antitumor effect (*13*). Infection of the tumor vasculature with retroviral vectors that express immunologic activators could provoke a local response, which might enhance tumor rejection.

REFERENCES AND NOTES

1. E. G. Nabel, G. Plautz, F. M. Boyce, J. C. Stanley, G. J. Nabel, *Science* **244**, 1342 (1989).
2. J. M. Wilson *et al.*, *ibid.*, p. 1344.
3. O. Danos and R. C. Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6460 (1988).
4. J. S. Reitman, R. W. Mahley, D. L. Fry, *Atherosclerosis* **43**, 119 (1982).
5. A. M. Dannenberg and M. Suga, in *Methods for Studying Mononuclear Phagocytes*, D. O. Adams, P. J. Edelson, H. S. Koren, Eds. (Academic Press, New York, 1981), pp. 375–395.
6. E. G. Nabel, G. Plautz, G. J. Nabel, unpublished observations.
7. T. J. C. VanBerkel, J. F. Nagelkerke, J. K. Kruijt, *FEBS Lett.* **132**, 61 (1981); J. M. Wilson, D. E. Johnston, D. M. Jefferson, R. C. Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* **85**, 4421 (1988); T. Kodama *et al.*, *Nature* **343**, 531 (1990); L. Rohrer, M. Freeman, T. Kodama, P. Penman, M. Krieger, *ibid.*, p. 570.
8. Animals were prepared for surgery, and the catheter was inserted as described in the legend to Fig. 1. Medium containing polybrene was instilled for 30 min. Examination of the artery at 4 weeks revealed no β-galactosidase activity but hyperplasia in the intimal layer.
9. P. M. Steele *et al.*, *Circ. Res.* **57**, 105 (1985); D. P. Faxon *et al.*, *Arteriosclerosis* **4**, 189 (1984); W. McBride, R. A. Large, L. D. Hillis, *N. Engl. J. Med.* **318**, 1734 (1988).
10. Detection of β-galactosidase was performed by means of the PCR with primers for β-galactosidase as follows: sense: TGG AGC GCC GAA ATC CCG AAT CTC TAT CGT; antisense: TAG CCA GCG CGG ATC ATC GGT CAG ACG ATT.
11. D. A. Dichek *et al.*, *Circulation* **80**, 1347 (1989).
12. J. R. Wilentz *et al.*, *ibid.* **75**, 636 (1987); L. Schwartz *et al.*, *N. Engl. J. Med.* **318**, 1714 (1988).
13. R. I. Tepper, P. K. Pattegale, P. Leder, *Cell* **57**, 503 (1989); E. R. Fearon *et al.*, *ibid.* **60**, 397 (1990).
14. J. Price, D. Turner, C. Cepko, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 156 (1987).
15. R. J. Mayer and J. H. Walker, in *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London, 1987), pp. 263–275.
16. Lymphocytes isolated from blood by buoyant density centrifugation were incubated in RPMI 1640 containing 10% fetal bovine serum, penicillin, streptomycin, and 5×10^{-5} M 2-mercaptoethanol alone (unstimulated) or with phorbol myristate acetate (PMA) (10 ng/ml) or PMA and phytohemagglutinin (PHA) (2 μg/ml) (stimulated). After filtration, 200-μl aliquots of supernatant were added to 3T3 cells containing a β-galactosidase (β-gal-at-gag) provirus (3T3-BAG) (*14*) in the presence of polybrene (8 μg/ml). 3T3-BAG cells were passaged up to four times, and filtered supernatant was tested for its ability to infect an appropriate target cell and express β-galactosidase. Presence of virus was assessed by infecting 3T3 cells at 10% confluence and staining with X-gal chromagen when cells had reached confluence and was estimated to be sensitive to <10 infectious particles per milliliter. No colonies were detected in normal pig plasma or lymphocytes, nor in any experimental animals, whereas supernatant from a helper-producing Ψ-AM subline [R. Mann, R. C. Mulligan, D. Baltimore, *Cell* **33**, 153 (1983); R. D. Cone and R. C. Mulligan, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6349 (1984)] shown previously to produce replication-competent helper virus, produced colonies too numerous to count. Reverse transcriptase was tested in 500 μl of supernatant from lymphocytes maintained as described above. Supernatants were concentrated by centrifugation at 13,000 rpm in a microfuge at 4°C for 60 min and resuspended in 50 μl of a reaction mixture containing polyadenylic acid (5 μg/ml) (Pharmacia), oligo(dT) primer (14 nucleotides) (1.57 μg/ml), 15 mM tris-Cl (pH 7.8), 7.5 mM KCl, 2 mM dithiothreitol, 5 mM $MgCl_2$, 5 mM $MnCl_2$, 0.05% NP-40 and 25 nM [^{32}P]dTTP (deoxythymidine triphosphate) (Amersham, 400 Ci/mM). Samples were incubated at 37° for 90 min, and 10 μl was placed onto Whatman DE81 paper. Filters were washed five times with 2× SSC (saline sodium citrate) for 5 min per wash, twice with 95% ethanol for 2 min and quantitated in a scintillation counter. Duplicate samples of supernatant from unstimulated, PMA-, or PMA- and PHA-stimulated lymphocytes at 24, 48, or 96 hours after culture showed no difference when normal uninfected pigs were compared to infected experimental animals and were not significantly above background (<100 cpm). Helper virus–positive Ψ-AM cell line supernatant and 0.2 unit of recombinant Moloney murine leukemia virus reverse transcriptase (Boehringer Mannheim) yielded values significantly above background, 4000 to 5000 or 6000 to 7000 cpm, respectively. Sensitivity of the reverse transcriptase assay was estimated to be 0.003 unit.

17. We thank A. McDermott for secretarial assistance, and F. Collins, J. Leiden, and J. Wilson for helpful discussions. G. Plautz is a recipient of a National Research Service Award (GM-13457). This work was supported in part by funds from the Department of Internal Medicine, University of Michigan Medical Center, and grants from NIH [Ai 29179 (G.J.N.) and DK 42706 (E.G.N.)].

11 April 1990; accepted 29 June 1990

5-Methylcytosine as an Endogenous Mutagen in the Human LDL Receptor and p53 Genes

WILLIAM M. RIDEOUT III, GERHARD A. COETZEE, ARIA F. OLUMI, PETER A. JONES*

Direct genomic sequencing revealed that cytosine residues known to have undergone a germ-line mutation in the low density lipoprotein receptor gene or somatic mutations in the p53 tumor suppressor gene were methylated in all normal human tissues analyzed. Thus, these mutations should be scored as transitions from 5-methylcytosine to thymine rather than from cytosine to thymine. Methylated cytosines occur exclusively at CpG dinucleotides, which, although markedly underrepresented in human DNA, are sites for more than 30 percent of all known disease-related point mutations. Thus, 5-methylcytosine functions as an endogenous mutagen and carcinogen in humans, in that methylation seems to increase the potential for mutation at cytosine residues at least by a factor of 10.

LESS THAN 1% OF THE BASES OF human DNA are 5-methylcytosine. Methyl groups that occur on both C residues in the double-stranded palindrome CpG account for more than 90% of the methylated C residues (*1*). Although the CpG sequence is underrepresented by a factor of five in the vertebrate genome (*2*), it is the site of a disproportionately high number of human germ-line point mutations. Estimates suggest that 35% of point mutations causing human genetic disorders have occurred at CpG dinucleotides (CpGs), and over 90% of these were transitions from C to T or corresponding G-to-A transitions (*3*). CpGs may also be overrepresented among sites of somatic mutation in tumor suppressor genes, such as the p53 or retinoblastoma genes, in which about 40% of reported point mutations are localized to CpGs (*4, 5*).

Methylation of CpGs in normal tissues might increase the probability of mutations at such sites because of the ability of 5-methylcytosine to undergo deamination, resulting in a thymine (*6*). Because of the inverse relation between the presence of 5-methylcytosine and gene expression (*7*), active genes might be expected to be unmethylated. Despite this inverse relation often observed between methylation and gene expression, many genes (such as HPRT and PGK) are expressed with methylation in their coding sequences (*8*). We therefore used ligation-mediated polymerase chain reaction (PCR) genomic sequencing (*9*) to analyze directly CpGs that have undergone either a germ-line mutation in the low density lipoprotein (LDL) receptor or somatic mutations in the p53 tumor suppressor genes (Fig. 1). These CpGs were all methylated in several human tissues obtained from five individuals.

We first obtained an overall indication of the methylation status of the LDL receptor and p53 genes by digesting DNA obtained from sperm or white blood cells (WBC)

Urological Cancer Research Laboratory, Kenneth Norris Jr. Comprehensive Cancer Center, University of Southern California, Los Angeles, CA 90033.

*To whom correspondence should be addressed.

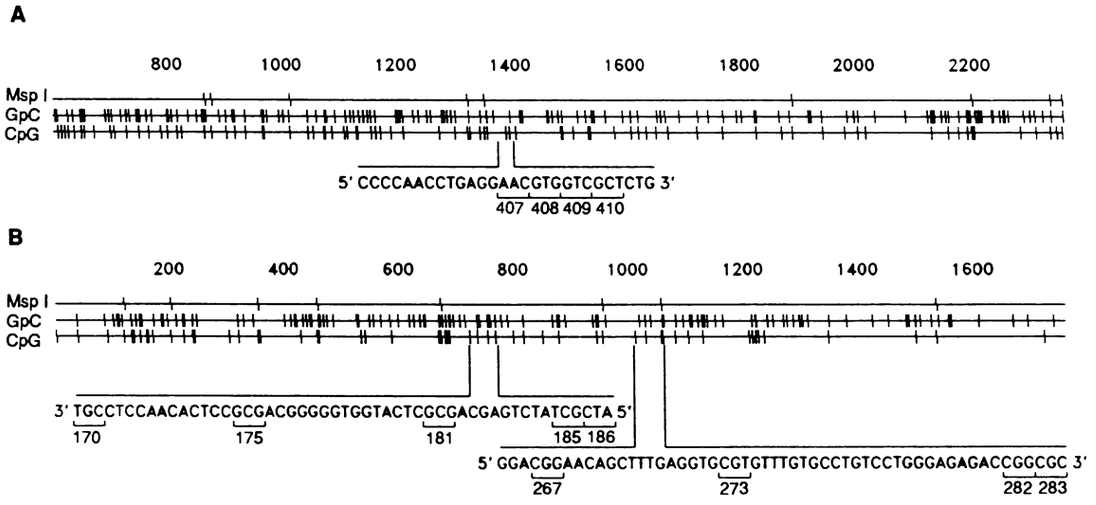

Fig. 1. Schematic representation of the cDNA sequences (nucleotide numbers as indicated) of (**A**) LDL receptor and (**B**) p53. Only a portion of the LDL receptor cDNA is shown (exon 9 spans nucleotides 1295 to 1466). The p53 coding sequence spans nucleotides 215 to 1396. Msp I recognition sequences (CCGG), GpC, and CpG dinucleotides are indicated. CpGs (compared to GpCs) are relatively less frequent in both sequences. The regions used for sequencing (see Fig. 3) are given, and the codons containing the relevant CpGs are indicated. In the region of p53 that contains codons 170 to 186, the sequence of the noncoding strand is given.

with either the methylation-insensitive enzyme Msp I (recognition sequence, CCGG) or the methylation-sensitive isoschizomer Hpa II. White blood cell DNA cut with Msp I and probed with an exon-9 fragment of the LDL receptor cDNA showed no large fragments (Fig. 2A); the low molecular weight fragments generated by Msp I digestion were not visible on the blot. Hpa II digestion of DNA from either WBC or sperm resulted in only a high molecular weight band (~13.5 kb), indicating that these CCGG sites were extensively methylated. Since the LDL receptor gene is expressed in lymphocytes (*10*), methylation in this region does not apparently silence gene expression. Similarly, comparison of Msp I and Hpa II digests of WBC and sperm DNA probed with the p53 cDNA indicated methylation at the five restriction sites (Fig. 2B) (*11*).

Because many C to T transitional mutation hot spots are not part of Msp I recognition sites, ligation-mediated PCR genomic sequencing (*9*) was used to focus on the CpG sites of interest. Methylation patterns in genomic DNA are readily detectable in this sequencing procedure because 5-methylcytosine does not react with hydrazine and does not appear in the cytosine ladder of the sequencing gel (*12*). An unmethylated cloned plasmid harboring the appropriate cDNA was analyzed in parallel. The CpG that contributes to codons 407 and 408 and the one contributing to codons 409 and 410 in the LDL receptor gene are both methylated in genomic DNA from WBC, sperm, and muscle but are unmethylated in plasmid DNA (Fig. 3A). Similar results were obtained from DNA extracted from blood and sperm of two other unrelated individuals, and DNA extracted from cultured skin fibroblasts (*13*). The CpG that spans codons 407 and 408 is mutated from G to A in one of the founder defective genes resulting in familial hypercholesterolemia in South African Afrikaners (*14*), changing Val to Met at amino acid 408. The occurrence of the 5-methylcytosine on the complementary strand in this position in human sperm is therefore consistent with the hypothesis that methylation has contributed to the chance for mutation at this site.

The CpG dinucleotides in the p53 gene at codons 170, 175, 181, and the one spanning codons 185 and 186 are all methylated in the germ line and in somatic tissues of all tested individuals (Fig. 3B) (*13*). The CGC (Arg) of codon 175 was recently reported to be mutated in four independent tumors (three colon tumors and one brain tumor); in all four cases, a change from CGC to CAC resulted in a change from Arg to His (*4*). This corresponded to a transition from 5-methylcytosine to T in the noncoding strand. We also assessed the methylation status of the CpG dinucleotide in codon 273, which similarly has undergone point mutations in three independent tumors (a brain tumor, a breast tumor, and a colon tumor), in two cases from CGT to CAT resulting in Arg to His, and in another case from CGT to TGT resulting in Arg to Cys (*4*). The CpGs in codon 273 and flanking codons (267, 282, and 283) were methylated in all tissues tested (Fig. 3C). Because p53 is considered to be a housekeeping gene and is presumably expressed in these tissues, the results also showed that methylation of these sites was not involved in gene suppression. Extensive methylation of such an actively expressed gene might predispose it to deamination-driven mutations that could alter the function of the gene product.

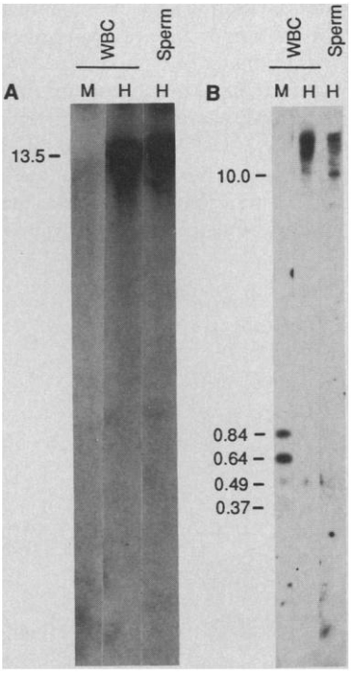

Fig. 2. Methylation analysis of (**A**) LDL receptor and (**B**) p53 genes. DNA extracted from the indicated sources was cut with Msp I (M) or Hpa II (H); digests were separated by electrophoresis, Southern (DNA) blotted, hybridized either to a 264-nucleotide Pst I fragment of LDL receptor cDNA containing exon 9 or to full-length p53 cDNA as described previously, and autoradiographed (*20*). WBC, white blood cells. Fragment sizes are indicated (kilobases).

In the LDL receptor and p53 genes, the CpG sites analyzed, some of which are known to be involved in human mutations, were methylated in normal tissues and showed no correlation between gene expression and methylation. Cytosine methylation, thought to contribute to prokaryotic mutations (*15*) and the extensive CG to TA transitions in *Neurospora crassa* (*16*), may also

be a major cause of human somatic mutations. A disproportionately high percentage of restriction enzyme polymorphisms in the human genome occurs in enzyme recognition sequences containing CpG (*17*). Germline mutations, possibly as a result of methylation, may be common at these sites as well. Thus, methylated cytosine may be commonly present at CpGs outside of CpG islands (*18*) and may be prone to mutation.

A role for cytosine methylation in germline mutations that cause disease has been proposed (*3*), but the contribution of 5-methylcytosine to somatic mutation resulting in human disease has not been recognized previously. Observation of CpG methylation at two sites in the p53 gene, accounting for 7 of 21 (33%) of observed point mutations in tumors reduced to homozygosity for chromosome 17p (*4*), suggests that deamination-driven point mutations may be of importance in tumorigenesis. Two of the mutations (at codon 248) involve another CpG that occurs in a CCGG site not directly sequenced, but methylated, since there was no cutting of the p53 gene by Hpa II (Fig. 2B). Thus, as many as 43% (9 of 21) of p53 somatic mutations may be due to the presence of 5-methylcytosine. Clearly, this is not the only mechanism of mutagenesis in the p53 gene (*4*), but it might be the most frequent single cause of point mutations. All the point mutations that occurred independently more than once were these transitions. Furthermore, C to T (or G to A) transition mutations at CpG sites have occurred in three of the eight point mutations described in the retinoblastoma tumor suppressor gene (*5*), as well as in the one point mutation recently identified in a candidate tumor suppressor gene (DCC gene) on chromosome 18q involved in colorectal cancers (*19*). Deamination-induced mutations presumably occur at many genomic sites in somatic cells, but are of little consequence to the organism unless they lead to altered function of gene products important in growth control, clonal expansion of the cell, and tumorigenesis.

The relevance of the methylation to mutations in p53 can perhaps best be appreciated by the following calculation. There are 82 CpGs in the 2362 nucleotides of the p53 double-stranded coding sequence (Fig. 1). Therefore, at most ~3.5% of the sequence is contributing to 33 to 43% of the point mutations, each of which is a transition from 5-methylcytosine to T (or a corresponding transition from G to A) only. Although other explanations might exist for this bias, we suggest that methylation of the C in the dinucleotide palindrome is the main, if not only, reason. Our observations suggest that 5-methylcytosine acts as an endogenous mutagen and that a high percentage of mutations in human tumor suppressor genes may be induced by deamination of methylated cytosines to form thymines.

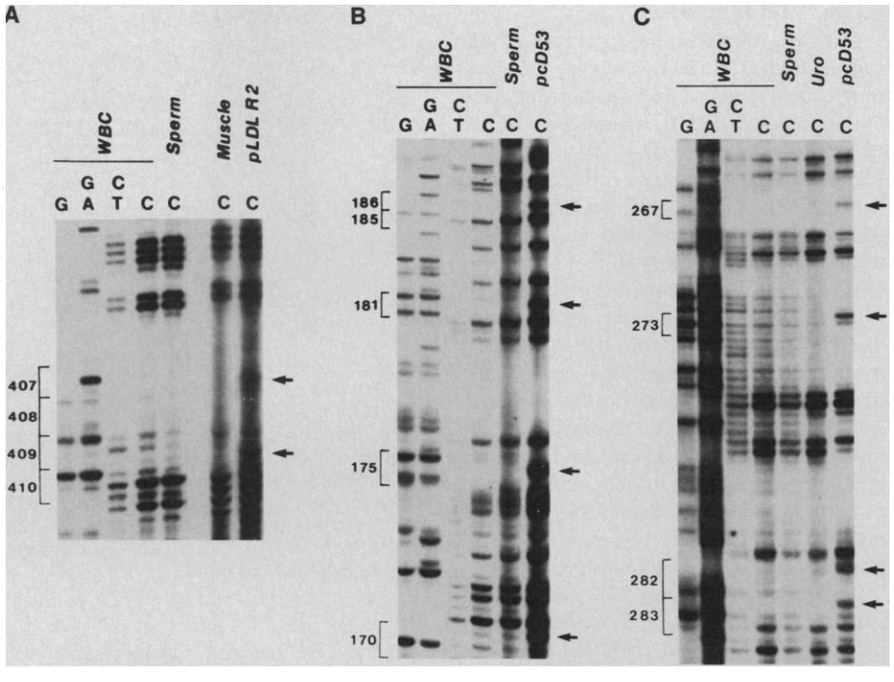

Fig. 3. Genomic sequence data for DNA isolated from the indicated sources. Arrows indicate methylated cytosine gaps in the C sequence ladders, and the corresponding unmethylated cytosine bands in the ladders from the cloned genes. (**A**) LDL receptor, (**B**) p53 codon 175 region, and (**C**) p53 codon 273 region. Base-specific chemical cleavage of DNA was followed by ligation-mediated PCR. The sequence ladders were visualized by autoradiography of sequence gels after end-labeling with ^{32}P and primer extension (*9*). For each site analyzed, a set of three oligonucleotides was used. For the LDL receptor (*21*) and p53 codon 273 regions (*22*), the coding strands were sequenced; for the p53 codon 175 region, the noncoding strand was sequenced (*23*). Relevant codons are indicated with square brackets. WBC, white blood cells; Uro, normal urothelium; pLDLRs, plasmid; pcD53, plasmid.

REFERENCES AND NOTES

1. A. P. Bird, *J. Mol. Biol.* **118**, 49 (1978); M. Ehrlich *et al.*, *Nucleic Acids Res.* **10**, 2709 (1982).
2. P. Grippo, M. Iaccarino, E. Parisi, E. Scarano, *J. Mol. Biol.* **36**, 195 (1968).
3. D. N. Cooper and Y. Youssoufian, *Hum. Genet.* **78**, 151 (1988).
4. J. M. Nigro *et al.*, *Nature* **342**, 705 (1989).
5. D. W. Yandell *et al.*, *N. Engl. J. Med.* **321**, 1689 (1989).
6. T. Lindahl, *Prog. Nucleic Acids Res. Mol. Biol.* **22**, 135 (1979); M. Ehrlich, K. F. Norris, R. Y.-H. Wang, K. C. Kuo, C. W. Gehrke, *Biosci. Rep.* **6**, 387 (1986).
7. W. Doerfler, *Annu. Rev. Biochem.* **52**, 93 (1983); P. A. Jones and J. D. Buckley, *Adv. Cancer Res.* **54**, 1 (1990).
8. S. F. Wolf *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 2806 (1984); D. H. Keith *et al.*, *Mol. Cell. Biol.* **6**, 4122 (1986).
9. P. R. Mueller and B. Wold, *Science* **246**, 780 (1989); G. P. Pfeifer, S. D. Steigerwald, P. R. Mueller, B. Wold, A. D. Riggs, *ibid.*, p. 810.
10. Y. K. Ho, M. S. Brown, D. W. Bilheimer, J. L. Goldstein, *J. Clin. Invest.* **58**, 1465 (1976).
11. M. Lübbert, C. W. Miller, J. Kahan, H. P. Koeffler, *Oncogene* **4**, 643 (1989).
12. G. M. Church and W. Gilbert, *Proc. Natl. Acad. Sci. U.S.A.* **81**, 1991 (1984).
13. W. Rideout, G. Coetzee, A. Olumi, unpublished data. CpGs in p53 were also methylated in urothelium, liver, muscle, and cultured skin fibroblasts.
14. E. Leitersdorf, D. R. Van der Westhuyzen, G. A. Coetzee, H. H. Hobbs, *J. Clin. Invest.* **84**, 954 (1989).
15. C. Coulondre, J. H. Miller, P. J. Farabaugh, W. Gilbert, *Nature* **274**, 775 (1978).
16. E. U. Selker and J. N. Stevens, *Proc. Natl. Acad. Sci. U.S.A.* **82**, 8114 (1985); *Mol. Cell. Biol.* **7**, 1032 (1987).
17. D. Barker, M. Schafer, R. White, *Cell* **36**, 131 (1984).
18. A. P. Bird, *Trends Genet.* **3**, 342 (1987).
19. E. R. Fearon *et al.*, *Science* **247**, 49 (1990).
20. Y. C. Tsai *et al.*, *Cancer Res.* **50**, 44 (1990).
21. Oligonucleotide primers used for PCR and sequencing of the LDL receptor codon 407 and 408 region: (i) 5′-CAGATCATTCTCTGGGAC-3′; (ii) 5′-TTCTCTGGGACAGGTCAGACCAGT-3′; (iii) 5′-GGGACAGGTCAGACCAGTAGATTCTATTGC-3′.
22. Primers used for p53 codon 273 region: (i) 5′-CTCGCTTAGTGCTCCCT-3′; (ii) 5′-GCAGCTCGTGGTGAGGCTCCC-3′; (iii) 5′-GCAGCTCGTGGTGAGGCTCCCCTTTCTTG-3′.
23. Primers used for p53 codon 175 region: (i) 5′-TGGGTTGATTCCACACC-3′; (ii) 5′-GCGTCCGCGCCATGGCCATCTACA-3′; (iii) 5′-GCGTCCGCGCCATGGCCATCTACAAGC-3′.

24. We thank D. W. Russell and L. Crawford for providing the pLDL R2 and pcD53 probes, respectively. Supported by the United Cancer Research Society and by the National Cancer Institute (grant R35 CA49758). G.A.C. was on sabbatical leave from the University of Cape Town, partially funded by the South African Medical Research Council. W.M.R. is supported by a training grant (T32 CA09569) from the National Cancer Institute.

30 April 1990; accepted 11 July 1990