MOLECULAR AND CELLULAR BIOLOGY, Apr. 1985, p. 885–889
0270-7306/85/040885-05$02.00/0

Vol. 5, No. 4

Characterization and Developmental Expression of a *Drosophila ras* Oncogene

BRIAN MOZER, ROBIN MARLOR, SUSAN PARKHURST, AND VICTOR CORCES*

Department of Biology, The Johns Hopkins University, Baltimore, Maryland 21218

Received 8 November 1984/Accepted 14 December 1984

We cloned a *Drosophila melanogaster ras* gene (Dmras64B) on the basis of its homology to the *ras* oncogen from Harvey murine sarcoma virus. This gene mapped at chromosomal position 64B on the left arm of the third chromosome. Sequencing of Dmras64B revealed extensive amino acid homology with the proteins encoded by the human and *Saccharomyces cerevisiae ras* genes. The coding region of the *Drosophila* gene is interrupted by two introns located in different positions with respect to its human counterpart. Dmras64B encodes three different RNAs (1.6, 2.1, and 2.6 kilobases long) that are constantly expressed throughout the development of the fly.

The transformed phenotype of tumor cells has been associated with the activation of normal cellular oncogenes by a variety of different processes. In the case of the *ras* genes, which were originally isolated from the Harvey and Kirsten murine sarcoma viruses, this activation results from somatic mutation, giving rise to a protein (p21) with an altered amino acid sequence (1, 14, 16–19, 21). Elevated expression of the *ras* proteins can also transform animal cells (2, 6). Nothing is known, however, about the molecular mechanisms by which the p21 *ras*-encoded protein causes the transformed phenotype.

To gain more insight into the biochemical and cellular roles of the *ras* gene family, we have begun a study of these genes in *Drosophila melanogaster*, taking advantage of the fact that they are so highly conserved during evolution (15). A Charon 4A *Drosophila* genomic library (obtained from T. Maniatis) was screened by using the *Bgl*II DNA fragment containing the Harvey murine sarcoma virus *ras* (Ha-*ras*) gene as a probe (3, 8). The hybridization was done in 30% formamide–0.75 M NaCl–0.075 M sodium citrate–0.2% polyvinylpyrrolidone–0.2% bovine serum albumin–0.2% Ficoll–1.0% sodium dodecyl sulfate–denatured calf thymus DNA (100 μg/ml) at 37°C for 12 h. The filters were then washed twice with 0.75 M NaCl–0.075 M sodium citrate at room temperature and twice with the same buffer at 60°C. Of the several positive clones obtained, one, designated λ-2A, was studied in detail. Figure 1 shows the restriction map of a *Pst*I subfragment, obtained from the λ clone, which contains all the sequences homologous to the Ha-*ras* oncogene. This DNA fragment was sequenced by using the protocol of Maxam and Gilbert (11) (see Fig. 1 for a diagram of the sequencing strategy), and the amino acid sequence of the different open reading frames was compared with that of the human and *Saccharomyces cerevisiae ras* oncogenes. We found three regions of homology interrupted by two stretches of nonhomologous sequences. These regions of nonhomologous sequences are likely to be introns because protein termination codons exist in all three possible reading frames. The precise location of the introns within the protein-coding regions was assigned on the basis of three different criteria: the existence of a consensus donor and acceptor splice site sequence, the maintenance of an open reading frame, and the maximization of the homology with human and yeast *ras* protein sequences.

Figure 2 shows the nucleotide sequence of the coding region of Dmras64B. The positions of the putative introns are indicated by open spaces in the corresponding amino acid sequence; we have also included several nucleotides from the bordering region of the introns to show the sequence of the consensus splice sites. The vertical arrows above the sequences indicate the locations of the introns in the human gene. The protein encoded by the *Drosophila ras* gene contains 195 amino acids and has a predicted molecular weight of 22,700. For the purpose of comparison with the yeast and human *ras* proteins, the amino acid sequences of both of the proteins encoded by the yeast *RAS1* and human *ras* genes (7, 13) are also shown in Fig. 2. It is apparent from the sequence comparison that the Dmras64B protein contains two extra amino acids at the amino-terminal end when compared with the human *ras* protein, but four amino acids fewer than the yeast *RAS1* gene product. For the next 90 amino acids, the *Drosophila* protein presents a high degree of homology (approximately 80%) with both yeast and

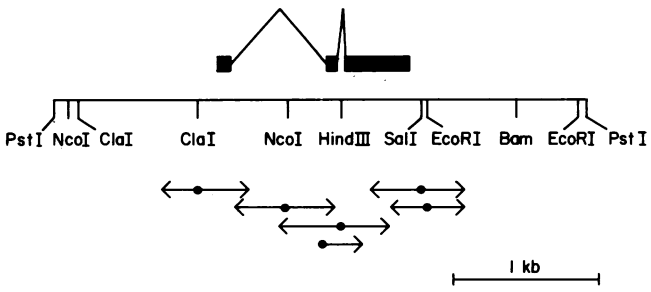

FIG. 1. Restriction map of the DNA sequences containing the Dmras64B gene. A Charon 4A *Drosophila* genomic library was screened by using a cloned Ha-*ras* gene as a probe (3, 8). Of the several positive clones obtained, one was analyzed in detail. By Southern blot analysis with the Ha-*ras* gene as a probe, it was found that all the *ras*-homologous sequences were contained within a 3.7-kilobase *Pst*I fragment. The restriction map of this fragment is shown above. Also shown is the position of the introns and coding exons deduced after a sequence analysis of this fragment. The lower part of the figure diagrams the sequencing strategy used to determine the nucleotide sequence of the Dmras64B gene.

* Corresponding author.

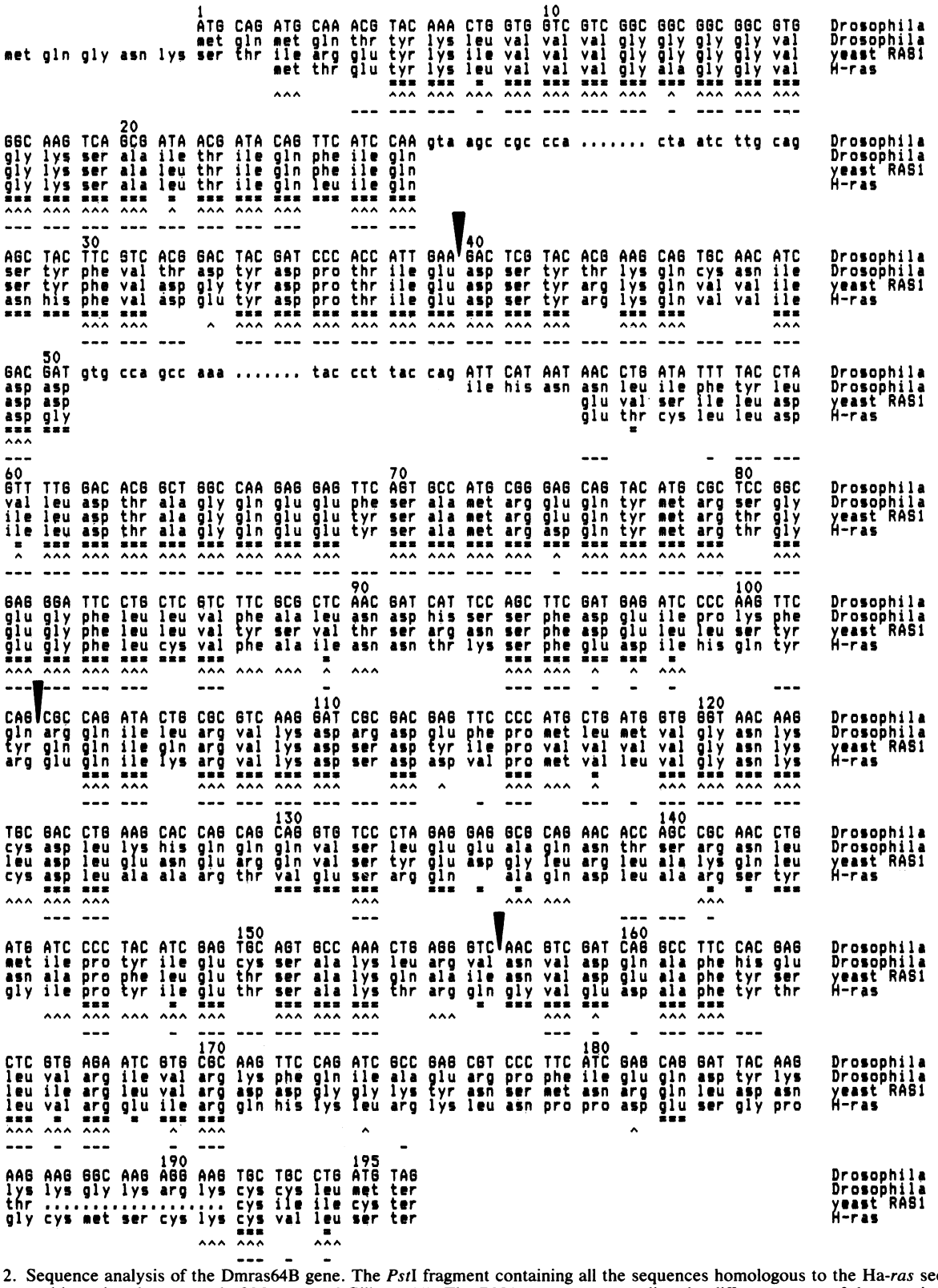

FIG. 2. Sequence analysis of the Dmras64B gene. The *Pst*I fragment containing all the sequences homologous to the Ha-*ras* sequences was sequenced by using the protocol of Maxam and Gilbert (11). The DNA sequences encoding the different exons of the protein and the corresponding amino acid sequence are shown. Also shown are the amino acid sequences for the yeast *RAS1* and human *ras* proteins (H-*ras*) for the purpose of comparison. Amino acids 192 to 305 from the yeast protein (indicated by dotted lines at the end of the *RAS1* protein sequence) have not been represented since this region does not have a corresponding region in the *Drosophila* and human genes (7, 13). The positions of the putative introns in the *Drosophila* gene are indicated by empty spaces in the amino acid sequence and by dotted lines in the nucleotide sequence. The first and last several nucleotides in the intron sequence are denoted by lowercase letters. The positions of the

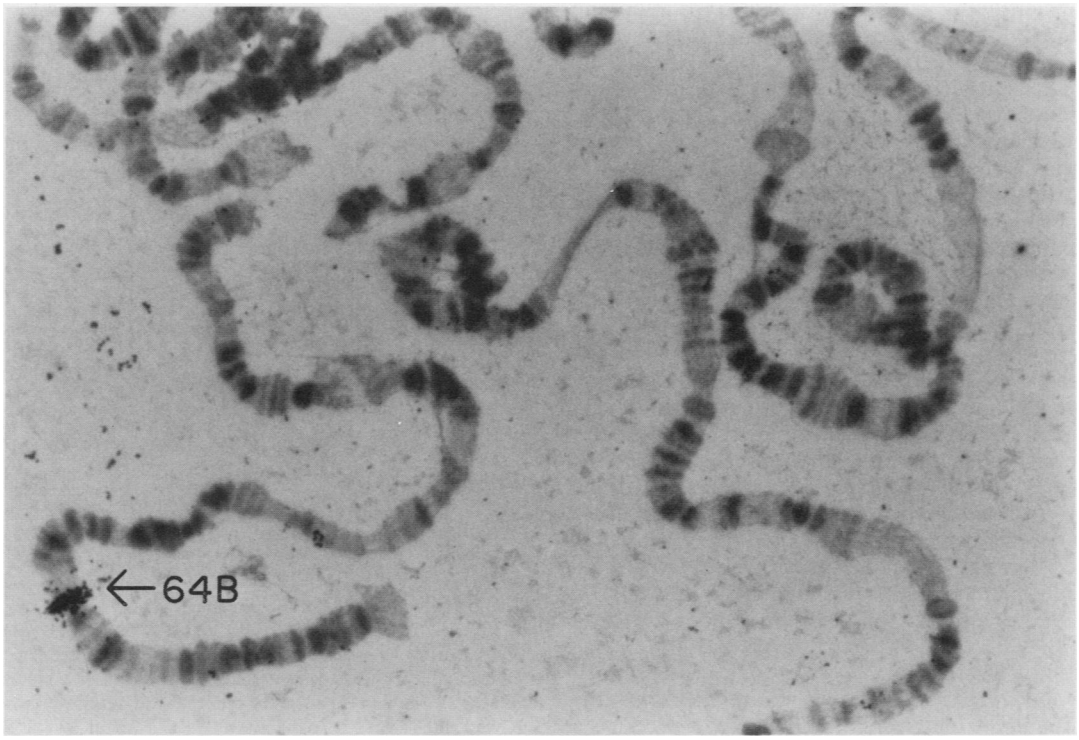

FIG. 3. In situ hybridization of Dmras64B to polytene chromosomes. Plasmid DNA containing the *Drosophila ras* gene was labeled with [^{125}I]dCTP by nick translation and hybridized to third-instar larvae polytene chromosomes. Grains can be seen at chromosomal position 64B on the left arm of the third chromosome.

human genes, except for the first 9 amino acids of the third exon. In this region, the Dmras64B protein contains three additional amino acids. The degree of relatedness between the *Drosophila* genes and the yeast and human genes breaks down after amino acid 90, and the homology is only patchy for the next 80 amino acids (between 55 and 60%). There is no homology between any of the *ras* genes for the region between amino acids 170 and 190, and the *Drosophila ras* protein ends with the conserved sequence CysAAX (as does its human homolog), where A is an aliphatic amino acid and X is the terminal one (7, 13).

It is interesting that the degree of homology between the *Drosophila* protein and either the human or the yeast *ras* protein is almost the same and very similar to the degree of homology that exists between yeast and human *ras* proteins. The distribution of this homology along the protein chain follows a specific pattern that has been observed when the different human *ras* proteins are compared with themselves and with those of yeasts (7, 13). This distribution supports the hypothesis of Powers et al. (13), which suggests that the degree of conservation of the amino acid sequence in different regions of the *ras* protein defines several domains with functional significances in the physiological role of the protein. The fact that the putative introns are located within these functional domains and the positions of the introns are different in the *Drosophila* and human *ras* genes suggests that the different exons of a protein may not correspond to functional units that are brought together during evolution to form a new protein.

These results differ from those described recently by Neuman-Silberberg et al. (12). They sequenced the same *Drosophila ras* gene and found no homology in the first portion of the protein-coding region when compared with the amino acid sequence of the human *ras* protein. We believe that the discrepancy in the results is due to the omission by Neuman-Silberberg et al. of two nucleotides in the DNA sequence of the Dmras64B gene in the region immediately upstream from the second exon of the protein. As a result of this, there is a change in the protein reading frame which causes a methionine residue to appear in the same reading frame as the *ras* protein. The DNA sequence that we have determined for the first intron (data not shown) contained a termination codon immediately before the second exon of the *Drosophila ras* protein when the proper reading frame was considered. We have found a stretch of 27 amino acids with more than 80% homology to the human and yeast *ras* proteins 646 kilobase pairs upstream from the second exon. This indicates that the sequence we have designated as the first exon truly represents the N-terminal region of the protein.

To determine the chromosomal location of the Dmras64B gene, plasmid DNA containing the *Pst*I fragment shown in

introns in the human gene are indicated by vertical arrows over the nucleotide sequence. Symbols: ===, amino acids conserved between Dmras64B and the yeast *RAS1* protein; ^^^, homology between *Drosophila* and human *ras* genes; ——, amino acids that have been conserved between the yeast and human *ras* proteins (7, 13). We have used the symbols =, ^ , and – to indicate conservative changes (gly and ala; glu and asp; lys and arg; or leu, val, and ile) in the amino acid composition of the *ras* proteins among the respective pairs of proteins to which we have referred.

Fig. 1 was labeled with [^{125}I]dCTP by nick translation and hybridized to polytene chromosomes from third-instar larvae salivary glands (4). Figure 3 shows the results of such an experiment. The hybridization is concentrated at chromosomal position 64B on the left arm of the third chromosome.

Since a yeast genome has two *ras* genes and the human genome contains three, we used genomic Southern blots to determine how many *ras* genes are present in the *Drosophila* genome. Total genomic DNA from *Drosophila* embryos was digested with different restriction enzymes, electrophoresed on agarose gels, transferred to nitrocellulose paper, and hybridized under different stringency conditions with labeled DNA from either Dmras64B or Ha-*ras* genes. The results of these experiments (data not shown) suggest that there are at least two, and most likely three, *ras*-like genes in the *Drosophila* genome. These results are in agreement with those of Neuman-Silberberg et al., who have recently demonstrated the existence of three different *Drosophila ras* genes (12). It may now be possible to delete one or all of the *ras* genes from *D. melanogaster* by using visible phenotypic markers located nearby; questions could then be asked about the effect that the absence of the *ras*-encoded protein has on the differentiation and development of the fly. In yeast cells it has been shown that deletion of both *ras* genes affects spore viability (10, 20).

The question of the role of the *ras* protein on the development of *D. melanogaster* is of special interest, since it is known that dominant mutations in protein p21 produce a transformed phenotype in mouse cells. To study the developmental expression of Dmras64B, we prepared total RNA from different stages of *Drosophila* development by lysing the tissues in 4 M guanidine isothiocyanate–0.2% lauroyl sarcosine–0.15 M mercaptoethanol–0.012 M EDTA–0.15 M Tris-HCl, (pH 7.5) in a Dounce homogenizer. After the addition of an equal volume of 0.1 M sodium acetate (pH 5.0) and three cycles of phenol extractions at 65°C followed by 10 min on ice, the RNA was precipitated with 2 volumes of ethanol (5). Polyadenylate-containing RNA was then prepared by chromatography on oligo deoxythymidylate-cellulose, and 5 μg of each RNA sample was electrophoresed on a 0.8% agarose-formaldehyde gel, transferred to a Gene Screen membrane (New England Nuclear Corp.), and hybridized with ^{32}P-labeled Dmras64B DNA (5). The results of this experiment are shown in Fig. 4.

There are three different RNAs homologous to the *Pst* fragment of Dmras64B, with sizes of 1.6, 2.1, and 2.6 kilobases, respectively. The amounts of these RNAs remain constant in the different stages during the development of the fly. Since the hybridization conditions used to detect these RNAs in the experiment described in the legend to Fig. 4 were very stringent, we do not think that some of these RNAs could be originating from other *ras* loci present elsewhere in the *Drosophila* genome. Furthermore, Dmras64B only hybridizes to one chromosome location under the same hybridization conditions.

These results suggest that the *ras*-encoded protein may not have a role in the induction of cellular differentiation or proliferation but that its function may be related to the maintenance of the proliferative state. Since it is possible to obtain proper expression of a cloned gene after the transformation into *Drosophila* embryos, the problem of the biochemical role of the *ras* protein can be easily approached by using this system. We are now in the process of obtaining a mutated Dmras64B gene which encodes an altered protein containing a valine residue instead of a glycine residue at position 14. This type of substitution produces the tumorigenic phenotype in mammalian systems (1, 14, 17–19, 21). The altered gene will be introduced into the *D. melanogaster* germ line by P-element-mediated transformation, and the effect of this dominant mutation on the development of the fly will be studied. If the transformed adult flies have a visible phenotype, it may be possible to obtain mutations in other genes that cause a reversion of the transformed phenotype. Following this kind of approach, it may be feasible to isolate genes whose protein products interact physically or biochemically with the *ras*-encoded protein and, therefore, to determine the biochemical pathways by which a mutation in the ras protein causes the tumorigenic phenotype.

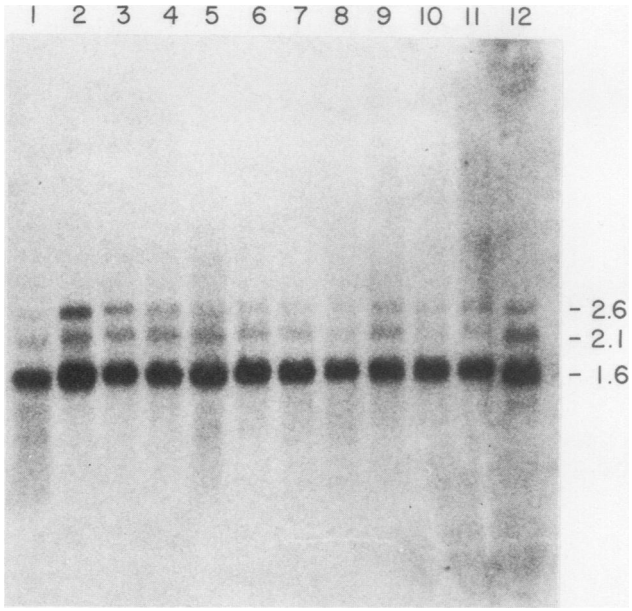

FIG. 4. Electrophoretic analysis of Dmras64B complementary RNA during *Drosophila* development. The different lanes in the gel contain: embryo RNA (lane 1); early and late first-instar larvae RNA (lanes 2 and 3, respectively); early and late second-instar larvae RNA (lanes 4 and 5, respectively); third-instar larvae RNA (lane 6); prepupae RNA (lane 7); RNA from different stages of pupal development 24 h apart (lanes 8 through 11); and RNA from adult flies (lane 12). The same filter was subsequently hybridized with a plasmid containing the 5C *Drosophila* actin gene (9) to control for the amount of RNA loaded on the gel (data not shown).

We thank I. Dawid for helpful criticisms of this manuscript.

This study was supported by Public Health Service Award GM32036 from the National Institutes of Health.

LITERATURE CITED

1. **Capon, D., P. Seeburg, J. McGrath, J. Hayflick, V. Edman, A. Levinson, and D. Goeddel.** 1983. Activation of Ki-*ras2* gene in human colon and lung carcinoma by different point mutations. Nature (London) **304**:507–513.
2. **Chang, E. H., M. E. Furth, E. M. Scolnick, and D. R. Lowy.** 1982. Tumorigenic transformation of mammalian cells induced by a normal human gene homologous to the oncogene of Harvey murine sarcoma virus. Nature (London) **297**:479–483.
3. **Chattopadhyay, S. K., E. H. Chang, M. R. Lander, R. W. Ellis, E. M. Scolnick, and D. R. Lowy.** 1982. Amplification and rearrangement of *onc* genes in mammalian species. Nature (London) **296**:361–363.
4. **Corces, V., R. Holmgren, R. Morimoto, and M. Meselson.** 1980. Four heat shock proteins in *Drosophila melanogaster* coded

within a 12 kilobase region in chromosome subdivision 67B. Proc. Natl. Acad. Sci. U.S.A. **77**:5390–5393.

5. **Corces, V., A. Pellicer, R. Axel, and M. Meselson.** 1981. Integration, transcription and control of a *Drosophila* heat shock gene in mouse cells. Proc. Natl. Acad. Sci. U.S.A. **78**:7038–7042.
6. **DeFeo, D., M. A. Gonda, H. A. Young, E. A. Chang, D. R. Lowy, E. M. Scolnick, and R. W. Ellis.** 1981. Analysis of two divergent rat genomic clones homologous to the transforming gene of Harvey murine sarcoma virus. Proc. Natl. Acad. Sci. U.S.A. **78**:3328–3332.
7. **Defeo-Jones, D., E. Scolnick, R. Koller, and R. Dhar.** 1983. *ras*-Related gene sequences identified and isolated from *Saccharomyces cerevisiae*. Nature (London) **306**:707–709.
8. **Ellis, R. W., D. DeFeo, T. Y. Shih, M. A. Gonda, H. A. Young, N. Tsuchida, D. R. Lowy, and E. M. Scolnick.** 1981. The p21 *src* genes of Harvey and Kirsten sarcoma viruses originate from divergent members of a family of normal vertebrate genes. Nature (London) **292**:506–511.
9. **Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson.** 1983. Transcripts of the six *Drosophila* actin genes accumulate in a stage and tissue specific manner. Cell **33**:115–123.
10. **Kataoka, T., S. Powers, C. McGill, O. Fasano, J. Strathern, J. Broach, and M. Wigler.** 1984. Genetic analysis of yeast *RAS1* and *RAS2* genes. Cell **37**:437–445.
11. **Maxam, A. M., and W. Gilbert.** 1980. Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol. **65**:499–560.
12. **Neuman-Silberberg, F. S., E. Schejter, F. M. Hoffman, and B. Shilo.** 1984. The *Drosophila ras* oncogenes: structure and nucleotide sequence. Cell **37**:1027–1033.
13. **Powers, S., T. Kataska, O. Fasano, M. Goldfarb, J. Strathern, J. Broach, and M. Wigler.** 1984. Genes in *S. cerevisiae* encoding proteins with domains homologous to the mammalian *ras* proteins. Cell **36**:607–612.
14. **Reddy, E., R. Reynolds, E. Santos, and M. Barbacid.** 1982. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. Nature (London) **300**:149–152.
15. **Shilo, B., and R. Weinberg.** 1981. DNA sequences homologous to vertebrate oncogenes are conserved in *Drosophila melanogaster*. Proc. Natl. Acad. Sci. U.S.A. **78**:6789–6792.
16. **Shimizu, K., D. Birnbaum, M. Ruley, D. Fasano, Y. Suard, L. Edlun, E. Taparowsky, M. Goldfarb, and M. Wigler.** 1983. The structure of the K-ras gene of the human lung carcinoma cell line Calu-1. Nature (London) **304**:497–500.
17. **Tabin, C., S. Bradley, C. Bargmann, R. Weinberg, A. Papageorge, E. Scolnick, R. Dhar, D. Lowy, and E. Chang.** 1982. Mechanism of activation of a human oncogene. Nature (London) **300**:143–148.
18. **Taparowsky, E., K. Shimizu, M. Goldfarb, and M. Wigler.** 1983. Structure and activation of the human N-*ras* gene. Cell **34**:581–586.
19. **Taparowsky, E., Y. Suard, O. Fasano, K. Shimizu, M. Goldfarb, and M. Wigler.** 1982. Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. Nature (London) **300**:762–765.
20. **Tatchell, K., D. T. Chaleff, D. DeFeo-Jones, and E. M. Scolnick.** 1984. Requirement of either of a pair of *ras*-related genes of *Saccharomyces cerevisiae* for spore vability. Nature (London) **309**:523–527.
21. **Yuasa, Y., S. Srivastava, C. Dunn, J. Rhim, E. Reddy, and S. Aaronson.** 1983. Acquisition of transforming properties by alternative point mutations within *c-bas/has* human proto-oncogene. Nature (London) **303**:775–779.