IV

CELLULAR PROTEINS

```
LOCUS       p53-php53c1  1317 bp ss-mRNA            PRI       22-OCT-1992
DEFINITION  Human p53 cellular tumor antigen mRNA, complete cds.
ACCESSION   X02469 M60950
KEYWORDS    antigen; tumor antigen.
SOURCE      Human, cDNA to mRNA, clone php53c1 from a cDNA library of SV40
            transformed fibroblasts.
REFERENCE   1  (bases 1 to 1317)
  AUTHORS   Zakut-Houri,R., Bienz-Tadmor,B., Givol,D. and Oren,M.
  TITLE     Human p53 cellular tumor antigen: cDNA sequence and expression in
            COS cells
  JOURNAL   EMBO J. 4, 1251-1255 (1985)
REFERENCE   2  (bases 1 to 1317)
  AUTHORS   Rideout,W.M.III., Coetzee,G.A., Olumi,A.F. and Jones,P.A.
  TITLE     5-methylcytosine as anendogenous mutagen in the human ldl receptor
            and p53 genes
  JOURNAL   Science 249, 1288-1290 (1990)
REFERENCE   3  (M14694 and M14695 shown in alignment below)
  AUTHORS   Harris,N., Brill,E., Shohat,O., Prokocimer,M., Wolf,D., Arai,N. and
            Rotter,V.
  TITLE     Molecular basis for heterogeneity of the human p53 protein
  JOURNAL   Mol. Cell. Biol. 6, 4650-4656 (1986)
REFERENCE   4  (K031991 shown in alignment below)
  AUTHORS   Harlow,E., Williamson,N.M., Ralston,R., Helfman,D.M. and Adams,T.E.
  TITLE     Molecular cloning and in vitro expression of a cDNA clone for human
            cellular tumor antigen p53
  JOURNAL   Mol. Cell. Biol. 5, 1601-1610 (1985)

COMMENT     Clone php53c1 is the wild-type form of p53, isolated from an SV40
            transformed fibroblast cell line.   It was first thought that p53,
            a cell-cycle regulatory protein, was an oncogene.  This has since
            been disproven, for it has been shown that the clones that were first
            isolated and tested were mutants.  In fact, the p53 gene is a tumor
            suppressor.  The mutated forms are thought to be trans-dominant
            over the wild-type p53.  The mutations cause a conformational change
            in the protein which facilitates its binding to the heat shock
            protein hsp70.  The two forms of p53 and hsp70 become bound together
            in long-lived unproductive complexes.  Since these complexes aggregate
            in the cytoplasm, the nucleus is deprived of p53.  With the
            negative regulatory effect of p53 reduced or eliminated, the cell
            is more easily able to proliferate.  However, a different mechanism of
            inactivation is employed by the viral oncoproteins: the large
            T-antigen of SV40 and the E1B 55kD protein of adenovirus.  Instead of
            the hsp70, mutant p53 complex trapping the wild-type p53 in inactive
            structures it is the viral oncoproteins that sequester the wild type
            p53 and inactivate it.   Human papillomavirus uses yet another
            mechanism.  The E6 protein of the oncogenic HPV types bind wild-type
            p53 and stimulate its destruction; although at this point there is no
            evidence that E6 proteins of noncogenic HPV types, such as HPV-6
            or HPV-11, can associate with wild-type p53 in vivo (Werness et al.
            Science 248, 76-79, Scheffner et al. Cell 63, 1129-1136).  It is however
            possible that the binding assay used to determine this fact was not
            sensitive enough to detect low affinity binding.  The degradation of
            p53 by HPV E6 is ATP dependant and the ubiquitin-dependant protease
            system has been suggested to be involved.  Since the large T antigen
            of SV40 also complexes with wild-type p53 and does not target it for
            degradation,  there must be an additional signaling mechanism not yet
            elucidated involved in the interaction with the HPV E6 protein.

            The p53 wild type php53c1 coding region and four variants are shown in
            the alignment below.  The p53 coding regions which are represented by
            the accession numbers M14694 and M14695 have both been isolated from
            the human transformed cell line SV-80.  The two forms differ from one
            another by a single base pair substitution and consequently accounts
            for the change in electrophoretic mobility of the two proteins.  Harris
            et al. believe that this heterogeneity is due to gene polymorphism.
```

```
            The coding region represented by accession number K03199 has been
            isolated from the human vulva carcinoma cell line A431.
BASE COUNT      295 a    408 c    352 g    262 t
ORIGIN      2 bp upstream of XbaI site; chromosome 17p13.
        1 gtctagagcc accgtccagg gagcaggtag ctgctgggct ccggggacac tttgcgttcg
       61 ggctgggagc gtgctttcca cgacggtgac acgcttccct ggattggcag ccagactgcc
      121 ttccgggtca ctgccATGga ggagccgcag tcagatccta gcgtcgagcc ccctctgagt
                p53 start ->
      181 caggaaacat tttcagacct atggaaacta cttcctgaaa acaacgttct gtcccccttg
      241 ccgtcccaag caatggatga tttgatgctg tccccggacg atattgaaca atggttcact
      301 gaagacccag gtccagatga agctcccaga atgccagagg ctgctccccc cgtGGCCCCT
                                                                   ->
      361 GCACcagcag ctcctacacc ggcGGCCCCT GCACcagccc cctcctggcc cctgtcatct
            <- direct rpt          ->        <- direct rpt
      421 tctgtccctt cccagaaaac ctaccagggc agctacggtt tccgtctggg cttcttgcat
      481 tctgggacag ccaagtctgt gacttgcacg tactcccctg ccctcaacaa gatgttttgc
      541 caactggcca agacctgccc tgtgcagctg tgggttgatt ccacaccccc gcccggcacc
      601 cgcgtccgcg ccatggccat ctacaagcag tcacagcaca tgacggaggt tgtgaggcgc
      661 tgcccccacc atgagcgctg ctcagatagc gatggtctgg cccctcctca gcatcttatc
      721 cgagtggaag gaaatttgcg tgtggagtat ttggatgaca gaaacacttt tcgacatagt
      781 gtggtggtgc cctatgagcc gcctgaggtt ggctctgact gtaccaccat ccactacaac
      841 tacatgtgta acagttcctg catgggcggc atgaaccgga ggcccatcct caccatcatc
      901 acactggaag actccagtgg taatctactg ggacggaaca gctttgaggt gcgtgtttgt
      961 gcctgtcctg ggagagaccg gcgcacagag gaagagaatc tccgcaagaa aggggagcct
     1021 caccacgagc tgcccccagg gagcactaag cgagcactgc ccaacaacac cagctcctct
     1081 ccccagccaa agaagaaacc actggatgga gaatatttca cccttcagat ccgtgggcgt
     1141 gagcgcttcg agatgttccg agagctgaat gaggccttgg aactcaagga tgcccaggct
     1201 gggaaggagc caggggggag cagggctcac tccagccacc tgaagtccaa aaagggtcag
     1261 tctacctccc gccataaaaa actcatgttc aagacagaag ggcctgactc agacTGA
                                                                     <- p53 end
//

P53.WTPHP53C1 ATGGAGGAGCCGCAGTCAGATCCTAGCGTCGAGCCCCCTCTGAGTCAGGAAACATTTTCAGACCTATGGA     70
p53.X54156    ----------------------------------------------------------------------     70
p53.M14695    ----------------------------------------------------------------------     70
p53.M14694    ----------------------------------------------------------------------     70
p53.K03199    ----------------------------------------------------------------------     70

P53.WTPHP53C1 AACTACTTCCTGAAAACAACGTTCTGTCCCCCTTGCCGTCCCAAGCAATGGATGATTTGATGCTGTCCCC    140
p53.X54156    ----------------------------------------------------------------------    140
p53.M14695    ----------------------------------------------------------------------    140
p53.M14694    ----------------------------------------------------------------------    140
p53.K03199    ----------------------------------------------------------------------    140

P53.WTPHP53C1 GGACGATATTGAACAATGGTTCACTGAAGACCCAGGTCCAGATGAAGCTCCCAGAATGCCAGAGGCTGCT    210
p53.X54156    ----------------------------------------------------------------------    210
p53.M14695    ----------------------------------------------------------------------    210
p53.M14694    ----------------------------------------------------------------------    210
p53.K03199    ----------------------------------------------------------------------    210

P53.WTPHP53C1 CCCCCCGTGGCCCCTGCACCAGCAGCTCCTACACCGGCGGCCCCTGCACCAGCCCCCTCCTGGCCCCTGT    280
p53.X54156    ----G-----------------------------------------------------------------    280
p53.M14695    -----------------------GA---------------------------------------------    280
p53.M14694    ----G------------------GA---------------------------------------------    280
p53.K03199    ----------------------------------------------------------------------    280

P53.WTPHP53C1 CATCTTCTGTCCCTTCCCAGAAAACCTACCAGGGCAGCTACGGTTTCCGTCTGGGCTTCTTGCATTCTGG    350
p53.X54156    ----------------------------------------------------------------------    350
p53.M14695    ----------------------------------------------------------------------    350
p53.M14694    ----------------------------------------------------------------------    350
p53.K03199    ----------------------------------------------------------------------    350

P53.WTPHP53C1 GACAGCCAAGTCTGTGACTTGCACGTACTCCCCTGCCCTCAACAAGATGTTTTGCCAACTGGCCAAGACC    420
p53.X54156    ----------------------------------------------------------------------    420
p53.M14695    ----------------------------------------------------------------------    420
p53.M14694    ----------------------------------------------------------------------    420
p53.K03199    ----------------------------------------------------------------------    420
```

```
P53.WTPHP53C1 TGCCCTGTGCAGCTGTGGGTTGATTCCACACCCCCGCCCGGCACCCGCGTCCGCGCCATGGCCATCTACA    490
p53.X54156    ----------------------------------------------------------------------    490
p53.M14695    ----------------------------------------------------------------------    490
p53.M14694    ----------------------------------------------------------------------    490
p53.K03199    ----------------------------------------------------------------------    490

P53.WTPHP53C1 AGCAGTCACAGCACATGACGGAGGTTGTGAGGCGCTGCCCCCACCATGAGCGCTGCTCAGATAGCGATGG    560
p53.X54156    ----------------------------------------------------------------------    560
p53.M14695    ----------------------------------------------------------------------    560
p53.M14694    ----------------------------------------------------------------------    560
p53.K03199    ----------------------------------------------------------------------    560

P53.WTPHP53C1 TCTGGCCCCTCCTCAGCATCTTATCCGAGTGGAAGGAAATTTGCGTGTGGAGTATTTGGATGACAGAAAC    630
p53.X54156    ----------------------------------------------------------------------    630
p53.M14695    ----------------------------------------------------------------------    630
p53.M14694    ----------------------------------------------------------------------    630
p53.K03199    ----------------------------------------------------------------------    630

P53.WTPHP53C1 ACTTTTCGACATAGTGTGGTGGTGCCCTATGAGCCGCCTGAGGTTGGCTCTGACTGTACCACCATCCACT    700
p53.X54156    ----------------------------------------------------------------------    700
p53.M14695    ----------------------------------------------------------------------    700
p53.M14694    ----------------------------------------------------------------------    700
p53.K03199    ----------------------------------------------------------------------    700

P53.WTPHP53C1 ACAACTACATGTGTAACAGTTCCTGCATGGGCGGCATGAACCGGAGGCCCATCCTCACCATCATCACACT    770
p53.X54156    ----------------------------------------------------------------------    770
p53.M14695    ----------------------------------------------------------------------    770
p53.M14694    ----------------------------------------------------------------------    770
p53.K03199    ----------------------------------------------------------------------    770

P53.WTPHP53C1 GGAAGACTCCAGTGGTAATCTACTGGGACGGAACAGCTTTGAGGTGCGTGTTTGTGCCTGTCCTGGGAGA    840
p53.X54156    ----------------------------------------------------------------------    840
p53.M14695    ----------------------------------------------------------------------    840
p53.M14694    ----------------------------------------------------------------------    840
p53.K03199    -----------------------------------------------A----------------------    840

P53.WTPHP53C1 GACCGGCGCACAGAGGAAGAGAATCTCCGCAAGAAAGGGGAGCCTCACCACGAGCTGCCCCCAGGGAGCA    910
p53.X54156    ----------------------------------------------------------------------    910
p53.M14695    ----------------------------------------------------------------------    910
p53.M14694    ----------------------------------------------------------------------    910
p53.K03199    ----------------------------------------------------------------------    910

P53.WTPHP53C1 CTAAGCGAGCACTGCCCAACAACACCAGCTCCTCTCCCCAGCCAAAGAAGAAACCACTGGATGGAGAATA    980
p53.X54156    ----------------------------------------------------------------------    980
p53.M14695    ----------------------------------------------------------------------    980
p53.M14694    ----------------------------------------------------------------------    980
p53.K03199    ----------------------------------------------------------------------    980

P53.WTPHP53C1 TTTCACCCTTCAGATCCGTGGGCGTGAGCGCTTCGAGATGTTCCGAGAGCTGAATGAGGCCTTGGAACTC   1050
p53.X54156    ----------------------------------------------------------------------   1050
p53.M14695    ----------------------------------------------------------------------   1050
p53.M14694    ----------------------------------------------------------------------   1050
p53.K03199    ----------------------------------------------------------------------   1050

P53.WTPHP53C1 AAGGATGCCCAGGCTGGGAAGGAGCCAGGGGGGAGCAGGGCTCACTCCAGCCACCTGAAGTCCAAAAAGG   1120
p53.X54156    ----------------------------------------------------------------------   1120
p53.M14695    ----------------------------------------------------------------------   1120
p53.M14694    ----------------------------------------------------------------------   1120
p53.K03199    ----------------------------------------------------------------------   1120

P53.WTPHP53C1 GTCAGTCTACCTCCCGCCATAAAAAACTCATGTTCAAGACAGAAGGGCCTGACTCAGACTGA   1182
p53.X54156    --------------------------------------------------------------   1182
p53.M14695    --------------------------------------------------------------   1182
p53.M14694    --------------------------------------------------------------   1182
p53.K03199    --------------------------------------------------------------   1182
```

```
LOCUS       Rb          4740 bp ss-mRNA           PRI       15-JUN-1989
DEFINITION  Human retinoblastoma susceptibility mRNA, complete cds.
ACCESSION   M15400
KEYWORDS    retinoblastoma susceptibility.
SOURCE      Human fetal retina, cDNA to mRNA, clone RB-[1,5].
REFERENCE   1  (bases 243 to 4740)
  AUTHORS   Lee,W.H., Bookstein,R., Hong,F., Young,L.J., Shew,J.Y. and Lee,E.
  TITLE     Human retinoblastoma susceptibility gene: Cloning, identification,
            and sequence
  JOURNAL   Science 235, 1394-1399 (1987)
REFERENCE   2  (bases 1 to 480)
  AUTHORS   Lee,E., Bookstein,R., Young,L.J., Lin,C.-J., Rosenfeld,M.G. and
            Lee,W.H.
  TITLE     Molecular mechanism of retinoblastoma gene inactivation in
            retinoblastoma cell line Y79
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 85, 6017-6021 (1988)

COMMENT     Draft entry and computer-readable copy of sequence in [1] kindly
            provided by R.Bookstein, 27-APR-1987.

            The authors [1] identified the retinoblastoma susceptibility (RB)
            gene encoding a messenger RNA (mRNA) of 4.6 kb on the basis
            of chromosomal location, homozygous deletion, and tumor-specific
            alterations in expression.  Transcription of the gene was abnormal
            in 6/6 retinoblastomas while normal in fetal retina and placenta.
            The sequence presented in this entry is derived from human fetal
            retina cDNA.  The authors [2] also examined cDNA from the
            retinoblastoma cell line Y79.  They found that the mRNA produced
            was a shortened transcript which reflects a deletion of exons 2-6
            and must have ocurred by a mechanism other than recombination of
            homologous sequences.

            The Retinoblastoma (Rb) susceptibility gene maps to chromosome
            13q14 and encodes a nuclear phosphoprotein of 105kDa.  Rb is a
            tumor suppressor gene; its nuclear location and ability to bind DNA
            suggests a role in transcriptional regulation.  On the side of the
            protein lies a complementary oncoprotein-binding pocket that is a
            common target for the three viral proteins: E1A of adenovirus, large
            T antigen of SV40, and E7 of HPV.  Munger et al. (EMBO 8, 4099-4105)
            determined that the high risk HPV types such as HPV16 and HPV18 bind
            to pRB with higher affinity than low risk types 6b and 11.  The amino
            acid locations of the pRB binding site of the HPV16 E7 protein have
            been mapped to a small stretch of amino acids.  In addition, pRb has
            shown to associate with a number of cellular host proteins at the
            onco-protein binding site.  pRB shows differential phosphorylation
            at different points within the cell cycle.  All three of the viral
            proteins mentioned above interact with the underphosphorylated
            form of pRB.

BASE COUNT     1508 a    887 c    862 g   1483 t
ORIGIN      1038 bp upstream of EcoRI site; chromosome 13q14.1-q14.2.
        1 ttccggtttt tctcagggga cgttgaaatt atttttgtaa cgggagtcgg gagaggacgg
       61 ggcgtgcccc gcgtgcgcgc gcgtcgtcct ccccggcgct cctccacagc tcgctggctc
      121 ccgccgcgga aaggcgtcat gccgcccaaa accccccgaa aaacggccgc caccgccgcc
      181 gctgccgccg cggaaccccc ggcaccgccg ccgccgcccc ctcctgagga ggacccagag
      241 caggacagcg gcccggagga cctgcctctc gtcaggcttg agtttgaaga aacagaagaa
      301 cctgatttta ctgcattatg tcagaaatta aagataccag atcatgtcag agagagagct
      361 tggttaactt gggagaaagt ttcatctgtg gatggagtat tgggaggtta tattcaaaag
      421 aaaaaggaac tgtggggaat ctgtatcttt attgcacgag ttgacctaga tgagatgtcg
      481 ttcactttac tgagctacag aaaaacatac gaaatcagtg tccataaatt ctttaactta
      541 ctaaaagaaa ttgataccag taccaaagtt gataatgcta tgtcaagact gttgaagaag
      601 tatgatgtat tgtttgcact cttcagcaaa ttggaaagga catgtgaact tatatatttg
      661 acacaaccca gcagttcgat atctactgaa ataaattctg cattggtgct aaaagtttct
      721 tggatcacat ttttattagc taaaggggaa gtattacaaa tggaagatga tctggtgatt
      781 tcatttcagt taatgctatg tgtccttgac tattttatta aactctcacc tcccatgttg
```

```
 841 ctcaaagaac catataaaac agctgttata cccattaatg gttcacctcg aacacccagg
 901 cgaggtcaga acaggagtgc acggatagca aaacaactag aaaatgatac aagaattatt
 961 gaagttctct gtaaagaaca tgaatgtaat atagatgagg tgaaaaatgt ttatttcaaa
1021 aattttatac cttttatgaa ttctcttgga cttgtaacat ctaatggact tccagaggtt
1081 gaaaatcttt ctaaacgata cgaagaaatt tatcttaaaa ataaagatct agatcgaaga
1141 ttatttttgg atcatgataa aactcttcag actgattcta tagacagttt tgaaacacag
1201 agaacaccac gaaaaagtaa ccttgatgaa gaggtgaata taattcctcc acacactcca
1261 gttaggactg ttatgaacac tatccaacaa ttaatgatga ttttaaattc tgcaagtgat
1321 caaccttcag aaaatctgat ttcctatttt aacaactgca cagtgaatcc aaaagaaagt
1381 atactgaaaa gagtgaagga tataggatac atctttaaag agaaatttgc taaagctgtg
1441 ggacagggtt gtgtcgaaat tggatcacag cgatacaaac ttggagttcg cttgtattac
1501 cgagtaatgg aatccatgct taaatcagaa gaagaacgat tatccattca aaattttagc
1561 aaacttctga atgacaacat tttcatatg tctttattgg cgtgcgctct tgaggttgta
1621 atggccacat atagcagaag tacatctcag aatcttgatt ctggaacaga tttgtctttc
1681 ccatggattc tgaatgtgct taatttaaaa gcctttgatt tttacaaagt gatcgaaagt
1741 tttatcaaag cagaaggcaa cttgacaaga gaaatgataa aacatttaga acgatgtgaa
1801 catcgaatca tggaatccct tgcatggctc tcagattcac ctttatttga tcttattaaa
1861 caatcaaagg accgagaagg accaactgat caccttgaat ctgcttgtcc tcttaatctt
1921 cctctccaga ataatcacac tgcagcagat atgtatcttt ctcctgtaag atctccaaag
1981 aaaaaaggtt caactacgcg tgtaaattct actgcaaatg cagagacaca agcaacctca
2041 gccttccaga cccagaagcc attgaaatct acctctcttt cactgtttta taaaaaagtg
2101 tatcggctag cctatctccg gctaaataca ctttgtgaac gccttctgtc tgagcaccca
2161 gaattagaac atatcatctg gaccctttttc cagcacaccc tgcagaatga gtatgaactc
2221 atgagagaca ggcatttgga ccaaattatg atgtgttcca tgtatggcat atgcaaagtg
2281 aagaatatag accttaaatt caaaatcatt gtaacagcat acaaggatct tcctcatgct
2341 gttcaggaga cattcaaacg tgttttgatc aaagaagagg agtatgattc tattatagta
2401 ttctataact cggtcttcat gcagagactg aaaacaaata ttttgcagta tgcttccacc
2461 aggcccccta ccttgtcacc aatacctcac attcctcgaa gcccttacaa gtttcctagt
2521 tcacccttac ggattcctgg agggaacatc tatatttcac ccctgaagag tccatataaa
2581 atttcagaag gtctgccaac accaacaaaa atgactccaa gatcaagaat cttagtatca
2641 attggtgaat cattcgggac ttctgagaag ttccagaaaa taaatcagat ggtatgtaac
2701 agcgaccgtg tgctcaaaag aagtgctgaa ggaagcaacc ctcctaaacc actgaaaaaa
2761 ctacgctttg atattgaagg atcagatgaa gcagatggaa gtaaacatct cccaggagag
2821 tccaaatttc agcagaaact ggcagaaatg acttctactc gaacacgaat gcaaaagcag
2881 aaaatgaatg atagcatgga tacctcaaac aaggaagaga aatgaggatc tcaggacctt
2941 ggtggacact gtgtacacct ctggattcat tgtctctcac agatgtgact gtataacttt
3001 cccaggttct gtttatggcc acatttaata tcttcagctc tttttgtgga tataaaatgt
3061 gcagatgcaa ttgtttgggt gagtcctaag ccacttgaaa tgttagtcat tgttatttat
3121 acaagattga aaatcttgtg taaatcctgc catttaaaaa gttgtagcag attgtttcct
3181 cttccaaagt aaaattgctg tgctttatgg atagtaagaa tggccctaga gtgggagtcc
3241 tgataaccca ggcctgtctg actactttgc cttcttttgt agcatatagg tgatgtttgc
3301 tcttgttttt attaatttat atgtatattt ttttaattta acatgaacac ccttagaaaa
3361 tgtgtcctat ctatcttcca aatgcaattt gattgactgc ccattcacca aaattatcct
3421 gaactcttct gcaaaaatgg atattattag aaattagaaa aaaattacta attttacaca
3481 ttagatttta ttttactatt ggaatctgat atactgtgtg cttgttttat aaaattttgc
3541 ttttaattaa ataaaagctg gaagcaaagt ataaccatat gatactatca tactactgaa
3601 acagatttca tacctcagaa tgtaaaagaa cttactgatt attttcttca tccaacttat
3661 gtttttaaat gaggattatt gatagtactc ttggttttta taccattcag atcactgaat
3721 ttataaagta cccatctagt acttgaaaaa gtaaagtgtt ctgccagatc ttaggtatag
3781 aggaccctaa cacagtatat cccaagtgca ctttctaatg tttctgggtc ctgaagaatt
3841 aagatacaaa ttaattttac tccataaaca gactgttaat tataggagcc ttaatttttt
3901 tttcatagag atttgtctaa ttgcatctca aaattattct gccctcctta atttgggaag
3961 gtttgtgttt tctctggaat ggtacatgtc ttccatgtat cttttgaact ggcaattgtc
4021 tatttatctt ttattttttt aagtcagtat ggtctaacac tggcatgttc aaagccacat
4081 tatttctagt ccaaaattac aagtaatcaa gggtcattat gggttaggca ttaatgtttc
4141 tatctgattt tgtgcaaaag cttcaaatta aaacagctgc attagaaaaa gaggcgcttc
4201 tcccctcccc tacacctaaa ggtgtattta aactatcttg tgtgattaac ttatttagag
4261 atgctgtaac ttaaaatagg ggatatttaa ggtagcttca gctagctttt aggaaaatca
4321 ctttgtctaa ctcagaatta tttttaaaaa gaaatctggt cttgttagaa aacaaaattt
4381 tattttgtgc tcatttaagt ttcaaactta ctattttgac agttattttg ataacaatga
4441 cactagaaaa cttgactcca tttcatcatt gtttctgcat gaatatcata caaatcagtt
4501 agtttttagg tcaagggctt actatttctg ggtcttttgc tactaagttc acattagaat
4561 tagtgccaga attttaggaa cttcagagat cgtgtattga gatttcttaa ataatgcttc
4621 agatattatt gctttattgc ttttttgtat tggttaaaac tgtacattta aaattgctat
```

```
4681 gttactattt tctacaatta atagtttgtc tattttaaaa taaattagtt gttaagagtc
```

Deleted in Colon Cancer

```
LOCUS       DCC          4608 bp ss-mRNA         PRI       04-FEB-1994
DEFINITION  Homo sapiens DCC mRNA.
ACCESSION   X76132
KEYWORDS    cell adhesion molecule; DCC gene; immunoglobulin gene superfamily;
            transmembrane protein; tumor suppressor gene.
SOURCE      Human normal adult and fetal brain, Clontech cDNA libraries
            CAT #HL1065b, HL1003b.
REFERENCE   1  (bases 1 to 4608)
  AUTHORS   Cho, K.
  TITLE     ;
  JOURNAL   direct submission
REFERENCE   2  (bases 1 to 4608)
  AUTHORS   Hedrick L., Cho K.R., Fearon E.R., Wu T.C., Kinzler K.W.,
            Vogelstein B.
  TITLE     The DCC gene product in cellular differentiation and colorectal
            tumorigenesis
  JOURNAL   Unpublished
COMMENT     Tumors were induced in nude mice after transplantation of the
            HPV-18 immortalized human keratinocyte cell line (1811) and
            treatment with the carcinogen nitrosomethylurea (NMU).  In these
            transformed cells, one allele of the deleted in colon cancer DCC
            tumor-suppressor gene was absent and the other allele was mutated
            (Klingelhutz et al.  Oncogene 8: 95-9).
            Related sequence: M32292
BASE COUNT     1282 a   1179 c   1058 g   1089 t
ORIGIN
        1 atggagaata gtcttagatg tgtttgggta cccaagctgg cttttgtact cttcggagct
       61 tccttgctca gcgcgcatct tcaagtaacc ggttttcaaa ttaaagcttt cacagcactg
      121 cgcttcctct cagaaccttc tgatgccgtc acaatgcggg gaggaaatgt cctcctcgac
      181 tgctccgcgg agtccgaccg aggagttcca gtgatcaagt ggaagaaaga tggcattcat
      241 ctggccttgg gaatggatga aaggaagcag caactttcaa atgggtctct gctgatacaa
      301 aacatacttc attccagaca ccacaagcca gatgagggac tttaccaatg tgaggcatct
      361 ttaggagatt ctggctcaat tattagtcgg acagcaaaag ttgcagtagc aggaccactg
      421 aggttccttt cacagacaga atctgtcaca gccttcatgg gagacacagt gctactcaag
      481 tgtgaagtca ttggggagcc catgccaaca atccactggc agaagaacca acaagacctg
      541 actccaatcc caggtgactc ccgagtggtg gtcttgccct ctggagcatt gcagatcagc
      601 cgactccaac cgggggacat tggaatttac cgatgctcag ctcgaaatcc agccagctca
      661 agaacaggaa atgaagcaga agtcagaatt ttatcagatc caggactgca tagacagctg
      721 tattttctgc aaagaccatc caatgtagta gccattgaag gaaaagatgc tgtcctggaa
      781 tgttgtgttt ctggctatcc tccaccaagt tttacctggt tacgaggcga ggaagtcatc
      841 caactcaggt ctaaaaagta ttctttattg ggtggaagca acttgcttat ctccaatgtg
      901 acagatgatg acagtggaat gtatacctgt gttgtcacat ataaaaatga gaatattagt
      961 gcctctgcag agctcacagt cttggttccg ccatggtttt taaatcatcc ttccaacctg
     1021 tatgcctatg aaagcatgga tattgagttt gaatgtacag tctctggaaa gcctgtgccc
     1081 actgtgaatt ggatgaagaa tggagatgtg gtcattccta gtgattattt tcagatagtg
     1141 ggaggaagca acttacggat acttggggtg gtgaagtcag atgaaggctt ttatcaatgt
     1201 gtggctgaaa atgaggctgg aaatgcccag accagtgcac agctcattgt ccctaagcct
     1261 gcaatcccaa gctccagtgt cctcccttcg gctcccagag atgtggtccc tgtcttggtt
     1321 tccagccgat ttgtccgtct cagctggcgc ccacctgcag aagcgaaagg gaacattcaa
     1381 actttcacgg tctttttctc cagagaaggt gacaacaggg aacgagcatt gaatacaaca
     1441 cagcctgggt cccttcagct cactgtggga aacctgaagc cagaagccat gtacaccttt
     1501 cgagttgtgg cttacaatga atggggaccg ggagagagtt ctcaacccat caaggtggcc
     1561 acacagcctg agttgcaagt tccagggcca gtagaaaacc tgcaagctgt atctacctca
     1621 cctacctcaa ttcttattac ctgggaaccc cctgcctatg caaacggtcc agtccaaggt
     1681 tacagattgt tctgcactga ggtgtccaca ggaaaagaac agaatataga ggttgatgga
     1741 ctatcttata aactggaagg cctgaaaaaa ttcaccgaat atagtcttcg attcttagct
     1801 tataatcgct atggtccggg cgtctctact gatgatataa cagtggttac actttctgac
     1861 gtgccaagtg ccccgcctca gaacgtctcc ctggaagtgg tcaattcaag aagtatcaaa
     1921 gttagctggc tgcctcctcc atcaggaaca caaaatggat ttattaccgg ctataaaatt
     1981 cgacacagaa agacgacccg caggggtgag atggaaacac tggagccaaa caacctctgg
     2041 tacctattca caggactgga gaaaggaagt cagtacagtt tccaggtgtc agccatgaca
     2101 gtcaatggta ctggaccacc ttccaactgg tatactgcag agactccaga gaatgatcta
     2161 gatgaatctc aagttcctga tcaaccaagc tctcttcatg tgaggcccca gactaactgc
     2221 atcatcatga gttggactcc tcccttgaac ccaaacatcg tggtgcgagg ttatattatc
```

```
2281 ggttatggcg ttgggagccc ttacgctgag acagtgcgtg tggacagcaa gcagcgatat
2341 tattccattg agaggttaga gtcaagttcc cattatgtaa tctccctaaa agcttttaac
2401 aatgccggag aaggagttcc tctttatgaa agtgccacca ccaggtctat aaccgatccc
2461 actgacccag ttgattatta tcctttgctt gatgatttcc ccacctcggt cccagatctc
2521 tccaccccca tgctcccacc agtaggtgta caggctgtgg ctcttaccca tgatgctgtg
2581 agggtcagct gggcagacaa ctctgtccct aagaaccaaa agacgtctga ggtgcgactt
2641 tacaccgtcc ggtggagaac cagcttttct gcaagtgcaa aatacaagtc agaagacaca
2701 acatctctaa gttacacagc aacaggcctc aaaccaaaca caatgtatga attctcggtc
2761 atggtaacaa aaaacagaag gtccagtact tggagcatga ctgcacatgc caccacgtat
2821 gaagcagccc ccacctctgc tcccaaggac tttacagtca ttactaggga agggaagcct
2881 cgtgccgtca ttgtgagttg gcagcctccc ttggaagcca atgggaaaat tactgcttac
2941 atcttatttt ataccttgga caagaacatc ccaattgatg actggattat ggaaacaatc
3001 agtggtgata ggcttactca tcaaatcatg gatctcaacc ttgatactat gtattacttt
3061 cgaattcaag cacgaaattc aaaaggagtg gggccactct ctgatcccat cctcttcagg
3121 actctgaaag tggaacaccc tgacaaaatg gctaatgacc aaggtcgtca tggagatgga
3181 ggttattggc cagttgatac taatttgatt gatagaagca ccctaaatga gccgccaatt
3241 ggacaaatgc accccccgca tggcagtgtc actcctcaga agaacagcaa cctgcttgtg
3301 atcattgtgg tcaccgttgg tgtcatcaca gtgctggtag tggtcatcgt ggctgtgatt
3361 tgcacccgac gctcttcagc ccagcagaga aagaaacggg ccacccacag tgctggcaaa
3421 aggaagggca gccagaagga cctccgaccc cctgatcttt ggatccatca tgaagaaatg
3481 gagatgaaaa atattgaaaa gccatctggc actgaccctg caggaaggga ctctcccatc
3541 caaagttgcc aagacctcac accagtcagc cacagccagt cagaaaccca actgggaagc
3601 aaaagcacct ctcattcagg tcaagacact gaggaagcag ggagctctat gtccactctg
3661 gagaggtcgc tggctgcacg ccgagccccc cgggccaagc tcatgattcc catggatgcc
3721 cagtccaaca atcctgctgt cgtgagcgcc atcccggtgc caacgctaga aagtgcccag
3781 tacccaggaa tcctcccgtc tcccacctgt ggatatcccc acccgcagtt cactctccgg
3841 cctgtgccat tcccaacact ctcagtggac cgaggtttcg gagcaggaag aagtcagtca
3901 gtgagtgaag gaccaactac ccaacaacca cctatgctgc ccccatctca gcctgagcat
3961 tctagcagcg aggaggcacc aagcagaacc atccccacag cttgtgttcg accaactcac
4021 ccactccgca gctttgctaa tcctttgcta cctccaccaa tgagtgcaat agaaccgaaa
4081 gtcccttaca caccactttt gtctcagcca gggcccactc ttcctaagac ccatgtgaaa
4141 acagcctccc ttgggttggc tggaaaagca agatcccctt tgcttcctgt gtctgtgcca
4201 acagcccctg aagtgtctga ggagagccac aaaccaacag aggattcagc caatgtgtat
4261 gaacaggatg atctgagtga acaaatggca agtttggaag gactcatgaa gcagcttaat
4321 gccatcacag gctcagcctt ttaacatgta tttctgaatg gatgaggtga attttccggg
4381 aactttgcag cataccaatt acccataaac agcacacctg tgtccaagaa ctctaaccag
4441 tgtacaggtc acccatcagg accactcagt taaggaagat cctgaagcag ttcagaagga
4501 ataagcattc cttctttcac aggcatcagg aattgtcaaa tgatgattat gagttcccta
4561 aacaaaagca aagatgcatt ttcactgcaa tgtcaaagtt tagctgct
```

E2F1

```
LOCUS       E2F1        2517 bp ss-mRNA            PRI       10-AUG-1992
DEFINITION  Homo sapiens (E2F-1) pRB-binding protein mRNA, complete cds.
ACCESSION   M96577
KEYWORDS    DNA-binding protein; pRB-binding protein; transcription factor E2F.
SOURCE      Homo sapiens fetal brain cDNA to mRNA.
REFERENCE   1  (bases 1 to 2517)
  AUTHORS   Helin,K., Lees,J.A., Vidal,M., Dyson,N.J., Harlow,E. and Fattaey,A.
  TITLE     A cDNA encoding a pRB-binding protein with properties of the
            transcription factor E2F
  JOURNAL   Cell 70, 337-350 (1992)
COMMENT     E2F, a cellular transcription factor, forms an inactive complex with
            pRB (retinoblastoma tumor suppressor protein).  HPV-16 E7 is one of
            several DNA virus oncoproteins that dissociates the E2F-pRB complex.
            This dissociation is coupled with enhanced E2F-dependent transcription.
            In the S-phase of the cell cycle, E2F complexes with cyclin A, p107,
            and cdk2 kinase.  Arroyo et al. (Mol and Cell Bio 13: 6537-46) report
            that the E7 protein of HPV-16 associates with the E2F-cyclin A complex.
            Arroyo et al. (Mol and Cell Bio 13: 6537-46) further note that a high
            risk HPV has a higher binding efficiency to the E2F-cyclin A complex
            than the low-risk HPV type. Pagano et al. (Oncogene 7: 1681-6) report
            that the E2F-cyclin A complex can occur in HPV-18 infected cells
            independent of pRB.
BASE COUNT      454 a    784 c    781 g    498 t
ORIGIN
        1 ggaattccgt ggccgggact ttgcaggcag cggcggccgg gggcggagcg ggatcgagcc
       61 ctcgccgagg cctgccgcca tgggcccgcg ccgccgccgc cgcctgtcac ccgggccgcg
      121 cgggccgtga gcgtcatggc cttggccggg gcccctgcgg gcggcccatg cgcgccggcg
      181 ctggaggccc tgctcggggc cggcgcgctg cggctgctcg actcctcgca gatcgtcatc
      241 atctccgccg cgcaggacgc cagcgccccg ccggctccca ccggccccgc ggcgcccgcc
      301 gccggcccct gcgaccctga cctgctgctc ttcgccacac cgcaggcgcc ccggcccaca
      361 cccagtgcgc cgcggcccgc gctcggccgc ccgccggtga agcggaggct ggacctggaa
      421 actgaccatc agtacctggc cgagagcagt gggccagctc ggggcagagg ccgccatcca
      481 ggaaaaggtg tgaaatcccc gggggagaag tcacgctatg agacctcact gaatctgacc
      541 accaagcgct tcctggagct gctgagccac tcggctgacg gtgtcgtcga cctgaactgg
      601 gctgccgagg tgctgaaggt gcagaagcgg cgcatctatg acatcaccaa cgtccttgag
      661 ggcatccagc tcattgccaa gaagtccaag aaccacatcc agtggctggg cagccacacc
      721 acagtgggcg tcggcggacg gcttgagggg ttgacccagg acctccgaca gctgcaggag
      781 agcgagcagc agctggacca cctgatgaat atctgtacta cgcagctgcg cctgctctcc
      841 gaggacactg acagccagcg cctggcctac gtgacgtgtc aggaccttcg tagcattgca
      901 gaccctgcag agcagatggt tatggtgatc aaagcccctc ctgagaccca gctccaagcc
      961 gtggactctt cggagaactt tcagatctcc cttaagagca aacaaggccc gatcgatgtt
     1021 ttcctgtgcc ctgaggagac cgtaggtggg atcagccctg ggaagacccc atcccaggag
     1081 gtcacttctg aggaggagaa cagggccact gactctgcca ccatagtgtc accaccacca
     1141 tcatctcccc cctcatccct caccacagat cccagccagt ctctactcag cctggagcaa
     1201 gaaccgctgt tgtcccggat gggcagcctg cgggctcccg tggacgagga ccgcctgtcc
     1261 ccgctggtgg cggccgactc gctcctggag catgtgcggg aggacttctc cggcctcctc
     1321 cctgaggagt tcatcagcct ttccccaccc cacgaggccc tcgactacca cttcggcctc
     1381 gaggagggcg agggcatcag agacctcttc gactgtgact ttggggacct cacccccctg
     1441 gatttctgac agggcttgga gggaccaggg tttccagagt agctcacctt gtctctgcag
     1501 ccctggagcc ccctgtccct ggccgtcctc ccagcctgtt tggaaacatt taatttatac
     1561 ccctctcctc tgtctccaga agcttctagc tctggggtct ggctaccgct aggaggctga
     1621 gcaagccagg aagggaagga gtctgtgtgg tgtgtatgtg catgcagcct acacccacac
     1681 gtgtgtaccg gggtgaatg tgtgtgagca tgtgtgtgtg catgtaccgg ggaatgaagg
     1741 tgaacataca cctctgtgtg tgcactgcag acacgcccca gtgtgtccac atgtgtgtgc
     1801 atgagtccat ctctgcgcgt ggggggggctc taactgcact ttcggccctt ttgctcgtgg
     1861 ggtcccacaa ggcccagggc agtgcctgct cccagaatct ggtgctctga ccaggccagg
     1921 tggggaggct ttggctggct gggcgtgtag gacggtgaga gcacttctgt cttaaaggtt
     1981 ttttctgatt gaagctttaa tggagcgtta tttatttatc gaggcctctt tggtgagcct
     2041 ggggaatcag caaaagggga ggaggggtgt ggggttgata ccccaactcc ctctaccctt
     2101 gagcaagggc aggggtccct gagctgttct tctgccccat actgaaggaa ctgaggcctg
     2161 ggtgatttat ttattgggaa agtgagggag ggagacagac tgactgacag ccatgggtgg
     2221 tcagatggtg gggtgggccc tctccagggg gccagttcag ggcccagctg ccccccagga
     2281 tggatatgag atgggagagg tgagtggggg accttcactg atgtgggcag gaggggtggt
     2341 gaaggcctcc cccagcccag accctgtggt ccctcctgca gtgtctgaag cgcctgcctc
```

```
2401 cccactgctc tgccccaccc tccaatctgc actttgattt gcttcctaac agctctgttc
2461 cctcctgctt tggttttaat aaatattttg atgacgttaa aaaaaggaat tcgatat
```

Epidermal Growth Factor Receptor

```
LOCUS       EGFR         5532 bp ss-mRNA            PRI       13-JUN-1985
DEFINITION  Human mRNA for precursor of epidermal growth factor receptor
ACCESSION   X00588
KEYWORDS    epidermal growth factor receptor; signal peptide.
SOURCE
REFERENCE   1  (bases 1 to 5532)
  AUTHORS   Ullrich A., Coussens L., Hayflick J.S., Dull T.J., Gray A.,
            Tam A.W., Lee J., Yarden Y., Libermann T.A., Schlessinger J.,
            Downward J., Mayes E.L., Whittle N., Waterfield M.D., Seeburg P.H.;
  TITLE     Human epidermal growth factor receptor cDNA sequence and aberrant
            expression of the amplified gene in A431 epidermoid carcinoma
            cells
  JOURNAL   Nature 309:418-425(1984).
COMMENT     Cohen et al. demonstrated that the BPV E5 protein activates the
            EGF receptor through complex formation.  They have further
            determined that this activation is specific to the cytoplasmic
            domain of the EGF receptor and is ligand independent (J Virol
            67: 5303-11).  Straight et al. (J Virol 67: 4521-32) demonstrated
            that E5 HPV-16 infected keratinocytes had a two to five fold
            increase in EGF receptors.

BASE COUNT     1472 a   1484 c   1337 g   1239 t
ORIGIN
        1 gccgcgctgc gccggagtcc cgagctagcc ccggcgccgc cgccgcccag accggacgac
       61 aggccacctc gtcggcgtcc gcccgagtcc ccgcctcgcc gccaacgcca caaccaccgc
      121 gcacggcccc ctgactccgt ccagtattga tcgggagagc cggagcgagc tcttcgggga
      181 gcagcgatgc gaccctccgg gacggccggg gcagcgctcc tggcgctgct ggctgcgctc
      241 tgcccggcga gtcgggctct ggaggaaaag aaagtttgcc aaggcacgag taacaagctc
      301 acgcagttgg gcacttttga agatcatttt ctcagcctcc agaggatgtt caataactgt
      361 gaggtggtcc ttgggaattt ggaaattacc tatgtgcaga ggaattatga tctttccttc
      421 ttaaagacca tccaggaggt ggctggttat gtcctcattg ccctcaacac agtggagcga
      481 attcctttgg aaaacctgca gatcatcaga ggaaatatgt actacgaaaa ttcctatgcc
      541 ttagcagtct tatctaacta tgatgcaaat aaaaccggac tgaaggagct gcccatgaga
      601 aatttacagg aaatcctgca tggcgccgtg cggttcagca acaaccctgc cctgtgcaac
      661 gtggagagca tccagtggcg ggacatagtc agcagtgact ttctcagcaa catgtcgatg
      721 gacttccaga accacctggg cagctgccaa aagtgtgatc caagctgtcc caatgggagc
      781 tgctggggtg caggagagga gaactgccag aaactgacca aaatcatctg tgcccagcag
      841 tgctccgggc gctgccgtgg caagtccccc agtgactgct gccacaacca gtgtgctgca
      901 ggctgcacag gcccccggga gagcgactgc ctggtctgcc gcaaattccg agacgaagcc
      961 acgtgcaagg acacctgccc cccactcatg ctctacaacc ccaccacgta ccagatggat
     1021 gtgaaccccg agggcaaata cagctttggt gccacctgcg tgaagaagtg tccccgtaat
     1081 tatgtggtga cagatcacgg ctcgtgcgtc cgagcctgtg gggccgacag ctatgagatg
     1141 gaggaagacg gcgtccgcaa gtgtaagaag tgcgaagggc cttgccgcaa agtgtgtaac
     1201 ggaataggta ttggtgaatt taaagactca ctctccataa atgctacgaa tattaaacac
     1261 ttcaaaaact gcacctccat cagtggcgat ctccacatcc tgccggtggc atttaggggt
     1321 gactccttca cacatactcc tcctctggat ccacaggaac tggatattct gaaaaccgta
     1381 aaggaaatca cagggttttt gctgattcag gcttggcctg aaaacaggac ggacctccat
     1441 gcctttgaga acctagaaat catacgcggc aggaccaagc aacatggtca gttttctctt
     1501 gcagtcgtca gcctgaacat aacatccttg ggattacgct ccctcaagga gataagtgat
     1561 ggagatgtga taatttcagg aaacaaaaat ttgtgctatg caaatacaat aaactggaaa
     1621 aaactgtttg ggacctccgg tcagaaaacc aaaattataa gcaacagagg tgaaaacagc
     1681 tgcaaggcca caggccaggt ctgccatgcc ttgtgctccc ccgagggctg ctggggcccg
     1741 gagcccaggg actgcgtctc ttgccggaat gtcagccgag gcagggaatg cgtggacaag
     1801 tgcaagcttc tggagggtga gccaagggag tttgtggaga actctgagtg catacagtgc
     1861 cacccagagt gcctgcctca ggccatgaac atcacctgca caggacgggg accagacaac
     1921 tgtatccagt gtgcccacta cattgacggc ccccactgcg tcaagacctg cccggcagga
     1981 gtcatgggag aaaacaacac cctggtctgg aagtacgcag acgccggcca tgtgtgccac
     2041 ctgtgccatc caaactgcac ctacggatgc actgggccag gtcttgaagg ctgtccaacg
     2101 aatgggccta agatcccgtc catcgccact gggatggtgg gggccctcct cttgctgctg
     2161 gtggtggccc tggggatcgg cctcttcatg cgaaggcgcc acatcgttcg gaagcgcacg
     2221 ctgcggaggc tgctgcagga gagggagctt gtggagcctc ttacacccag tggagaagct
     2281 cccaaccaag ctctcttgag gatcttgaag gaaactgaat tcaaaaagat caaagtgctg
     2341 ggctccggtg cgttcggcac ggtgtataag ggactctgga tcccagaagg tgagaaagtt
     2401 aaaattcccg tcgctatcaa ggaattaaga gaagcaacat ctccgaaagc caacaaggaa
     2461 atcctcgatg aagcctacgt gatggccagc gtggacaacc cccacgtgtg ccgcctgctg
```

```
2521 ggcatctgcc tcacctccac cgtgcaactc atcacgcagc tcatgccctt cggctgcctc
2581 ctggactatg tccgggaaca caaagacaat attggctccc agtacctgct caactggtgt
2641 gtgcagatcg caaagggcat gaactacttg gaggaccgtc gcttggtgca ccgcgacctg
2701 gcagccagga acgtactggt gaaaacaccg cagcatgtca agatcacaga ttttgggctg
2761 gccaaactgc tgggtgcgga agagaaagaa taccatgcag aaggaggcaa agtgcctatc
2821 aagtggatgg cattggaatc aattttacac agaatctata cccaccagag tgatgtctgg
2881 agctacgggg tgaccgtttg ggagttgatg acctttggat ccaagccata tgacggaatc
2941 cctgccagcg agatctcctc catcctggag aaaggagaac gcctccctca gccacccata
3001 tgtaccatcg atgtctacat gatcatggtc aagtgctgga tgatagacgc agatagtcgc
3061 ccaaagttcc gtgagttgat catcgaattc tccaaaatgg cccgagaccc ccagcgctac
3121 cttgtcattc agggggatga aagaatgcat ttgccaagtc ctacagactc caacttctac
3181 cgtgccctga tggatgaaga agacatggac gacgtggtgg atgccgacga gtacctcatc
3241 ccacagcagg gcttcttcag cagcccctcc acgtcacgga ctcccctcct gagctctctg
3301 agtgcaacca gcaacaattc caccgtggct tgcattgata gaaatgggct gcaaagctgt
3361 cccatcaagg aagacagctt cttgcagcga tacagctcag accccacagg cgccttgact
3421 gaggacagca tagacgacac cttcctccca gtgcctgaat acataaacca gtccgttccc
3481 aaaaggcccg ctggctctgt gcagaatcct gtctatcaca atcagcctct gaaccccgcg
3541 cccagcagag acccacacta ccaggacccc cacagcactg cagtgggcaa ccccgagtat
3601 ctcaacactg tccagcccac ctgtgtcaac agcacattcg acagccctgc ccactgggcc
3661 cagaaaggca gccaccaaat tagcctggac aaccctgact accagcagga cttctttccc
3721 aaggaagcca agccaaatgg catctttaag ggctccacag ctgaaaatgc agaataccta
3781 agggtcgcgc cacaaagcag tgaatttatt ggagcatgac cacggaggat agtatgagcc
3841 ctaaaaatcc agactctttc gatacccagg accaagccac agcaggtcct ccatcccaac
3901 agccatgccc gcattagctc ttagacccac agactggttt tgcaacgttt acaccgacta
3961 gccaggaagt acttccacct cgggcacatt ttgggaagtt gcattccttt gtcttcaaac
4021 tgtgaagcat ttacagaaac gcatccagca agaatattgt ccctttgagc agaaatttat
4081 ctttcaaaga ggtatatttg aaaaaaaaaa aaaaagtata tgtgaggatt tttattgatt
4141 ggggatcttg gagtttttca ttgtcgctat tgatttttac ttcaatgggc tcttccaaca
4201 aggaagaagc ttgctggtag cacttgctac cctgagttca tccaggccca actgtgagca
4261 aggagcacaa gccacaagtc ttccagagga tgcttgattc cagtggttct gcttcaaggc
4321 ttccactgca aaacactaaa gatccaagaa ggccttcatg gccccagcag gccggatcgg
4381 tactgtatca agtcatggca ggtacagtag gataagccac tctgtccctt cctgggcaaa
4441 gaagaaacgg aggggatgaa ttcttcctta gacttacttt tgtaaaaatg tccccacggt
4501 acttactccc cactgatgga ccagtggttt ccagtcatga gcgttagact gacttgtttg
4561 tcttccattc cattgttttg aaactcagta tgccgcccct gtcttgctgt catgaaatca
4621 gcaagagagg atgacacatc aaataataac tcggattcca gcccacattg gattcatcag
4681 catttggacc aatagcccac agctgagaat gtggaatacc taaggataac accgcttttg
4741 ttctcgcaaa aacgtatctc ctaatttgag gctcagatga aatgcatcag gtcctttggg
4801 gcatagatca gaagactaca aaaatgaagc tgctctgaaa tctcctttag ccatcacccc
4861 aacccccaa aattagtttg tgttacttat ggaagatagt tttctccttt tacttcactt
4921 caaaagcttt ttactcaaag agtatatgtt ccctccaggt cagctgcccc caaacccccct
4981 ccttacgctt tgtcacacaa aaagtgtctc tgccttgagt catctattca agcacttaca
5041 gctctggcca caacagggca ttttacaggt gcgaatgaca gtagcattat gagtagtgtg
5101 aattcaggta gtaaatatga aactagggtt tgaaattgat aatgctttca caacatttgc
5161 agatgtttta gaaggaaaaa agttccttcc taaaataatt tctctacaat tggaagattg
5221 gaagattcag ctagttagga gcccattttt tcctaatctg tgtgtgccct gtaacctgac
5281 tggttaacag cagtcctttg taaacagtgt tttaaactct cctagtcaat atccacccca
5341 tccaatttat caaggaagaa atggttcaga aaatattttc agcctacagt tatgttcagt
5401 cacacacaca tacaaaatgt tccttttgct tttaaagtaa tttttgactc ccagatcagt
5461 cagagcccct acagcattgt taagaaagta tttgattttt gtctcaatga aaataaaact
5521 atattcattt cc
```

ERBB2

```
LOCUS       ERBB2        4473 bp ss-mRNA            PRI       30-JUN-1987
DEFINITION  Human c-erb-B-2 mRNA.
ACCESSION   X03363
KEYWORDS    c-myc proto-oncogene; cell surface glycoprotein; erbB gene;
            erbB oncogene; glycoprotein; growth factor; growth factor receptor;
            kinase; neu oncogene; transmembrane protein; tyrosine kinase.
SOURCE      Homo sapiens cell line MKN-7
REFERENCE   1  (bases 1 to 4473)
  AUTHORS   Yamamoto,T., Ikawa,S., Akiyama,T., Semba,K., Nomura,N.,
            Miyajima,N., Saito,T. and Toyoshima,K.
  TITLE     Similarity of protein encoded by the human c-erb-B-2 gene to
            epidermal growth factor receptor
  JOURNAL   Nature 319, 230-234 (1986)
COMMENT     Milde-Langosch et al. (Verh Dtsch Ges Pathol 75: 363-5)
            demonstrated an increase in expression of c-erbB2 in
            4/11 HPV-positive endometrial carcinomas and in 3/19
            HPV-positive CIN3 lesions.

            The c-erb-B-2 protein shows similarity to the epidermal growth
            factor receptor.

               EMBL features not translated to GenBank features:
               key        from     to       description

               SITE        376    384       pot. glycosylation site
               SITE        544    558       pot. glycosylation site
               SITE        733    741       pot. glycosylation site
               SITE        949    957       pot. glycosylation site
               SITE       1762   1770       pot. glycosylation site
               SITE       1885   1893       pot. glycosylation site
               SITE       2059   2067       pot. glycosylation site
               SITE       2353   3132       aa 727-986, seq. homologous to EGF
                                            receptor kinase domain
               SITE       2446   2454       pot. glycosylation site
               SITE       4455   4460       put. polyA signal
               POLYA      4473   4473       polyA site
BASE COUNT      902 a   1383 c   1329 g    859 t
ORIGIN
        1 aaggggaggt aaccctggcc cctttggtcg gggccccggg cagccgcgcg ccccttccca
       61 cggggccctt tactgcgccg cgcgcccggc ccccacccct cgcagcaccc cgcgccccgc
      121 gccctcccag ccgggtccag ccggagccat ggggccggag ccgcagtgag caccatggag
      181 ctggcggcct tgtgccgctg gggctcctc ctcgccctct tgccccccgg agccgcgagc
      241 acccaagtgt gcaccggcac agacatgaag ctgcggctcc ctgccagtcc cgagacccac
      301 ctggacatgc tccgccacct ctaccagggc tgccaggtgg tgcagggaaa cctggaactc
      361 acctacctgc ccaccaatgc cagcctgtcc ttcctgcagg atatccagga ggtgcagggc
      421 tacgtgctca tcgctcacaa ccaagtgagg caggtcccac tgcagaggct gcggattgtg
      481 cgaggcaccc agctctttga ggacaactat gccctggccg tgctagacaa tggagacccg
      541 ctgaacaata ccacccctgt cacaggggcc tccccaggag gcctgcggga gctgcagctt
      601 cgaagcctca cagagatctt gaaaggaggg gtcttgatcc agcggaaccc ccagctctgc
      661 taccaggaca cgattttgtg gaaggacatc ttccacaaga acaaccagct ggctctcaca
      721 ctgatagaca ccaaccgctc tcgggcctgc acccctgtt ctccgatgtg taagggctcc
      781 cgctgctggg gagagagttc tgaggattgt cagagcctga cgcgcactgt ctgtgccggt
      841 ggctgtgccc gctgcaaggg gccactgccc actgactgct gccatgagca gtgtgctgcc
      901 ggctgcacgg gccccaagca ctctgactgc ctggcctgcc tccacttcaa ccacagtggc
      961 atctgtgagc tgcactgccc agccctggtc acctacaaca cagacacgtt tgagtccatg
     1021 cccaatcccg agggccggta tacattcggc gccagctgtg tgactgcctg tccctacaac
     1081 tacctttcta cggacgtggg atcctgcacc ctcgtctgcc ccctgcacaa ccaagaggtg
     1141 acagcagagg atggaacaca gcggtgtgag aagtgcagca agccctgtgc ccgagtgtgc
     1201 tatggtctgg gcatggagca cttgcgagag gtgagggcag ttaccagtgc caatatccag
     1261 gagtttgctg gctgcaagaa gatctttggg agcctggcat ttctgccgga gagctttgat
     1321 ggggacccag cctccaacac tgccccgctc cagccagagc agctccaagt gtttgagact
     1381 ctggaagaga tcacaggtta cctatacatc tcagcatggc cggacagcct gcctgacctc
     1441 agcgtcttcc agaacctgca agtaatccgg ggacgaattc tgcacaatgg cgcctactcg
     1501 ctgaccctgc aagggctggg catcagctgg ctggggctgc gctcactgag ggaactgggc
     1561 agtggactgg ccctcatcca ccataacacc cacctctgct tcgtgcacac ggtgccctgg
     1621 gaccagctct ttcggaaccc gcaccaagct ctgctccaca ctgccaaccg gccagaggac
```

```
1681 gagtgtgtgg gcgagggcct ggcctgccac cagctgtgcg cccgagggca ctgctggggt
1741 ccagggccca cccagtgtgt caactgcagc cagttccttc ggggccagga gtgcgtggag
1801 gaatgccgag tactgcaggg gctccccagg gagtatgtga atgccaggca ctgtttgccg
1861 tgccaccctg agtgtcagcc ccagaatggc tcagtgacct gttttggacc ggaggctgac
1921 cagtgtgtgg cctgtgccca ctataaggac cctcccttct gcgtggcccg ctgccccagc
1981 ggtgtgaaac ctgacctctc ctacatgccc atctggaagt ttccagatga ggagggcgca
2041 tgccagcctt gccccatcaa ctgcacccac tcctgtgtgg acctggatga caagggctgc
2101 cccgccgagc agagagccag ccctctgacg tccatcatct ctgcggtggt tggcattctg
2161 ctggtcgtgg tcttgggggt ggtctttggg atcctcatca agcgacggca gcagaagatc
2221 cggaagtaca cgatgcggag actgctgcag gaaacggagc tggtggagcc gctgacacct
2281 agcggagcga tgcccaacca ggcgcagatg cggatcctga aagagacgga gctgaggaag
2341 gtgaaggtgc ttggatctgg cgcttttggc acagtctaca agggcatctg gatccctgat
2401 ggggagaatg tgaaaattcc agtggccatc aaagtgttga gggaaaacac atcccccaaa
2461 gccaacaaag aaatcttaga cgaagcatac gtgatggctg gtgtgggctc cccatatgtc
2521 tcccgccttc tgggcatctg cctgacatcc acggtgcagc tggtgacaca gcttatgccc
2581 tatggctgcc tcttagacca tgtccgggaa aaccgcggac gcctgggctc ccaggacctg
2641 ctgaactggt gtatgcagat tgccaagggg atgagctacc tggaggatgt gcggctcgta
2701 cacagggact tggccgctcg gaacgtgctg gtcaagagtc ccaaccatgt caaaattaca
2761 gacttcgggc tggctcggct gctggacatt gacgagacag agtaccatgc agatgggggc
2821 aaggtgccca tcaagtggat ggcgctggag tccattctcc gccggcggtt cacccaccag
2881 agtgatgtgt ggagttatgg tgtgactgtg tgggagctga tgacttttgg ggccaaacct
2941 tacgatggga tcccagcccg ggagatccct gacctgctgg aaaaggggga gcggctgccc
3001 cagcccccca tctgcaccat tgatgtctac atgatcatgg tcaaatgttg gatgattgac
3061 tctgaatgtc ggccaagatt ccgggagttg gtgtctgaat tctcccgcat ggccagggac
3121 ccccagcgct ttgtggtcat ccagaatgag gacttgggcc cagccagtcc cttggacagc
3181 accttctacc gctcactgct ggaggacgat gacatggggg acctggtgga tgctgaggag
3241 tatctggtac cccagcaggg cttcttctgt ccagaccctg ccccgggcgc tgggggcatg
3301 gtccaccaca ggcaccgcag ctcatctacc aggagtggcg gtggggacct gacactaggg
3361 ctggagccct ctgaagagga ggcccccagg tctccactgg caccctccga aggggctggc
3421 tccgatgtat ttgatggtga cctgggaatg gggcagcca aggggctgca aagcctcccc
3481 acacatgacc ccagccctct acagcggtac agtgaggacc ccacagtacc cctgccctct
3541 gagactgatg gctacgttgc ccccctgacc tgcagccccc agcctgaata tgtgaaccag
3601 ccagatgttc ggccccagcc cccttcgccc cgagagggcc ctctgcctgc tgcccgacct
3661 gctggtgcca ctctggaaag gcccaagact ctctccccag ggaagaatgg ggtcgtcaaa
3721 gacgtttttg cctttggggg tgccgtggag aacccgagt acttgacacc ccagggagga
3781 gctgcccctc agccccaccc tcctcctgcc ttcagcccag ccttcgacaa cctctattac
3841 tgggaccagg acccaccaga gcgggggct ccacccagca ccttcaaagg gacacctacg
3901 gcagagaacc cagagtacct gggtctggac gtgccagtgt gaaccagaag gccaagtccg
3961 cagaagccct gatgtgtcct cagggagcag ggaaggcctg acttctgctg gcatcaagag
4021 gtgggagggc cctccgacca cttccagggg aacctgccat gccaggaacc tgtcctaagg
4081 aaccttcctt cctgcttgag ttcccagatg gctggaaggg gtccagcctc gttggaagag
4141 gaacagcact ggggagtctt tgtggattct gaggccctgc ccaatgagac tctagggtcc
4201 agtggatgcc acagcccagc ttggcccttt ccttccagat cctgggtact gaaagcctta
4261 gggaagctgg cctgagaggg gaagcggccc taagggagtg tctaagaaca aaagcgaccc
4321 attcagagac tgtccctgaa acctagtact gccccccatg aggaaggaac agcaatggtg
4381 tcagtatcca ggctttgtac agagtgcttt tctgtttagt ttttactttt tttgttttgt
4441 ttttttaaag atgaaataaa gacccagggg gag
```

```
LOCUS       KRAS         5775 bp ds-DNA             PRI       05-JUN-1991
DEFINITION  Human K-ras oncogene, complete cds.
ACCESSION   M54968 M38506
KEYWORDS    K-ras oncogene.
SOURCE      Human tumor DNA.
REFERENCE   1  (bases 1 to 5775)
  AUTHORS   Kahn,S., Yamamoto,F.-I., Almoguera,C., Winter,E., Forrester,K.,
            Jordano,J. and Perucho,M.
  TITLE     The c-K-ras gene and human cancer (review)
  JOURNAL   Anticancer Res. 7, 639-652 (1987)
BASE COUNT     1739 a    974 c   1105 g   1957 t
ORIGIN
        1 tcctaggcgg cggccgcggc ggcggaggca gcagcggcgg cggcagtggc ggcggcgaag
       61 gtggcggcgg ctcggccagt actcccggcc cccgccattt cggactggga gcgagcgcgg
      121 cgcaggcact gaaggcggcg gcggggccag aggctcagcg gctcccaggt gcgggagaga
      181 ggcctgctga aaatgactga atataaactt gtggtagttg gagcttgtgg cgtaggcaag
      241 agtgccttga cgatacagct aattcagaat cattttgtgg acgaatatga tccaacaata
      301 gaggattcct acaggaagca agtagtaatt gatggagaaa cctgtctctt ggatattctc
      361 gacacagcag gtcaagagga gtacagtgca atgagggacc agtacatgag gactggggag
      421 ggctttcttt gtgtatttgc cataaataat actaaatcat ttgaagatat tcaccattat
      481 agagaacaaa ttaaaagagt taaggactct gaagatgtac ctatggtcct agtaggaaat
      541 aaatgtgatt tgccttctag aacagtagac acaaaacagg ctcaggactt agcaagaagt
      601 tatggaattc cttttattga aacatcagca aagacaagac agggtgttga tgatgccttc
      661 tatacattag ttcgagaaat tcgaaaacat aaagaaaaga tgagcaaaga tggtaaaaag
      721 aagaaaaaga agtcaaagac aaagtgtgta attatgtaaa tacaatttgt actttttttct
      781 taaggcatac tagtacaagt ggtaattttt gtacattaca ctaaattatt agcatttgtt
      841 ttagcattac ctaatttttt tcctgctcca tgcagactgt tagcttttac cttaaatgct
      901 tattttaaaa tgacagtgga agttttttt tcctcgaagt gccagtattc ccagagtttt
      961 ggtttttgaa ctagcaatgc ctgtgaaaaa gaaactgaat acctaagatt tctgtcttgg
     1021 ggtttttggt gcatgcagtt gattacttct tatttttctt accaagtgtg aatgttggtg
     1081 tgaaacaaat taatgaagct tttgaatcat ccctattctg tgttttatct agtcacataa
     1141 atggattaat tactaatttc agttgagacc ttctaattgg tttttactga aacattgagg
     1201 gacacaaatt tatgggcttc ctgatgatga ttcttctagg catcatgtcc tatagtttgt
     1261 catccctgat gaatgtaaag ttacactgtt cacaaaggtt ttgtctcctt tccactgcta
     1321 ttagtcatgg tcactctccc caaaatatta tattttttct ataaaaagaa aaaaatggaa
     1381 aaaaattaca aggcaatgga aactattata aggccatttc cttttcacat tagataaatt
     1441 actataaaga ctcctaatag ctttttcctg ttaaggcaga cccagtatga atgggattat
     1501 tatagcaacc attttggggc tatatttaca tgctactaaa tttttataat aattgaaaag
     1561 attttaacaa gtataaaaaa attctcatag gaattaaatg tagtctccct gtgtcagact
     1621 gctctttcat agtataactt taaatctttt cttcaacttg agtctttgaa gatagtttta
     1681 attctgcttg tgacattaaa agattatttg ggccagttat agcttattag gtgttgaaga
     1741 gaccaaggtt gcaagccagg ccctgtgtga accttgagct ttcatagaga gtttcacagc
     1801 atggactgtg tgccccacgg tcatccgagt ggttgtacga tgcattggtt agtcaaaaat
     1861 ggggagggac tagggcagtt tggatagctc aacaagatac aatctcactc tgtggtggtc
     1921 ctgctgacaa atcaagagca ttgcttttgt ttcttaagaa aacaaactct tttttaaaaa
     1981 ttacttttaa atattaactc aaaagttgag attttggggt ggtggtgtgc caagacatta
     2041 attttttttt taaacaatga agtgaaaaag ttttacaatc tctaggtttg gctagttctc
     2101 ttaacactgg ttaaattaac attgcataaa cacttttcaa gtctgatcca tatttaataa
     2161 tgctttaaaa taaaaataaa aacaatcctt ttgataaatt taaaatgtta cttattttaa
     2221 aataaatgaa gtgagatggc atggtgaggt gaaagtatca ctggactagg ttgttggtga
     2281 cttaggttct agataggtgt cttttaggac tctgattttg aggacatcac ttactatcca
     2341 tttcttcatg ttaaaagaag tcatctcaaa ctcttagttt ttttttttta cactatgtga
     2401 tttatattcc atttacataa ggatacactt atttgtcaag ctcagcacaa tctgtaaatt
     2461 tttaacctat gttacaccat cttcagtgcc agtcttgggc aaaattgtgc aagaggtgaa
```

```
2521 gtttatattt gaatatccat tctcgtttta ggactcttct tccatattag tgtcatcttg
2581 cctccctacc ttccacatgc cccatgactt gatgcagttt taatacttgt aattccccta
2641 accataagat ttactgctgc tgtggatatc tccatgaagt tttcccactg agtcacatca
2701 gaaatgccct acatcttatt ttcctcaggg ctcaagagaa tctgacagat accataaagg
2761 gatttgacct aatcactaat tttcaggtgg tggctgatgc tttgaacatc tctttgctgc
2821 ccaatccatt agcgacagta ggatttttca accctggtat gaatagacag aaccctatcc
2881 agtggaagga gaatttaata aagatagtgc agaaagaatt ccttaggtaa tctataacta
2941 ggactactcc tggtaacagt aatacattcc attgttttag taaccagaaa tcttcatgca
3001 atgaaaaata ctttaattca tgaagcttac tttttttttt ttggtgtcag agtctcgctc
3061 ttgtcaccca ggctggaatg cagtggcgcc atctcagctc actgcaacct tccatcttcc
3121 caggttcaag cgattctcgt gcctcggcct cctgagtagc tgggattaca ggcgtgtgca
3181 ctacactcaa ctaatttttg tatttttagg agagacgggg tttcacctgt tggccaggct
3241 ggtctcgaac tcctgacctc aagtgattca cccaccttgg cctcataaac ctgttttgca
3301 gaactcattt attcagcaaa tatttattga gtgcctacca gatgccagtc accgcacaag
3361 gcactgggta tatggtatcc ccaaacaaga gacataatcc cggtccttag gtactgctag
3421 tgtggtctgt aatatcttac taaggccttt ggtatacgac ccagagataa cacgatgcgt
3481 attttagttt tgcaaagaag gggtttggtc tctgtgccag ctctataatt gttttgctac
3541 gattccactg aaactcttcg atcaagctac tttatgtaaa tcacttcatt gttttaaagg
3601 aataaacttg attatattgt ttttttattt ggcataactg tgattctttt aggacaatta
3661 ctgtacacat taaggtgtat gtcagatatt catattgacc caaatgtgta atattccagt
3721 tttctctgca taagtaatta aaatatactt aaaaattaat agttttatct gggtacaaat
3781 aaacagtgcc tgaactagtt cacagacaag ggaaacttct atgtaaaaat cactatgatt
3841 tctgaattgc tatgtgaaac tacagatctt tggaacactg tttaggtagg gtgttaagac
3901 ttgacacagt acctcgtttc tacacagaga aagaaatggc catacttcag gaactgcagt
3961 gcttatgagg ggatatttag gcctcttgaa tttttgatgt agatgggcat ttttttaagg
4021 tagtggttaa ttacctttat gtgaactttg aatggtttaa caaaagattt gtttttgtag
4081 agattttaaa gggggagaat tctagaaata aatgttacct aattattaca gccttaaaga
4141 caaaaatcct tgttgaagtt tttttaaaaa aagactaaat tacatagact taggcattaa
4201 catgtttgtg gaagaatata gcagacgtat attgtatcat ttgagtgaat gttcccaagt
4261 aggcattcta ggctctattt aactgagtca cactgcatag gaatttagaa cctaactttt
4321 ataggttatc aaaactgttg tcaccattgc acaattttgt cctaatatat acatagaaac
4381 tttgtggggc atgttaagtt acagtttgca caagttcatc tcatttgtat tccattgatt
4441 tttttttttc ttctaaacat tttttcttca aaacagtata tataactttt tttaggggat
4501 tttttttaga cagcaaaaaa ctatctgaag atttccattt gtcaaaaagt aatgatttct
4561 tgataattgt gtagtgaatg ttttttagaa cccagcagtt accttgaaag ctgaatttat
4621 atttagtaac ttctgtgtta atactggata gcatgaattc tgcattgaga aactgaatag
4681 ctgtcataaa atgctttctt tcctaaagaa agatactcac atgagttctt gaagaatagt
4741 cataactaga ttaagatctg tgttttagtt taatagtttg aagtgcctgt ttgggataat
4801 gataggtaat ttagatgaat ttaggggaaa aaaaagttat ctgcagttat gttgagggcc
4861 catctctccc cccacacccc cacagagcta actgggttac agtgttttat ccgaaagttt
4921 ccaattccac tgtcttgtgt tttcatgttg aaaatacttt tgcatttttc ctttgagtgc
4981 caatttctta ctagtactat ttcttaatgt aacatgttta cctggcctgt cttttaacta
5041 tttttgtata gtgtaaactg aaacatgcac attttgtaca ttgtgctttc ttttgtgggt
5101 catatgcagt gtgatccagt tgttttccat catttggttg cgctgaccta ggaatgttgg
5161 tcatatcaaa cattaaaaat gaccactctt ttaatgaaat taacttttaa atgtttatag
5221 gagtatgtgc tgtgaagtga tctaaaattt gtaatatttt tgtcatgaac tgtactactc
5281 ctaattattg taatgtaata aaaatagtta cagtgactat gagtgtgtat ttattcatgc
5341 aaatttgaac tgtttgcccc gaaatggata tggatacttt ataagccata gacactatag
5401 tataccagtg aatcttttat gcagcttgtt agaagtatcc ttttattttc taaaaggtgc
5461 tgtggatatt atgtaaaggc gtgtttgctt aaacaatttt ccatatttag aagtagatgc
5521 aaaacaaatc tgcctttatg acaaaaaaat aggataacat tatttattta tttcctttta
5581 tcaataaggt aattgataca caacaggtga cttggtttta ggcccaaagg tagcagcagc
5641 aacattaata atggaaataa ttgaatagtt agttatgtat gttaatgcca gtcaccagca
5701 ggctatttca aggtcagaag taatgactcc atacatatta tttatttcta taactacatt
```

```
5761 taaatcatta ccagg
```

```
LOCUS       MCP1         3227 bp ds-DNA            PRI       15-JUN-1994
DEFINITION  Human MCP-1 gene for monocyte chemoattractant protein-1.
ACCESSION   D26087
KEYWORDS    monocyte chemoattractant protein-1.
SOURCE      Homo sapiens (library: EMBL3 SP6/T7) placenta DNA.
REFERENCE   1  (sites)
  AUTHORS   Rollins,B.J., Stier,P., Ernst,T. and Wong,G.G.
  TITLE     The human homolog of the JE gene encodes a monocyte secretory
            protein
  JOURNAL   Mol. Cell. Biol. 9, 4687-4695 (1989)
REFERENCE   2  (sites)
  AUTHORS   Shyy,Y.-J., Li,Y.-S. and Kolattukudy,P.E.
  TITLE     Structure of human monocyte chemotactic protein gene and its
            regulation by TPA
  JOURNAL   Biochem. Biophys. Res. Commun. 169, 346-351 (1990)
REFERENCE   3  (bases 1 to 3227)
  AUTHORS   Ueda,A., Okuda,K., Ohno,S., Shirai,A., Igarashi,T., Matsunaga,K.,
            Fukushima,J., Kawamoto,S., Ishigatsubo,Y. and Okubo,T.
  TITLE     NF-kB and Sp1 regulate transcription of human monocyte
            chemoattractant protein-1 gene
  JOURNAL   J. Immunol. (1993) In press
COMMENT     Rosl et al. (J Virol 68: 2142-50) demonstrated that MCP-1 protein
            is not expressed in HPV-18 positive cervical carcinoma cells, despite
            the fact that the gene which encodes it, the JE gene, has no mutations.
            Conversely, in non-malignant cells the tumor necrosis factor alpha
            induces MCP-1 expression which results in HPV-18 repression.

            Submitted (06-Dec-1993) to DDBJ by: Atsuhisa Ueda
            First Department of Internal Medicine
            Yokohama City University School of Medicine
            3-9 Fukuura, Kanazawa-ku
            Yokohama 236
            Japan
            Phone: 045-787-2630
            Fax:   045-786-3444.

            NCBI gi: 516772
BASE COUNT      850 a    793 c    732 g    852 t
ORIGIN
        1 ctgcagcgat gaccaaggtt catgcagagg atcttagtgg tggggtcaga ccccgggagg
       61 aatgaagaaa gcattattca ccaagaggag cttttccatt ctttatctat gagttgatag
      121 agaggaggac ccggggtaac tgaggattct ggacagcatc agagcattga ccctcatttt
      181 ccccatagcc cctctggggg cctttccctt gtgtgtcccc aagcgagagt ccaaccaagg
      241 tttgtgccag agcctaaccc aggcttgtgc cgagatgttc ccagcacagc cccatgtgag
      301 agctccctgg ctccgggccc agtatctgga atgcaggctc cagccaaatg cattctcttc
      361 tacgggatct gggaacttcc aaagctgcct cctcagagtg ggaatttcca ctcacttctc
      421 tcacgccagc actgacctcc cagcgggggga gggcatcttt tcttgacaga gcagaagtgg
      481 gaggcagaca gctgtcactt tccagaagac tttcttttct gattcatacc cttcaccttc
      541 cctgtgttta ctgtctgata tatgcaaagg ccaagtcact ttccagagat gacaactcct
      601 tcctgaagta gagacatgct tccaacactc agaagcctat gtgaacactc agccagcaaa
      661 gctggaagtt tttctctgtg accatgggct aattggtctc cttctctgga ttgtggctta
      721 tcagataaaa acaagtgagt catgccacag gatgtctata agcccattga ttctgggatt
      781 ctatgagtga tgctgatatg actaagccag gagagactta tttaaagatc tcagcatctt
      841 tcagcttgtt aacctagaga aaacccgaag catgactgga ttataaaggg aaattgaatg
      901 cggtccacca agttcatggt aaaggatgca ctaacagatt agagagaggt ttcccctgat
      961 atgaggaaaa cttcttggaa gatgaggtga gatggcctag gaagaaattc ctacacaaag
     1021 ttgcacagtc tctagtcctg gaaacatttt attcattgga taagaatgga ttgaggcatg
     1081 agcagaggac tgagaccaca cagagaagtt tcaacactgg ttggggagaa aaggagtaac
     1141 tagtgagatt caggcagaac agaataaggc tcctcaagag gcacaagcaa agcaggcctc
     1201 gagttgattt gttctctctt catcctgctt tttgtaattc caccagagtc tgaaatggcc
     1261 actccataga gtctctgctc tgggattctc caggaaacca atatccatca tgagacatca
     1321 agtctagtcc caggaagaag agattctgga atggaaacat cctgggtggg agtctcagca
     1381 catctactat tctgtctgag ttactggaca aataacttca gttttaacct aacgaaagct
     1441 gggttggttg gaggactggg caggcagcgc ttggaaagta tgtcagcacc atacctgact
     1501 ccctgaatgc actcaacaat gccattactg accacttact agaaataaaa cagtcatttg
```

```
1561 ttgaatacaa cccgtttctt tttacaagtg tagtgaaaag tgttttcttt caagaaaccc
1621 catgcattta tagacattgc ctcagtgacc ctttatgaaa gaagtcacta gtctttgtat
1681 gcccattggg caagggcacc gcaaggctca gaaggaggag gcagtgggct aggagaatcg
1741 agagatcaga attttaaact cagcccagcc attaacatgc ctcaagtact cctatcatat
1801 ttgtaagaga caacagttca ctgaaatgaa ttctaaggtc tttgggtttt tatcagtgtg
1861 cttctgtagt ttctgaggaa atctaaggca caactgagga atgaagtcag gctttccaat
1921 tcccgaaata ctcctccact gcttactcat gtcccttgga aattaagaag gaagccagga
1981 gcatagctgc cataaccagg gatgaacttc ttgtccactg ctgcctgcta tgctagcaac
2041 agcctcctaa ctcataatga cttagccatg aggaatgttt ctagattctc ctttagctgt
2101 ctgcccattt ggaagatgct gaggacagag agaggaccca agcaggcaac tagttggagg
2161 acttgtacac gtttccttcc agcagtatgt cagagaggtg gcagcccact ggggacaggg
2221 ctgcctgggt tctgtgctcg aggggacctt gagcaggcta tttaaccctt cttgcctcag
2281 ttgcctgatc tataacatga aaattagcaa tccctactag ataaagttgg ggaatttaca
2341 gagttaatat ttgtaaaggt ctgagaatat tcctggcaga gtaagcactc tgtgagtatg
2401 acactggcat ttcttctgca gcactacatg ctgtctatgc ctttgtccaa gtctgaaacc
2461 ctagaactct tagaattcag ttcaatgttt acacaatcct acagttctgc taggcttcta
2521 tgatgctact attctgcatt tgaatgagca aatggattta atgcattgtc agggagccgg
2581 ccaaagcttg agagctcctt cctggctggg aggccccttg gaatgtggcc tgaaggtaag
2641 ctggcagcga gcctgacatg ctttcatcta gtttcctcgc ttccttcctt ttctgcagtt
2701 ttcgcttcac agaaagcaga atccttaaaa ataaccctct tagttcacat ctgtggtcag
2761 tctgggctta atggcacccc atcctcccca tttgctcatt tggtctcagc agtgaatgga
2821 aaaagtgtct cgtcctgacc ccctgcttcc ctttcctact tcctggaaat ccacaggatg
2881 ctgcatttgc tcagcagatt taacagccca cttatcactc atggaagatc cctcctcctg
2941 cttgactccg ccctctctcc ctctgcccgc tttcaataag aggcagagac agcagccaga
3001 ggaaccgaga ggctgagact aacccagaaa catccaattc tcaaactgaa gctcgcactc
3061 tcgcctccag catgaaagtc tctgccgccc ttctgtgcct gctgctcata gcagccacct
3121 tcattcccca agggctcgct cagccaggta aggccccctc ttcttctcct tgaaccacat
3181 tgtcttctct ctgagttatc atggaccatc caagcagacg tggtacc
```

```
LOCUS       NF1          3238 bp ss-mRNA            PRI       02-NOV-1990
DEFINITION  Human hepatic nuclear factor 1 (HNF1) mRNA, complete cds, clones
            HCL10, HCL12, HCL17, and HCL20.
ACCESSION   M57732 J04771
KEYWORDS    hepatic nuclear factor 1; transcription factor.
SOURCE      Homo sapiens liver NF1 cDNA to mRNA.
REFERENCE   1  (bases 1 to 3238)
  AUTHORS   Bach,I., Galcheva-Gargova,Z.I., Mattei,M.-G., Simon-Chazottes,D.,
            Guenet,J.-L., Cereghini,S. and Yaniv,M.
  TITLE     Cloning of human hepatic nuclear factor 1 (HNF1) and chromosomal
            localization of its gene in man and mouse
  JOURNAL   Genomics 8, 155-164 (1990)
COMMENT     Chong et al. (Nucleic Acids Res 18: 465-70) report that the HPV-16
            enhancer is activated by NF1, AP-1 and TEF-2, and that it occurs
            without cooperation.  Taniguchi et al. (Virology 195: 500-10)
            demonstrated reduced enhancer activity upon mutation of any of the
            three NF-1 binding sites (GCCAA).

BASE COUNT      651 a   1100 c    955 g    531 t      1 others
ORIGIN
        1 cgtggccctg tggcagccga gccatggttt ctaaactgag ccagctgcag acggagctcc
       61 tggcggccct gctcgagtca gggctgagca aagaggcact gatccaggca ctgggtgagc
      121 cggggcccta cctcctggct ggagaaggcc ccctggacaa gggggagtcc tgcggcggcg
      181 gtcgagggga gctggctgag ctgcccaatg ggctgggggа gactcggggc tccgaggacg
      241 agacggacga cgatggggaa gacttcacgc cacccatcct caaagagctg gagaacctca
      301 gccctgagga ggcggcccac cagaaagccg tggtggagac ccttctgcag gaggacccgt
      361 ggcgtgtggc gaagatggtc aagtcctacc tgcagcagca caacatccca cagcgggagg
      421 tggtcgatac cactggcctc aaccagtccc acctgtccca acacctcaac aagggcactc
      481 ccatgaagac gcagaagcgg gccgccctgt acacctggta cgtccgcaag cagcgagagg
      541 tggcgcagca gttcacccat gcagggcagg gagggctgat tgaagagccc acaggtgatg
      601 agctaccaac caagaagggg cggaggaacc gtttcaagtg gggcccagca tcccagcaga
      661 tcctgttcca ggcctatgag aggcagaaga accctagcaa ggaggagcga gagacgctag
      721 tggaggagtg caatagggcg gaatgcatcc agagaggggt gtccccatca caggcacagg
      781 ggctgggctc caacctcgtc acggaggtgc gtgtctacaa ctggtttgcc aaccggcgca
      841 aagaagaagc cttccggcac aagctggcca tggacacgta cagcgggccc cccccagggc
      901 caggcccggg acctgcgctg cccgctcaca gctcccctgg cctgcctcca cctgccctct
      961 cccccagtaa ggtccacggt gtgcgctntg gacagcctgc gaccagtgag actgcagaag
     1021 taccctcaag cagcggcggt cccttagtga cagtgtctac acccctccac caagtgtccc
     1081 ccacgggcct ggagcccagc cacagcctgc tgagtacaga agccaagctg gtctcagcag
     1141 ctgggggccc cctccccсct gtcagcaccc tgacagcact gcacagcttg gagcagacat
     1201 ccccaggcct caaccagcag ccccagaacc tcatcatggc ctcacttcct ggggtcatga
     1261 ccatcgggcc tggtgagcct gcctccctgg gtcctacgtt caccaacaca ggtgcctcca
     1321 ccctggtcat cggcctggcc tccacgcagg cacagagtgt gccggtcatc aacagcatgg
     1381 gcagcagcct gaccaccctg cagcccgtcc agttctccca gccgctgcac ccctcctacc
     1441 agcagccgct catgccacct gtgcagagcc atgtgaccca gagccccttc atggccacca
     1501 tggctcagct gcagagcccc cacgccctct acagccacaa gcccgaggtg gcccagtaca
     1561 cccacacggg cctgctcccg cagactatgc tcatcaccga caccaccaac ctgagcgccc
     1621 tggccagcct cacgcccacc aagcaggtct tcacctcaga cactgaggcc tccagtgagt
     1681 ccgggcttca cacgccggca tctcaggcca ccaccctcca cgtccccagc caggaccctg
     1741 ccggcatcca gcacctgcag ccggcccacc ggctcagcgc cagccccaca gtgtcctcca
     1801 gcagcctggt gctgtaccag agctcagact ccagcaatgg ccagagccac ctgctgccat
     1861 ccaaccacag cgtcatcgag accttcatct ccacccagat ggcctcttcc tcccagtaac
     1921 cacggcacct gggccctggg gcctgtactg cctgcttggg gggtgatgag ggcagcagcc
     1981 agccctgcct ggaggacctg agcctgccga gcaaccgtgg cccttcctgg acagctgtgc
     2041 ctcgctcccc actctgctct gatgcatcag aaagggaggg ctctgaggcg ccccaacccg
     2101 tggaggctgc tcggggtgca caggaggggg tcgtggagag ctaggagcaa agcctgttca
     2161 tggcagatgt aggagggact gtcgctgctt cgtgggatac agtcttctta cttggaactg
     2221 aagggggcgg cctatgactt gggcaccccc agcctgggcc tatggagagc cctgggaccg
     2281 ctacaccact ctggcagcca cacttctcag gacacaggcc tgtgtagctg tgacctgctg
     2341 agctctgaga ggccctggat cagcgtggcc ttgttctgtc accaatgtac ccaccgggcc
     2401 actccttcct gccccaactc cttccagcta gtgacccaca tgccatttgt actgacccca
     2461 tcacctactc acacaggcat ttcctgggtg gctactctgt gccagagcct ggggctctaa
     2521 ctgcctgagc ccagggaggc cgaagctaac agggaaggca ggcagggctc tcctggtctt
     2581 cccatcccca gcgattccct ctcccaggcc ccatgacctc cagctttcct gtatttcttc
     2641 ccaagagcat gatgcctctg aggccagcct ggcctcctgc ctctactggg aaggctactt
```

```
2701 cggggctggg aagtcgtcct tactcctgtg ggagcctcgc aacccgtgcc aagtccaggt
2761 cctggtgggg cagctcctct gtctcgagcg ccctgcagac cctgcccttg tttggggcag
2821 gagtagctga gctcacaagg cagcaaggcc cgagcagctg agcagggccg gggaactggc
2881 caagctgagg tgcccaggag aagaaagagg tgaccccagg gcacaggagc tacctgtgtg
2941 gacaggacta acactcagaa gcctgggtgc ctggctggct gagggcagtt cgcagccacc
3001 ctgaggagtc tgaggtcctg agcactgcca ggagggacaa aggagcctgt gaacccagga
3061 caagcatggt cccacatccc tgggcctgct gctgagaacc tggccttcag tgtaccgcgt
3121 ctaccctggg attcaggaaa aggcctgggg tgacccggca ccccctgcag cttgtagcca
3181 gccggggcga gtggcacgtt tatttaactt ttagtaaagt caaggagaaa tgcggtgg
```

```
LOCUS       OCT1         3824 bp ss-mRNA            PRI       18-NOV-1993
DEFINITION  Human octamer binding transcription factor 1 (OTF1) mRNA, complete
            cds.
ACCESSION   L20433
KEYWORDS    octamer binding transcription factor 1.
SOURCE      Homo sapiens (library: lambda pSH4K) cDNA to mRNA.
REFERENCE   1  (bases 1 to 3824)
  AUTHORS   Bhargava,A.K., Zhen,L.I. and Weissman,S.M.
  TITLE     Differential expression of four members of the POU family of
            proteins in activated and TPA-treated Jurkat T-cells
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 93, 10260-10264 (1994)
COMMENT     Hoppe-Seyler et al. (J Virol 65: 5613-8) demonstrated that Oct-1
            can repress transcription of HPV-18.  Mutagenesis experiments
            suggest that this effect is mediated without direct DNA binding.
            Morris et al. determined that HPV-16 transcription is repressed
            by Oct-1 only in non-cervical cells.  Conversely, in cervical
            cells Oct-2 is expressed which transactivates HPV-16 transcription.
BASE COUNT      896 a   1005 c   1067 g    856 t
ORIGIN
        1 gcggggctag agctgtcgga gaagcgggac cgcgaggccg gcgcgcggcg ctctgcgcgg
       61 tcagagggag cgcctggcag cagcaggagc agcagcagca gcccgcggcg gggccgccgc
      121 cagccgccgc gaccgccgcg gctgcagcct ccgaagggag gccgggtgag ccggcgtacg
      181 cactttcccg cggactttcg gagtgtttgt ggatatacat gccaagccgc cacgatgatg
      241 tccatgaaca gcaagcagcc tcactttgcc atgcatccca ccctccctga gcacaagtac
      301 ccgtcgctgc actccagctc cgaggccatc cggcgggcct gcctgcccac gccgccgctg
      361 cagagcaacc tcttcgccag cctggacgag acgctgctgg cgcgggccga ggcgctggcg
      421 gccgtggaca tcgccgtgtc ccagggcaag agccatcctt tcaagccgga cgccacgtac
      481 cacacgatga acagcgtgcc gtgcacgtcc acttccacgg tgcctctggc gcaccaccac
      541 caccaccacc accaccacca ggcgctcgaa cccggcgatc tgctggacca catctcctcg
      601 ccgtcgctcg cgctcatggc cggcgcgggc ggcgcgggcg cggcggccgg cggcggcggc
      661 gcccacgacg gcccgggggg cggtggcggc ccgggcggcg gcggcggccc gggcggcggc
      721 ggccccgggg gaggcggcgg tggcggcccg gggggcggcg gcggcggccc gggcggcggg
      781 ctcctgggcg gctccgcgca ccctcacccg catatgcaca gcctgggcca cctgtcgcac
      841 cccgcggcgg cggccgccat gaacatgccg tccgggctgc cgcaccccgg gctggtggcg
      901 gcggcggcgc accacggcgc ggcagcggca gcggcggcgg cgtcggccgg gcaggtggca
      961 gcggcatcgg cggcggcggc cgtggtgggc gcagcgggcc tggcgtccat ctgcgactcg
     1021 gacacggacc cgcgcgagct cgaggcgttc gcggagcgct tcaagcagcg gcgcatcaag
     1081 ctgggcgtga cgcaggccga cgtgggctcg gcgctggcca acctcaagat cccgggcgtg
     1141 ggctcactca gccagagcac catctgcagg ttcgagtcgc tcacgctctc gcacaacaac
     1201 atgatcgcgc tcaagcccat cctgcaggcg tggctcgagg aggccgaggg cgcccagcgc
     1261 gagaaaatga acaagcctga gctcttcaac ggcggcgaga agaagcgcaa gcggacttcc
     1321 atcgccgcgc ccgagaagcg ctccctcgag gcctacttcg ccgtgcagcc ccggccctcg
     1381 tccgagaaga tcgccgccat cgccgagaaa ctggacctca aaaagaacgt ggtgcgggtg
     1441 tggttttgca accagagaca gaagcagaag cggatgaaat tctctgccac ttactgaggg
     1501 ggctgggagg tgtcgggcgg gacagaatgg ggagctgagg aggcattttt gggggggcttt
     1561 cctctgcttg cctcccctcg gatttggagt gtccgttatc ctgcctgcat ttggggagtc
     1621 ccttctcgct ctctttcctc cacccattct ctgattttcc tgcctttgct gtcccctagc
     1681 cttgaggact ggggtgctgg gtgtggggat tggagtatag ggtaggggag aagggggggga
     1741 gcattcgggg gagtggggag tggggggaag gaaagcggag acccgagcag gggttttaag
     1801 gagcaggatg gttctggggt ttgggtgggg ggagacgcgg gaagggtagg aaaatggact
     1861 gtttctgacc agagacactt acctaaatat cctggggacc aaggaactat gtacaaaaac
     1921 aaacctacca accaccaaaa actagacaaa taaagacaaa ctaaaacaaa acagaacaaa
     1981 agcaaaggaa aatgctttag aaattttaac tccggggagc cataatctgc aacttcattt
     2041 tcccccatag aagagaaaaa agagcaccac cattattacc acctccccaa ccctacacgc
     2101 acgaactgag tcgaaaaacg aaaaccaaac gagcgagaag ttgaagttct gggtatcaaa
     2161 gctagttgtt ctgtctgcgt gtttaatttt tccctctctc acctccaccc catccatatc
     2221 ctctttattt cctccgttcc aatgagaggc ctatggctgc tctccaatcc cgggaagtga
     2281 gtgggagcac agctgaaaag agagggtcag ggggaggctg gctgcttgct taggtggaat
     2341 ccaacttttc ccgtggccct gcctatactc tggtggcctg gtcctgttgg ggtgggggtc
     2401 tttggagaga agggcatagt ctttgagcta ctaaaaagca gaattccgga gcttcgagat
     2461 atcttattct aggaaaatga aacaatttta acaacagttt tttttcctct tatgtcgaag
     2521 atctagtttt agacaatttc aaaataagct tttcccactc atagaacttt aacttgccct
     2581 ttcagtttta tcttttttttt agagagaggt ttaaactact gatttttcct gttgattcaa
     2641 atagactaat ggggtgaaag ttattaggag agatactctc tcctgttttc tccactgaac
```

Oct1

```
2701 gagactcatc ttgctcttct aggtcccgtt tcttcctctc ttggaggaca tgaaattata
2761 gaaatgttga gaagttcctg ctttcttttg cggtaggact tggctgtgag aaaatcacct
2821 aaatcccaga aaagaggaag acagatttaa agtgccccca cccccatttg tttcaaagag
2881 gtctgcatgt tgggcgaaaa cagaacaact gtgtttcctt ttacttgttc ttattattca
2941 agagtcattt attacagggg ataaatgttg ggtagcaaga actttaattt gcactaccag
3001 tctcccaaat agaaaatcat gtatagtatt tcatagtaat aatcaggtac cttacaagct
3061 gctggtggat tttaaaaaat taagatagtt gaaggtggtt aggtaaaatg cctgctttgt
3121 gtacaagata ctctttggat ctctcgtaga gatggtttgt taccatcctt taatcataac
3181 taaaacattg aaaacagaac aaatgagaaa agaaaaaaaa cctgccgatt aacaagactg
3241 aaatcatgca tgatctgaaa ggtgtggaaa gaaacacaat taggtctcac tctggttagg
3301 cattatttat ttaattatgt tgtatatcat tgtttgcagg gcaaacattc tatgcatttg
3361 aaactgagca ctaaactggg ctagctttct ggtagaccgt tttgtggcta gtgcgatttc
3421 acagtctact gcctgtttcc actgaaaaca tttttgtcat attcttgtat tcaaagaaaa
3481 caggaaaaaa gttattgtaa atattttatt taatgcacac attcacacag tggtaacaga
3541 ctgccagtgt tcatcctgaa atgtctcacg gattgatcta cctgtctatg tatgtctgct
3601 gagctttctc cttggttatg tttttctct tttacctttc tcctccctta cttctatcag
3661 aaccaattct atgcgccaaa tacaacaggg ggatgtgtcc cagtacactt acaaaataaa
3721 acataactga aagaagagca gttttatgat ttgggtgcgt ttttgtgttt atactgggcc
3781 aggtcctggt agaacctttc aacaaacaac caaacaaaaa aaaa
```

```
LOCUS       PDGFR        5570 bp ss-mRNA            PRI       28-SEP-1992
DEFINITION  Human platelet-derived growth factor (PDGF) receptor mRNA, complete
            cds.
ACCESSION   J03278
KEYWORDS    cell surface glycoprotein; glycoprotein; kinase; tyrosine kinase.
SOURCE      Human skin fibroblast cell (SK5), cDNA to mRNA, clone phPDGFR.
REFERENCE   1  (bases 1 to 5570)
  AUTHORS   Gronwald,R.G.K., Grant,F.J., Haldeman,B.A., Hart,C.E., O'Hara,P.J.,
            Hagen,F.S., Ross,R., Bowen-Pope,D.F. and Murray,M.J.
  TITLE     Cloning and expression of a cDNA coding for the human
            platelet-derived growth factor receptor: evidence for more than one
            receptor class
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 85, 3435-3439 (1988)
COMMENT     Nilson and DiMaio (Mol Cell Biol 13: 4137-4145, Cohen et al. J Virol
            67:5303-11) reported that the BPV-1 E5 protein activates PDGF beta
            receptors through complex formation resulting in fibroblast
            transformation. Cohen et al. demonstrated that this activation is
            specific to the transmembrane domain of the PDGF receptor and that
            it is ligand independent (Cohen et al. J Virol 67: 5303-11).

            Draft entry and computer-readable sequence [1] kindly submitted by
            R.G.K.Gronwald, 06-APR-1988.
BASE COUNT     1195 a   1676 c   1530 g   1169 t
ORIGIN      Chromosome 5q31-q32.
        1 ggcccctcag ccctgctgcc cagcacgagc ctgtgctcgc cctgcccaac gcagacagcc
       61 agacccaggg cggcccctct ggcggctctg ctcctcccga aggatgcttg gggagtgagg
      121 cgaagctggg cgctcctctc ccctacagca gccccttcc tccatccctc tgttctcctg
      181 agccttcagg agcctgcacc agtcctgcct gtccttctac tcagctgtta cccactctgg
      241 gaccagcagt ctttctgata actgggagag ggcagtaagg aggacttcct ggaggggtg
      301 actgtccaga gcctggaact gtgcccacac cagaagccat cagcagcaag gacaccatgc
      361 ggcttccggg tgcgatgcca gctctggccc tcaaaggcga gctgctgttg ctgtctctcc
      421 tgttacttct ggaaccacag atctctcagg gcctggtcgt cacaccccg gggccagagc
      481 ttgtcctcaa tgtctccagc accttcgttc tgacctgctc gggttcagct ccggtggtgt
      541 gggaacggat gtcccaggag cccccacagg aaatggccaa ggcccaggat ggcaccttct
      601 ccagcgtgct cacactgacc aacctcactg ggctagacac gggagaatac ttttgcaccc
      661 acaatgactc ccgtggactg gagaccgatg agcggaaacg gctctacatc tttgtgccag
      721 atcccaccgt gggcttcctc cctaatgatg ccgaggaact attcatcttt ctcacggaaa
      781 taactgagat caccattcca tgccgagtaa cagacccaca gctggtggtg acactgcacg
      841 agaagaaagg ggacgttgca ctgcctgtcc cctatgatca ccaacgtggc ttttctggta
      901 tctttgagga cagaagctac atctgcaaaa ccaccattgg ggacagggag gtggattctg
      961 atgcctacta tgtctacaga ctccaggtgt catccatcaa cgtctctgtg aacgcagtgc
     1021 agactgtggt ccgccagggt gagaacatca ccctcatgtg cattgtgatc gggaatgagg
     1081 tggtcaactt cgagtggaca tacccccgca aagaaagtgg gcggctggtg gagccggtga
     1141 ctgacttcct cttggatatg ccttaccaca tccgctccat cctgcacatc cccagtgccg
     1201 agttagaaga ctcggggacc tacacctgca atgtgacgga gagtgtgaat gaccatcagg
     1261 atgaaaaggc catcaacatc accgtggttg agagcggcta cgtgcggctc ctgggagagg
     1321 tgggcacact acaatttgct gagctgcatc ggagccggac actgcaggta gtgttcgagg
     1381 cctacccacc gcccactgtc ctgtggttca aagacaaccg caccctgggc gactccagcg
     1441 ctggcgaaat cgccctgtcc acgcgcaacg tgtcggagac ccggtatgtg tcagagctga
     1501 cactggttcg cgtgaaggtg gcagaggctg gccactacac catgcgggcc ttccatgagg
     1561 atgctgaggt ccagctctcc ttccagctac agatcaatgt ccctgtccga gtgctggagc
     1621 taagtgagag ccaccctgac agtggggaac agacagtccg ctgtcgtggc cggggcatgc
     1681 cccagccgaa catcatctgg tctgcctgca gagacctcaa aaggtgtcca cgtgagctgc
     1741 cgcccacgct gctggggaac agttccgaag aggagagcca gctggagact aacgtgacgt
     1801 actgggagga ggagcaggag tttgaggtgg tgagcacact gcgtctgcag cacgtggatc
     1861 ggccactgtc ggtgcgctgc acgctgcgca acgctgtggg ccaggacacg caggaggtca
     1921 tcgtggtgcc acactccttg ccctttaagg tggtggtgat ctcagccatc ctggccctgg
     1981 tggtgctcac catcatctcc cttatcatcc tcatcatgct ttggcagaag aagccacgtt
     2041 acgagatccg atggaaggtg attgagtctg tgagctctga cggccatgag tacatctacg
     2101 tggaccccat gcagctgccc tatgactcca cgtgggagct gccgcgggac cagcttgtgc
     2161 tgggacgcac cctcggctct ggggcctttg ggcaggtggt ggaggccacg gctcatggcc
     2221 tgagccattc tcaggccacg atgaaagtgg ccgtcaagat gcttaaatcc acagcccgca
     2281 gcagtgagaa gcaagccctt atgtcggagc tgaagatcat gagtcacctt gggccccacc
     2341 tgaacgtggt caacctgttg gggcctgca ccaaaggagg acccatctat atcatcactg
     2401 agtactgccg ctacggagac ctggtggact acctgcaccg caacaaacac accttcctgc
```

```
2461 agcaccactc cgacaagcgc cgcccgccca gcgcggagct ctacagcaat gctctgcccg
2521 ttgggctccc cctgcccagc catgtgtcct tgaccgggga gagcgacggt ggctacatgg
2581 acatgagcaa ggacgagtcg gtggactatg tgcccatgct ggacatgaaa ggagacgtca
2641 aatatgcaga catcgagtcc tccaactaca tggcccctta cgataactac gttccctctg
2701 cccctgagag gacctgccga gcaactttga tcaacgagtc tccagtgcta agctacatgg
2761 acctcgtggg cttcagctac caggtggcca atggcatgga gtttctggcc tccaagaact
2821 gcgtccacag agacctggcg gctaggaacg tgctcatctg tgaaggcaag ctggtcaaga
2881 tctgtgactt tggcctggct cgagacatca tgcgggactc gaattacatc tccaaaggca
2941 gcaccttttt gcctttaaag tggatggctc cggagagcat cttcaacagc ctctacacca
3001 ccctgagcga cgtgtggtcc ttcgggatcc tgctctggga gatcttcacc ttgggtggca
3061 ccccttaccc agagctgccc atgaacgagc agttctacaa tgccatcaaa cggggttacc
3121 gcatggccca gcctgcccat gcctccgacg agatctatga gatcatgcag aagtgctggg
3181 aagagaagtt tgagattcgg cccccttct cccagctggt gctgcttctc gagagactgt
3241 tgggcgaagg ttacaaaaag aagtaccagc aggtggatga ggagtttctg aggagtgacc
3301 acccagccat ccttcggtcc caggcccgct tgcctgggtt ccatggcctc cgatctcccc
3361 tggacaccag ctccgtcctc tatactgccg tgcagcccaa tgagggtgac aacgactata
3421 tcatccccct gcctgacccc aaacccgagg ttgctgacga gggcccactg gagggttccc
3481 ccagcctagc cagctccacc ctgaatgaag tcaacacctc ctcaaccatc tcctgtgaca
3541 gcccccctgga gccccaggac gaaccagagc cagagcccca gcttgagctc caggtggagc
3601 cggagccaga gctggaacag ttgccggatt cggggtgccc tgcgcctcgg gcggaagcag
3661 aggatagctt cctgtagggg gctggcccct accctgccct gcctgaagct cccccccctgc
3721 cagcacccag catctcctgg cctggcctga ccgggcttcc tgtcagccag gctgccctta
3781 tcagctgtcc ccttctggaa gctttctgct cctgacgtgt tgtgccccaa accctggggc
3841 tggcttagga ggcaagaaaa ctgcaggggc cgtgaccagc cctctgcctc cagggaggcc
3901 aactgactct gagccagggt tcccccaggg aactcagttt tcccatatgt aagatgggaa
3961 agttaggctt gatgacccag aatctaggat tctctccctg gctgacaggt ggggagaccg
4021 aatccctccc tgggaagatt cttggagtta ctgaggtggt aaattaactt ttttctgttc
4081 agccagctac ccctcaagga atcatagctc tctcctcgca ctttttatcc acccaggagc
4141 tagggaagag accctagcct ccctggctgc tggctgagct agggcctagc cttgagcagt
4201 gttgcctcat ccagaagaaa gccagtctcc tccctatgat gccagtccct gcgttccctg
4261 gcccgagctg gtctggggcc attaggcagc ctaattaatg ctggaggctg agccaagtac
4321 aggacacccc cagcctgcag cccttgccca gggcacttgg agcacacgca gccatagcaa
4381 gtgcctgtgt ccctgtcctt caggcccatc agtcctgggg ctttttcttt atcaccctca
4441 gtcttaatcc atccaccaga gtctagaagg ccagacgggc cccgcatctg tgatgagaat
4501 gtaaatgtgc cagtgtggag tggccacgtg tgtgtgccag tatatggccc tggctctgca
4561 ttggacctgc tatgaggctt tggaggaatc cctcaccctc tctgggcctc agtttcccct
4621 tcaaaaaatg aataagtcgg acttattaac tctgagtgcc ttgccagcac taacattcta
4681 gagtattcca ggtggttgca catttgtcca gatgaagcaa ggccatatac cctaaacttc
4741 catcctgggg gtcagctggg ctcctgggag attccagatc acacatcaca ctctggggac
4801 tcaggaacca tgccccttcc ccaggccccc agcaagtctc aagaacacag ctgcacaggc
4861 cttgacttag agtgacagcc ggtgtcctgg aaagccccaa gcagctgccc cagggacatg
4921 ggaagaccac gggacctctt tcactaccca cgatgacctc cgggggtatc ctgggcaaaa
4981 gggacaaaga gggcaaatga gatcacctcc tgcagcccac cactccagca cctgtgccga
5041 ggtctgcgtc gaagacagaa tggacagtga ggacagttat gtcttgtaaa agacaagaag
5101 cttcagatgg taccccaaga aggatgtgag aggtggccgc ttggagtttg cccctcaccc
5161 accagctgcc ccatccctga ggcagcgctc catgggggta tggttttgtc actgcccaga
5221 cctagcagtg acatctcatt gtccccagcc cagtgggcat tggaggtgcc aggggagtca
5281 gggttgtagc caagacgccc ccgcacgggg agggttggga agggggtgca ggaagctcaa
5341 cccctctggg caccaaccct gcattgcagg ttggcacctt acttccctgg gatccccaga
5401 gttggtccaa ggagggagag tgggttctca atacggtacc aaagatataa tcacctaggt
5461 ttacaaatat ttttaggact cacgttaact cacatttata cagcagaaat gctattttgt
5521 atgctgttaa gtttttctat ctgtgtactt ttttttaagg gaaagatttt
```

```
LOCUS       PP2A72       2338 bp ss-mRNA            PRI       20-AUG-1993
DEFINITION  Homo sapiens protein phosphatase 2A 72 kDa regulatory subunit mRNA,
            complete cds.
ACCESSION   L12146
KEYWORDS    phosphoprotein phosphatase 2A; phosphoprotein phosphohydrolase;
            protein phosphatase 2A 72 kDa regulatory subunit;
            regulatory subunit.
SOURCE      Homo sapiens (library: lambda ZAP; Stratagene) female heart muscle
            cDNA to mRNA.
REFERENCE   1  (bases 1 to 2338)
  AUTHORS   Hendrix,P., Mayer-Jaekel,R.E., Cron,P., Goris,J., Hofsteenge,J.,
            Merlevede,W. and Hemmings,B.A.
  TITLE     Structure and expression of a 72-kDa regulatory subunit of protein
            phosphatase 2A. Evidence for different size forms produced by
            alternative splicing
  JOURNAL   J. Biol. Chem. 268, 15267-15276 (1993)
COMMENT     PP2A consists of regulatory subunits and one catalytic subunit.
            In their review, "Protein Phosphatase 2A and and the Regulation
            of Human Papillomavirus Gene Activity, Schegget and Noordaa
            (in Human Papillomaviruses, 121-9) report that PP2A is
            inactivated by high expression of its regulatory subunit,
            PR55Beta.  This inhibition is linked to the activation of the LCR
            region of HPV-16.  Thus, the authors suggest that the regulation
            of the HPV-16 promoter is correlated to celluar protein
            phosphorylation.
BASE COUNT      694 a    497 c    526 g    621 t
ORIGIN
        1 gaattccgca gctgtgttta taacttgaaa attatgagca ggtattggga gccacagact
       61 gtatttggga acagctgcat gtcaggacaa gcatagattg tggtctgtgg aggctgggag
      121 agcacaatgc ttgaatttca tttctttggg agataagaag aaaaaatacc ttcctacctg
      181 acaagattga caacttgcta aatctgatta tcttcagaag tgtgctggac agttgcgaag
      241 taaacattta cttttcaaag atgatgatca aggaaacatc tctacgaagg gacccggatt
      301 taaggggaga gctagctttc ctggcaaggg gctgtgattt tgttctccct tcacggttta
      361 agaagcggct gaagtcattt cagcagacac agattcaaaa taaaccagaa aagaaacctg
      421 gaacaccact cccacctcca gccacctctc caagtagtcc ccgacctctc tccccggttc
      481 cccatgtgaa taatgttgtg aatgcgccat tgtccataaa cattccacgg ttctactttc
      541 ctgaaggact cccagatacc tgtagtaatc atgaacaaac tctaagcaga attgaaactg
      601 ctttcatgga tattgaagaa cagaaagcag acatttatga aatggggaaa attgcaaagg
      661 tctgtggctg tcctctctat tggaaagccc ccatgttcag ggctgcaggg ggagagaaga
      721 caggatttgt gacagcacag tcattcattg ccatgtggag aaagttgctg aataaccatc
      781 atgatgatgc ctctaaattc atctgtcttc tagcaaagcc caactgcagc tctctagaac
      841 aggaggattt catccctcta cttcaggatg tggtggatac ccaccctggt ctcacgttcc
      901 tgaaagatgc tccagaattc cactcccgct acatcaccac ggttattcag agaatattct
      961 acacagtcaa cagatcttgg agtggaaaaa ttacttcgac agagataaga aaaagcaact
     1021 ttttgcaaac cctagcactt ttggaagaag aggaagatat aaaccaaatt acagattact
     1081 tctcctatga acatttctat gttatttatt gtaaattctg ggaactagat actgatcacg
     1141 acctctacat cagccaggcc gatctgtctc gatacaatga ccaggcttca tcaagcagga
     1201 ttattgaaag gatattctct ggtgcagtaa caaggggaaa aacaatacag aaagagggaa
     1261 gaatgagcta tgcagatttt gtttggtttt tgatctctga agaagacaaa aggaatccta
     1321 ccagcattga gtattggttc cgctgcatgg atgtggatgg agacggtgta ctctccatgt
     1381 atgagctgga gtacttctat gaggagcagt gtgaacggat ggaagccatg ggaattgagc
     1441 ccttgccatt ccatgattta ctgtgccaga tgcttgacct agtgaagcca gctgttgatg
     1501 gcaaaataac tctaagagat ctgaagaggt gcagaatggc tcacatcttc tatgacactt
     1561 tctttaatct ggagaaatac ttagaccatg aacagagaga tccctttgcg gtccagaagg
     1621 atgttgagaa cgatgggcct gagccctcag actgggaccg gtttgccgct gaggagtatg
     1681 agacgcttgt tgcagaggaa tctgcccaag cacaattcca ggaaggcttt gaagattatg
     1741 aaacagatga acctgcctct ccctctgaat ttggaaacaa aagcaataaa atattaagtg
     1801 caagccttcc agagaaatgt ggaaagcttc aatcagtgga tgaagaatag ctgccggtgt
     1861 ctacaatgaa acgaagatgt gtattttaaa tgtttctttc ttgtgaagag atgttctcgt
     1921 ttgcatactg ctttttaaag actttgattt ctccaagtgt gtatcatctg cactaggaac
     1981 tttgttttta agcaataggt ctggatacac atttaactta ggaggctcct ccaatttgcc
     2041 tcaaacctct tacggagctt ctcctcagaa gtggtaccat cgccttccaa agtcagcact
     2101 ctacactctt gaatgtacca aggatctctt ggcgacagta ccaagcacgg ttctctacac
     2161 aggtgactga agttgcctct gtgttggctg gcatccctga gtcccctccg gctcctatgg
```

PP2A 72kDa Regulatory Subunit

```
2221 agcctagaag aaaccttcac ttgcagaaaa cttgagtcag aaaattctgg aacttgaaaa
2281 agtagtaagg gcctccagaa ttgacttagc cctagtaata aaagcactgc cggaattc
```

```
LOCUS       PP2ACAT      2966 bp ds-DNA             PRI       04-MAR-1991
DEFINITION  Human protein phosphatase 2A catalytic subunit-alpha gene, complete
            cds.
ACCESSION   M60483 J05297
KEYWORDS    protein phosphatase-2A catalytic subunit-alpha.
SOURCE      Human placenta leukocyte DNA.
REFERENCE   1  (bases 1 to 2966)
  AUTHORS   Khew-Goodall,Y., Mayer,R.E., Maurer,F., Stone,S.R. and
            Hemmings,B.A.
  TITLE     Structure and transcriptional regulation of protein phosphatase 2A
            catalytic subunit genes
  JOURNAL   Biochemistry 30, 89-97 (1991)
COMMENT     PP2A consists of regulatory subunits and one catalytic subunit.
            In their review, "Protein Phosphatase 2A and and the Regulation
            of Human Papillomavirus Gene Activity, Schegget and Noordaa
            (in Human Papillomaviruses, 121-9) report that PP2A is
            inactivated by high expression of its regulatory subunit,
            PR55Beta.  This inhibition is linked to the activation of the LCR
            region of HPV-16.  Thus, the authors suggest that the regulation
            of the HPV-16 promoter is correlated to celluar protein
            phosphorylation.

            Although this is a genomic sequence, the introns have been omitted.
            Intronic sequences have been requested. The data is presented as
            submitted by the author.

            NCBI gi: 190223
BASE COUNT      747 a    728 c    732 g    759 t
ORIGIN
        1 aaccaccggc gaggagcggg gcgcgtggaa gcgagccgcg gtccgaggcc caaagaaaag
       61 cccaagcctc gcccccgcca tcgcgcccga cgagacacct aggtccgggg acgggtgtgt
      121 gccgcggaag tcaggtgcac tgcgcagcac tcccccggta ggtacacgct cctccaccta
      181 cgagtgacct aattacaagg tgccagccgc gcccagaggt gggggtggtt aatccaagcg
      241 gccactcgct gcccgttcct gcccccaaag atgacggaaa cccacacgat tacagagccg
      301 cagcacccca gatgagccac ggggtcgcaa ttctcgtttc cgtgatcgga ctgccaggcc
      361 ccaggtgagg agctgagttc atcaccagag cggccttccc aggggaacca gttacaggct
      421 gccagtggcc ccggcttcca tccggtctgc gcctgcgcgc ggcccaagcc ctcgcctctc
      481 ctggaatagt gctcagggat tagtccggtt cgccgctgtg ccactgcgca tgctccagct
      541 ccatccttcc cttcccccac caccccgccc tccgggagcc acgcccaaaa agtcaaggcg
      601 cttcagttac cagccggcta cgtggcctgc gctttgaccc ccagtttgcg ccccaactcc
      661 ggtcgtgcgg ccgcccgggg agggctctgc agttgcgcag cttgctcccc ggcccttttc
      721 ccctccgctc cccgccgcct cctgacgccg ggcgtgacgt caccacgccc ggcggccgcc
      781 attacagaga gccgagctct ggagcctcag cgagcggagg aggaggcgca gggccgacgg
      841 ccgagtactg cggtgagagc cagcgggcca gcgccagcct caacagccgc cagaagtaca
      901 cgaggaaccg gcggcggcgt gtgcgtgtag gcccgtgtgc gggcggcggc gcgggaggag
      961 cgcggagcgg cagccggctg gggcgggtgg catcatggac gagaaggtgt tcaccaagga
     1021 gctggaccag tggatcgagc agctgaacga gtgcaagcag ctgtccgagt cccaggtcaa
     1081 gagcctctgc gagaaggcta aagaaatcct gacaaaagaa tccaacgtgc aagaggttcg
     1141 atgtccagtt actgtctgtg gagatgtgca tgggcaattt catgatctca tggaactgtt
     1201 tagaattggt ggcaaatcac cagatacaaa ttacttgttt atgggagatt atgttgacag
     1261 aggatattat tcagttgaaa cagttacact gcttgtagct cttaaggttc gttaccgtga
     1321 acgcatcacc attcttcgag ggaatcatga gagcagacag atcacacaag tttatggttt
     1381 ctatgatgaa tgtttaagaa aatatggaaa tgcaaatgtt tggaaatatt ttacagatct
     1441 ttttgactat cttcctctca ctgccttggt ggatgggcag atcttctgtc tacatggtgg
     1501 tctctcgcca tctatagata cactggatca tatcagagca cttgatcgcc tacaagaagt
     1561 tccccatgag ggtccaatgt gtgacttgct gtggtcagat ccagatgacc gtggtggttg
     1621 gggtatatct cctcgaggag ctggttacac ctttgggcaa gatatttctg agacatttaa
     1681 tcatgccaat ggcctcacgt tggtgtctag agctcaccag ctagtgatgg agggatataa
     1741 ctggtgccat gaccggaatg tagtaacgat tttcagtgct ccaaactatt gttatcgttg
     1801 tggtaaccaa gctgcaatca tggaacttga cgatactcta aaatactctt tcttgcagtt
     1861 tgacccagca cctcgtagag gcgagccaca tgttactcgt cgtaccccag actacttcct
     1921 gtaatgaaat tttaaacttg tacagtattg ccatgaacca tatatcgacc taatggaaat
     1981 gggaagagca acagtaactc caaagtgtca gaaaatagtt aacattcaaa aaacttgttt
     2041 tcacatggac caaaagatgt gccatataaa aatacaaagc ctcttgtcat caacagccgt
     2101 gaccacttta gaatgaacca gttcattgca tgctgaagcg acattgttgg tcaagaaacc
```

```
2161 agtttctggc atagcgctat ttgtagttac ttttgctttc tctgagagac tgcagataat
2221 aagatgtaaa cattaacacc tcgtgaatac aatttaactt ccatttagct atagctttac
2281 tcagcatgac tgtagataag gatagcagca aacaatcatt ggagcttaat gaacattttt
2341 aaaaataatt accaaggcct cccttctact tgtgagtttt gaaattgttc tttttatttt
2401 cagggatacc gtttaattta attatatgat ttgtctgcac tcagtttatt ccctactcaa
2461 atctcagccc catgttgttc tttgttattg tcagaacctg gtgagttgtt ttgaacagaa
2521 ctgttttttc cccttcctgt aagacgatgt gactgcacaa gagcactgca gtgtttttca
2581 taataaactt gtgaactaag aactgagaag gtcaaatttt aattgtatca atgggcaaga
2641 ctggtgctgt ttattaaaaa agttaaatca attgagtaaa ttttagaatt tgtagacttg
2701 taggtaaaat aaaaatcaag ggcactacat aacctctctg gtaactcctt gacattcttc
2761 agattaactt caggatttat ttgtatttca catattacaa tttgtcacat tgttggtgtg
2821 cactttgtgg gttcttcctg catattaact tgtttgtaag aaaggaaatc tgtgctgctt
2881 cagtaagact taattgtaaa accatataac ttgagattta agtctttggg ttgtgtttta
2941 ataaaacagc atgttttcag gtagag
```

```
LOCUS       PP2APR55     2131 bp ss-mRNA            PRI       09-MAY-1991
DEFINITION  Human protein phosphatase 2A alpha subunit mRNA, complete cds.
ACCESSION   M64929 J05328
KEYWORDS    protein phosphatase-2A subunit-alpha; regulatory subunit.
SOURCE      Human lung fibroblast cell line WI38, cDNA to mRNA.
REFERENCE   1  (bases 1 to 2131)
  AUTHORS   Mayer,R.E., Hendrix,P., Cron,P., Matthies,R., Stone,S.R., Goris,J.,
            Merlevede,W., Hofsteenge,J. and Hemmings,B.A.
  TITLE     Structure of the 55 kDa regulatory subunit of protein phosphatase
            2A: Evidence for a neuronal specific isoform
  JOURNAL   Biochemistry 30, 3589-3597 (1991)
COMMENT     PP2A consists of regulatory subunits and one catalytic subunit.
            In their review, "Protein Phosphatase 2A and and the Regulation
            of Human Papillomavirus Gene Activity, Schegget and Noordaa
            (in Human Papillomaviruses, 121-9) report that PP2A is
            inactivated by high expression of its regulatory subunit,
            PR55Beta.  This inhibition is linked to the activation of the LCR
            region of HPV-16.  Thus, the authors suggest that the regulation
            of the HPV-16 promoter is correlated to celluar protein
            phosphorylation.

            NCBI gi: 190421
BASE COUNT      665 a    421 c    460 g    585 t
ORIGIN
        1 ccgccgccat ccgccctctc tacccccca tccccaggtg aggggggtga gttcaggaag
       61 cggagacccc gaggaaccca gcagggtcac catttgcagc gcaacatggc aggagctgga
      121 ggagggaatg atattcagtg gtgtttttct caggtgaaag gagcagtaga tgatgatgta
      181 gcagaagcag atataatttc tacagtagaa tttaatcatt ctggagaatt actagcaaca
      241 ggagataaag gtggtagagt tgtcatcttt caacaggagc aggagaacaa aatccagtct
      301 catagcagag gagaatataa tgtttacagc accttccaga gccatgaacc agagtttgac
      361 tacttgaaaa gtttagaaat agaagaaaag atcaataaaa ttaggtggtt accccagaaa
      421 aatgctgctc agtttttatt gtctaccaat gataaaacaa taaaattatg gaaaatcagt
      481 gaaagggaca aaagaccaga agggtataac ttgaaagagg aggatggaag gtatagagat
      541 cctactacag ttactacact acgagtgcca gtctttaggc ctatggatct aatggttgag
      601 gccagtccac gaagaatatt tgccaatgct catacatatc acatcaactc aatttctatt
      661 aatagtgatt atgaaacata tttatctgca gatgatttgc ggattaatct ttggcatctg
      721 gaaattacag acaggagttt taacattgtg gatatcaagc ctgccaatat ggaagagcta
      781 acagaggtga ttacagcagc agaatttcat ccaaacagct gtaacacatt tgtatacagc
      841 agcagtaaag gaactattcg gctatgtgac atgagggcat ctgccctctg tgatagacat
      901 tctaaattgt ttgaagaacc tgaagatccc agtaacaggt cattttttttc cgaaatcatc
      961 tcctctattt cggatgtaaa attcagccat agtggtcgat atatgatgac tagagactat
     1021 ttgtcagtca aaatttggga cttaaatatg gaaaacaggc ctgtggaaac ataccaggtg
     1081 catgaatacc tcagaagtaa actctgttca ctgtatgaaa atgactgcat atttgacaaa
     1141 tttgaatgtt gttggaatgg atctgacagt gttgtcatga ctggatctta caataatttc
     1201 ttcagaatgt ttgacagaaa cacaaagcga gacataaccc tagaagcatc gcgggaaaac
     1261 aataagcctc gcacagttct gaagcctcgc aaagtctgtg caagtggcaa gcgaaagaaa
     1321 gatgaaataa gtgttgacag cctagacttc aataagaaaa tccttcacac agcctggcac
     1381 cccaaggaaa atatcattgc cgtagctact acaaacaatc tgtatatatt tcaagacaaa
     1441 gtgaattagg gttggcattc ctagcagaag aacccacttc ctgcttagtt gagatagttg
     1501 aatctagcat tcgttcctat aaaagagaga ggtccattgt ggcgcccctt tccagtgttt
     1561 gacagtgtgc cattcgacaa cacattgtta tagctacatg gagaaagctc tgtggattca
     1621 tcactgtggt gttctccatg tctgctagcc atttaggtaa gggtagggca cttttaattt
     1681 aaatgacttc ttgcaccatc ttgcctaatg gactagattg gactgtatca acattgattt
     1741 actccacttt ttatgccttc cattgtgatg acgtcaaaca cagtgaaagc cttcagtcat
     1801 gctatgggat ttaattgtgt atcctcatta ctgtatcatt tgtggggtac accccttccc
     1861 cctttttttta aattaaatac agctcattct tactgtggct tgtagcattc ctcctcttct
     1921 ggcctcctgg actgctcccc ttcatctctt acccttgccc cctccacccg gtcttggtgg
     1981 tggtatatta aaaaaagaaa gaatgaaagc acacaaaatg agtcagtttg gggtcagtgg
     2041 tataaagggg gtatatgttg caaacaaatg ttttagtaac agttggctgt aatcactcct
     2101 cgccgtgtct ggcactgaaa ataaggaaaa g
```

PP2A PR65 Regulatory Subunit

```
LOCUS       PP2APR65     2205 bp ss-mRNA            PRI       22-OCT-1992
DEFINITION  Human protein phosphatase 2A regulatory subunit alpha-isotype
            (alpha-PR65) mRNA, complete cds.
ACCESSION   J02902
KEYWORDS    protein phosphatase-2A regulatory alpha-subunit.
SOURCE      Human HeLa cell, cDNA to mRNA, clone lambda-HHPR65-3.
REFERENCE   1  (bases 1 to 2205)
  AUTHORS   Hemmings,B.A., Adams-Pearson,C., Maurer,F., Mueller,P., Goris,J.,
            Merlevede,W., Hofsteenge,J. and Stone,S.R.
  TITLE     Alpha and beta-forms of the 65-kDa subunit of protein phosphatase
            2A have a similar 39 amino acid repeating structure
  JOURNAL   Biochemistry 29, 3166-3173 (1990)
COMMENT     PP2A consists of regulatory subunits and one catalytic subunit.
            In their review, "Protein Phosphatase 2A and and the Regulation
            of Human Papillomavirus Gene Activity, Schegget and Noordaa
            (in Human Papillomaviruses, 121-9) report that PP2A is
            inactivated by high expression of its regulatory subunit,
            PR55Beta.  This inhibition is linked to the activation of the LCR
            region of HPV-16.  Thus, the authors suggest that the regulation
            of the HPV-16 promoter is correlated to celluar protein
            phosphorylation.

            Draft entry and printed sequence [1] kindly submitted by
            B.A.Hemmings, 23-MAR-1990, for release after publication.

            NCBI gi: 189427
BASE COUNT      447 a    676 c    639 g    443 t
ORIGIN
        1 gaattccggt tctcactctt gacgttgtcc agctccagca ccttggcaac tcccccagct
       61 tggacggccg gcccgccgct ccatggggga gtcatctgag cacagctgct ggccgcagtc
      121 tgacaggaaa gggacggagc caagatggcg gcggccgacg gcgacgactc gctgtacccc
      181 atcgcggtgc tcatagacga actccgcaat gaggacgttc agcttcgcct caacagcatc
      241 aagaagctgt ccaccatcgc cttggccctt ggggttgaaa ggacccgaag tgagcttctg
      301 cctttcctta cagataccat ctatgatgaa gatgaggtcc tcctggccct ggcagaacag
      361 ctgggaacct tcactaccct ggtgggaggc ccagagtacg tgcactgcct gctgccaccg
      421 ctggagtcgc tggccacagt ggaggagaca gtggtgcggg acaaggcagt ggagtcctta
      481 cgggccatct cacacgagca ctcgccctct gacctggagg cgcactttgt gccgctagtg
      541 aagcggctgg cgggcggcga ctggttcacc tcccgcacct cggcctgcgg cctcttctcc
      601 gtctgctacc cccgagtgtc cagtgctgtg aaggcggaac ttcgacagta cttccggaac
      661 ctgtgctcag atgacacccc catggtgcgg cgggccgcag cctccaagct gggggagttt
      721 gccaaggtgc tggagctgga caacgtcaag agtgagatca tccccatgtt ctccaacctg
      781 gcctctgacg agcaggactc ggtgcggctg ctggcggtgg aggcgtgcgt gaacatcgcc
      841 cagcttctgc cccaggagga tctggaggcc ctggtgatgc ccactctgcg ccaggccgct
      901 gaagacaagt cctgggccgt ccgctacatg gtggctgaca agttcacaga gctccagaaa
      961 gcagtggggc ctgagatcac caagacagac ctggtccctg ccttccagaa cctgatgaaa
     1021 gactgtgagg ccgaggtgag ggccgcagcc tcccacaagg tcaaagagtt ctgtgaaaac
     1081 ctctcagctg actgtcggga gaatgtgatc atgtcccaga tcttgccctg catcaaggag
     1141 ctggtgtccg atgccaacca acatgtcaag tctgccctgg cctcagtcat catgggtctc
     1201 tctcccatct tgggcaaaga caacaccatc gagcacctct tgcccctctt cctggctcag
     1261 ctgaaggatg agtgccctga ggtacggctg aacatcatct ctaacctgga ctgtgtgaac
     1321 gaggtgattg gcatccggca gctgtcccag tccctgctcc ctgccattgt ggagctggct
     1381 gaggacgcca agtggcgggt gcggctggcc atcattgagt acatgcccct cctggctgga
     1441 cagctgggag tggagttctt tgatgagaaa cttaactcct tgtgcatggc ctggcttgtg
     1501 gatcatgtat atgccatccg cgaggcagcc accagcaacc tgaagaagct agtggaaaag
     1561 tttgggaagg agtgggccca tgccacaatc atccccaagg tcttggccat gtccggagac
     1621 cccaactacc tgcaccgcat gactacgctc ttctgcatca atgtgctgtc tgaggtctgt
     1681 gggcaggaca tcaccaccaa gcacatgcta cccacggttc tgcgcatggc tggggacccg
     1741 gttgccaatg tccgcttcaa tgtggccaag tctctgcaga agatagggcc catcctggac
     1801 aacagcacct tgcagagtga agtcaagccc atcctagaga agctgaccca ggaccaggat
     1861 gtggacgtca aatactttgc ccaggaggct ctgactgttc tgtctctcgc ctgatgctgg
     1921 aagaggagca aacactggcc tctggtgtcc accctccaac ccccacaagt ccctctttgg
     1981 ggagacactg gggggccttt ggctgtcact ccctgtgcat ggtctgaccc caggcccctt
     2041 cccccagcac ggttcctcct ctccccagcc tgggaagatg tctcactgtc cacctcccaa
     2101 cggctagggg agcacggggt tggacaggac agtgaccttg ggaggaaggg gctactccgc
     2161 catccttaaa agccatggag ccggaggtgg caattcaccg aattc
```

```
LOCUS       TEF          4443 bp ds-DNA            PRI       28-JUL-1992
DEFINITION  Transcriptional enhancer factor DNA, complete CDS.
ACCESSION   M63896
KEYWORDS    trans-acting transcriptional activator; transcription enhancer.
SOURCE      Homo sapiens (library: ZAP-II random primed cDNA) DNA.
REFERENCE   1  (bases 1 to 4443)
  AUTHORS   Xiao,J.-H., Davidson,I., Matthes,H., Garnier,J.-M. and Chambon,P.
  TITLE     Cloning, expression, and transcriptional properties of the human
            enhancer factor TEF-1
  JOURNAL   Cell 65, 551-568 (1991)
BASE COUNT     1226 a   1040 c    969 g   1208 t
ORIGIN
        1 cgccgcccgc cgcgggcgcc caccaagcac tttgcagact cgcttccacc ctgcgggcca
       61 ttccgcgcgg cggggcccgg gcccggggcg gccgcgtcca ggcacaggcc atgcagtgac
      121 gcccccccac ccctccacct ttgcccggac ggcgggcagc agcccagcgc gccagccggc
      181 cccggggcag gagcggtgct aggcaggggt ggggtggccg ggcccaggga ccgggagccg
      241 gggagggagc cgggcaccga gcagagggcg ggggaagcgg cgccgaagtt tgcctcggac
      301 tcgccgggcg ctgcggtggc tccctgggcc gaggactgtt gctgccgctg ccgccgccgc
      361 ttcattgcac attcaagtgg aaaattttca ggagtcagca gaaacattgt gtccaaaaaa
      421 gactgagtcg cagttaccac caaacccagg aggagactct ccctggaaaa cttcccttcc
      481 ctttcggttt attttcttga aaaggctcca ggcttcggct tggaaaatcc caccgccaaa
      541 attgagccca gcagctggag cggcagtgag agccctgccg aaaacatgga aaggatgagt
      601 gactctgcag ataagccaat tgacaatgat gcagaagggg tctggagccc cgacatcgag
      661 caaagctttc aggaggccct ggctatctat ccaccatgtg ggaggaggaa aatcatctta
      721 tcagacgaag gcaaaatgta tggtaggaat gaattgatag ccagatacat caaactcagg
      781 acaggcaaga cgaggaccag aaaacaggtg tctagtcaca ttcaggttct tgccagaagg
      841 aaatctcgtg attttcattc caagctaaag gatcagactg caaaggataa ggccctgcag
      901 cacatggcgg ccatgtcctc agcccagatc gtctcggcca ctgccattca taacaagctg
      961 gggctgcctg ggattccacg cccgaccttc ccaggggcgc cggggttctg gccgggaatg
     1021 attcaaacag ggcagccagg atcctcacaa gacgtcaagc cttttgtgca gcaggcctac
     1081 cccatccagc cagcggtcac agcccccatt ccagggtttg agcctgcatc ggccccagct
     1141 ccctcagtcc ctgcctggca aggtcgctcc attggcacaa ccaagcttcg cctggtggaa
     1201 ttttcagctt ttctcgagca gcagcgagac ccagactcgt acaacaaaca cctcttcgtg
     1261 cacattgggc atgccaacca ttcttacagt gacccattgc ttgaatcagt ggacattcgt
     1321 cagatttatg acaaatttcc tgaaaagaaa ggtggcttaa aggaactgtt tggaaagggc
     1381 cctcaaaatg ccttcttcct cgtaaaattc tgggctgatt taaactgcaa tattcaagat
     1441 gatgctgggg ctttttatgg tgtaaccagt cagtacgaga gttctgaaaa tatgacagtc
     1501 acctgttcca ccaaagtttg ctcctttggg aagcaagtag tagaaaaagt agagacggag
     1561 tatgcaaggt ttgagaatgg ccgatttgta taccgaataa accgctcccc aatgtgtgaa
     1621 tatatgatca acttcatcca caagctcaaa cacttaccag agaaatatat gatgaacagt
     1681 gttttggaaa acttcacaat tttattggtg gtaacaaaca gggatacaca agaaactcta
     1741 ctctgcatgg cctgtgtgtt tgaagtttca aatagtgaac acggagcaca acatcatatt
     1801 tacaggcttg taaaggactg aacatggtta tttatatata tagatatctg tatatacaca
     1861 cacacatatg tgcacacaca cactctctct ccattatcga acgactgact gtaaacctca
     1921 ccacacaggg tggtgccctg gccccgaggt caccccgact tttctaaatc ttgtttgagt
     1981 gaagtcattt tttcatgtgt tcatactatc attgtagctg tgaagttctg gtacagttgt
     2041 aaaaagagaa attgagttgt ttctctatgt tcttcagatg tgcagcccac aattcctcgg
     2101 gaaaggtgaa cctgaacaac ccaagtctct ctctgcagag ccctgtttct aattgtggta
     2161 gaaaatattg agacagagca tttgccatgg gacatttaca gcctttatac aaatgtattt
     2221 agttctcttt tttccaacat aaaattcttg ttttaagata caagtaaaat taatctttaa
     2281 atataaatgt aaattagtac acaaaactaa gaatctttag acttatcttt gtaactaatt
     2341 agggtggaag ttatgaaaga atgtaattca ctaaattatt ttttaaatga aaccttttttt
     2401 tttctttttg aaaccaaatg ttaaactata gccttaagaa atgcttggta gaagtgtcct
     2461 aatgagacaa atttgtactt ttatcctcaa ggttaacact aatctcctaa tccattaaac
     2521 tcttgaacag gtattacaaa ggaagaaaac ttcacccctt atccttaaca tatatagtat
     2581 atttaaaaaa tataaaattg tattgtacta atgtgatgat ggattattta atgaaaaaga
     2641 aaaaatggct ctttttgcaa taagtagata catactgaaa aaatctaaac ttacaatgtt
     2701 tatagtcttg tgtgtgcagt tatattttat atggacgacc aaattttttta ttaagatgag
     2761 taaatatttg aaccactgaa ttttaataac aaaattttaa aattggcatg aatacggaat
     2821 actgcactgt gagatgcaaa gtatacagaa tctgtggctg ggagaaaatt tcatcaaata
     2881 gacaagtaaa aggctcatca gttttagcat ctctgctccc cagaaaattg taagcatcct
     2941 caccagcctg tggatacatt ctttatttct agtgacccaa tatgcatatt aacctgctat
     3001 aactagggct atatgtgtag gtatgtgtat acatatacac aaatgcacat atagagttaa
     3061 cacatttagt gaacacttgt ttagtgtcac tcagtttgct aggtgctgat atgtacgtat
```

```
3121 atctcaatgt gtctgtagac ttagatacat cctcttgaag cacatccatt tctttagcgt
3181 ctctcagtaa gttacagtac ttgtttgact taggtttaag aggcccagct acctatctct
3241 gaccttttca aataggctca tttgggagat tcttttgcca ggagagattc aactttccaa
3301 tctaagtatt ccagagcatt gcccaggcag agttggtttg atgtggccag atgttttgag
3361 ttatttccct taagtgtttc actggggaga gaacagggag tgctcctcca gcttcccaaa
3421 gaaatatgtt tttgtaagtg gtaggaacat gtgcacacaa tagaacatga aataagtttt
3481 ttaacttgta aaacatgtca agatttttcc accaagctag aaaataaaaa acttagttct
3541 accacatcca attaacttac acaccccctt ccctgtctca acacctgctt tgaccctgct
3601 tttctattat tacatcagtc agcatcttgt ggtccctaac atgaggatgt ggctggctcg
3661 tgggaaacag caaaacacta agcctgacct ctcccaaatt gggaagacca gaggagaaag
3721 tgcaaaactg tccccatttg gaatgcccat tccttctaga aaccagttgg acagtgctcc
3781 tctgcccttc ataaacagac tactgttggg tccctgattc caggctggcc tgtgaaggat
3841 tgccccaggt gtcccctttc acggttgtca catttacagt gacttctgtt gaacacccct
3901 cttagggatg tttcttttgc tcttatttcc tgcatctttc caattgggaa gccccatcct
3961 ctcccaggac caggagttta tgaccaggcg agcacaaatg gctaaaagca agctgtccta
4021 gaacttcagt gggagagctg tctggttcat attctaccca ggaatggtac ttttcagtgc
4081 agccaggagg gctcttggga tttcctttcc aaagcacaaa aatactggga cccaagaaga
4141 acagctagag gacaactctg ttggcacaga gacggggaca gcccagtctg ctgacctcac
4201 agggtcagtg ggcccccctg gtgcttcacc acctgcatcc tcttgctcag aatgcctttg
4261 cagttgagtt ttctgggttt ctatgattga ccttgaggtt tactccttgc tcttacaaca
4321 tttctaagga tttttaaaag tttacttctt gtcttgttct tctaaagctt tctccaggac
4381 agatattttc cctgtcttaa ccactggtcc agtcatccca gtgggcttct ctttgtctct
4441 ccc
```

```
LOCUS       TGFA          867 bp ss-mRNA           PRI       19-APR-1994
DEFINITION  Human (cell line 1027 F57) transforming growth factor-alpha mRNA,
            complete cds.
ACCESSION   K03222
KEYWORDS    growth factor; transforming gene; transforming growth factor-alpha.
SOURCE      Human renal carcinoma cell line 1027 F57, cDNA to mRNA, clone
            pTGF-Cl.
REFERENCE   1  (bases 1 to 867)
  AUTHORS   Derynck,R., Roberts,A.B., Winkler,M.E., Chen,E.Y. and Goeddel,D.V.
  TITLE     Human transforming growth factor-alpha: Precursor structure and
            expression in E. coli
  JOURNAL   Cell 38, 287-297 (1984)
COMMENT     Pietenpol et al. (Cell 6: 777-785) report that HPV-16 and HPV-18
            infection inhibits growth suppresion by TGF-B1 in
            human foreskin keratinocytes.  The authors suggest that the inhibitory
            capability of TGF-B1 is linked to the down-regulation of c-myc and
            that HPV-16 E7 expression interferes with this pathway.

            [1] also sequenced exons 1 and 2 of this gene from the genomic DNA
            library of Lawn et al. (see separate entries).
BASE COUNT      175 a    240 c    254 g    198 t
ORIGIN      Chromosome 2p13
        1 aggctggaga gcctgctgcc cgcccgcccg taaaatggtc ccctcggctg gacagctcgc
       61 cctgttcgct ctgggtattg tgttggctgc gtgccaggcc ttggagaaca gcacgtcccc
      121 gctgagtgca gacccgcccg tggctgcagc agtggtgtcc cattttaatg actgcccaga
      181 ttcccacact cagttctgct tccatggaac ctgcaggttt ttggtgcagg aggacaagcc
      241 agcatgtgtc tgccattctg ggtacgttgg tgcacgctgt gagcatgcgg acctcctggc
      301 cgtggtggct gccagccaga agaagcaggc catcaccgcc ttggtggtgg tctccatcgt
      361 ggccctggct gtccttatca tcacatgtgt gctgatacac tgctgccagg tccgaaaaca
      421 ctgtgagtgg tgccgggccc tcatctgccg gcacgagaag cccagcgccc tcctgaaggg
      481 aagaaccgct tgctgccact cagaaacagt ggtctgaaga gcccagagga ggagtttggc
      541 caggtggact gtggcagatc aataaagaaa ggcttcttca ggacagcact gccagagatg
      601 cctgggtgtg ccacagacct tcctacttgg cctgtaatca cctgtgcagc cttttgtggg
      661 ccttcaaaac tctgtcaaga actccgtcgg cttggggtta ttcagtgtga cctagagaag
      721 aaatcagcgg accacgattt caagacttgt taaaaaagaa ctgcaaagag acggactcct
      781 gttcacctag gtgaggtgtg tgcagcagtt ggtgtctgag tccacatgtg tgcagttgtc
      841 ttctgccagc catggattcc aggccgt
```

```
LOCUS       TNF          7112 bp ds-DNA             PRI       15-JUN-1989
DEFINITION  Human tumor necrosis factor and lymphotoxin genes, complete cds.
ACCESSION   M16441
KEYWORDS    lymphotoxin; tumor necrosis factor.
SOURCE      Human placenta DNA, clone pTNF186.
REFERENCE   1  (bases 1 to 7112)
  AUTHORS   Nedospasov,S.A., Shakhov,A.N., Turetskaya,R.L., Mett,V.A.,
            Azizov,M.M., Georgiev,G.P., Korobko,V.G., Dobrynin,V.N.,
            Filippov,S.A., Bystrov,N.S., Boldyreva,E.F., Chuvpilo,S.A.,
            Chumakov,A.M., Shingarova,L.N. and Ovchinnikov,Y.A.
  TITLE     Tandem arrangement of genes coding for tumor necrosis factor
            (TNF-alpha) and lymphotoxin (TNF-beta) in the human genome
  JOURNAL   Cold Spring Harb. Symp. Quant. Biol. 51, 611-624 (1986)
COMMENT     Kyo et al. (Virology 200: 130-9) demonstrated that tumor necrosis
            factor (TNF) alpha represses transcription of the early HPV-16
            genes through responsive elements in the LCR. Malejczyk et al.
            (Int J Cancer 56: 593-8) reported this inhibition for weakly
            tumorigenic cell lines, but not for highly tumorigenic cell lines.
            Further Malejczyk et al. demonstrated that TNF-alpha receptor
            expression is reduced in the higly tumorigenic lines and that this
            reduction may inhibit the autocrine TNF-alpha-mediated growth
            limitation.

            Draft entry and computer-readable sequence for [1] kindly submitted
            by C.V.Jongeneel, 02-OCT-1988.
BASE COUNT     1676 a   2005 c   1865 g   1566 t
ORIGIN      1 bp upstream of EcoRI site; chromosome 6p21.3.
        1 gaattctcga aacttccttt gtagaaaact ttggaaggtg tctgccacat tgatcctgga
       61 atgtgtgttt atttggggtt atataaatct gttctgtgga agccacctga agtcaggaag
      121 agatggaggg catccttcag gagtgagatg agacctcatc atacttgact gtccagcatc
      181 atctctgagt gaggggacca aaaaatttat cttccaaact aggacacttt caagagtgga
      241 agggggatcc attaatattt tcacctggac aagaggcaaa caccagaatg tccccgatga
      301 aggggatata taatggacct tcttgatgtg aaacctgcca gatgggctgg aaagtccgta
      361 tactgggaca agtatgattt gagttgtttg ggacaaggac aggggtacaa gagaaggaaa
      421 tgggcaaaga gagaagcctg tactcagcca agggtgcaga gatgttatat atgattgctc
      481 ttcagggaac cgggcctcca gctcacaccc cagctgctca accacctcct ctctgaattg
      541 actgtccctt ctttggaact ctaggcctga ccccactccc tggccctccc agcccacgat
      601 tcccctgacc cgactccctt tcccagaact cagtcgcctg aacccccagc ctgtggttct
      661 ctcctaggcc tcagcctttc ctgcctttga ctgaaacagc agtatcttct aagccctggg
      721 ggcttccccg ggccccagcc ccgacctaga acccgcccgc tgcctgccac gctgccactg
      781 ccgcttcctc tataaaggga cctgagcgtc cgggcccagg ggctccgcac agcaggtgag
      841 gctctcctgc cccatctcct tgggctgccc gtgcttcgtg ctttggacta ccgcccagca
      901 gtgtcctgcc ctctgcctgg gcctcggtcc ctcctgcacc tgctgcctgg atccccggcc
      961 tgcctgggcc tgggccttgg tgggtttggt tttggtttcc ttctctgtct ctgactctcc
     1021 atctgtcagt ctcattgtct ctgtcacaca ttctctgttt ctgccatgat tcctctctgt
     1081 tcccttcctg tctctctctg tctccctctg ctcaccttgg ggtttctctg actgcatctt
     1141 gtccccttct ctgtcgatct ctctctcggg ggtcgggggg tgctgtctcc cagggcggga
     1201 ggtctgtctt ccgccgcgtg ccccgccccg ctcactgtct ctctctctct ctctctttct
     1261 ctgcaggttc tccccatgac accacctgaa cgtctcttcc tcccaagggt gtgtggcacc
     1321 accctacacc tcctccttct ggggctgctg ctggttctgc tgcctggggc ccaggtgagg
     1381 cagcaggaga atgggggctg ctggggtggc tcagccaaac cttgagccct agagcccccc
     1441 tcaactctgt tctcccctag gggctccctg gtgttggcct cacaccttca gctgcccaga
     1501 ctgcccgtca gcaccccaag atgcatcttg cccacagcac cctcaaacct gctgctcacc
     1561 tcattggtaa acatccacct gacctcccag acatgtcccc accagctctc ctcctacccc
     1621 tgcctcagga acccaagcat ccaccccctct cccccaactt cccccacgct aaaaaaaaca
     1681 gagggagccc actcctatgc ctccccctgc catcccccag gaactcagtt gttcagtgcc
     1741 cacttcctca gggattgaga cctctgatcc agacccctga tctcccaccc ccatccccta
     1801 tggctcttcc taggagaccc cagcaagcag aactcactgc tctggagagc aaacacggac
     1861 cgtgccttcc tccaggatgg tttctccttg agcaacaatt ctctcctggt ccccaccagt
     1921 ggcatctact tcgtctactc ccaggtggtc ttctctggga aagcctactc tcccaaggcc
     1981 acctcctccc cactctacct ggcccatgag gtccagctct tctcctccca gtacccccttc
     2041 catgtgcctc tcctcagctc ccagaagatg gtgtatccag ggctgcagga accctggctg
     2101 cactcgatgt accacggggc tgcgttccag ctcacccagg gagaccagct atccacccac
     2161 acagatggca tcccccacct agtcctcagc cctagtactg tcttctttgg agccttcgct
     2221 ctgtagaact tggaaaaatc cagaaagaaa aaataattga tttcaagacc ttctccccat
```

```
2281 tctgcctcca ttctgaccat ttcaggggtc gtcaccacct ctcctttggc cattccaaca
2341 gctcaagtct tccctgatca agtcaccgga gctttcaaag aaggaattct aggcatccca
2401 ggggaccaca cctccctgaa ccatccctga tgtctgtctg gctgaggatt tcaagcctgc
2461 ctaggaattc ccagcccaaa gctgttggtc ttgtccacca gctaggtggg gcctagatcc
2521 acacacagag gaagagcagg cacatggagg agcttggggg atgactagag gcagggaggg
2581 gactatttat gaaggcaaaa aaattaaatt atttatttat ggaggatgga gagaggggaa
2641 taatagaaga acatccaagg agaaacagag acaggcccaa gagatgaaga gtgagagggc
2701 atgcgcacaa ggctgaccaa gagagaaaga agtaggcatg agggatcaca gggccccaga
2761 aggcagggaa aggctctgaa agccagctgc cgaccagagc cccacacgga ggcatctgca
2821 ccctcgatga agcccaataa acctcttttc tctgaaatgc tgtctgcttg tgtgtgtgtg
2881 tctgggagtg agaacttccc agtctatcta aggaatggag ggagggacag agggctcaaa
2941 gggagcaaga gctgtgggga gaacaaaagg ataagggctc agagagcttc agggatatgt
3001 gatggactca ccaggtgagg ccgccagact gctgcagggg aagcaaagga gaagctgaga
3061 agatgaagga aaagtcaggg tctggagggg cgggggtcag ggagctcctg ggagatatgg
3121 ccacatgtag cggctctgag gaatgggtta caggagacct ctggggagat gtgaccacag
3181 caatgggtag gagaatgtcc agggctatga aagtcgagta tggggacccc cccttaacga
3241 agacagggcc atgtagaggg ccccagggag tgaaagagcc tccaggacct ccaggtatgg
3301 aatacagggg acgtttaaga agatatggcc acacactggg gccctgagaa gtgagagctt
3361 catgaaaaaa atcagggacc ccagagttcc ttggaagcca agactgaaac cagcattatg
3421 agtctccggg tcagaatgaa agaagagggc ctgccccagt ggggtctgtg aattcccggg
3481 ggtgatttca ctccccgggg ctgtcccagg cttgtccctg ctacccgcac ccagcctttc
3541 ctgaggcctc aagcctgcca ccaagccccc agctccttct ccccgcaggg cccaaacaca
3601 ggcctcagga ctcaacacag cttttccctc caaccccgtt ttctctccct caacggactc
3661 agctttctga agcccctccc agttctagtt ctatcttttt cctgcatcct gtctggaagt
3721 tagaaggaaa cagaccacag acctggtccc caaaagaaat ggaggcaata ggttttgagg
3781 ggcatgggga cggggttcag cctccagggt cctacacaca aatcagtcag tggcccagaa
3841 gacccccctc ggaatcggag cagggaggat ggggagtgtg aggggtatcc ttgatgcttg
3901 tgtgtcccca actttccaaa tccccgcccc cgcgatggag aagaaaccga gacagaaggt
3961 gcagggccca ctaccgcttc ctccagatga gctcatgggt ttctccacca aggaagtttt
4021 ccgctggttg aatgattctt tccccgccct cctctcgccc cagggacata taaaggcagt
4081 tgttggcaca cccagccagc agacgctccc tcagcaagga cagcagagga ccagctaaga
4141 gggagagaag caactacaga cccccccctga aaacaaccct cagacgccac atcccctgac
4201 aagctgccag gcaggttctc ttcctctcac atactgaccc acggcttcac cctctctccc
4261 ctggaaagga caccatgagc actgaaagca tgatccggga cgtggagctg gccgaggagg
4321 cgctccccaa gaagacaggg gggccccagg gctccaggcg gtgcttgttc ctcagcctct
4381 tctccttcct gatcgtggca ggcgccacca cgctcttctg cctgctgcac tttggagtga
4441 tcggccccca gagggaagag gtgagtgcct ggccagcctt catccactct cccacccaag
4501 gggaaatgga gacgcaagag agggagagag atgggatggg tgaaagatgt gcgctgatag
4561 ggagggatgg agagaaaaaa acgtggagaa agacggggat gcagaaagag atgtggcaag
4621 agatggggaa gagagagaga gaaagatgga gagacaggat gtctggcaca tggaaggtgc
4681 tcactaagtg tgtatggagt gaatgaatga atgaatgaat gaacaagcag atatataaat
4741 aagatatgga gacagatgtg gggtgtgaga agagagatgg gggaagaaac aagtgatatg
4801 aataaagatg gtgagacaga aagagcggga aatatgacag ctaaggagag agatggggga
4861 gataaggaga gaagaagata gggtgtctgg cacacagaag acactcaggg aaagagctgt
4921 tgaatgcctg gaaggtgaat acacagatga atggagagag aaaaccagac acctcagggc
4981 taagagcgca ggccagacag gcagccagct gttcctcctt taagggtgac tccctcgatg
5041 ttaaccattc tccttctccc caacagttcc ccagggacct ctctctaatc agccctctgg
5101 cccaggcagt cagtaagtgt ctccaaacct ctttcctaat tctgggtttg ggtttggggg
5161 tagggttagt accggtatgg aagcagtggg ggaaatttaa agttttggtc ttgggggagg
5221 atggatggag gtgaaagtag gggggtattt tctaggaagt ttaagggtct cagcttttc
5281 ttttctctct cctcttcagg atcatcttct cgaaccccga gtgacaagcc tgtagcccat
5341 gttgtaggta agagctctga ggatgtgtct tggaacttgg agggctagga tttggggatt
5401 gaagcccggc tgatggtagg cagaacttgg agacaatgtg agaaggactc gctgagctca
5461 agggaagggt ggaggaacag cacaggcctt agtgggatac tcagaacgtc atggccaggt
5521 gggatgtggg atgacagaca gagaggacag gaaccggatg tggggtgggc agagctcgag
5581 ggccaggatg tggagagtga accgacatgg ccacactgac tctcctctcc ctctctccct
5641 ccctccagca aaccctcaag ctgaggggca gctccagtgg ctgaaccgcc gggccaatgc
5701 cctcctggcc aatggcgtgg agctgagaga taaccagctg gtggtgccat cagagggcct
5761 gtacctcatc tactcccagg tcctcttcaa gggccaaggc tgcccctcca cccatgtgct
5821 cctcacccac accatcagcc gcatcgccgt ctcctaccag accaaggtca acctcctctc
5881 tgccatcaag agcccctgcc agagggagac cccagagggg gctgaggcca agccctggta
5941 tgagcccatc tatctgggag gggtcttcca gctggagaag ggtgaccgac tcagcgctga
6001 gatcaatcgg cccgactatc tcgactttgc cgagtctggg caggtctact ttgggatcat
6061 tgccctgtga ggaggacgaa catccaacct tcccaaacgc ctcccctgcc ccaatccctt
```

```
6121 tattaccccc tccttcagac accctcaacc tcttctggct caaaaagaga attgggggct
6181 tagggtcgga acccaagctt agaactttaa gcaacaagac caccacttcg aaacctggga
6241 ttcaggaatg tgtggcctgc acagtgaagt gctggcaacc actaagaatt caaactgggg
6301 cctccagaac tcactggggc ctacagcttt gatccctgac atctggaatc tggagaccag
6361 ggagcctttg gttctggcca gaatgctgca ggacttgaga agacctcacc tagaaattga
6421 cacaagtgga ccttaggcct tcctctctcc agatgtttcc agacttcctt gagacacgga
6481 gcccagccct ccccatggag ccagctccct ctatttatgt ttgcacttgt gattatttat
6541 tatttattta ttatttattt atttacagat gaatgtattt atttgggaga ccggggtatc
6601 ctgggggacc caatgtagga gctgccttgg ctcagacatg ttttccgtga aaacggagct
6661 gaacaatagg ctgttcccat gtagcccccct ggcctctgtg ccttcttttg attatgtttt
6721 ttaaaatatt tatctgatta agttgtctaa acaatgctga tttggtgacc aactgtcact
6781 cattgctgag cctctgctcc ccaggggagt tgtgtctgta atcgccctac tattcagtgg
6841 cgagaaataa agtttgctta gaaaagaaac atggtctcct tcttggaatt aattctgcat
6901 ctgcctcttc ttgtgggtgg gaagaagctc cctaagtcct ctctccacag gctttaagat
6961 ccctcggacc cagtcccatc cttagactcc tagggccctg gagaccctac ataaacaaag
7021 cccaacagaa tattccccat cccccaggaa acaagagcct gaacctaatt acctctccct
7081 cagggcatgg gaatttccaa ctctgggaat tc
```

```
LOCUS       YY1          2353 bp ss-mRNA            PRI       25-MAY-1993
DEFINITION  Homo sapiens GLI-Krupple related protein (YY1) mRNA, complete cds.
ACCESSION   M77698
KEYWORDS    GLI-Krupple related protein.
SOURCE      Homo sapiens cDNA to mRNA.
REFERENCE   1  (bases 1 to 2353)
  AUTHORS   Shi,Y., Seto,E., Chang,L.-S. and Shenk,T.
  TITLE     Transcriptional repression by YY1, a human GLI-Krupple related
            protein, and relief of repression by adenovirus E1a protein
  JOURNAL   Cell 67, 377-388 (1991)
COMMENT     YY1 has been shown to be involved in transcriptional repression of
            both HPV-16 and HPV-18 (May,M., Dong,X.-P., Beyer-Finkler,E.,
            Stubenrauch,F., Fuchs,P., and Pfister,H., EMBO J.:13, 1460-1466
            (1994); Bauknecht,T., Angel,P., Royer,H.-D., and zur Hausen,H.,
            EMBO J.:11, 4607-4617 (1992)).  It is thought to interfere with
            transcriptional initiation by binding to sites contained within a
            transcriptional silencer domain in the LCRs of these HPV types.
            In HPV-16, this silencer domain is contained within a naturally
            occurring deletion mutation which was found by May et al. to be
            present in 40% of the HPV-16 DNA in a lymph node metastasis of a
            cervical cancer.  May et al. contend that this deletion may
            represent a means by which expression of oncogenic genes is
            de-regulated in episomal copies of HPV-16 DNA.  Since at least
            one-third of HPV-16-positive cervical cancers contain only
            extra-chromosomal viral DNA, this deletion may be an important
            alternative to the destruction of the E2 reading frame, and thereby
            the E2-dependent negative regulation of E6/E7 expression, that
            occurs during integration of the viral DNA into the host genome.
            Mutations of YY-1 binding sites in HPV-18 DNA has been shown to
            affect transcriptional activity differently in malignant and
            non-malignant cell lines.  In non-malignant cells, YY1 binding-site
            mutations led to increased transcriptional activity, whereas in
            certain malignant cell lines, transcriptional activity was
            apparently reduced by the mutations.  This may point to more
            complex interactions between YY1 and other cellular factors.
BASE COUNT      625 a    598 c    608 g    522 t
ORIGIN
        1 cgccgagacg agcagcggcc gagcgagcgc gggcgcgggc gcaccgaggc gagggaggcg
       61 gggaagcccc gccgccgccg ccccgcccgc cccttccccc gccgcccgcc ccctctcccc
      121 ccgcccgctc gccgccttcc tccctctgcc ttccttcccc acggccggcc gcctcctcgc
      181 ccgcccgccc gcagccgagg agccgaggcc gccgcggccg tggcggcgga gccctcagcc
      241 atggcctcgg gcgacaccct ctacatcgcc acggacggct cggagatgcc ggccgagatc
      301 gtggagctgc acgagatcga ggtggagacc atcccggtgg agaccatcga gaccacagtg
      361 gtgggcgagg aggaggagga ggacgacgac gacgaggacg gcggcggtgg cgaccacggc
      421 ggcgggggcg gccacgggca cgccggccac caccaccacc accatcacca ccaccaccac
      481 ccgcccatga tcgctctgca gccgctggtc accgacgacc cgacccaggt gcaccaccac
      541 caggaggtga tcctggtgca gacgcgcgag gaggtggtgg gcggcgacga ctcggacggg
      601 ctgcgcgccg aggacggctt cgaggatcag attctcatcc cggtgcccgc gccggccggc
      661 ggcgacgacg actacattga acaaacgctg gtcaccgtgg cggcggccgg caagagcggc
      721 ggcggcggct cgtcgtcgtc gggaggcggc cgcgtcaaga agggcggcgg caagaagagc
      781 ggcaagaaga gttacctcag cggcggggcc ggcgcggcgg gcgggcgcgg cgccgacccg
      841 ggcaacaaga agtgggagca gaagcaggtg cagatcaaga ccctggaggg cgagttctcg
      901 gtcaccatgt ggtcctcaga tgaaaaaaaa gatattgacc atgagacagt ggttgaagaa
      961 cagatcattg gagagaactc acctcctgat tattcagaat atatgacagg aaagaaactt
     1021 cctcctggag gaatacctgg cattgacctc tcagatccca aacaactggc agaatttgct
     1081 agaatgaagc caagaaaaat taaagaagat gatgctccaa gaacaatagc ttgccctcat
     1141 aaaggctgca caaagatgtt cagggataac tcggccatga gaaaacatct gcacacccac
     1201 ggtcccagag tccacgtctg tgcagaatgt ggcaaagctt ttgttgagag ttcaaaacta
     1261 aaacgacacc aactggttca tactggagag aagccctttc agtgcacgtt cgaaggctgt
     1321 gggaaacgct tttcactgga cttcaatttg cgcacacatg tgcgaatcca taccggagac
     1381 aggccctatg tgtgcccctt cgatggttgt aataagaagt ttgctcagtc aactaacctg
     1441 aaatctcaca tcttaacaca tgctaaggcc aaaaacaacc agtgaaaaga agagagaaga
     1501 cccttctcga ccacgggaag catcttccag aagtgtgatt gggaataaat atgcctctcc
     1561 tttgtatatt atttctagga agaattttaa aaatgaatcc tacacaccta agggacatgt
```

```
1621 tttgataaag tagtaaaaat taaaaaaaaa aaactttact aagatgacat tgctaagatg
1681 ctctatcttg ctctgtaatc tcgtttcaaa aacacagtgt ttttgtaaag tgtggtccca
1741 acaggaggac aattcatgaa cttcgcatca aaagacaatt ctttatacaa cagtgctaaa
1801 aatgggactt cttttcacat tcttataaat atgaagctca cctgttgctt acaatttttt
1861 taattttgta ttttccaagt gtgcatattg tacacttttt tggggatatg cttagtaatg
1921 ctacgtgtga tttttctgga ggttgataac tttgcttgca gtagattttc tttaaaagaa
1981 tgggcagtta catgcatact tcaaaagtat tttcctgtaa aaaaaaaaaa agttatatag
2041 gttttgtttg ctatcttaat tttggttgta ttctttgatg ttaacacatt ttgtataatt
2101 gtatcgtata gctgtattga atcatgtagt atcaaatatt agatgtgatt taatagtgtt
2161 aatcaattta aacccatttt agtcactttt tttttccaaa aaaatactgc cagatgctga
2221 tgttcagtgt aatttctttg cctgttcagt tacagaaagt ggtgctcagt tgtagaatgt
2281 attgtacctt ttaacacctg atgtgtacat cccatgtaac agaaagggca acaataaaat
2341 agcaatccta aag
```

```
LOCUS       c-myc        8082 bp ds-DNA             PRI       10-AUG-1992
DEFINITION  Human (Lawn) c-myc proto-oncogene, complete coding sequence and
            flanks.
ACCESSION   J00120
KEYWORDS    Alu repeat; c-myc proto-oncogene; myc oncogene; proto-oncogene;
            repeat region; transforming gene.
SOURCE      Human DNA (genomic library of Lawn et al.), clones lambda-M1 [1],
            and pUC9-myc [2].
REFERENCE   1  (bases 3507 to 7559)
  AUTHORS   Colby,W.W., Chen,E.Y., Smith,D.H. and Levinson,A.D.
  TITLE     Identification and nucleotide sequence of a human locus homologous
            to the v-myc oncogene of avian myelocytomatosis virus MC29
  JOURNAL   Nature 301, 722-725 (1983)
REFERENCE   2  (bases 1 to 8082)
  AUTHORS   Gazin,C., Dupont,S., de Dinechin,D., Hampe,A., Masson,J.M.,
            Martin,P., Stehelin,D. and Galibert,F.
  TITLE     Nucleotide sequence of the human c-myc locus: provocative open
            reading frame within the first exon
  JOURNAL   EMBO J. 3, 383-387 (1984)
COMMENT     In several cases of HPV-positive cancer, Couturier et al.
            (J Virol 65: 4534-8) determined that HPV sequences integrate
            near the myc locus.  In several instances, this integration
            disrupted the integrity of the proto-oncogene sequence or its
            regulation.

            The myc gene is the cellular homologue of the transforming gene
            carried by the avian myelocytomatosis virus MC29. Unlike the ras
            proto-oncogenes which obtain transforming potential through
            mutations within their coding exons (namely mutations within codon
            12), the myc gene identified as the cause of Burkitt lymphomas
            acquires its transforming potential through defects of either
            transcriptional or translational control. Thus it is not an altered
            gene product that induces tumors, but a normal product that is
            present either in the wrong quantity or at the wrong time in the
            life cycle of the cell.

            [2] notes an open reading frame upstream of the c-myc coding exons
            with an 'atg' start codon at bases 2304-2306 and a 'tag' stop codon
            at bases 2868-2870. However other researchers have used c-myc and
            v-myc DNA sequences to probe for mRNA's with homology to c-myc in
            various human cell lines and none of them have noted any mRNA's
            beginning upstream of bp 2328 (see other human c-myc entries).

            The t(8;14) translocation site in the Burkitt lymphoma cell line
            BL22 occurs between bp 1316 and 1317 of this sequence.

BASE COUNT     1850 a   2115 c   2135 g   1982 t
ORIGIN      198 bp upstream of Sau96A site, on chromosome 8 (q24).
        1 agcttgtttg gccgttttag ggtttgttgg aatttttttt tcgtctatgt acttgtgaat
       61 tatttcacgt ttgccattac cggttctcca tagggtgatg ttcattagca gtggtgatag
      121 gttaattttc accatctctt atgcggttga atagtcacct ctgaaccact ttttcctcca
      181 gtaactcctc tttcttcgga ccttctgcag ccaacctgaa agaataacaa ggaggtggct
      241 ggaaacttgt tttaaggaac cgcctgtcct tcccccgctg gaaaccttgc acctcggacg
      301 ctcctgctcc tgcccccacc tgacccccgc cctcgttgac atccaggcgc gatgatctct
      361 gctgccagta gagggcacac ttactttact ttcgcaaacc tgaacgcggg tgctgcccag
      421 agagggggcg gagggaaaga cgctttgcag caaaatccag catagcgatt ggttgctccc
      481 cgcgtttgcg gcaaaggcct ggaggcagga gtaatttgca atccttaaag ctgaattgtg
      541 cagtgcatcg gatttggaag ctactatatt cacttaacac ttgaacgctg agctgcaaac
      601 tcaacgggta ataacccatc ttgaacagcg tacatgctat acacacaccc ctttcccccg
      661 aattgttttc tcttttggag gtggtggagg gagagaaaag tttacttaaa atgcctttgg
      721 gtgagggacc aaggatgaga agaatgtttt ttgtttttca tgccgtggaa taacacaaaa
      781 taaaaaatcc cgagggaata tacattatat attaaatata gatcatttca gggagcaaac
      841 aaatcatgtg tggggctggg caactagctg agtcgaagcg taaataaaat gtgaatacac
      901 gtttgcgggt tacatacagt gcactttcac tagtattcag aaaaaattgt gagtcagtga
      961 actaggaaat taatgcctgg aaggcagcca aattttaatt agctcaagac tccccccccc
     1021 ccccaaaaaa aggcacggaa gtaatactcc tctcctcttc tttgatcaga atcgatgcat
```

c-myc

```
1081 tttttgtgca tgaccgcatt tccaataata aaaggggaaa gaggacctgg aaaggaatta
1141 aacgtccggt ttgtccgggg aggaaagagt taacggtttt tttcacaagg gtctctgctg
1201 actcccccgg ctcggtccac aagctctcca cttgcccctt ttaggaagtc cggtcccgcg
1261 gttcgggtac cccctgcccc tcccatattc tcccgtctag cacctttgat ttctcccaaa
1321 cccggcagcc cgagactgtt gcaaaccggc gccacagggc gcaaagggga tttgtctctt
1381 ctgaaacctg gctgagaaat tgggaactcc gtgtgggagg cgtgggggtg ggacggtggg
1441 gtacagactg gcagagagca ggcaacctcc ctctcgccct agcccagctc tggaacaggc
1501 agacacatct cagggctaaa cagacgcctc ccgcacgggg ccccacggaa gcctgagcag
1561 gcggggcagg aggggcggta tctgctgctt tggcagcaaa ttggggggact cagtctgggt
1621 ggaaggtatc caatccagat agctgtgcat acataatgca taatacatga ctccccccaa
1681 caaatgcaat gggagtttat tcataacgcg ctctccaagt atacgtggca atgcgttgct
1741 gggttatttt aatcattcta ggcatcgttt tcctccttat gcctctatca ttcctcccta
1801 tctacactaa catcccacgc tctgaacgcg cgcccattaa tacccttctt tcctccactc
1861 tccctgggac tcttgatcaa agcgcggccc tttccccagc cttagcgagg cgccctgcag
1921 cctggtacgc gcgtggcgtg gcggtgggcg cgcagtgcgt tctctgtgtg gagggcagct
1981 gttccgcctg cgatgattta tactcacagg acaaggatgc ggtttgtcaa acagtactgc
2041 tacggaggag cagcagagaa agggagaggg tttgagaggg agcaaaagaa aatggtaggc
2101 gcgcgtagtt aattcatgcg gctctcttac tctgtttaca tcctagagct agagtgctcg
2161 gctgcccggc tgagtctcct ccccaccttc cccaccctcc ccaccctccc cataagcgcc
2221 cctcccgggt tcccaaagca gagggcgtgg gggaaaagaa aaaagatcct ctctcgctaa
2281 tctccgccca ccggcccttt ataatgcgag ggtctggacg gctgaggacc cccgagctgt
2341 gctgctcgcg gccgccaccg ccgggccccg gccgtccctg gctcccctcc tgcctcgaga
2401 agggcagggc ttctcagagg cttggcggga aaaagaacgg agggagggat cgcgctgagt
2461 ataaaagccg gttttcgggg ctttatctaa ctcgctgtag taattccagc gagaggcaga
2521 gggagcgagc gggcggccgg ctagggtgga agagccgggc gagcagagct gcgctgcggg
2581 cgtcctggga agggagatcc ggagcgaata gggggcttcg cctctggccc agccctcccg
2641 ctgatccccc agccagcggt ccgcaaccct tgccgcatcc acgaaacttt gcccatagca
2701 gcgggcgggc actttgcact ggaacttaca acacccgagc aaggacgcga ctctcccgac
2761 gcggggaggc tattctgccc atttggggac acttccccgc cgctgccagg acccgcttct
2821 ctgaaaggct ctccttgcag ctgcttagac gctggatttt tttcgggtag tggaaaacca
2881 ggtaagcacc gaagtccact tgccttttaa tttatttttt tatcacttta atgctgagat
2941 gagtcgaatg cctaaatagg gtgtcttttc tcccattcct gcgctattga cacttttctc
3001 agagtagtta tggtaactgg ggctggggtg gggggtaatc cagaactgga tcggggtaaa
3061 gtgacttgtc aagatgggag aggagaaggc agagggaaaa cgggaatggt ttttaagact
3121 accctttcga gatttctgcc ttatgaatat attcacgctg actcccggcc ggtcggacat
3181 tcctgcttta ttgtgttaat tgctctctgg gttttggggg gctgggggtt gctttgcggt
3241 gggcagaaag ccccttgcat cctgagctcc ttggagtagg gaccgcatat cgcctgtgtg
3301 agccagatcg ctccgcagcc gctgacttgt ccccgtctcc gggagggcat ttaaatttcg
3361 gctcaccgca tttctgacag ccggagacgg acactgcggc gcgtcccgcc cgcctgtccc
3421 cgcggcgatt ccaacccgcc ctgatccttt taagaagttg gcatttggct ttttaaaaag
3481 caataataca atttaaaacc tgggtctcta gaggtgttag gacgtggtgt tgggtaggcg
3541 caggcagggg aaaagggagg cgaggatgtg tccgattctc ctggaatcgt tgacttggaa
3601 aaaccagggc gaatctccgc acccagccct gactcccctg ccgcggccgc cctcgggtgt
3661 cctcgcgccc gagatgcgga ggaactgcga ggagcggggc tctgggcggt tccagaacag
3721 ctgctaccct tggtggggtg gctccggggg aggtatcgca gcggggtctc tggcgcagtt
3781 gcatctccgt attgagtgcg aagggaggtg cccctattat tatttgacac ccccccttgta
3841 tttatggagg ggtgttaaag cccgcggctg agctcgccac tccagccggc gagagaaaga
3901 agaaaagctg gcaaaaggag tgttggacgg gggcggtact gggggtgggg acgggggcgg
3961 tggagaggga aggttgggag gggctgcggt gccggcgggg gtaggagagc ggctagggcg
4021 cgagtgggaa cagccgcagc ggagggggccc cggcgcggag cggggttcac gcagccgcta
4081 gcgcccaggc gcctctcgcc ttctccttca ggtggcgcaa aactttgtgc cttggatttt
4141 ggcaaattgt tttcctcacc gccacctccc gcggcttctt aagggcgcca gggccgattt
4201 cgattcctct gccgctgcgg ggccgactcc cgggctttgc gctccgggct cccgggggag
4261 cgggggctcg gcgggcacca agccgctggt tcactaagtg cgtctccgag atagcagggg
4321 actgtccaaa gggggtgaaa gggtgctccc tttattcccc caccaagacc acccagccgc
4381 tttaggggat agctctgcaa ggggagaggt tcgggactgt ggcgcgcact gcgcgctgcg
4441 ccaggtttcc gcaccaagac ccctttaact caagactgcc tcccgctttg tgtgccccgc
4501 tccagcagcc tcccgcgacg atgcccctca acgttagctt caccaacagg aactatgacc
4561 tcgactacga ctcggtgcag ccgtatttct actgcgacga ggaggagaac ttctaccagc
4621 agcagcagca gagcgagctg cagcccccgg cgcccagcga ggatatctgg aagaaattcg
4681 agctgctgcc caccccgccc ctgtccccta gccgccgctc cgggctctgc tcgccctcct
4741 acgttgcggt cacacccttc tcccttcggg gagacaacga cggcggtggc gggagcttct
4801 ccacggccga ccagctggag atggtgaccg agctgctggg aggagacatg gtgaaccaga
4861 gtttcatctg cgacccggac gacgagacct tcatcaaaaa catcatcatc caggactgta
```

```
4921 tgtggagcgg cttctcggcc gccgccaagc tcgtctcaga gaagctggcc tcctaccagg
4981 ctgcgcgcaa agacagcggc agcccgaacc ccgcccgcgg ccacagcgtc tgctccacct
5041 ccagcttgta cctgcaggat ctgagcgccg ccgcctcaga gtgcatcgac ccctcggtgg
5101 tcttccccta ccctctcaac gacagcagct cgcccaagtc ctgcgcctcg caagactcca
5161 gcgccttctc tccgtcctcg gattctctgc tctcctcgac ggagtcctcc ccgcagggca
5221 gccccgagcc cctggtgctc catgaggaga caccgcccac caccagcagc gactctggta
5281 agcgaagccc gcccaggcct gtcaaaagtg ggcggctgga tacctttccc attttcattg
5341 gcagcttatt taacgggcca ctcttattag gaaggagaga tagcagatct ggagagattt
5401 gggagctcat cacctctgaa accttgggct ttagcgtttc ctcccatccc ttccccttag
5461 actgcccatg tttgcagccc ccctccccgt ttgtctccca cccctcagga atttcattta
5521 ggtttttaaa ccttctggct tatcttacaa ctcaatccac ttcttcttac ctcccgttaa
5581 cattttaatt gccctggggc ggggtggcag ggagtgtatg aatgaggata agagaggatt
5641 gatctctgag agtgaatgaa ttgcttccct cttaacttcc gagaagtggt gggatttaat
5701 gaactatcta caaaaatgag gggctgtgtt tagaggctag gcagggcctg cctgagtgcg
5761 ggagccagtg aactgcctca agagtgggtg ggctgaggag ctgggatctt ctcagcctat
5821 tttgaacact gaaaagcaaa tccttgccaa agttggactt tttttttttct tttattcctt
5881 cccccgccct cttggacttt tggcaaaact gcaatttttt ttttttttatt tttcatttcc
5941 agtaaaatag ggagttgcta aagtcatacc aagcaatttg cagctatcat ttgcaacacc
6001 tgaagtgttc ttggtaaagt ccctcaaaaa taggaggtgc ttgggaatgt gctttgcttt
6061 gggtgtgtcc aaagcctcat taagtcttag gtaagaattg gcatcaatgt cctatcctgg
6121 gaagttgcac ttttcttgtc catgccataa cccagctgtc tttcccttta tgagactctt
6181 accttcatgg tgagaggagt aagggtggct ggctagattg gttctttttt tttttttttc
6241 ctttttaag acggagtctc actctgtcac taggctggag tgcagtggcg caatcaacct
6301 ccaacccccct ggttcaagag attctcctgc ctcagcctcc caagtagctg ggactacagg
6361 tgcacaccac catgccaggc taatttttgt aattttagta gagatggggt ttcatcgtgt
6421 tggccaggat ggtctctcct gacctcacga tccgcccacc tcggcctccc aaagtgctgg
6481 gattacaggt gtgagccagg gcaccaggct tagatgtggc tctttgggga gataattttg
6541 tccagagacc tttctaacgt attcatgcct tgtatttgta cagcattaat ctggtaattg
6601 attattttaa tgtaaccttg ctaaaggagt gatttctatt tcctttctta aagaggagga
6661 acaagaagat gaggaagaaa tcgatgttgt ttctgtggaa aagaggcagg ctcctggcaa
6721 aaggtcagag tctggatcac cttctgctgg aggccacagc aaacctcctc acagcccact
6781 ggtcctcaag aggtgccacg tctccacaca tcagcacaac tacgcagcgc ctccctccac
6841 tcggaaggac tatcctgctg ccaagagggt caagttggac agtgtcagag tcctgagaca
6901 gatcagcaac aaccgaaaat gcaccagccc caggtcctcg gacaccgagg agaatgtcaa
6961 gaggcgaaca cacaacgtct tggagcgcca gaggaggaac gagctaaaac ggagcttttt
7021 tgccctgcgt gaccagatcc cggagttgga aaacaatgaa aaggccccca aggtagttat
7081 ccttaaaaaa gccacagcat acatcctgtc cgtccaagca gaggagcaaa agctcatttc
7141 tgaagaggac ttgttgcgga aacgacgaga acagttgaaa cacaaacttg aacagctacg
7201 gaactcttgt gcgtaaggaa aagtaaggaa aacgattcct tctaacagaa atgtcctgag
7261 caatcaccta tgaacttgtt tcaaatgcat gatcaaatgc aacctcacaa ccttggctga
7321 gtcttgagac tgaaagattt agccataatg taaactgcct caaattggac tttgggcata
7381 aaagaacttt tttatgctta ccatcttttt tttttcttta acagatttgt atttaagaat
7441 tgtttttaaa aaattttaag atttacacaa tgtttctctg taaatattgc cattaaatgt
7501 aaataacttt aataaaacgt ttatagcagt tacacagaat ttcaatccta gtatatagta
7561 cctagtatta taggtactat aaaccctaat ttttttttatt taagtacatt ttgcttttta
7621 aagttgattt ttttctattg tttttagaaa aaataaaata actggcaaat atatcattga
7681 gccaaatctt aagttgtgaa tgttttgttt cgtttcttcc ccctcccaac caccaccatc
7741 cctgtttgtt ttcatcaatt gccccttcag agggcggtct taagaaaggc aagagttttc
7801 ctctgttgaa atgggtctgg gggccttaag gtctttaagt tcttggaggt tctaagatgc
7861 ttcctggaga ctatgataac agccagagtt gacagttaga aggaatggca gaaggcaggt
7921 gagaaggtga gaggtaggca aaggagatac aagaggtcaa aggtagcagt taagtacaca
7981 aagaggcata aggactgggg agttgggagg aaggtgagga agaaactcct gttactttag
8041 ttaaccagtg ccagtcccct gctcactcca aacccaggaa tt
```

Cyclin A

```
LOCUS       cyclinA      1181 bp ds-DNA            PRI       21-JUN-1993
DEFINITION  Human cyclin A (CCNA) gene sequence.
ACCESSION   M96390
KEYWORDS    cyclin A.
SOURCE      Homo sapiens DNA.
REFERENCE   1  (bases 1 to 1181)
  AUTHORS   Nikaido,T. and Yamamoto,M.
  TITLE     p34cdc2 and cyclin A expression may be suppressed by RB and/or p53
  JOURNAL   Unpublished (1992)
COMMENT     In the S-phase of the cell cycle, cyclin A complexes with E2F (a
            cellular transcription factor), p107, and cdk2 kinase.  Arroyo et
            al. (Mol. Cell Biol. 13: 6537-46) report that the E7 protein of
            HPV-16 associates with the E2F-cyclin A complex.  Arroyo et al.
            (Mol. Cell Biol. 13: 6537-46) further note that the high-risk
            HPV has a higher binding efficiency to the E2F-cyclin A complex than
            the low-risk HPV type. Steinmann et al. (Oncogene 9: 387-94)
            report a 20-fold increase in the level of cyclinA as a result of
            HPV-16 infection in human keratinocytes. Pagano et al. (Oncogene
            7: 1681-6) report that the E2F-cyclin A complex can occur in HPV-18
            infected cells independent of pRB.
BASE COUNT      302 a    279 c    262 g    338 t
ORIGIN
        1 aagctttgta tattcttata tttatatata aatataaaaa tttgttaaag gcacgtatag
       61 ttaagagagt tttattttaa taaggtcata ttgtttttac tatgtttaaa aaactttact
      121 ctgaaaggaa cataattata tctaggtcac tagaacgtca ttgtgttttt tgttggttgc
      181 acagcttggg gaaaaataga aaaaaattaa tgactgattt gaatattttg taatgcactg
      241 ctatttatta tatatatcaa cagtagttca aggtgccatc ttaaattaat tgcatcttca
      301 ttaggaaaaa taaaaagcat aaaacacaat ttctggttac tatgaataaa cgcctaaatg
      361 ttaagatgac attacagtct tgacacttga gtactgtatt actatgtgag ctccgtgtta
      421 aataatttat gcacattatt taatcctaac aaccatatga ctgtagttat tagtccctat
      481 taacacataa gaaaacggag aatcggagat actgaaaaac gtgccccaga ttttagacct
      541 ttggaaaaag tcacttaagc taactagacg tcccagagct aaaggctggg caacccaaat
      601 gatagtcgcc aaagtttaat tccgtttaat tccctaaaag gcttagagtc agcttcggac
      661 agcctcgctc actaggtggc tcagcttaaa ataatcggaa gcgtcgggcc ctaaatccta
      721 cctctccccg ccccgcgcag gcgttttctc ccgccccagc cagtttgttt ctccctcctg
      781 ccccgcccct gctcagtttc ctttggttta cccttcactc gccctgaccc tgtcgccttg
      841 aatgacgtca aggccgcgag cgctttcatt ggtccatttc aatagtcgcg ggatacttga
      901 actgcaagaa cagccgccgc tccagcgggc tgctcgctgc atctctgggc gtctttggct
      961 cgccacgctg ggcagtgcct gcctgcgcct ttcgcaacct cctcggccct gcgtggtctc
     1021 gagctgggtg agcgagcggg cgggctggta ggctggcctg ggctgcgacc ggcggctacg
     1081 actattcttt ggccgggtcg gtgcgagtgg tcggctgggc agagtgcacg ctgcttggcg
     1141 ccgcaggctg atcccgccgt ccactcccgg gagcagtgat g
```

```
LOCUS       GCORTICOIDR  2518 bp ds-DNA          PRI       15-JUN-1990
DEFINITION  Human glucocorticoid receptor gene, partial.
ACCESSION   M32284
KEYWORDS    glucocorticoid receptor.
SOURCE      Human blood leukocyte DNA.
REFERENCE   1  (bases 1 to 2518)
  AUTHORS   Zong,J., Ashraf,J. and Thompson,E.B.
  TITLE     The promoter and first, untranslated exon of the human
            glucocorticoid receptor gene are GC rich but lack consensus
            glucocorticoid receptor element sites
  JOURNAL   Mol. Cell. Biol. 10, 5580-5585 (1990)
COMMENT     Authorin copy of sequence [Unpublished (1990) Univ. of TX Med.
            Branch, Galveston TX 77550] kindly submitted by B.E.Thompson,
            22-FEB-1990.
            Mittal et al. (J Virol 67: 5656-9) identified three glucocorticoid
            response elements in the LCR of HPV-16 at position 7640, 7385, and
            7474.  This element, which binds the glucocorticoid receptor, is a
            partial palindrome and consists of the consensus TGTACANNNTGTCAT
            (Chan et al. J Virol 63: 3261-9).  In the presence of glucocorticoids,
            up-regulation of the P$_97$ promoter is observed.

BASE COUNT      403 a    802 c    882 g    431 t
ORIGIN
        1 aggattaggt ggagctgcgg cagcctcccg cccgtgtcag gagctggcaa gcgatgtcta
       61 cctgtgtgtg cgcaaaagtt acctccccaa accctaaacc cacacagcac aacctttccc
      121 agatcacaaa aatcataatc tgtgtctgca caaggtagga ggctcggtcc cggcatcgtc
      181 gaagccttcc cgacgcggcg agctggggaa gggagctggg gcgggggctt cccgcacggg
      241 cacccctcgc cccacgccct ctcctttctc aggacggacc acgagttccc ttccccttgg
      301 actgaggggg aagctcctaa caggaacatc tgtagggagt tgaacgctgg cattttaaag
      361 ctgcctgtat tttgttttat ttgtaggggc aggggtccta tgaacgtgat agggtgagca
      421 acgcacagag tcgagggcag caaatgtcaa gattcggggg tggggcctgc accgggaact
      481 tggacgcggg ccctggccgg ggtggaagaa gaggtcagga gttcggaagg gggctatatt
      541 tcgccagcaa cttactattt cgcctgcaac ttgcttttaa gcctgccgcc ccctgctttc
      601 cttaatcata ataataaaaa aaaagtgcaa agaaatccag ctcgctggag gttttgcatt
      661 tggcgtgcaa cttccttcga gtgtgagcac attgggcggg aggggtgggg gttgaacttg
      721 gcaggcggcg cctccttctg ccgccgccgc cgcctcgcag actcggggaa gagggtgggg
      781 gacggtcggg gcgcggggga gggtgggttc tgctttgcaa cttctctccc agtgcgagag
      841 cgcggcggcg gcagctgaag acccggccgc ccagatgatg cggtggtggg ggacctgccg
      901 gcacgccgac tccccccggg cccaaagtac gtatgcgccg acccccgcta tcccgtccct
      961 tccctgaagc ctccccagag ggcgtgtcag gccgcccggc cccgagcgcg gccgagacgc
     1021 tgcggcaccg tttccgtgca accccgtagc ccctttcgaa gtgacacact tcacgcaact
     1081 cggcccggcg gcggcggcgc gggccactca cgcagctcag ccgcgggagg cgccccggct
     1141 cttgtggccc gcccgctgtc accgcagggg cactcggcgc ttgccgccaa ggggcagagc
     1201 gagctcccga gtgggtctgg agccgcggag ctgggcgggg gcgggaagga ggtagcgaga
     1261 aaagaaactg gagaaactcg gtggccctct taacgccgcc ccagagagac caggtcggcc
     1321 cccgccctgc cgccgccacc ctttttcctg gggagttggg ggcggggggc gaagcgcggc
     1381 gcaccgggcg gggcggccac gccaggggac gcgggcgtgc aggcgcgtcg gggccggggt
     1441 ggcggggccc gcgcggaggg cgtgggggca gggaccgcgg gcgcccctgc agttgccaag
     1501 cgtcaccaac aggttgcatc gttccccgcg ccgccgcgcg gcccctcggg cggggagcgg
     1561 ccgggggtgg agtgggagcg cgtgtgtgcg agtgtgtgcc gccgtggcgc cgcctccacg
     1621 ccgctccccg ctcggtcccg ctcgctcgcc caggccgggc tgccctttcg tgtccgcgct
     1681 ctcttccctc cgccgccgcc tcctccattt tgcgagctcg tgtctgtgac gggagcccga
     1741 gtcaccgcct gccgtcgggg acggattctg tgggtggaag gagacccgca ccggagcggc
     1801 cgaagcagct gggaccggga cgggcacgcg cgcccggaag ccccgacccg cggagcccgg
     1861 cgcggggcgg agggctggct tgtcagctgg gcaatgggag actttcttaa atagggctct
     1921 ccccccaccc atggagaaag gggcggctgt ttacttcctt tttttagaaa aaaaaaatat
     1981 atttccctcc tgctccttct gcgttcacaa gctaagttgt ttatctcggc tgcggcggga
     2041 actgcggacg gtggcgggcg agcggctcct ctgccagagg taagaagcga ggcgggaggg
     2101 ggccggggcg cgctcgctcc cccgaggtgc cgctgggacc ggagacaact cgggggccgc
     2161 cgcgggagcc tacaaacttt tattagcctc ggggagtggg ggtgggggc tggcaagggc
     2221 cgggcgacgg tgacgaaagg gtagcgcgcg ggtgacagcg ctggcctctt cctctccctc
     2281 cgccggcgtc cctggccggg ccgaggggga ggaacctgac ctcggacggc gagcggagcc
     2341 ctgtcgaact ccggggcttc gagcctctca ttctcgcggg aatcctggcc tcttttctcc
     2401 ccctagtgtc ccctttcctc caagggggtc gcccgacacc cgttttcgtg gtgagactaa
     2461 gccgcgtctg aattttactc gcccgaatat ttcacccacc ccgcccgagc gcgagccc
```

```
LOCUS       junB         2136 bp ds-DNA             PRI       30-MAR-1994
DEFINITION  Human transactivator (jun-B) gene, complete cds.
ACCESSION   M29039
KEYWORDS    c-myc proto-oncogene; transactivator.
SOURCE      Homo sapiens DNA.
REFERENCE   1  (bases 1 to 2136)
  AUTHORS   Schuette,J., Viallet,J., Nau,M., Segal,S., Fedorko,J. and Minna,J.
  TITLE     jun-B inhibits the transforming and transactivating activities of
            c-jun
  JOURNAL   Cell 59, 987-997 (1989)
COMMENT     Authorin copy of sequence [1] kindly submitted by J.D.Minna
            10-OCT-1989.

            Thierry et al. ( J Virol 66: 3740-8 ) demonstrated that JunB
            binds AP-1 target sites within the LCR of the HPV-18
            P$_105$ promoter during human keratinocyte infection.  Mutation
            of these two target sites destroyed P$_105$ promoter activity
            in these cells.  Bossy-Wetzel et al. (Genes Dev 6: 2340-51) reported
            that junB or c-jun are overexpressed during the aggressive
            fibromatoses or fibrosarcoma stage of BPV infection.  They also
            correlated this overexpression to changes in fibroblast cell shape
            and anchorage dependence.

BASE COUNT      387 a    716 c    671 g    362 t
ORIGIN
        1 cgcgagccgc ctcctcccct tccccacgct cgaggagggg ggcgcggggg cccggctccg
       61 gcgacggcca atcggagcgc acttccgtgg ctgactagcg cggtataaag gcgtgtggct
      121 caggctgagc ggctgggacc ttgagagcgg ccaggccagc ctcggagcca gcagggagct
      181 gggagctggg ggaaacgacg ccaggaaagc tatcgcgcca gagagggcga cgggggctcg
      241 ggaagcctga cagggctttt gcgcacagct gccggctggc tgctacccgc ccgcgccagc
      301 ccccgagaac gcgcgaccag gcacccagtc cggtcaccgc agcggagagc tcgccgctcg
      361 ctgcagcgag gcccggagcg gccccgcagg gaccctcccc agaccgcctg ggccgcccgg
      421 atgtgcacta aaatggaaca gcccttctac cacgacgact catacacagc tacgggatac
      481 ggccgggccc ctggtggcct ctctctacac gactacaaac tcctgaaacc gagcctggcg
      541 gtcaacctgg ccgaccccta ccggagtctc aaagcgcctg ggctcgcgg acccggccca
      601 gagggcggcg gtggcggcag ctacttttct ggtcagggct cggacaccgg cgcgtctctc
      661 aagctcgcct cttcggagct ggaacgcctg attgtcccca acagcaacgg cgtgatcacg
      721 acgacgccta cacccccggg acagtacttt tacccccgcg ggggtggcag cggtggaggt
      781 gcagggggcg cagggggcgg cgtcaccgag gagcaggagg gcttcgccga cggctttgtc
      841 aaagccctgg acgatctgca caagatgaac cacgtgacac cccccaacgt gtccctgggc
      901 gctaccgggg ggcccccggc tgggcccggg ggcgtctacg ccggcccgga gccacctccc
      961 gtttacacca acctcagcag ctactcccca gcctctgcgt cctcgggagg cgccggggct
     1021 gccgtcggga ccgggagctc gtacccgacg accaccatca gctacctccc acacgcgccg
     1081 cccttcgccg gtggccaccc ggcgcagctg ggcttgggcc gcggcgcctc caccttcaag
     1141 gaggaaccgc agaccgtgcc ggaggcgcgc agccgggacg ccacgccgcc ggtgtccccc
     1201 atcaacatgg aagaccaaga gcgcatcaaa gtggagcgca agcggctgcg gaaccggctg
     1261 gcggccacca agtgccggaa gcggaagctg gagcgcatcg cgcgcctgga ggacaaggtg
     1321 aagacgctca aggccgagaa cgcggggctg tcgagtaccg ccggcctcct ccgggagcag
     1381 gtggcccagc tcaaacagaa ggtcatgacc cacgtcagca acggctgtca gctgctgctt
     1441 ggggtcaagg gacacgcctt ctgaacgtcc cctgcccctt tacggacacc ccctcgcttg
     1501 gacggctggg cacacgcctc ccactggggt ccagggagca ggcggtgggc acccaccctg
     1561 ggacctaggg gcgccgcaaa ccacactgga ctccggccct cctaccctgc gcccagtcct
     1621 tccacctcga cgtttacaag cccccccttc cactttttt tgtatgtttt ttttctgctg
     1681 gaaacagact cgattcatat tgaatataat atatttgtgt atttaacagg gaggggaaga
     1741 gggggcgatc gcggcggagc tggccccgcc gcctggtact caagcccgcg gggacattgg
     1801 gaaggggacc cccgcccccct gccctcccct ctctgcaccg tactgtggaa aagaaacacg
     1861 cacttagtct ctaaagagtt tattttaaga cgtgtttgtg tttgtgtgtg tttgttcttt
     1921 ttattgaatc tatttaagta aaaaaaaaat tggttcttta ttaatttctg ttgtcttttt
     1981 ttccaagctg ggagggcggg gggaaaaaaa aagcactggt ttgcccccag ctcagtgctg
     2041 ttggtggctc ggtcctgtat gtgtccccct cgtcggttcg gcgcaggcat cttgtggtcc
     2101 cagcccagga gtcccaccct tcccgcgtcc ccagat
```

```
LOCUS       p107Rb       3960 bp ss-mRNA            PRI       06-MAY-1993
DEFINITION  Human retinoblastoma related protein (p107) mRNA, complete cds.
ACCESSION   L14812
KEYWORDS    cell cycle regulation protein; retinoblastoma protein;
            tumor suppressor.
SOURCE      Homo sapiens cDNA to mRNA.
REFERENCE   1  (sites)
  AUTHORS   Ewen,M.E., Xing,Y., Bentley-Lawrence,J. and Livingston,D.M.
  TITLE     Molecular cloning, chromosome mapping and expression of the cDNA
            for p107 a retinoblastoma gene product-related protein
  JOURNAL   Cell 66, 1155-1164 (1991)
REFERENCE   2  (bases 1 to 3960)
  AUTHORS   Zhu,L., van den Heuvel,S., Helin,K., Fattaey,A., Ewen,M.,
            Livingston,D., Dyson,N. and Harlow,E.
  TITLE     Inhibition of cell proliferation by p107, a relative of the
            retinoblastoma protein
  JOURNAL   Genes Dev. (1993) In press
COMMENT     p107, an Rb related protein, has been shown to complex with
            Rb (Carlotti et al. J Gen Virol 74: 2479-86 ).  Davies et al.
            (J Virol 67: 2521-8) demonstrated that the HPV-16 E7 protein
            binds p107 and that this binding involves the Rb target site.
            B-myb, a gene involved in cell cycle progression, is
            inappropriately transcribed during G1 in HPV-16 E7 transactivated
            cells.  Lam et al (EMBO J 13: 871-8) demonstrated that B-myb
            transcription is regulated through interactions with p107.

BASE COUNT     1257 a    771 c    869 g   1063 t
ORIGIN
        1 cgggtagcgc gcctgggagg gagaaagaag tcgggggccg tggcgcgcag cccgcggggc
       61 ctgaagggat gttcgaggac aagccccacg ctgagggggc ggcggtggtc gccgcagccg
      121 gggaggcgct acaggccctg tgccaggagc tgaacctgga cgaggggagc gcggccgaag
      181 ccctggacga ctttactgcc atccgaggca actacagcct agagggagaa gttacacact
      241 ggttggcatg ttcattatat gttgcatgcc gcaaaagcat tattcccacg gttggaaagg
      301 gtatcatgga aggcaactgt gtttcactta ccagaatact acgttcagct aaattaagtt
      361 taatacaatt ttttagtaaa atgaagaaat ggatggacat gtcaaatcta ccacaagaat
      421 ttcgtgaacg tatagaaagg ctagagagaa attttgaggt gtctactgta atattcaaaa
      481 aatatgagcc aatttttttta gatatatttc aaaatccata tgaagaacca ccaaagttac
      541 cacgaagccg gaagcagagg aggattcctt gcagtgttaa ggatctgttt aatttctgtt
      601 ggacactttt tgtttatact aagggtaatt ttcggatgat tggggatgac ttagtaaact
      661 cttatcattt acttctatgc tgcttggatc tgatttttgc caatgcgatt atgtgcccaa
      721 atagacaaga cttgctaaat ccatcattta aaggtttacc atctgatttt catactgctg
      781 actttacggc ttctgaagag ccaccctgca tcattgctgt actgtgtgaa ctgcatgatg
      841 gacttctcgt agaagcaaaa ggaataaagg agcactactt taagccatat atttcaaaac
      901 tctttgacag gaagatatta aaaggagaat gcctcctgga cctttcaagt tttactgata
      961 atagcaaagc agtgaataag gagtatgaag agtatgttct aactgttggt gattttgatg
     1021 agaggatctt tttgggagca gacgcagaag aggaaattgg aacacctcga aagttcactc
     1081 gtgacacccc attagggaaa ctgacagcac aggctaatgt ggagtataac cttcaacagc
     1141 actttgaaaa aaaaaggtca tttgcacctt ctaccccact gaccggacgg agatatttac
     1201 gagaaaaaga agcagtcatt actcctgttg catcagccac ccaaagtgtg agccggttac
     1261 agagtattgt ggctggtctg aaaaatgcac caagtgacca acttataaat atttttgaat
     1321 cttgtgtgcg taatcctgtt gaaaacatta tgaaaatact aaaaggaata ggagagactt
     1381 tctgtcaaca ctatactcaa tcaacagatg aacagccagg atctcacata gactttgctg
     1441 taaacagact aaagctggca gaaattttgt attataaaat actagagact gtaatggttc
     1501 aggaaacacg aagacttcat ggaatggaca tgtcagttct tttagagcaa gatatatttc
     1561 atcgttcctt gatggcttgt tgtttggaaa ttgtgctctt tgcctatagc tcacctcgta
     1621 cttttccttg gattattgaa gttctcaact tgcaaccatt ttacttttat aaggttattg
     1681 aggtggtgat ccgctcagaa gaggggctct caagggacat ggtgaaacac ctaaacagca
     1741 ttgaagaaca gattttggag agtttagcat ggagtcacga ttctgcactg tgggaggctc
     1801 tccaggtttc tgcaaacaaa gttcctacct gtgaagaagt tatattccca aataactttg
     1861 aaacaggaaa tggaggaaat gtgcagggac atcttcccct gatgccaatg tctcctctaa
     1921 tgcacccaag agtcaaggaa gttcgaactg acagtgggag tcttcgaaga gatatgcaac
     1981 cattgtctcc aatttctgtc catgaacgct acagttctcc taccgcaggg agtgctaaga
     2041 gaagactctt tggagaggac cccccaaagg aaatgcttat ggacaagatc ataacagaag
     2101 gaacaaaatt gaaaatcgct ccttcttcaa gcattactgc tgaaaatgta tcaattttac
     2161 ctggtcaaac tcttctaaca atggccacag ccccagtaac aggaacaaca ggacataaag
```

```
2221 ttacaattcc attacatggt gtcgcaaatg atgctggaga gatcacactg atacctcttt
2281 ccatgaatac aaatcaggag tccaaagtca agagtcctgt atcacttact gctcattcat
2341 taattggtgc ttctccaaaa cagaccaatc tgactaaagc acaagaggta cattcaactg
2401 gaataaacag gccaaagaga actgggtcct tagcactatt ttacagaaag gtctatcatt
2461 tggcaagtgt acgcttacgt gatctatgtc taaaactgga tgtttcaaat gagttacgaa
2521 ggaagatatg gacgtgtttt gaattcactt tagttcactg tcctgatcta atgaaagaca
2581 ggcatttgga tcagctcctc ctttgtgcct tttatatcat ggcaaaggta acaaaagaag
2641 aaagaacttt tcaagaaatt atgaaaagtt ataggaatca gccccaagct aatagtcacg
2701 tatatagaag tgttctgctg aaaagtattc caagagaagt tgtggcatat aataaaaata
2761 taaatgatga ctttgaaatg atagattgtg acttagaaga tgctacaaaa acacctgact
2821 gttccagtgg accagtgaaa gaggaaagaa gtgatcttat aaaattttac aatacaatat
2881 atgtaggaag agtgaagtca tttgcactga aatacgactt ggcgaatcag gaccatatga
2941 tggatgctcc accactctct ccttttccac atattaaaca acagccaggc tcaccacgcc
3001 gcatttccca gcagcactcc atttatattt ccccgcacaa gaatgggtca ggccttacac
3061 caagaagcgc tctgctgtac aagttcaatg gcagcccttc taagagtttg aaagatatca
3121 acaacatgat aaggcaaggt gagcagagaa ccaagaagcg agtaatagcc atcgatagtg
3181 atgcagaatc ccctgccaaa cgcgtctgtc aagaaaatga tgacgtttta ctgaaacgac
3241 tacaggatgt tgtcagtgaa agagcaaatc attaatgttg ttcttgtttc tatgataaaa
3301 gcactttcag attgttctgc agaaagttgg agctctgtcc ttcaaacctt ttagccctat
3361 agatgataaa tatcactggg ttataagaaa aaattgcaca aaaattatgt gctttttaaa
3421 atatttatcc aaaatgtagt tgacagagat gtattttgag ttggattgga aaggaatatt
3481 ttaagtgcct tttaaaaata ctaatagtcc ggccaggcgc tgtggctcac gcctctaatc
3541 ccaggacttt gggaggccaa ggcgggcaga tcaccggagt caggagttcg agaccagcct
3601 gaccaacatg gagaaacccc atctctacta aaaatacaaa attagccggg tggtgtggcg
3661 catgcctata atcccagcta cttgggaggc tgaggcagaa ttgcttgaac ccaggaagcg
3721 gaggttgtgg tgagccaagg ttgcgccact gcactccagc ctgggcaaca agagtaaaac
3781 tccatctcaa aaaatatata tatatatata aatagggaat ttttttaat gtttgctcct
3841 tgagttttca agatgaaata aggagaaacc ccataacttt ttagctctct tttaaaaata
3901 aatgtctcct tctgtgttct gtaatatgag gataaataat ctgcttttga tagcaaaaaa
```