DNA sequences homologous to vertebrate oncogenes are conserved in Drosophila melanogaster
(acute leukemia viruses/evolution)

Ben-Zion Shilo* and Robert A. Weinberg

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Communicated by David Baltimore, August 4, 1981

ABSTRACT Sequences homologous to the oncogene sequences of acute RNA tumor viruses have been shown to be highly conserved within vertebrates. In the present work, eight different oncogene DNA sequences have been used as probes to search for homologous sequences in the DNA of organisms of other phyla. Five of these probes hybridized to the DNA of Drosophila melanogaster. Abelson leukemia virus probe detected a single homologous DNA fragment in Drosophila DNA. In contrast, probes prepared from the genomes of Harvey, avian, and feline sarcoma viruses and avian myelocytomatosis virus hybridized with multiple homologous sequences in Drosophila DNA. The identification of sequences homologous to vertebrate oncogenes in invertebrates demonstrates both a high degree of conservation of these genes and a wide distribution among divergent species. It seems likely that sequences homologous to vertebrate oncogenes play a crucial role in metazoan metabolism.

The acutely oncogenic retroviruses constitute a class of 15–20 virus types that are able to induce rapidly a variety of leukemias and sarcomas in appropriate hosts (1–3). These strains arose upon passage of slowly leukemogenic viruses through rodents, cats, and chickens. The new viruses that emerged differ from the original, slowly leukemogenic parent viruses in two respects. First, most of these viruses lack portions of the genetic information of the parental virus and, hence, have become defective in replication. Second, in place of the deleted genetic sequences, these viruses have acquired new genetic information which has, in many cases, been identified as being responsible for their acutely oncogenic properties (1–9).

The source of this new genetic information has been traced in many instances to the genome of the host animal through which the parent virus was originally passed (4–6, 9–14). Thus, each of the transforming genes that have been acquired by the virus can be associated with a counterpart sequence in the genome of the host. These transforming genes have been termed "oncogenes," whereas their normal cellular counterparts have been referred to as "proto-oncogenes" (6).

In one well-studied case, it has been shown that the protein encoded by the proto-oncogene appears similar to the induced oncogene protein (14–16). Such a result raises questions regarding the mechanism of transformation by oncogenes and the role of their gene products in normal cellular metabolism. One possibility is that the association of the proto-oncogene with the genome of a retrovirus results in transformation due to higher levels of expression of the gene (14–20). It remains possible as well that changes in structural sequences of the proto-oncogenes give rise to transforming capacity in the resulting oncogene.

A variety of observations have shown that DNA and protein sequences of some well-characterized proto-oncogenes are highly conserved within vertebrates (4, 5, 14, 15, 21). Therefore, it appears that proto-oncogenes, like many other cellular genes, fulfill functions that are essential in vertebrates at the level of cellular metabolism or tissue differentiation. A clue to the function of proto-oncogenes was provided by experiments showing that the products of certain proto-oncogenes may interact with one another in a cascade pathway of protein kinases (22, 23), whose function may be related to regulation of cellular energy metabolism.

We wished to get a broader view of the evolution of proto-oncogenes by looking for the most primitive organism in which these genes could be detected. The presence of these genes in organisms very dissimilar from vertebrates would suggest a role of these genes in cellular metabolism rather than in vertebrate-specific tissue differentiation. Moreover, the genetic systems of these less complex organisms could be exploited to gain new insights into the function of proto-oncogenes.

MATERIALS AND METHODS

Sources and Length of the Oncogene Specific Probes. The oncogene specific probes were kindly provided to us by the following scientists: (i) Abelson leukemia virus probe (clone pA4Bu9: 3-kilobase pair (kb) insert; J. Wang and D. Baltimore); (ii) avian myelocytomatosis virus (MC29) (clone MyC3-Pst: 1.5-kb insert; J. M. Bishop); (iii) avian sarcoma virus (ASV) (Pst II E fragment: 800-bp insert; J. M. Bishop); (iv) ST/feline sarcoma virus (FeSV) (clone Pst-3: 700-bp insert; C. Sherr); (v) Moloney murine sarcoma virus (clone p440: 440-bp insert; S. Coff and D. Baltimore); (vi) Kirsten murine sarcoma virus (clone HiiHi-3: 1-kbp insert; R. Ellis and E. Scolnick); (vii) Avian erythroleukemia virus (clone pA4E6 Pst II: 2.5-kbp insert; J. M. Bishop); and (viii) Harvey murine sarcoma virus (HaSV) (clone B5-9: 450-bp insert; R. Ellis and E. Scolnick). This last-named probe is not reactive with rat 30 sequences (R. Ellis, personal communication).

The clones contain either a fragment of the oncogene sequence or the entire oncogene and are not reactive with the leukemia virus sequences. All probes were cloned and propagated in the bacterial plasmid pBR322.

Cellular DNA. Mouse DNA was extracted from NIH3T3 cells. The other DNAs were kindly provided by the following investigators: Drosophila melanogaster fly DNA, M. Hoffman; A4E6 Drosophila cell line DNA, L. Cherbas; Lytechinus pictus sea urchin DNA, J. Ruderman; Caenorhabditis elegans ne-

Abbreviations: kb, kilobase(s); kbp, kilobase pair(s); HaSV, Harvey murine sarcoma virus; MC29, avian myelocytomatosis virus; ASV, avian sarcoma virus; FeSV, feline sarcoma virus.

* Present address: Dept. of Virology, Weizmann Institute, Rehovot, Israel.
mate DNA, V. Ambros; Saccharomyces cerevisiae yeast DNA, N. Neff; and Dicyostelium discoideum slime mold DNA, R. Kessin.

Nucleic Acid Hybridization. DNA was digested with restriction endonucleases, fractionated by agarose gel electrophoresis, and transferred to nitrocellulose filter paper by the Southern procedure (24). The nitrocellulose filters to which DNA was bound were prehybridized for 3–6 hr at 40°C in 35% (vol/vol) formamide/0.75 M NaCl/75 mM sodium citrate/65 mM KH₂PO₄/5 mM EDTA/0.1% polyvinyl pyrrolidone/0.1% Ficoll/1% bovine serum albumin containing 500 μg of boiled salmon sperm DNA per ml. Hybridization was done at 40°C for 12–24 hr in 35% formamide/0.75 M NaCl/75 mM sodium citrate/65 mM KH₂PO₄/5 mM EDTA/0.02% polyvinyl pyrrolidone/0.02% Ficoll/0.2% bovine serum albumin containing 100 μg of salmon sperm DNA per ml, 10% (wt/vol) Dextran sulfate, and nick translated, ³²P-labeled probe (25) (5 × 10⁶ cpm; specific activity of about 1 × 10⁹ cpm of DNA per μg). After hybridization, the filters were washed in 0.3 M NaCl/30 mM sodium citrate/25 mM KH₂PO₄/1 mM EDTA/0.1% NaDodSO₄. Initially the filters were washed at 55°C and then exposed to x-ray film. The filter was then incubated in a wash solution held at 60°C and reexposed to film. The filter was then re-washed at 68°C and exposed a final time to x-ray film. The filters were washed each time for 1–2 hr in 0.3 M NaCl/30 mM sodium citrate/25 mM KH₂PO₄/1 mM EDTA/0.1% NaDodSO₄.

RESULTS

We wished to use retrovirus-derived onc probes to search for homologous sequences in the DNA of a variety of species using the Southern blot procedure (24). The use of such probes to detect distantly related sequences is successful only if a limited degree of sequence divergence has occurred during the evolution of these genes. Excessive sequence divergence results in less stable hybrids, and consequently a reduction in hybridization signal may occur. To compensate for such divergence, we incubated probes with filter-bound DNAs at a formamide concentration lower than normally used, in order to favor formation of less stable hybrids.

A second factor compensated as well for a decrease in hybridization signal. The lower sequence complexity of many nonvertebrate genomes made it possible to apply a greater number of genome equivalents of DNA to each gel channel used in the Southern analysis. This resulted in turn in a proportionate enhancement of any signal detected by hybridization.

Cloned, sequence-specific probes for eight different vertebrate oncogenes were provided by several investigators. The DNA of plasmids containing these sequences was labeled by nick translation (25) and used initially to probe nitrocellulose filters carrying DNA extracted from mouse or Drosophila melanogaster cells. These DNAs had been cleaved with various restriction enzymes prior to electrophoresis and gel-filter transfer. Each probe consisted only of oncogene sequences and, therefore, carried no DNA sequence of a parental leukemia virus. Probes derived from Moloney murine sarcoma virus, Kirsten sarcoma virus, and avian erythroblastosis virus detected homologous fragments in mouse DNA, but no hybridization with Drosophila DNA was detected (data not shown). However, the other five probes hybridized not only with mouse DNA but also with DNA extracted from Drosophila.

The simplest pattern of hybridization with Drosophila DNA was observed when using the oncogene probe of Abelson murine leukemia virus. A single band was observed after digestion of the Drosophila DNA with BamHI, EcoRI, or HindIII (Fig. 1). The single band that appeared after BamHI digestion corresponds to a DNA fragment of only 2 kbp, even though the probe contained 3 kbp of Abelson oncogene sequences. Thus, it is likely that only a portion of the Abelson virus oncogene is conserved in Drosophila DNA. However, within the sequences that are conserved, the degree of mismatch to the probe is limited because hybridization could be readily detected even after a subsequent high-temperature wash of the filter. The Abelson proto-oncogene in mice is highly spliced and has a size of at least 30 kilobase (kb) (26). The homologous gene in Drosophila is either unspliced or contains only small introns.

In contrast to the Abelson virus probe, the probes prepared from HaSV, MC29, ASV, and FeSV detected multiple DNA fragments in Drosophila DNA (Figs. 2, 3). The number of bands that appeared varied from 3 to 10, depending upon the probe and restriction enzyme used. Because the oncogene sequence in each of these four probes is less than 1.5 kbp, the multiplicity of bands cannot be accounted for solely by cleavage of a single stretch of homologous sequence into several DNA fragments. Rather, each one of these four oncogene probes appears to detect a family of genes present in Drosophila DNA. The pattern of the bands is different for each probe, indicating that these four different gene families are distinct from one another. Within a given gene family, differences in the intensity of hybridization are observed. For example, an EcoRI digest of Drosophila DNA was probed with a 450-bp HaSV-specific sequence, and three bands were resolved (Fig. 2). A fragment of 6 kbp annealed intensely. An intermediate degree of hybridization to a fragment of 9 kbp was detected, and faint hybridization appeared with a DNA fragment of 5 kbp. The differences in intensity are likely to be a reflection of the degree of divergence of the different members of the family from the mammalian oncogene sequence. The intensity of the darker bands in Drosophila DNA was greater than that of the major band.

![Figure 1: Hybridization of Abelson leukemia virus DNA probe with Drosophila and mouse DNAs. Drosophila melanogaster DNA was extracted from flies and digested with different restriction enzymes. Four micrograms was loaded on each channel and fractionated by agarose gel electrophoresis. The DNA was transferred to nitrocellulose filter and hybridized with 32P-labeled Abelson virus DNA. DNA was digested by endonuclease BamHI (channel a), EcoRI (channel b), HindIII (channel c), or Pat I (channel d). DNA from NIH3T3 mouse cell (10 μg) was digested with BamHI (channel e). Size markers are in kbp. This exposure was obtained after a filter wash at 68°C.](image-url)
The ability to detect sequences in *Drosophila* DNA that are homologous to five different vertebrate oncogenes encouraged us to search for homologies in the DNA of organisms from other phyla and kingdoms including sea urchin, nematode, yeast, and slime mold. No homologies were detected when yeast or slime mold DNA was probed with these five oncogenes. With nematode DNA, however, one faint band was observed after hybridization to HaSV probe and to Abelson virus probe (data not shown), three bands were detected as a result of hybridization to MC29 probe (Fig. 3), and no hybridization to the FeSV probe was observed. Sea urchin DNA demonstrated only faint hybridization to the Abelson virus probe (data not shown) and no hybridization to the other probes. The failure to detect sequences in sea urchin is likely a consequence of its genomic complexity, which is 1 order of magnitude higher than that of *Drosophila* DNA. This reduces the number of genome equivalents that can be applied to the gel and, consequently, lowers the intensity of the signal observable upon Southern analysis.

**DISCUSSION**

In this work we have shown that five different vertebrate oncogene probes hybridized with homologous sequences in *Drosophila melanogaster* DNA. The results cannot be explained by the reactivity of the pBR322 plasmid DNA sequences that are present in all of the labeled probes used in the hybridizations. Various results point to this conclusion. For example, analysis of a single preparation of EcoRI-cleaved *Drosophila* DNA yielded a different pattern of hybridization with five of the probes used and no hybridization with the other three probes. Moreover, when one of the filters was used sequentially with two different probes, distinct patterns of hybridization were observed. Therefore, the observed hybridization was specific to the oncogene sequences in each probe.

The DNA sequences of several vertebrate oncogenes have been shown to be present within mammals and birds (4, 5, 21). Such findings implied that some of these genes evolved more than 300 million years ago and may fulfill functions essential for vertebrate cell metabolism or differentiation. In this work, we show that five vertebrate oncogene sequences can be detected in *Drosophila* DNA. We conclude that the common precursors of these genes were already evolved 800 million to a billion years ago at a period prior to the divergence of the Annelid-arthropod from the Echinoderm-Chordate superphylum. These sequences must play roles that are crucial to metazoan cellular or organismal physiology. Four of the oncogenes which we found to be present in *Drosophila* DNA are known to encode kinases in vertebrate cells (28–31). Some of these gene products have also been reported to be members of a cascade of interacting kinases which may be important in regulating energy metabolism (22, 23). This would suggest that those oncogenes and their nonchordate homologs code for centrally important functions of cellular metabolism which may be found in all eukaryotic cells.

We have been unable to detect sequences that are homologous to those of vertebrate oncogenes in unicellular eukaryotic organisms such as yeast by using the Southern technique. Such findings would have indicated that these sequences play a role in essential metabolic functions unrelated to processes of differentiation occurring in multicellular organisms.

The ability to detect sequences in *Drosophila* that are homologous to vertebrate oncogenes suggests that they should be conserved in all organisms of the Echinoderm-Chordate and the Annelid-Arthropod superphylum. Even though sea urchin is evolutionarily closer to vertebrates than are arthropods, oncogene homologs were not detected in sea urchin DNA by the Southern blotting procedure. This is likely due to the higher sequence complexity, which is 1 order of magnitude higher than that of *Drosophila* DNA. This reduces the number of genome equivalents that can be applied to the gel and, consequently, lowers the intensity of the signal observable upon Southern analysis.

**FIG. 2.** Hybridization of HaSV DNA probe with *Drosophila* and mouse DNAs. DNAs were analyzed as in Fig. 1. HaSV probe was hybridized with DNA extracted from *D. melanogaster* flies and digested with restriction enzymes. Four micrograms of DNA was loaded on each channel: a, BamHI digest; b, EcoRI; c, HindIII, d, Pst I. The same probe was incubated with DNA extracted from A4E *Drosophila* cells grown in culture. Four micrograms of A4E DNA was applied to each gel channel; e, BamHI digest; f, EcoRI; g, HindIII; h, 10 μg of EcoRI-cleaved mouse NIH3T3 cell DNA. This exposure was obtained after a filter wash at 68°C.

**FIG. 3.** Hybridization of MC29, FeSV, and ASV DNA probes with *Drosophila* nematode DNAs. DNAs were analyzed as in Fig. 1. MC29 probe was incubated with DNA extracted from *Drosophila* flies. Four micrograms was loaded on each channel: a, BamHI digest; b, EcoRI; c, HindIII, d, Pst I; e, the same probe hybridized with 5 μg of *C. elegans* nematode DNA that had been digested with HindIII. FeSV probe was hybridized with *Drosophila* DNA cleaved with HindIII (channel f) or with Pst I (channel g). ASV probe was hybridized with *Drosophila* DNA cleaved with EcoRI (channel h) or with HindIII (channel i). These exposures were obtained after filter washes at the following temperatures: 68°C (channels a–d) and 80°C (channels e–i).
complexity of the sea urchin genome. Nevertheless, it may be possible to study the oncogene homologs of organisms having high genomic complexity by isolating these genes in the form of molecular clones from phage vector libraries. Such isolation would make possible a detailed study of these sequences whose role as functional genes is at present only speculative. Thus, the spectrum of organisms in which proto-oncogene structure and function could be studied would be significantly broadened.

We thank J. Wang, D. Baltimore, E. Scolnick, J. M. Bishop, C. Sherr, S. Goff, R. Ellis, and D. Lowy for providing oncogene DNA clones; L. Parada for help in plasmid preparations; M. Hoffmann, L. Cherbas, J. Ruderman, V. Ambros, N. Neff, and R. Kessin for providing cellular DNA; and J. Toole for critical reading of the manuscript. B.-Z.S. was a Chaim Weizmann Postdoctoral Fellow. This work was supported by National Cancer Institute Grant 26717 to R.A.W. and Grant 14051 to S. E. Luria.