

Sequence and Structure of the Coding Region of the Human *H-ras-1* Gene from T24 Bladder Carcinoma Cells

Ottavio Fasano, Elizabeth Taparowsky, John Fiddes, Michael Wigler, and Mitchell Goldfarb

Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, U.S.A.

Summary: We have molecularly cloned and sequenced cDNA to the transcript of *H-ras-1*, the transforming gene of the T24 human bladder carcinoma cell line. The transcript derives from at least five exons in the *H-ras-1* gene, and RNA splicing occurs at sites typical of exon-intron junctions. T24 *H-ras-1* RNA has an AUG-initiated open reading frame of 567 nucleotides, which can encode a protein of mass comparable to the apparent molecular weight of the T24 *H-ras-1* gene product. The T24 *H-ras-1* gene product is nearly identical to v-*H-ras* p21, the transforming protein encoded by the genome of Harvey sarcoma virus. We discuss the implications of this sequence conservation in the structure-function relationships of ras proteins. **Key Words:** cDNA—*H-ras-1* gene—T24 bladder carcinoma cells—Gene sequence.

The genomes of higher organisms contain a set of genes which are homologous to the oncogenes carried in acutely pathogenic retroviruses (1-7). This suggests that neoplasia can result from the faulty activation of cellular homologs to the viral oncogenes. There are now experimental findings consistent with this hypothesis. Certain human tumor cell lines contain genes that can induce growth transformation of NIH 3T3 murine fibroblasts *in vitro* when these genes are introduced by the calcium phosphate method of DNA-mediated

gene transfer (8-13). Some of these transforming genes have been identified as human homologs to retroviral oncogenes. The activated transforming gene in many human lung and colon carcinoma cell lines is the human homolog to v-*K-ras*, the oncogene of Kirsten sarcoma virus (12,14). The transforming gene in the T24 (and the related EJ) human bladder carcinoma cell lines is *H-ras-1*, the human homolog to the oncogene of Harvey murine sarcoma virus (v-*H-ras*) (12,15,16). v-*H-ras* and v-*K-ras* are virally transduced rat cellular genes which encode related but distinct 21,000 dalton proteins (3).

We report the molecular cloning of cDNA templated by the 1.1-kilobase (kb) transcript (17) of the T24 bladder carcinoma transforming gene,

Received December 16, 1982; revised February 18, 1983; accepted February 22, 1983.

Address correspondence and reprint requests to Mitchell Goldfarb, Cold Spring Harbor Laboratory, P.O. Box 100, Cold Spring Harbor, NY 11724, U.S.A.

H-ras-1. This cDNA has been sequenced in its entirety. A comparison of cDNA to genomic sequences has allowed us to define the exons and introns within the T24 *H-ras-1* gene. The predicted amino acid sequence of the *H-ras-1* gene product is nearly identical to the v-*H-ras* encoded protein, differing at only three of 189 residues. This virtually precise gene conservation through evolution suggests that the normal functions of the *H-ras* gene product cannot tolerate even limited changes in primary structure.

Furthermore, the structural differences between *H-ras* and *K-ras* proteins must be critical to the differential functions of these related gene products.

MATERIALS AND METHODS

Materials

Restriction endonucleases (Bethesda Research Laboratories and New England Biolabs) were used according to suppliers' instructions. Polynucleotide kinase, T4 DNA ligase, *E. coli* DNA polymerase I, and large fragment Klenow DNA polymerase I were purchased from N. E. Biolabs, AMV reverse transcriptase from Life Sciences Inc., nuclease S1 from Sigma, and DNase I from Worthington Biochemical. Synthetic oligonucleotide *Eco*RI and *Sal*I linkers were purchased from Collaborative Research. Radioisotope-labeled nucleotides were obtained from Amersham.

Bacterial Strains

All plasmids were transformed into and maintained in *E. coli* strain DH-1.

Plasmid Strains

pT24 is a pBR322 clone containing the biologically active T24 *H-ras-1* gene within a 6.2-kb pair *Bam*HI restriction endonuclease fragment. The gene was transferred to pBR322 from recombinant bacteriophage λ T24 (19).

Preparation of DNA

Plasmid DNAs and eukaryotic genomic DNA were prepared as described elsewhere (10,20).

Construction and Screening of cDNA Library

T24a1-1 polyadenylated RNA (see Results) was the template for construction of cDNA molecules. The entire construction and cloning protocol has been employed by Helfman et al. (21), and will be described in detail by Fiddes and Hanahan (unpublished observations). A brief description is given in the text. Transformed colonies were screened by filter hybridization (22) for plasmid sequences homologous to T24 *H-ras-1*. For hybridization probe, the 6.2-kb pair *Bam*HI restriction fragment containing T24 *H-ras-1* was labeled with 32 P by nick translation (23) with DNA polymerase I, DNase I, and α - 32 P-dNTPs.

Southern Blot Filter Hybridization

Restriction endonuclease cleaved DNAs were electrophoresed through 1.0% agarose gels, and then transferred to nitrocellulose filter paper by the method of Southern (24). DNA hybridization probe labeled to a specific activity of 2×10^8 dpm/ μ g DNA by nick translation (23) was hybridized to filter-bound DNA as described elsewhere (25).

DNA Sequencing

DNA sequencing of 32 P-end-labeled DNA fragments was performed with the chemical base modification and cleavage procedures of Maxam and Gilbert (26). All segments of cDNA were sequenced on both strands either by sequencing in opposing directions with 3'-end-labeled DNA fragments or by sequencing in parallel directions with both 3'- and 5'-end-labeled DNA fragments (see Fig. 2). Only

portions of the T24 H-ras-1 gene were sequenced in both directions.

RESULTS AND DISCUSSION

Isolation of cDNA Molecular Clones Homologous to the T24 Transforming Gene

We have estimated that the H-ras-1 transcript represents approximately 0.01% of total polyadenylated RNA in T24 bladder carcinoma cells (17). An NIH 3T3 cell line transformed with the T24 H-ras-1 gene, T24a1-1, has several-fold more specific transcript than does the T24 tumor cell itself (data not shown). We therefore chose T24a1-1 RNA as our template for cDNA clones.

cDNA molecular clones in plasmid vector pBR322 were derived by the method of Fiddes and Hanahan (unpublished observations) as described by Helfman et al. (21). Briefly, cDNA was synthesized by oligothymidilate-primed reverse transcription of polyadenylated cytoplasmic RNA from T24a1-1 cells and, after alkaline hydrolysis of RNA, second-strand DNA synthesis was self-primed with reverse transcriptase. *SalI* linkers were ligated to the blunt end of the double-stranded hairpin cDNA molecules. The hairpin loops were digested with endonuclease S1, and *EcoRI* linkers were ligated to the other end of the cDNA molecules. The DNA was cleaved with *SalI* and *EcoRI* restriction endonucleases and ligated to the large *EcoRI/SalI* fragment of pBR322 DNA. The chimeric plasmids were used in a high efficiency transformation protocol (18) to derive ampicillin-resistant *E. coli* colonies, which were then screened for sequences homologous to the H-ras-1 gene by colony filter hybridization. We used a ³²P-labeled 6.2-kbp *Bam*HI restriction endonuclease DNA fragment containing the entire bladder carcinoma-transforming gene (19) as hybridization probe. From a total of approximately 5×10^4 colonies, we obtained three colonies containing plasmids with nucleotide sequences homologous to the probe.

Detection of H-ras-1 Sequences in cDNA Clones

The three plasmids (RS-3, RS-4, RS-6) containing cDNA sequences homologous to the T24 transforming gene were analyzed by gel electrophoresis and Southern filter hybridization. RS-3 contains a 1.1-kbp *EcoRI/SalI* cDNA insert, whereas RS-4 and RS-6 carry inserts of 0.4 kbp and 1.0 kbp, respectively. All three inserts are homologous to sequences in the T24-transforming gene (data not shown). However, these data do not exclude the possibility that these colonies contain cDNA sequences derived from the NIH 3T3 murine homolog to the H-ras gene. We therefore used cDNA clone RS-6 as hybridization probe to filter-blotted restriction endonuclease-cleaved DNAs from human T24 bladder carcinoma cells, NIH 3T3 cells, and NIH 3T3 transformants T24a5-4 and T24a2-2 containing the T24 H-ras-1 gene.

The results of these hybridizations are shown in Fig. 1. RS-6 hybridizes to both human

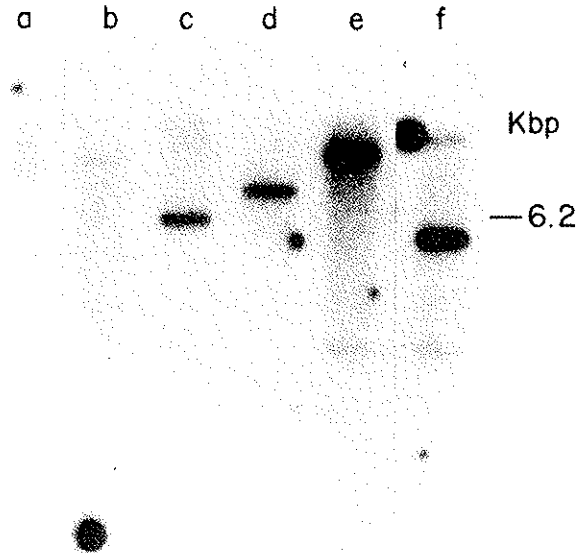


FIG. 1. Hybridization of cDNA clone RS-6 to genomic DNAs. Six micrograms of each genomic DNA were cleaved with restriction endonuclease, electrophoresed through a neutral pH 1.0% agarose gel, and blotted to nitrocellulose (24). The filter-bound DNA was hybridized to RS-6 cDNA-containing plasmid which had been labeled with ³²P by nick translation. Lane a, NIH 3T3 DNA with *Bam*HI endonuclease; b, NIH 3T3 with *Bgl*II; c, T24 with *Bam*HI; d, T24 with *Bgl*II; e, T24a5-4 with *Bam*HI; f, T24a2-2 with *Bam*HI.

T24 and murine NIH 3T3 DNAs (lanes a-d), but the intensity of hybridization to human DNA is far greater. RS-6 also hybridizes to newly acquired sequences in transformants of NIH 3T3 induced by the T24 *H-ras-1* gene (lanes e,f). Therefore, we conclude that RS-6 contains cDNA sequences from the T24 *H-ras-1* transcript. Indeed, coding sequences in RS-6 show complete concordance with the sequences in the coding exons of the cloned human genomic *H-ras-1* gene (see below). The hybridization between RS-6 and mouse DNA sequences may be due to the conservation of *H-ras* sequences among mammalian species.

Composite Nucleotide Sequence of cDNA Clones

We have sequenced the cDNA inserts of RS-6, RS-4, and portions of RS-3. Restriction endonuclease DNA fragments of the cDNA inserts were each ^{32}P -labeled at the terminal nucleotide of one end of one DNA strand, and labeled DNA was subsequently sequenced by the Maxam/Gilbert base modification and cleavage procedure. Figure 2 shows the sites for ^{32}P -labeling and the direction of sequencing. The sequences of the three cDNA inserts are overlapping (Fig. 2). The RS-4 insert contains a subset of the sequences in RS-6, and the over-

lapping sequences contain no base mismatches. RS-3 contains 150 base pairs with perfect homology to RS-6, demonstrating that RS-3 contains cDNA to the human *H-ras-1* transcript. To one side of this overlap, the RS-3 insert contains 63 nucleotides terminating in polyadenylic acid. This presumably represents the 3' end of a *H-ras-1* transcript. On the other side of the overlap are mouse sequences unrelated to *H-ras* (data not shown). The presence in RS-3 of mouse sequences together with human *H-ras-1* sequences may represent spurious DNA-DNA linkage established during the ligase reactions in the cDNA synthesis and cloning procedures.

There is also some question as to the origin of some of the sequences in RS-6. RS-6 cDNA contains the entire *H-ras-1* coding sequence (see below) along with 138 base pairs 5' to the ATG initiator codon. However, RS-6 sequences more than five base pairs 5' to the initiator ATG do not derive from *H-ras-1* genomic sequences directly upstream from the initiation codon, and indeed are not homologous to any portion of the T24 *H-ras-1* transforming gene contained within the 6.2-kbp *Bam*HI fragment (data not shown). This may be another example of cDNA cloning artifacts or may indicate that initiation of the RNA transcript, which served as template for our

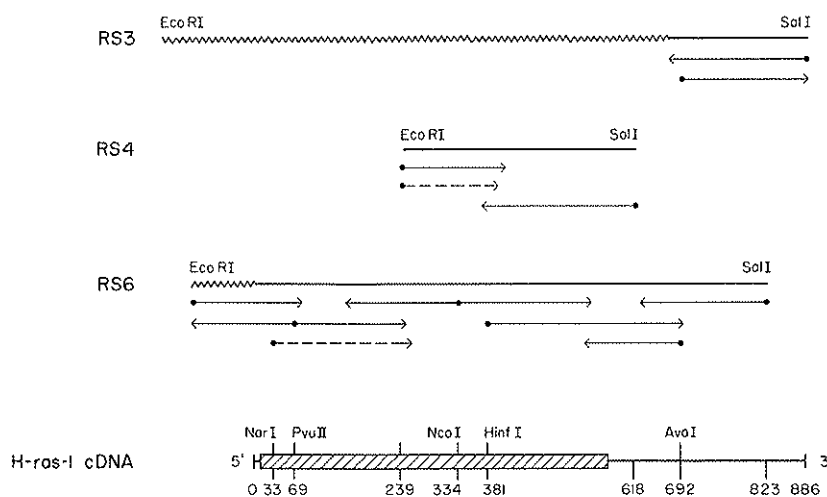


FIG. 2. Scheme for sequencing T24 *H-ras-1* cDNA. The cDNA inserts of plasmid clones RS3, RS4, and RS6 (flanked by *Eco*RI and *Sal*I cleavage sites) were sequenced by the method of Maxam and Gilbert (26). Sites for 5' ^{32}P -end labeling with polynucleotide kinase or 3' ^{32}P -end labeling with DNA polymerase I are shown as circles, and the subsequent direction and extent of sequencing is indicated by either dotted arrows (5' labeled sequence) or solid arrows (3'-labeled). The cDNA inserts are aligned in accordance with their overlapping *H-ras-1* cDNA sequences. **Bottom:** deduced composite structure of T24 *H-ras-1* cDNA; the map positions of naturally occurring and linker-constructed restriction endonuclease cleavage sites used for sequencing are indicated. Zig-zag, sequences unrelated to *H-ras-1*; solid line, T24 *H-ras-1* cDNA; striped bar, translational open reading frame for *H-ras-1* cDNA.

H-ras-1 cDNA 5'..GAGCG

```

v-K-ras      --- --T --G --- --A --T --- --A --T --A --T AGT --C --A --- --- --- --C T-- --G
v-H-ras      --- --A --- --C --- --T --- --- --- --- --T AGA --C --- --A --- --- --- --C --- ---
H-ras-1 cDNA ATG ACG GAA TAT AAG CTG GTG GTG GTG GGC GCC GTC GGT GTG GGC AAG AGT GCG CTG ACC65
H-ras-1 cDNA met thr glu tyr lys leu val val val gly ala val gly val gly lys ser ala leu thr
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      --A --- --A --T --- --T --C --- --- --- --T --- --T --T --T --G --- C-- --C --- ---
v-H-ras      --- --- --- --- --- --- --- --- --- --- --G --T --T --- --- --- --- --C --- ---
H-ras-1 cDNA ATC CAG CTG ATC CAG AAC CAT TTT GTG GAC GAA TAC GAC CCC ACT ATA GAG GAT TCC TAC125
H-ras-1 cDNA ile gln leu ile gln asn his phe val asp glu tyr asp pro thr ile glu asp ser tyr
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      A-- --A --A --A --A --- --- --A --A --- --C --T --C --- --T --T --C --C --A A-A --T
v-H-ras      --- --A --- --A --- --- --- --- --- --- --- --T T-A C-- --- --- T-A --C --A A-A --T
H-ras-1 cDNA CGG AAG CAG GTG GTC ATT GAT GGG GAG ACG TGC CTG TTG GAC ATC CTG GAT ACC GCC GGC185
H-ras-1 cDNA arg lys gln val val ile asp gly glu thr cys leu leu asp ile leu asp thr ala gly
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      --A --- --- --- --T --A --- A-- --- --- --- --- --- A-A --T --- --- --- --T --T ---
v-H-ras      --A --A --- --T --T --- --- --- --- --- --- --- --A --- --- --- --- --C --- ---
H-ras-1 cDNA CAG GAG GAG TAC AGC GCC ATG CCG GAC CAG TAC ATG CGC ACC GGG GAG GGC TTC CTG TGT245
H-ras-1 cDNA gln glu glu tyr ser ala met arg asp gln tyr met arg thr gly glu gly phe leu cys
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      --A --- --- --A --T --T --T --A --A --- --A --T --T --- --T --T --A --A --A T-A
v-H-ras      --A --- --- --- --- --- --- --- --- --- --- --- --A --- --- --- --- --- --- ---
H-ras-1 cDNA GTG TTT GCC ATC AAC AAC ACC AAG TCT TTT GAG GAC ATC CAG CAG TAC AGG GAG CAG ATC305
H-ras-1 cDNA val phe ala ile asn asn thr lys ser phe glu asp ile his gln tyr arg glu gln ile
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      --- A-A --A --- --- --T --A --T --- --T --- --C --A --A --- --T --- --- --- T--
v-H-ras      --G --- --- --A --T --A --- --T --- --A --- --- --- --- --C --- --- --- --- ---
H-ras-1 cDNA AAA CGG GTG AAG GAC TCG GAT GAC GTG CCC ATG GTG CTG GTG GGG AAC AAG TGT GAC CTG365
H-ras-1 cDNA lys arg val lys asp ser asp asp val pro met val leu val gly asn lys cys asp leu
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      C-- T-T A-A --A --A --C A-G AAA --- --- --- --G T-A --A A-G --T --T --G --T --A
v-H-ras      --- -GT --- --- --T --G --- --- --- --- --- --- --T --T --C --- --T --- ---
H-ras-1 cDNA GCT GCA CGC ACT GTG GAA TCT CCG CAG GCT CAG GAC CTC GCC CGA AGC TAC GGC ATC CCC425
H-ras-1 cDNA ala ala arg thr val glu ser arg gln ala gln asp leu ala arg ser tyr gly ile pro
v-H-ras      * gly * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      pro ser * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      -T-- --T --- --- --A --G --- --A A-A --- A-- --- --- --- --T --T --T --A --- ---
v-H-ras      --- --T --A --A --A --- --- --- --- --- --- --T --A --- --- --- --- --A C-A --A
H-ras-1 cDNA TAC ATC GAG ACC TCG GCC AAG ACC CCG CAG GGA GTG GAG GAT GCC TTC TAC ACG TTG GTG485
H-ras-1 cDNA tyr ile glu thr ser ala lys thr arg gln gly val glu asp ala phe tyr thr leu val
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      phe * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      A-A --- --- --A --- T-- -GA T-A AAA --- --A A-C -G- AAA GAA --A A-- -C- CCT GG- T-T
v-H-ras      --- --- --T --- --- --T --A --- --- --- --- --- --G --- --- --- --- --- --T ---
H-ras-1 cDNA CGT GAG ATC CCG CAG CAC AAG CTG CCG AAG CTG AAC CCT CCT GAT GAG AGT GGC CCC GGC545
H-ras-1 cDNA arg glu ile arg gln his lys leu arg lys leu asn pro pro asp glu ser gly pro gly
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

v-K-ras      GTA --AA --T AAA --A --C --T A-A ATG -A-
v-H-ras      --- --- --- --- --- --- --- --- --- --G --- --- --- --- --- --- ---
H-ras-1 cDNA TGC ATG AGC TGC AAG TGT GTG CTC TCC TGA575
H-ras-1 cDNA cys met ser cys lys cys val leu ser ter
v-H-ras      * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *
v-K-ras      val lys ile lys * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *

```

615
695
775
855

H-ras-1 cDNA CGCAGCACAGCTCAGGACATGGAGGTGCCGATGCAGGAAGGAGGTGCAGACGGAAGGAGGGAAGGACCGGAAG⁶⁵⁵

H-ras-1 cDNA CAAGGAAGGAAGGAAGGGCTGCTGGAGCCGATCACCCTGGGACCGTGGCCGAGGTGACTGCAGACCCTCCAGGGAGG⁷³⁵

H-ras-1 cDNA CTGTGCACAGACTGTCTTGAACATCCCAATGCCACCGGAACCCAGCCCTTAGCTCCCTCCAGGCCTCTGTGGGCC⁸¹⁵

H-ras-1 cDNA TTGTGGCCAGATGGGATCACAGTAAATTATTGGATGGTCTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3'

FIG. 3. Nucleotide sequence of T24 H-ras-1 cDNA and a comparison of its coding sequence to those of viral ras genes. The nucleotide sequence of H-ras-1 cDNA is given together with the amino acid sequence of the predicted protein product. The nucleotide coding sequences for v-H-ras and v-K-ras and their predicted protein amino acid sequences are data derived by others (34,35). Nucleotides in viral ras genes that are common to those in T24 H-ras-1 cDNA are indicated by dashed lines; amino acids in viral ras proteins common to those in T24 H-ras-1 protein are indicated by asterisks; ter, termination codon.

cDNA clone, began outside the 6.2-kbp *Bam*HI fragment.

Figure 3 presents the nucleotide sequence of the T24-H-*ras*-1 transcript as deduced from the composite nucleotide sequences of the cDNA clones. The sequence contains an open reading frame of 567 nucleotides which encodes a protein of 189 amino acid residues and with a predicted mass of 21,350 daltons. The ATG codon that initiates this open reading frame is most likely the site for translation initiation, since the size of the resulting protein is close to the apparent molecular weight of the T24 H-*ras*-1 gene product observed in transformed cells (12,15,19,27,28).

Beyond the predicted polypeptide chain terminator is a 3' untranslated region of 285 nucleotides followed by a polyadenylate track. The transcript lacks the sequences AATAAA or AATTAAA near the 3' end, two sequences frequently found near the 3' end of messenger RNA and believed to signal transcription termination and polyadenylation (29,30). There is the sequence AGTAAA, 17 to 22 bases from the poly A track of the H-*ras*-1 transcript. This sequence may signal polyadenylation for this transcript, and such a sequence is also found at the transcription termination site in the mouse mammary tumor virus provirus (31). Another

feature of the 3' untranslated region is a 42-base purine-rich region made up of the sequence elements GGA and GGAA repeated in tandem. We do not know what role, if any, this unusual sequence may play.

Comparison of T24 H-*ras*-1 cDNA and Genomic Sequences

In order to map the transcriptional unit of the T24 H-*ras*-1 gene, we have sequenced portions of the molecularly cloned gene (19). As shown in Fig. 4, these sequences include the four protein-encoding exons, part of an exon corresponding to the 3' untranslated region of the H-*ras*-1 transcript, and the gene's exon-intron boundaries. On the basis of cDNA and genomic restriction endonuclease cleavage data (not shown), we suspect that the 3' untranslated region of the H-*ras*-1 transcript derives from a single exon. The map positions of the coding exons closely correspond to those of the four exons which Chang et al. (32) have detected by analysis of heteroduplexes between H-*ras*-1 and viral H-*ras*. We do not know the position of exon(s) that contribute 5' untranslated sequences to the H-*ras*-1 transcript.

The sequences at the exon-intron junctions

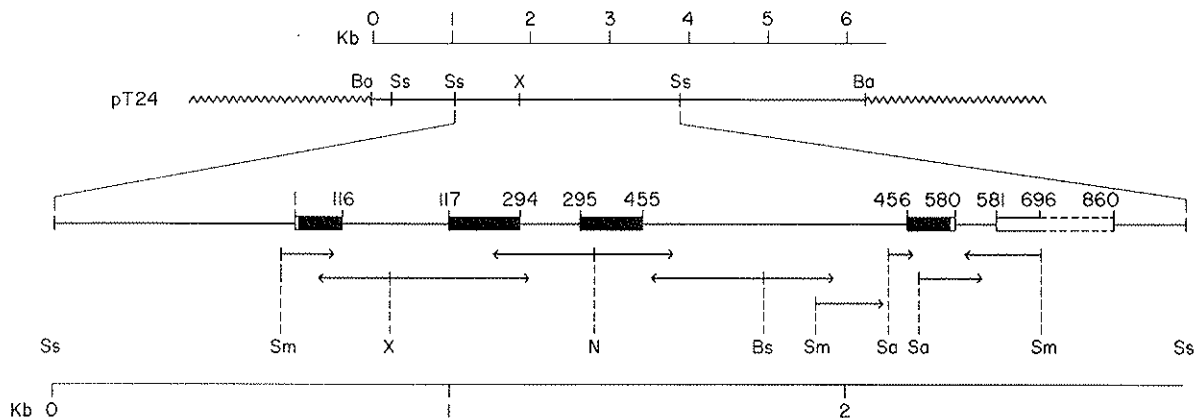


FIG. 4. The T24 H-*ras*-1 transcriptional unit. pT24 contains the biologically active T24 H-*ras*-1 gene. The upper portion of the figure shows the *Bam*HI (Ba), *Sst*I (Ss), and *Xba*I (X) sites in pT24. The lower portion of the figure shows the exon-intron configuration of the gene, as determined by nucleotide sequencing from the indicated restriction endonuclease cleavage sites. Filled segments indicate protein-coding sequences, blank segments indicate noncoding exon sequences, the dashed blank segment is the region which is likely homologous to the 3' terminus of the H-*ras*-1 transcript. Ss, *Sst*I; Sm, *Sma*I; X, *Xba*I; N, *Nco*I; Bs, *Bst*E II; Sa, *Sau*3AI. There are other sites for restriction endonucleases *Sma*I, *Sau*3AI, and *Bst*E II not shown in the figure. The numbers bracketing exons indicate the corresponding nucleotide position in the cDNA (see Fig. 3).

TABLE 1. Sequences at exon-intron boundaries in T24 H-ras-1

5' EXON/-----INTRON-----/EXON 3'	
GAG/GTGAGCC (116) CCTGCAG/GAT (117)
CAG/GTGAACC (294) CTCTCAG/GGA (295)
CAG/GTNAGGC (455) TTNCNG/GGA (456)
CAG/GTGAGGG (580) CGCCCAG/CAC (581)

The nucleotide sequences at exon-intron junctions were obtained using the sequencing scheme illustrated in Fig. 4. Undetermined nucleotides are designated by N. Numbers in parentheses designate the nucleotide positions in the cDNA corresponding to the ends of exons.

of the H-ras-1 gene are shown in Table 1. All of the junctions contain the consensus sequences for RNA splicing (5' exon . . . AG/GT . . . intron . . . PyPyPyNCAG/N . . . 3' exon) (33).

The first coding exon of T24 H-ras-1 has a valine codon (GTC) at position 12. The same gene in normal human placenta has a glycine codon (GGC) at this position. It is this structural mutation which confers on T24 H-ras-1 its transforming capability (19,27,28). There may be other coding differences between T24 and normal H-ras-1, but these differences are not essential for transformation.

Comparison of Human H-ras-1 cDNA and Viral ras Sequences

The ras genes of Harvey and Kirsten sarcoma viruses (v-H-ras and v-K-ras) are distinct genes of rat origin which are both highly conserved in higher organisms (3). We have compared the published sequences of v-H-ras (34) and v-K-ras (35) to that of T24 H-ras-1 cDNA (Fig. 3). T24 H-ras-1 and v-H-ras coding sequences have diverged by 12%. Most of these base changes are silent, and the two H-ras genes encode proteins that differ at only three of their 189 amino acid residues. One of these amino differences is at position 12: T24 H-ras-1 encodes valine and v-H-ras encodes arginine. However, normal human H-ras-1 encodes

glycine at this position (19,27,28), as does the normal rat H-ras gene (E. M. Scolnick, personal communication).

Another significant amino acid difference is at position 59: T24 H-ras-1 encodes alanine, whereas both v-H-ras and v-K-ras encode threonine. This threonine is the site of phosphorylation in viral p21 (36). The near identity between human and rat ras gene products, despite 40% of possible silent mutations in the ras genes, demonstrates that the normal function of H-ras p21 cannot withstand even limited alterations in amino acid composition.

The high degree of H-ras gene conservation also suggests that virtually all differences between Harvey and Kirsten ras gene products are essential for the differential function of these genes. As shown in Fig. 4, the major differences in v-H-ras and v-K-ras gene products are 1) 19 of the 24 amino acids at their carboxyl termini, and 2) 5 of 8 amino acids from positions 121 to 128. By contrast, all but eight of the remaining 157 amino acids of v-H-ras and v-K-ras proteins are identical. These discrete conserved and diverged sequence domains may mediate common and differential activities of multifunctional ras proteins.

Acknowledgments: This work was supported by grants from NIH and the American Business Cancer Research Foundation. O. F. is on leave from I Istituto di Chimica Biologica, II Facoltà di Medicina, Università di Napoli, Italy. M. G. is a Damon Runyon-Walter Winchell Cancer Research Fellow.

REFERENCES

1. Spector DH, Varmus HE, Bishop JM. Nucleotide sequences related to the transforming gene of avian sarcoma virus are present in DNA of uninfected vertebrates. *Proc Natl Acad Sci USA* 1978; 75:4102-6.
2. Collet MS, Brugge JS, Erikson RL. Characterization of a normal avian cell protein related to the avian sarcoma virus transforming gene product. *Cell* 1978; 15:1363-9.
3. Ellis RW, Defeo D, Shih TY, et al. The p21 src genes of Harvey and Kirsten sarcoma viruses originate from divergent members of a family of normal vertebrate genes. *Nature* 1981; 292:506-11.
4. Roussel M, Saule S, Lagrou C, Rommens C, Beug H, Graf T, Stehelin D. Three new types of viral oncogene of cellular origin specific for haematopoietic cell transformation. *Nature* 1979; 281:452-5.

5. Yoshida M, Kawai S, Toyoshima K. Uninfected avian cells contain structurally unrelated progenitors of viral sarcoma genes. *Nature* 1980; 287:653-4.
6. Tronick SR, Robbins KC, Canaani E, Devare SG, Anderson PR, Aaronson SA. Molecular cloning of Moloney murine sarcoma virus: arrangement of virus-related sequences within the normal mouse genome. *Proc Natl Acad Sci USA* 1979; 76:6314-8.
7. Goff SP, Gilboa E, Witte ON, Baltimore D. Structure of the Abelson murine leukemia virus genome and the homologous cellular gene: studies with cloned viral DNA. *Cell* 1980; 22:777-85.
8. Krontiris TG, Cooper GM. Transforming activity of human tumor DNAs. *Proc Natl Acad Sci USA* 1981; 78:1181-4.
9. Murray M, Shilo B, Shih C, Cowing D, Hsu HW, Weinberg RA. Three different human tumor cell lines contain different oncogenes. *Cell* 1981; 25:355-61.
10. Perucho M, Goldfarb MP, Shimizu K, Lama C, Fogh J, Wigler MH. Human tumor-derived cell lines contain common and different transforming genes. *Cell* 1981; 27:467-76.
11. Lane MA, Sauten A, Cooper GM. Activation of related transforming genes in mouse and human mammary carcinomas. *Proc Natl Acad Sci USA* 1981; 78:5185-9.
12. Der CJ, Krontiris TG, Cooper GM. Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses. *Proc Natl Acad Sci USA* 1982; 79:3637-40.
13. Lane MA, Sauten A, Cooper GM. Stage-specific transforming genes of human and mouse B- and T-lymphocyte neoplasms. *Cell* 1982; 28:873-80.
14. Shimizu K, Goldfarb MP, Suard Y, et al. Three human transforming genes are related to the viral ras oncogenes. *Proc Natl Acad Sci USA*, 1983 (in press).
15. Parada LF, Tabin CJ, Shih C, Weinberg RA. Human EJ bladder carcinoma oncogene is homologue of Harvey sarcoma virus ras gene. *Nature* 1982; 297:474-8.
16. Santos E, Tronick SR, Aaronson SA, Pulciani S, Barbacid M. T24 human bladder carcinoma oncogene is an activated form of the normal human homologue of Balb- and Harvey-MSV transforming genes. *Nature* 1982; 297:474-8.
17. Goldfarb M, Shimizu K, Perucho M, Wigler MH. Isolation and preliminary characterization of a human transforming gene from T24 bladder carcinoma cells. *Nature* 1982; 296:404-9.
18. Hanahan D. *J. Mol. Biol.*, 1983 (in press).
19. Taparowsky E, Suard Y, Fasano O, Shimizu K, Goldfarb M, Wigler MH. Activation of the T24 bladder carcinoma transforming gene is linked to a single amino acid change. *Nature* 1982; 300:762-5.
20. Tanaka T, Weisblum B. Construction of a Colicin EI-R factor composite plasmid *in vitro*: means for amplification of deoxyribonucleic acid. *J Bacteriol* 1975; 121:354-62.
21. Helfman D, Feramisco J, Fiddes J, Thomas GP, Hughes S. Identification of clones that encode chicken tropomyosin by direct immunological screening of a cDNA expression library. *Proc Natl Acad Sci USA* 1983; 80:31-5.
22. Grunstein M, Hogness D. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. *Proc Natl Acad Sci USA* 1975; 72:3961-5.
23. Maniatis T, Jeffrey A, Kleid DG. Nucleotide sequence of the rightward operator of phage λ . *Proc Natl Acad Sci USA* 1975; 72:1184-8.
24. Southern EM. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol* 1975; 98:503-17.
25. Wigler MH, Sweet R, Sim GK, Wold B, Pellicer A, Lacy E, Maniatis T, Silverstein S, Axel R. Transformation of mammalian cells with genes from prokaryotes and eucaryotes. *Cell* 1979; 16:777-85.
26. Maxam A, Gilbert W. Sequencing end-labeled DNA with base-specific chemical cleavages. In: Grossman L, Moldave K, eds. *Methods in Enzymology, Vol 65*. New York: Academic Press, 1980:499-580.
27. Tabin CJ, Bradley SM, Bargmann CI, et al. Mechanism of activation of a human oncogene. *Nature* 1982; 300:143-9.
28. Reddy EP, Reynolds RK, Santos E, Barbacid M. A point mutation is responsible for the acquisition of transforming properties by the T24 human bladder carcinoma oncogene. *Nature* 1982; 300:149-52.
29. Proudfoot NJ, Brownlee GG. 3'-Non-coding region sequences in eukaryotic messenger RNA. *Nature* 1976; 263:211-4.
30. Unterman RD, Lynch KR, Nakhasi HL, et al. Cloning and sequence of several $\alpha_2\mu$ -globulin cDNAs. *Proc Natl Acad Sci USA* 1981; 78:3478-82.
31. Donehower LA, Huang AL, Hager GL. Regulatory and coding potential of the mouse mammary tumor virus long terminal redundancy. *J. Virology* 1981; 37:226-38.
32. Chang EH, Gonda MA, Ellis RW, Scolnick EM, Lowy DR. Human genome contains four genes homologous to transforming genes of Harvey and Kirsten murine sarcoma viruses. *Proc Natl Acad Sci USA* 1982; 79:4848-52.
33. Sharp PA. Speculation on RNA splicing. *Cell* 1981; 23:643-6.
34. Dahr R, Ellis RW, Shih TY, et al. Nucleotide sequence of the p21 transforming protein of Harvey murine sarcoma virus. *Science* 1982; 217:934-7.
35. Tsuchida N, Ryder T, Ohtsubo E. Nucleotide sequence of the oncogene encoding the p21 transforming protein of Kirsten murine sarcoma virus. *Science* 1982; 217:937-9.
36. Shih TY, Stokes PE, Smythers GW, Dhar R, Oroszlan S. Characterization of the phosphorylation sites and the surrounding amino acid sequences of the p21 transforming proteins coded for by the Harvey and Kirsten strains of murine sarcoma virus. *J Biol Chem* 1982; 257:11767-73.