization with the isolated DNA cistrons (5) give a much slower rate of hybridization than DNA reassociation. This retardation may be due to the small size of RNA fragments or the great secondary structure exhibited by ribosomal RNA. The data reported here (see Fig. 1) do not indicate a retardation of hybridization of rapidly labeled RNA. For the purpose of estimation, we have tentatively assumed that the RNA–DNA hybridization reactions have the same rate constants as the DNA–DNA reassociation reactions.

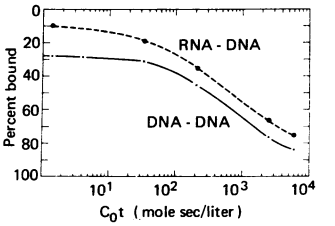

FIG. 1. The reassociation of mouse DNA and rapidly-labeled RNA. Uniformly labeled L-cell [^{14}C]DNA (0.02 mg/ml) and rapidly labeled [^{32}P]RNA from mouse embryo (0.1 mg/ml) were incubated at 60°C in 0.36 M phosphate buffer with unlabeled mouse DNA (6.85 mg/ml). The unreacted RNA was digested by RNase and deproteinized, and the solution was fractionated on hydroxyapatite. The ^{32}P was counted by Cerenkov radiation. The samples were then precipitated with trichloroacetic acid, the DNA was collected on filters, and the ^{14}C was determined.

DNA and RNA labeling and preparation

DNA was isolated by a modified Marmur method (6) including pronase digestion. Unlabeled bulk DNA was prepared from skinned, eviscerated mice and labeled DNA was prepared from L cells grown in tissue culture for 4–8 generations in the presence of [^{14}C]thymidine. A specific activity of about 10,000 cpm/μg was attained. Measurements of the reassociation (hydroxyapatite assay) of a mixture of these DNAs failed to detect any difference between L cell and mouse DNA in the relative quantities of repeated and non-repeated DNA (7). For measurements of reassociation or hydridization, the DNA was sheared to relatively uniform, small (about 500-nucleotide, single-stranded) fragments by passage through a needle valve at 50,000 psi using a specially built high-pressure pump (1).

Unlabeled RNA was prepared from newborn mice by the method of Scherrer and Darnell (8), using the hot-phenol step. Contaminating DNA was removed by DNAse treatment (50 μg/ml for 30 min, 37°C in 0.005 M $MgCl_2$–2.0 M CH_3COOK). Resulting DNA oligonucleotides were removed by three precipitations of the RNA with 75% ethanol at −10°C overnight (9). The DNAse treatment and precipitation cycles were repeated 4 times. Pulse-labeled RNA was prepared by a similar method from 17-day embryos removed 1 hr after an intraperitoneal injection (of the mother) with 10 mCi of [^{32}P]PO_4. This RNA was treated only once with DNAse.

Stability of RNA preparations

The long incubations at high concentrations make special demands on the stability of the RNA against various kinds of degradation. Stability was checked by measurement of the CCl_3COOH solubility of the RNA at the end of the incubations. Nuclease contamination appears to have been minimized by the methods of purification, which included passage of the RNA through a bed of Dowex-50 resin. The newborn-mouse RNA preparation was found to be free of contaminating nucleases upon incubation of the solution for 50 hr at 70°C with ^{32}P-labeled RNA. At the end of the incubation, 9% of the labeled RNA was acid soluble. Such a small amount of deterioration was not considered to be crucial. However, we found that RNA became CCl_3COOH soluble in neutral 2 × SSC (0.3 M NaCl–0.03 M sodium citrate) in minutes at 100°C or hours at 60°C. If the pH was adjusted to 5.5, or if the incubations were done in neutral phosphate buffer, the RNA was stable for the extended periods.

Preparation of labeled nonrepeated DNA

DNA from mouse L cells, labeled and sheared as described above, was denatured in 100°C for 3 min in phosphate buffer, incubated to a Cot of 1500, and passed over hydroxyapatite (Bio-Rad) at 60°C in 0.12 M phosphate buffer. 54% of the DNA has been reassociated and was bound to the column. Under these conditions, the (nonreassociated) fraction that did not bind is nearly pure nonrepeated DNA, as indicated by subsequent tests. This nonrepeated DNA was concentrated by absorption to a nitrocellulose filter and eluted in a small volume of water at 90°C.

RNA–DNA hybrid separation on Sephadex

A series of preliminary experiments were done to estimate a suitable ratio of RNA to DNA that would permit efficient hybrid formation with nonrepeated DNA sequences. For this purpose, DNA (sheared and denatured) was incubated to Cot 100 and fractionated on hydroxyapatite. The nonreassociated, essentially nonrepeated, fraction (68%) was incubated with fetal [^{32}P]RNA at various ratios. In each case, the DNA was present at 10 mg/ml in a volume of 0.15 ml of 6 × SSC, and the mixture was denatured at 100°C and incubated at 60°C for 16 hr. After incubation, the mixture was treated with RNase (15 min, 37°C, 5 μg/ml)[10] and chromatographed on Sephadex G-100 using a 25 × 1 cm column with a buffer flow rate of 30 ml/hr. The hybridized RNA appears in the excluded volume well separated from the digested RNA. When the RNA was digested with RNAse in the absence of DNA and the DNA subsequently added, no radioactivity was detected in the excluded volume. When the procedure was performed with salmon-sperm DNA instead of mouse DNA, there was no detectable hybridization.

Hydroxyapatite preparation of RNA–DNA hybrids

Under the conditions used in these measurements (0.12 M phosphate buffer and 60°C), a significant fraction of our RNA

† A Cot of 1 results if DNA is incubated for 1 hr at a concentration of 83 μg/ml, which corresponds to an absorbance of about 2.0 at 260 nm.

TABLE 1. *Reassociation with mouse DNA of the [^{14}C]DNA released from RNA–DNA hybrids*

Cot	DNA bound to column (%)		Increment of DNA bound (%)	
	Non-repeated	Unfractionated mouse	Non-repeated	Unfractionated mouse
2×10^{-1}	5.3	(18)		
25	13.6	28	8.3	10
2.5×10^{3}	48.3	61	34.7	33

The [^{14}C]DNA that formed hybrids with mRNA was released from the RNA by RNase, purified, and concentrated by absorption to a nitrocellulose filter and elution with water. 0.08 ml of 1.2 M phosphate buffer was added to 0.675 ml of this [^{14}C]DNA solution (2.2 μg DNA/ml). A 0.225-ml aliquot of this solution was incubated at 60°C for 100 hr (*first line*). 2 mg of unlabeled, unfractionated mouse DNA was added to the remaining solution; half of this was incubated for 10 hr and the other half for 100 hr. The amount of ^{14}C- and unlabeled-DNA that bound to the hydroxyapatite column was determined. The value in parentheses is the expected reassociation of unfractionated mouse DNA at this Cot.

preparations binds to hydroxyapatite in the absence of DNA. In one series of experiments, this problem was circumvented by using labeled DNA which only binds to hydroxyapatite when the DNA is double stranded. These measurements were done with low concentrations of DNA and very high concentrations of RNA in order to estimate the fraction of the nonrepeated sequences that are transcribed into RNA. For this purpose, 0.14 μg of nonrepeated [^{14}C]DNA was mixed, in 1 ml of 0.24 M phosphate buffer (the equivalent Cot‡ correction factor for this buffer is 2.93) with 64 mg of RNA from newborn mice. The mixture was denatured (100°C, 3 min) and incubated at 70°C for 50 hr. The incubation mixture was diluted into a large volume of 0.12 M buffer at 70°C and passed over a 25-ml column of hydroxyapatite 4 times; the bound DNA was collected after each passage and pooled. The bound fraction was eluted with 0.48 M phosphate buffer.

Combined assay of DNA reassociation and RNA–DNA hybrids

A series of samples were assayed on hydroxyapatite for hybrid formation and DNA reassociation. In this set of measurements, the samples consisted of a mixture of 0.1 mg/ml of ^{32}P-labeled fetal RNA, 6.9 mg/ml of mouse DNA, and 0.02 mg/ml of L-cell ^{14}C-labeled DNA which were incubated for different periods of time in 0.36 M buffer. (The equivalent Cot correction factor for 0.36 M buffer is 4.48.) After incubation, the unhybridized RNA was digested with RNase to prevent the binding of unhybridized RNA to hydroxyapatite. The RNase was removed by deproteinizing the sample with an equal volume of phenol, centrifuging, and passing the supernatant over a bed of Dowex-50 resin. The sample was then heated to 60°C and passed over hydroxyapatite in 0.12 M buffer; the bound fraction was eluted by raising the temperature to 100°C. The [^{32}P]RNA was assayed by Cerenkov radiation and the [^{14}C]DNA assayed after acid precipitation and collection on membrane filters. There was no detectable hybridization of the ^{32}P-labeled RNA with salmon-sperm DNA after comparable incubation when assayed by this method.

‡ Equivalent Cot is Cot multiplied by a factor that corrects for the difference in rate of reassociation due to salt concentration relative to the rate in 0.12 M phosphate buffer (11).

RESULTS

Estimate of the fraction of the unique DNA transcribed

After a high concentration of RNA was incubated with labeled nonrepeated DNA (as described in the *Methods* section under preparation of RNA–DNA hybrids), approximately 8% of the DNA was bound to hydroxyapatite. Since under our conditions, only $^2/_3$ of the genome is nonrepeated DNA, the bound DNA corresponds to about 5.6% of the total genome. This fraction was eluted from hydroxyapatite (0.48 M phosphate buffer, 70°C) so that the hybrids were retained in double-stranded form. Recovery of this fraction permitted further testing to show that it was composed of RNA–DNA hybrids and that nonrepeated DNA was involved.

The melting temperature of the RNA–DNA hybrid was determined by binding another aliquot of the incubation mixture to a hydroxyapatite column at 70°C and determining the radioactive DNA eluted when the temperature was raised by 5°C increments up to 100°C. The T_m of the hybrids was 77°C, which is about 5–7°C below the T_m of well-matched DNA strands eluted from hydroxyapatite under these conditions. There was very little melting above 85°C, where most of the DNA-strand pairs would melt. Since RNA–DNA hybrids have been shown to melt 5°C below DNA–DNA pairs (4), this result indicates that most, if not all, of the DNA had been in hybrid form.

The recovered hybrid fraction was given prolonged treatment with RNase (50 μg/ml, 18 hr, 37°C), while being dialyzed against 0.03 M phosphate buffer. After removal of the RNase, $^3/_4$ of the original 8% of [^{14}C]DNA which bound to the hydroxyapatite no longer would bind under the same conditions, indicating that at least this much DNA had been in hybrid form. When an aliquot of this single-stranded DNA was reincubated with another sample of RNA from the same preparation, and assayed under the same conditions as before, 37% was bound. Thus, this DNA, as expected, was enriched in the fraction that would hybridize with RNA. However, the fact that 100% of the DNA was not bound on the second hybridization indicated that the incubation time and RNA concentration were not sufficient to get complete hybridization

TABLE 2. *The reassociation of nonrepeated DNA with rapidly labeled RNA*

DNA Cot	DNA:RNA Ratio	Estimated ratio of DNA to mRNA	[^{32}P]RNA reacted (%)
1600	10	500	42
1600	1000	50,000	52.8

For purposes of calculation, the mRNA is estimated to be 2% of the total. These results probably underestimate the quantity of RNA homologous to the nonrepeated DNA because the reaction did not go to completion.

of the transcribed nonrepeated DNA. Thus, the 8% value for the nonrepeated DNA that is hybridizable with RNA is a minimum estimate of the nonrepeated DNA that is transcribed.

Another portion of the recovered RNase-treated DNA was incubated with a high concentration of unfractionated, unlabeled mouse DNA. The reaction of the [^{14}C]DNA was much less than that of the unfractionated mouse DNA at a low Cot as expected, since most of the repeated DNA sequences had been removed (Table 1). At a higher Cot, the [^{14}C]DNA that had been hybridized reacted to the same extent as the nonrepeated fraction of native mouse DNA: 35% versus 33%. Thus, the majority of the hybridizing DNA was nonrepeated DNA.

Fraction of the rapidly labeled RNA transcribed from nonrepeated DNA

A series of hybridization experiments were done with rapidly labeled RNA prepared from 17-day fetal mice. The first set of experiments determined the effect of the ratio of RNA to DNA on the extent of hybridization. Table 2 presents the results of a representative pair of measurements, in which the extent of RNA hybridization of nonrepeated DNA was assayed with sephadex after RNase treatment. It would appear that there is no great reduction in the fraction of RNA hybridized at the higher concentration of RNA. Some of the nonrepeated stretches of DNA may be fully occupied (saturated) at this RNA:DNA ratio, but they must be in a minority. Of course, the extent of reaction, in terms of amount of labeled RNA recovered in hybrids per mg of DNA, is greater at the higher RNA concentration. For practical reasons, therefore, we decided to carry out the extensive series of measurements at an intermediate ratio of 70 mg of DNA per mg of RNA.

Fig. 1 shows the kinetics of the hybridization of rapidly labeled fetal RNA with unfractionated mouse DNA. The reassociation of the DNA and the hybridization were both measured on the same set of samples as described in the *Methods* section (*Combined assay of DNA reassociation and RNA–DNA hybrids*). The earliest sample was taken after most of the repeated sequences had reacted. Further samples were taken until most of the nonrepeated DNA had reacted. Under these conditions the RNA–DNA-hybridization curve generally parallels that of the DNA-reassociation curve. In this set of measurements, the DNA is in excess and the kinetics of the reaction are determined by the concentration of the DNA sequences. The major part of the hybridization occurs in the late part of the reaction, just when the nonrepeated DNA reassociates. This result shows clearly that, in mouse fetal tissue, the major part of the rapidly labeled RNA is transcribed from nonrepeated DNA sequences.

The fraction of the RNA sequences that are transcribed from repeated DNA sequences is probably underestimated in this experiment as a result of the RNase treatment. Many of the hybrids with repeated DNA sequences can be expected to be only partially base paired, and, for many of these hybrids, the RNA may be RNase digestible to some extent.

The data of Fig. 1 imply that as much as 80% of the rapidly labeled RNA in mouse embryo is transcribed from nonrepeated DNA sequences. It appears that nearly as large a fraction of the RNA has formed hybrids as the fraction of DNA that reassociated.

This result also gives an estimate of the amount of ribosomal RNA precursor synthesized in mouse embryo. Surely, the ribosomal RNA cistrons would be saturated with ribosomal RNA at the DNA:RNA ratio of 70:1. As a result, only a small part of the radioactive ribosomal-precursor RNA would hybridize in this experiment. Thus, it appears that less than 20% of the rapidly labeled RNA is ribosomal precursor.

DISCUSSION

Each of the three methods tested in this paper indicates that nonrepeated DNA sequences are expressed. This is, of course, not a surprising conclusion. It is reasonable to assume that many of the structural genes are present in one copy, and that their DNA sequences are not repeated. Nevertheless, the complexity of the mammalian genome is great enough that even such an apparently elementary fact must be experimentally established. Similar measurements must be made in other tissues and circumstances, since, in the case of whole fetal and newborn mice, both the extent of the DNA and the amount of the rapidly-labeled RNA that is transcribed from nonrepeated DNA sequences are surprisingly large.

It appears that a minimum of 6% of the nonrepeated DNA is transcribed in the newborn mouse. This corresponds to 12% of one strand of DNA; presumably, only one strand of the DNA is transcribed. Since the nonrepeated DNA is about $^2/_3$ of the total DNA, this corresponds to a minimum of about 8% of the whole genome. The reaction on which this estimate is based apparently did not go to completion, and this figure is very likely an underestimate. When the DNA that had been hybridized was incubated again with RNA, under the conditions in which it originally hybridized, only $^1/_3$ of it reacted. This is our only estimate of the degree to which the hybridization reaction fails to be completed under these conditions and it may be inaccurate. However, it is quite possible that single-copy DNA amounting to $^1/_4$ of the genome is transcribed in newborn mice.

When this work was undertaken, no satisfactory prediction could be made as to the extent that the nonrepeated portion of the genome is transcribed in a given tissue. It was conceivable that only a few of the many millions of potential genes would be expressed at a given time. For this reason, the rapidly growing and complex embryo was chosen for these initial studies. It was assumed that a relatively large number of genes would be expressed, and, thus, the hybridization measurements would have a greater chance of success. This view was borne out. The complexity, or potential information content, of the transcribed DNA is large. The estimate of 25% of the genome amounts to about 8×10^8 nucleotide pairs.

This amount of DNA in structural genes would code for about two million different hemoglobin-sized proteins. Direct evidence would be valuable regarding whether all these RNA sequences are translated and for the existence of so many diverse proteins. No cellular requirements or functions are known for so many protein molecules. We are, therefore, led to raise the possibility that some unknown process of hybridization might have occurred. The high thermal stability indicates that the hybrids were base paired over an extended length. However, direct measurement of the length of base pairing is called for to support the radical conclusion that so many diverse base sequences are actually transcribed.

1. Britten, R. J., and D. E. Kohne, *Science*, **111**, 529 (1968).
2. Gelderman, A. H., A. V. Rake, and R. J. Britten, *Carnegie Inst. Washington Yearb.*, **67**, 320 (1969).
3. Davidson, E. H., and B. R. Hough, *Proc. Nat. Acad. Sci. USA*, **63**, 342 (1969).
4. Britten, R. J., and J. Smith, *Carnegie Inst. Washington Yearb.*, **68**, 378 (1970).
5. Kohne, D. E., *Biophys. J.*, **8**, 1104 (1969).
6. Marmur, J., *J. Mol. Biol.*, **3**, 108 (1961).
7. Britten, R. J., and A. V. Rake, *Carnegie Inst. Washington Yearb.*, **67**, 325 (1969).
8. Scherrer, K. M., and J. K. Darnell, *Biochem. Biophys. Res. Commun.*, **7**, 486 (1962).
9. Barlow, J., and Mathias, A. P., *Procedures in Nucleic Acid Research*, eds. J. L. Contoni and D. R. Davis (1966), p. 448.
10. Attardi, G., L. P. Huang, and S. Kabat, *Proc. Nat. Acad. Sci. USA*, **53**, 1490 (1965).
11. Britten, R. J., and J. Smith, *Carnegie Inst. Washington Yearb.*, **68**, 385 (1970).