MOLECULAR AND CELLULAR BIOLOGY, Apr. 1988, p. 1853–1856
0270-7306/88/041853-04$02.00/0

Vol. 8, No. 4

Developmentally Regulated Expression of a Human "Finger"-Containing Gene Encoded by the 5′ Half of the *ret* Transforming Gene

MASAHIDE TAKAHASHI,[1]* YUTAKA INAGUMA,[1] HIROSHI HIAI,[1] AND FUMIKO HIROSE[2,3]

Laboratory of Experimental Pathology[1] and Laboratory of Cell Biology,[2] Aichi Cancer Center Research Institute, Chikusa-ku, Nagoya 464, and Laboratory of Cancer Cell Biology, Research Institute for Disease Mechanism and Control, Nagoya University School of Medicine, Showa-ku, Nagoya 466,[3] Japan

Received 27 October 1987/Accepted 18 January 1988

We isolated and sequenced a cDNA clone of the human gene encoded by the 5′ half of the *ret* transforming gene. The nucleotide sequence indicates that it encodes a protein with "finger" structures which represent putative metal- and nucleic acid-binding domains. Transcription of this gene was detected at high levels in a variety of human and rodent tumor cell lines, mouse testis, and embryos. In addition, a unique transcript was observed in testis RNA. When the expression of the unique transcript was examined at different stages of spermatogenesis, a striking increase in mRNA levels accompanied progression from meiotic prophase pachytene spermatocytes to postmeiotic round spermatids. This finger-containing gene may thus function in male germ cell development.

The *ret* transforming gene was activated by recombination between two unlinked human DNA segments during transfection of NIH 3T3 cells (23). The nucleotide sequence of the *ret* cDNA indicated that the *ret* transforming gene was generated by the fusion of a truncated tyrosine kinase with previously unlinked amino-terminal sequences (22). It has recently been reported that recombination events also resulted in activation of transforming potential of several other protein kinase genes, including *met* (14), *trk* (12), *ros* (2), and *raf-1* (8, 21). In each case, a unique cellular sequence was fused to the amino terminus of a 5′-truncated gene with homology to the *src* kinase region. These fused amino-terminal sequences may modulate transforming activities of kinase genes or simply function as promoters.

Here we report the isolation and sequencing of a cDNA clone corresponding to the normal gene encoded by the 5′ sequences of the hybrid *ret* transforming gene. The amino-terminal sequences of *ret* contain a tandem repeat of a sequence which represents most of the consensus sequence of a putative metal- and nucleic acid-binding domain termed a "finger" structure (13). Thus, we have designated this finger-containing gene *rfp* (*ret* finger protein). Northern (RNA) blot analyses revealed that *rfp* is expressed at high levels in a variety of human and rodent tumor cell lines, mouse testis, and embryos.

Transcription of the amino-terminal sequences of the *ret* transforming gene was detected in all human tumor cell lines tested (22). To further characterize the nature of this gene, a cDNA library was constructed in λ gt10 from poly(A)$^+$ RNA of a THP-1 human monocytic leukemia cell line. A library of 1.3×10^5 recombinant bacteriophage was screened with a cDNA fragment from the 5′ end of the hybrid *ret* transforming gene. Five positive plaques were detected. The longest insert obtained was subcloned into plasmid pUC19 and sequenced by the chain termination method (18). The restriction map of the *rfp* cDNA is shown in Fig. 1. The sequence of 1,782 nucleotides contains an open reading frame of 1,539 nucleotides which encode a protein of 513 amino acids (Fig. 2). Two in-frame stop codons are found at nucleotides 4 and 31 in the 5′ untranslated region. The termination codon TGA at nucleotide 1774 is followed by 6 nucleotides of 3′ untranslated region. There is no other open reading frame which can encode a long polypeptide (>50 amino acids). Comparison of the sequence of the *rfp* cDNA with that of the *ret* cDNA (22) showed that the first 315 amino acids of *rfp* were fused with the 5′ truncated *ret* tyrosine kinase gene (Fig. 2).

The amino acid sequence of *rfp* was compared with those of other proteins by using the National Biomedical Research Foundation protein data bank. Although *rfp* was not significantly homologous to any known proteins, we detected the consensus sequence characteristic of finger structures, which have been identified in several species, including the *Xenopus* 5S gene transcription factor (TFIIIA) (13), the *Drosophila* loci *Krüppel* (17), *serendipity* (25), and *hunchback* (24), yeast *ADR1* (7), and mouse *mkr1* and *mkr2* (4). The *rfp* cDNA contains two finger domains (amino acids 16 to 37 and 53 to 74, underlined in Fig. 2) separated by 15 amino acids, although the last histidine residue of the consensus sequence in both fingers and the leucine residue in the second finger are absent (Fig. 3A). In the first finger domain, the last histidine residue is replaced with a cysteine residue which could be involved in the formation of a variant finger loop. The homology between the finger domains of this gene and those of the proteins mentioned above varies from 15 to 35%. These genes contain different numbers of finger domains with unique amino- and carboxy-terminal sequences.

In addition to the tandemly repeated finger domains, we

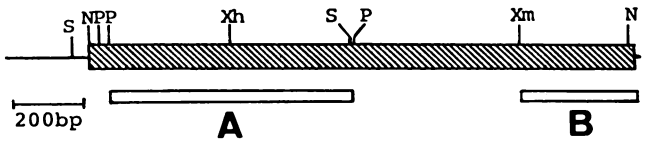


FIG. 1. Restriction map of *rfp* cDNA. Coding (▨) and noncoding (——) sequences for the *rfp* gene are indicated; ▭, cDNA fragments used for subsequent analyses. Restriction endonuclease sites are *Nco*I (N), *Pst*I (P), *Sac*I (S), *Xho*I (Xh), and *Xmn*I (Xm). bp, Base pairs.

* Corresponding author.

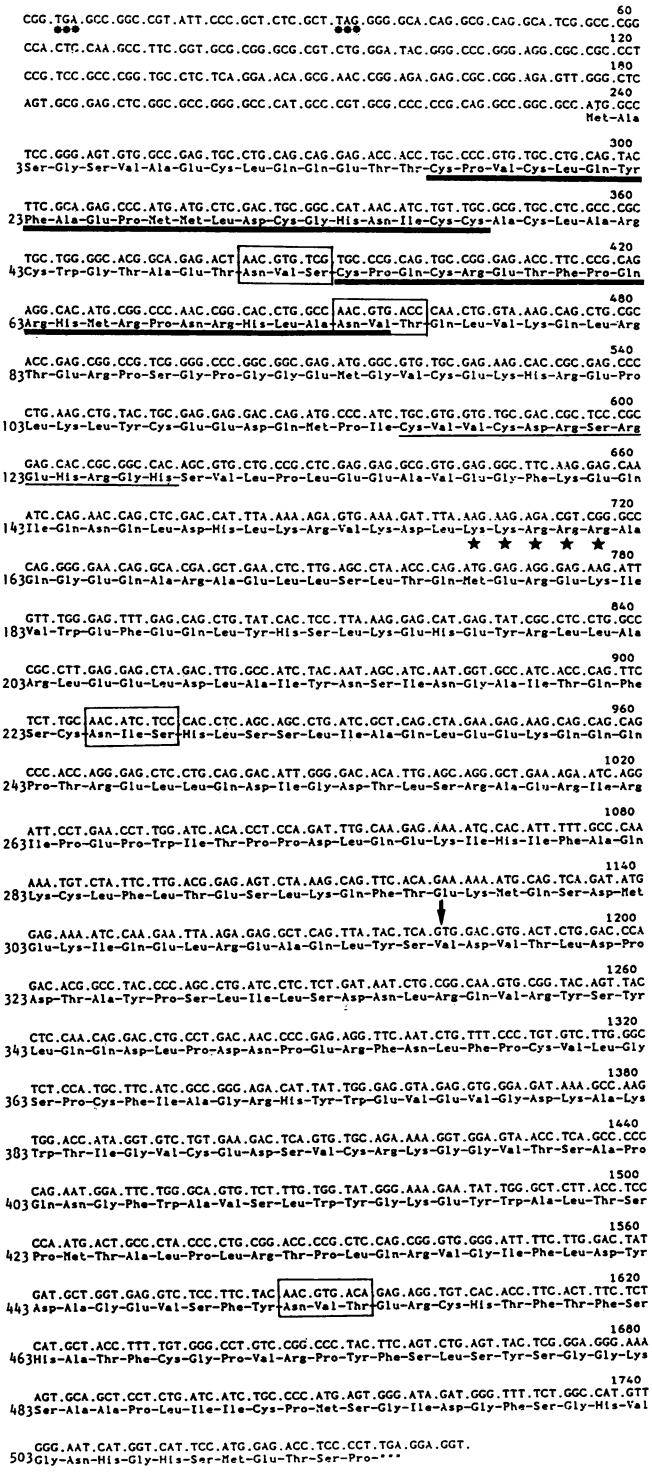


FIG. 2. Nucleotide sequence of *rfp* cDNA. The nucleotide sequence and the deduced amino acid sequence are shown. Nucleotides are numbered above the sequence, and amino acids are numbered in the left-hand margin. Two in-frame stop codons upstream of the putative initiation codon are indicated by the asterisks. The potential finger domains are underlined by thick and thin solid lines. A stretch of basic amino acids is indicated by the stars. Four possible glycosylation sites are enclosed in boxes. An arrow indicates the recombination site in the hybrid *ret* transforming gene.

detected another potential finger domain in *rfp* between amino acids 115 and 127, consisting of the sequence Cys-XX-Cys-XXXXX-His-XX-His (where X may be any amino acid; Fig. 2 and 3B). This kind of Cys- and His-containing sequence is conserved in a variety of nucleic acid-binding proteins (1). The predicted *rfp* amino acid sequence also includes a stretch of five basic amino acids (Lys-Lys-Arg-Arg-Arg, amino acids 157 to 161). This sequence is similar to the nuclear location signals of simian virus 40 large-T (9, 11) and polyoma large-T (16) antigens (Fig. 3C). In particular, the lysine residue conserved in both simian virus 40 and polyoma signals is found at a corresponding position (Lys-158) of *rfp*.

Miller et al. (13) proposed that the conserved Cys and His residues of the consensus sequence of finger domains form a metal binding core and that the sequences between these residues make a finger tip which might interact with specific nucleic acid sequences. The presence of finger domains and a potential nuclear location signal thus suggests that the *rfp* gene may encode a nuclear protein which binds nucleic acids.

We have previously reported that the *rfp* gene is transcribed as 2.4- and 3.4-kilobase (kb) RNAs in all human tumor cell lines tested (22). To find out whether rodent cells also express *rfp*, poly(A)$^+$ RNAs were prepared from mouse and rat tumor cell lines and analyzed by Northern blotting. As shown in Fig. 4A, a *Pst*I fragment of the *rfp* cDNA clone (designated "A"; Fig. 1) hybridized to a 2.4-kb transcript in all five rodent tumor cell lines tested. Another cDNA fragment (designated "B"; Fig. 1) hybridized to the same transcript in these cell lines (data not shown).

To further investigate the expression of *rfp* in normal mouse tissues, total RNAs from 11 organs of adult mice were examined with fragment A as a probe. They included thymus, testis, spleen, heart, ovary, kidney, brain, muscle, lymph node, lung, and liver. This probe detected *rfp* transcripts only in testis RNA as a broad 2.2- to 2.8-kb band (Fig. 4B). To confirm this result, poly(A)$^+$ RNAs from testis, kidney, and liver were hybridized with the same probe. Two

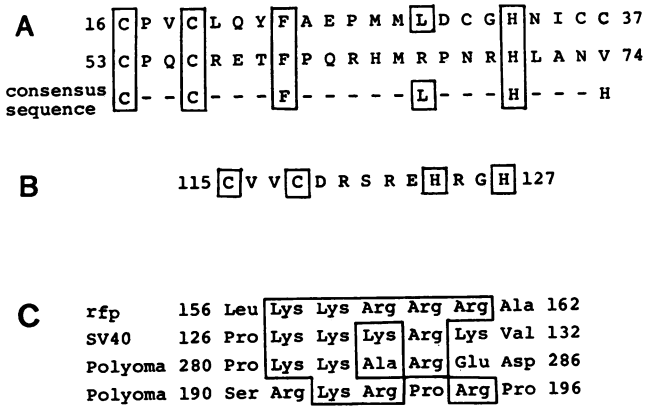
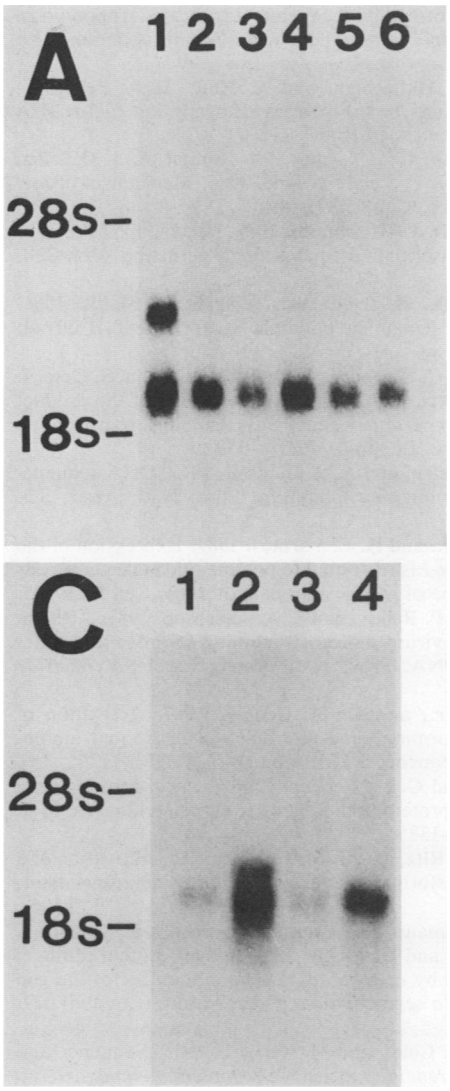


FIG. 3. Detection of finger domains and a potential nuclear location signal in the *rfp* gene. (A) The amino acid sequences of tandemly repeated finger domains (amino acids 16 to 37 and 53 to 74) in the *rfp* gene are aligned with the consensus sequence (13) of finger proteins. Identical amino acids are enclosed in boxes. (B) Another potential finger domain (amino acids 115 to 127) is shown. The Cys and His residues which may form a metal-binding core are boxed. (C) The homologous region of the *rfp* amino acid sequence is aligned with those of the nuclear location signals of simian virus 40 (SV40) large-T (9, 11) and polyoma large-T (16) antigens. Identical amino acids are enclosed in boxes.



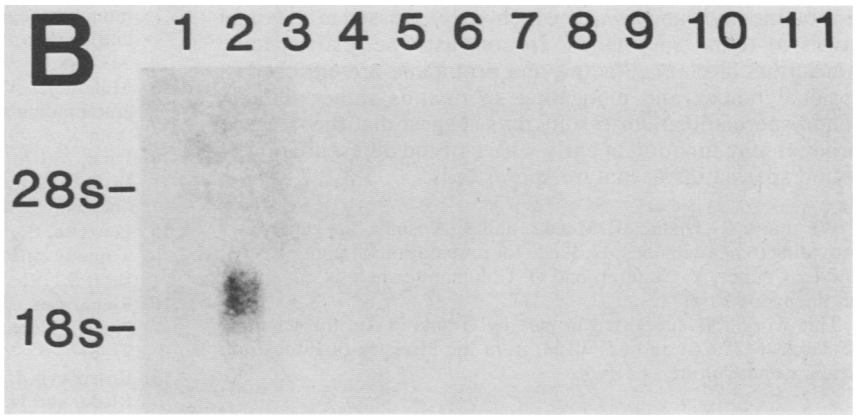


FIG. 4. Northern blot analyses with the *rfp* cDNA probe. A 5-μg portion of poly(A)$^+$ RNAs (A, C) and a 15-μg portion of total RNAs (B) were analyzed by Northern blot hybridization with fragment A of the *rfp* cDNA (Fig. 1). Poly(A)$^+$ RNAs were isolated from six tumor cell lines (A) and 4 BALB/c mouse tissues (C), and total RNAs were isolated from 11 BALB/c mouse tissues (B). Lanes in panel A: 1, THP-1 human monocytic leukemia; 2, T9 rat gliosarcoma; 3, EA285 rat glioma; 4, TR481 rat neurinoma; 5, M1 mouse myeloblastic leukemia; 6, P388D1 mouse monocytic leukemia. Lanes in panel B: 1, thymus; 2, testis; 3, spleen; 4, heart; 5, ovary; 6, kidney; 7, brain; 8, muscle; 9, lymph node; 10, lung; 11, liver. Lanes in panel C: 1, liver; 2, testis; 3, kidney; 4, 11.5-day embryos.

distinct bands of 2.4 and 2.8 kb were detected in testis RNA (Fig. 4C, lane 2), indicating that these two bands closely migrated as a broad band in the total RNA. Although we could not detect any transcript in the total RNAs from liver and kidney, low levels of expression of the 2.4-kb transcript were observed in poly(A)$^+$ RNAs of these tissues (Fig. 4C, lanes 1 and 3). The level of *rfp* message in testis was approximately 20-fold higher than that in liver or kidney. In addition, the 2.4-kb transcript of *rfp* was highly expressed in 11.5-day mouse embryos (Fig. 4C, lane 4). The high levels of *rfp* expression in a variety of tumor cell lines, mouse testis, and embryos suggest the possibility that mitotically and meiotically dividing cells contain high amounts of *rfp* mRNA.

To investigate whether the *rfp* gene is expressed in spermatogenic cells of mouse testis, we fractionated them by the gradient sedimentation method described by Chandley et al. (3). Three major fractions were prepared. The predominant cell types in the three fractions were pachytene spermatocytes (meiotic prophase), round spermatids (early postmeiotic cells), and residual bodies (cytoplasmic fraction from elongating spermatids) and elongating spermatids. The purity of these cell types was greater than 90%. Total RNAs from cells of each fraction were probed with the *rfp* cDNA fragment A. Although the probe hybridized to 2.4- and 2.8-kb transcripts in RNAs from the pachytene spermatocyte and round spermatid fractions, the level of the 2.8-kb transcript in round spermatids was approximately four- to fivefold higher than that in pachytene spermatocytes (Fig. 5). The expression of both transcripts strikingly decreased in residual bodies and elongating spermatids. When the same filter was rehybridized with v-*abl* probe (20) as a control, a 4.7-kb testis-specific *abl* transcript was observed in fractions enriched for round spermatids and residual bodies as previously described (15). A faint 4.7-kb band was also detected in the pachytene spermatocyte fraction but may represent a small amount of round spermatid contamination.

It is interesting that, like *rfp*, several genes including β-actin (26), α-tubulin (5), protamine (10), and protooncogenes c-*abl* (15), *int-1* (19), and c-*mos* (6) are expressed in a differentiation-specific pattern in germ cells. The expression pattern of the 2.8-kb *rfp* transcript is similar to those of

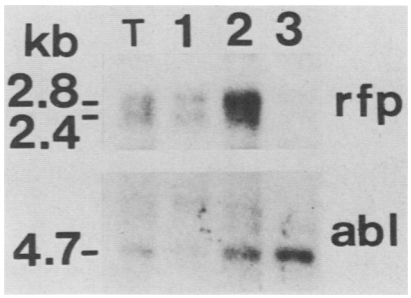


FIG. 5. Level of *rfp* transcript in spermatogenic cells of mouse testis. Total RNAs (3 μg) from fractionated male germ cells of 6-week-old ICR mice were hybridized with fragment A (Fig. 1) or v-*abl* probe (20). Lanes: T, total testis RNA; 1, pachytene spermatocyte RNA; 2, round spermatid RNA; 3, residual body and elongating spermatid RNA.

α-tubulin, *int-1*, and c-*mos*, which are expressed at maximal levels in round spermatids. In contrast, the testis-specific transcripts of c-*abl*, β-actin, and protamine are enriched in residual bodies and elongating spermatids rather than in round spermatids. Our results thus suggest that the *rfp* gene product may function at early stages of the differentiation of round spermatids to mature sperm cells.

We thank R. Ueda, M. Maeda, and J. Yoshida for generously providing tumor cell lines; H. Ikeda for preparation of mouse RNAs; and G. Cooper, V. Stanton, and D. Goldman for helpful comments on this manuscript.

This work was supported in part by Grants-in-Aid for scientific research (61770264 and 62570168) from the Ministry of Education, Science and Culture of Japan.

LITERATURE CITED

1. **Berg, J. M.** 1986. Potential metal-binding domains in nucleic acid binding proteins. Science **232**:485–487.
2. **Birchmeier, C., D. Birnbaum, G. Waitches, O. Fasano, and M. Wigler.** 1986. Characterization of an activated human *ros* gene. Mol. Cell. Biol. **6**:3109–3116.
3. **Chandley, A. C., Y. Hotta, and H. Stern.** 1977. Biochemical analysis of meiosis in the male mouse. Chromosoma **62**: 243–253.
4. **Chowdhury, K., U. Deutsch, and P. Gruss.** 1987. A multigene family encoding several "finger" structures is present and differentially active in mammalian genomes. Cell **48**:771–778.
5. **Distel, R., K. C. Kleene, and N. B. Hecht.** 1984. Haploid expression of a mouse testis α-tubulin gene. Science **224**:68–70.
6. **Goldman, D. S., A. A. Kiessling, C. F. Millette, and G. M. Cooper.** 1987. Expression of c-*mos* RNA in germ cells of male and female mice. Proc. Natl. Acad. Sci. USA **84**:4509–4513.
7. **Hartshorne, T. A., H. Blumberg, and E. T. Young.** 1986. Sequence homology of the yeast regulatory protein ADR1 with *Xenopus* transcription factor TFIIIA. Nature (London) **320**: 283–287.
8. **Ishikawa, F., F. Takaku, M. Nagao, and T. Sugimura.** 1987. Rat c-*raf* oncogene activation by a rearrangement that produces a fused protein. Mol. Cell. Biol. **7**:1226–1232.
9. **Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith.** 1984. Sequence requirements for nuclear location of simian virus 40 large-T antigen. Nature (London) **311**:33–38.
10. **Kleene, K. C., R. J. Distel, and N. B. Hecht.** 1984. Translational regulation and deadenylation of a protamine mRNA during spermatogenesis in the mouse. Dev. Biol. **105**:71–79.
11. **Lanford, R. E., and J. S. Butel.** 1984. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. Cell **37**:801–813.
12. **Martin-Zanca, D., S. H. Hughes, and M. Barbacid.** 1986. A human oncogene formed by the fusion of truncated tropomyosin and protein tyrosine kinase sequences. Nature (London) **319**: 743–748.
13. **Miller, J., A. D. McLachlan, and A. Klug.** 1985. Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J. **4**:1609–1614.
14. **Park, M., M. Dean, C. S. Cooper, M. Schmidt, S. J. O'Brien, D. G. Blair, and G. F. Vande Woude.** 1986. Mechanism of *met* oncogene activation. Cell **45**:895–904.
15. **Ponzetto, C., and D. J. Wolgemuth.** 1985. Haploid expression of a unique c-*abl* transcript in the mouse male germ line. Mol. Cell. Biol. **5**:1791–1794.
16. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44**: 77–85.
17. **Rosenberg, U. B., C. Schroder, A. Preiss, A. Kienlin, S. Cote, I. Riede, and H. Jackle.** 1986. Structural homology of the product of the *Drosophila Krüppel* gene with *Xenopus* transcription factor IIIA. Nature (London) **319**:336–339.
18. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
19. **Shackleford, G. M., and H. E. Varmus.** 1987. Expression of the proto-oncogene *int-1* is restricted to postmeiotic male germ cells and the neural tube of mid-gestational embryos. Cell **50**:89–95.
20. **Srinivasan, A., E. P. Reddy, and S. A. Aaronson.** 1981. Abelson murine leukemia virus: molecular cloning of infectious integrated proviral DNA. Proc. Natl. Acad. Sci. USA **78**:2077–2081.
21. **Stanton, V. P., Jr., and G. M. Cooper.** 1987. Activation of human *raf* transforming genes by deletion of normal amino-terminal coding sequences. Mol. Cell. Biol. **7**:1171–1179.
22. **Takahashi, M., and G. M. Cooper.** 1987. *ret* transforming gene encodes a fusion protein homologous to tyrosine kinases. Mol. Cell. Biol. **7**:1378–1385.
23. **Takahashi, M., J. Ritz, and G. M. Cooper.** 1985. Activation of a novel human transforming gene, *ret*, by DNA rearrangement. Cell **42**:581–588.
24. **Tautz, D., R. Lehmann, H. Schnurch, R. Schuh, E. Seifert, A. Kienlin, K. Jones, and H. Jackle.** 1987. Finger protein of novel structure encoded by *hunchback*, a second member of the gap class of *Drosophila* segmentation genes. Nature (London) **327**: 383–389.
25. **Vincent, A., H. V. Colot, and M. Rosbash.** 1985. Sequence and structure of the *serendipity* locus of *Drosophila melanogaster*: a densely transcribed region including a blastoderm-specific gene. J. Mol. Biol. **186**:149–166.
26. **Waters, S. H., R. J. Distel, and N. B. Hecht.** 1985. Mouse testes contain two size classes of actin mRNA that are differentially expressed during spermatogenesis. Mol. Cell. Biol. **5**:1649–1654.