Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 3286–3289, April 1991
Cell Biology

Patterns of gene expression that characterize the colonic mucosa in patients at genetic risk for colonic cancer

LEONARD H. AUGENLICHT*†, JOHN TAYLOR‡, LEIGH ANDERSON§, AND MARTIN LIPKIN¶

*Albert Einstein Cancer Center, Bronx, NY 10467; ‡Argonne National Laboratory, Argonne, IL 60439; §LSB Corporation, Rockville, MD 20850; and ¶Memorial Sloan–Kettering Cancer Center, New York, NY 10021

Communicated by Frank Lilly, January 14, 1991

ABSTRACT **We have used a computer-driven scanning and image-processing system to identify a panel of 30 cDNA clones whose pattern of expression in individual biopsy specimens distinguishes the flat, normal-appearing colonic mucosa of patients in two genetic groups at high risk for development of colorectal cancer from that of normal colonic mucosa in low-risk individuals. The two high-risk groups, familial adenomatous polyposis and hereditary nonpolyposis colon cancer, are indistinguishable based on the pattern of expression of the 30 selected clones. This suggests that the extensive pleiotropic effects of the inherited loci, which may play an important role in the mechanism of increased risk and early onset of the disease, are similar in these populations.**

Patients with familial adenomatous polyposis (FAP) inherit a defective gene (*APC*, adenomatous polyposis coli) located on chromosome 5q21 predisposing them for development of multiple benign colonic adenomas that are at very high risk for progression to colonic cancer (1, 2). The genetic events in progression from the earliest adenomas to carcinoma have been studied in detail in both FAP patients with the 5q21 inherited abnormality and patients with "sporadic" colonic cancer (3–5): they involve the accumulation of multiple genetic abnormalities that include a high frequency of mutations of members of the *ras* gene family (6–8), deletions and mutations on chromosome 17p of the p53 gene (8–12), and deletions on 18q of the *DCC* gene (8, 11–13). However, these abnormalities, while common, are not always present in carcinomas, nor do they occur in a rigidly defined order during progression (8, 11). There are additional deletions detected on every chromosome in human colonic carcinomas, the frequency and pattern of which are highly variable (11), and amplifications of the c-*myc* gene, rare in common sporadic colorectal carcinomas, are frequently found in highly aggressive unique subtypes of colorectal tumors (14). Therefore, human colonic cancer develops with profoundly heterogeneous molecular alterations.

Little attention has been given to the molecular changes that develop in the histologically normal flat mucosa of individuals with FAP or in the unaffected mucosa of other groups at increased risk for colonic cancer. The somatic alterations documented during progression have thus far been detected only after neoplasms develop, because the genetic changes are selected by clonal outgrowth and become highly enriched in the tumors that can be recognized and sampled (1, 6, 15).

Using an alternative approach that can assay the relative level of expression of each of many thousands of cloned sequences expressed in small biopsy specimens of colonic tissue, we reported that flat mucosa of FAP patients who are at increased risk for colonic cancer contained a higher frequency of alterations in gene expression when compared with low-risk mucosa than either the benign adenomas or carcinomas that subsequently arise (16). In this report, we have extended the work to identify a panel of sequences whose pattern of expression characterizes colonic mucosa of two groups at high genetic risk for colon cancer: FAP and hereditary nonpolyposis colon cancer (HNPCC) (17–19).

MATERIALS AND METHODS

Computerized Scanning. The methods used have been described in detail (16, 20, 21). The approach is made necessary by the very small amounts of poly(A)$^+$ RNA (<50 ng) that can be isolated from small colonic biopsy samples of normal mucosa. In brief, a labeled cDNA probe is prepared from the poly(A)$^+$ RNA from each biopsy and is hybridized to an arrayed replica of a reference library of cDNA sequences cloned from the HT29 human colon carcinoma cell line. After autoradiographic exposure, films are scanned, the images are digitized, background and other corrections are incorporated by image processing, and a numerical index of hybridization based on maximal pixel intensity is assigned to each of 379 clones analyzed (16, 20). Hybridization data for each clone are expressed as the ratio to the mean level of hybridization to all 379 clones, therefore standardizing for small differences in probe specific activity, hybridization and washing stringency, and film exposure between experiments. Changes in gene expression detected by this method can be confirmed by quantitative dot blot and Northern blots (21).

Population Groups. The population groups investigated have been described (16–20). The data were generated from biopsy samples of the normal-appearing flat mucosa of 7 patients with FAP and 12 patients from families with HNPCC, and from the normal mucosa of 6 individuals with no colonic or other cancer in their families for several generations.

RESULTS

Using a computerized scanning and image-processing system to assay the relative level of expression of each of large numbers of cloned cDNA sequences in very small colonic biopsy samples, we previously compared the expression of each of 379 clones in high-risk colonic mucosa of FAP patients to low-risk mucosa of patients with no colonic or other cancer for several generations: a >3-fold increase or decrease in mean expression of each of ≈25% of the sequences was detected (16). From this data base, we selected 30 clones that characterized the FAP high-risk mucosa in terms of mean level, and range, of expression (Fig. 1). Comparison of the distribution of expression for the FAP samples to the low-risk samples by a rank sum test demon-

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: FAP, familial adenomatous polyposis; HNPCC, hereditary nonpolyposis colon cancer.

†To whom reprint requests should be addressed at: Department of Oncology, Montefiore Medical Center, 111 East 210th Street, Bronx, NY 10467.

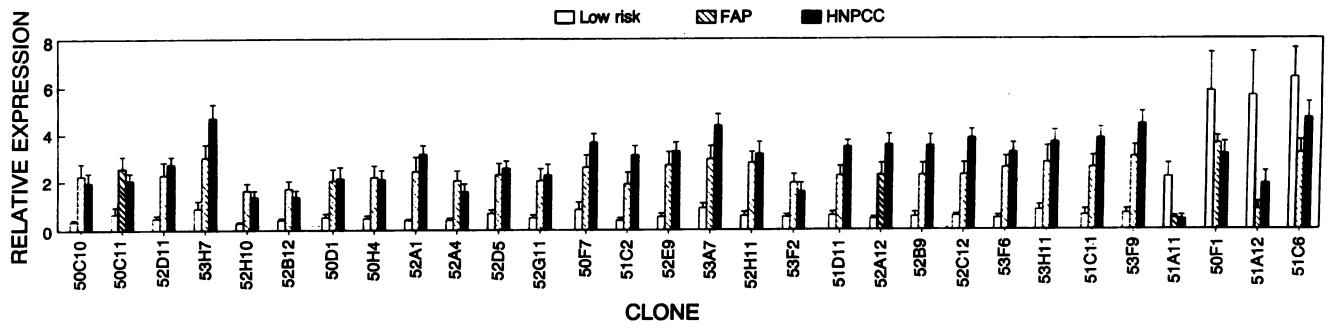

FIG. 1. Mean level of expression of selected clones in low-risk mucosa compared to FAP and HNPCC high-risk mucosa. Relative levels of expression of each of 379 cloned cDNA sequences were determined for each biopsy sample. Values for tissues of each risk group are plotted for 30 of these clones as mean + SEM. For each clone, the three bars correspond to (left to right) low-risk mucosa (open), FAP high-risk mucosa (hatched), HNPCC high-risk mucosa (solid).

strated that 13 of the 30 clones differed with $P < 0.01$, and 16 with $P < 0.05$. This suggested that the selection of these 30 clones out of the 379 screened to distinguish the FAP high-risk mucosa, and the differences in expression, was not due to random variation. To pursue this, a second screen was done with biopsy specimens of flat mucosa from HNPCC families, another population group at increased genetic risk for development of colorectal cancer (17–19). These results are also presented in Fig. 1. Statistical analysis of the HNPCC data compared with the low-risk data (rank sum test) demonstrated that the distribution of expression for 25 of 30 of the clones differed with $P < 0.01$ and for another 2 of the 30 with $P < 0.05$.

The data of Fig. 1 suggest that the two high-risk groups with increased genetic susceptibility to colonic cancer have similar patterns of expression of these 30 cloned gene sequences. Linear regression analysis comparing the data for the two high-risk groups is presented in Fig. 2. The correlation coefficient is 0.77, which is highly significant at $P < 0.01$. Therefore, FAP and HNPCC are indistinguishable in the expression of this panel of 30 cDNA clones. This may be related to the recent observation in a pedigree with familial polyps that the inherited locus maps to the same location (5q21) as the FAP gene (5). However, whether or not the FAP gene is the locus of the HNPCC inherited defect in the families in this study, the data of Fig. 2 suggest that pleiotropic effects of the inherited gene are similar in FAP and these HNPCC families.

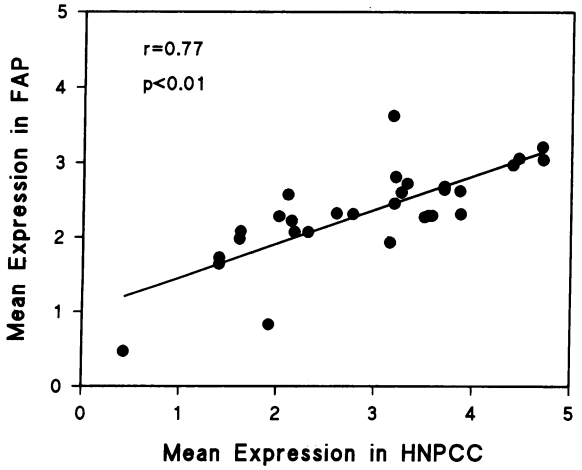

FIG. 2. Comparison of gene expression in FAP and HNPCC high-risk flat mucosa. The 30 selected clones shown in Fig. 1 were compared for mean level of expression in the high-risk groups. The line represents a linear regression analysis of the data; r is the correlation coefficient.

To evaluate the potential that these markers might discriminate levels of risk for colonic neoplasia, the data for each biopsy sample and each clone were considered individually. Examples of comparisons made are shown in Fig. 3. For each of three clones, two of which increase with risk (50C10 and 52D11) and one of which decreases (51A11), the relative hybridization value for each sample studied from different individuals is shown in each of the risk categories. For each clone, a value can be defined (broken line) that distinguishes most of the high-risk biopsy samples from most of the low-risk samples, although in each case, there is overlap between high and low risk. For example, for 50C10, two FAP and two HNPCC biopsy samples would be misclassified as low-risk, although none of the low-risk samples would be misclassified. Similar analyses were carried out with each biopsy sample for each of the 30 clones (Fig. 4).

In Fig. 4, the 30 clones are ordered (represented by solid circles across the top) as they are in Fig. 1. For each biopsy sample listed on the left, a shaded area is inserted in the appropriate clone column if that clone scores the sample as high-risk, as illustrated by the examples in Fig. 3. On the right, the number of clones that are scored as high-risk for each sample is tabulated. For 6 of 7 of the FAP biopsy samples and 12 of 12 HNPCC biopsies, well above 20 of the 30 clones correctly classify the samples, while for 6 of 6 low-risk biopsy samples, the number is substantially less than

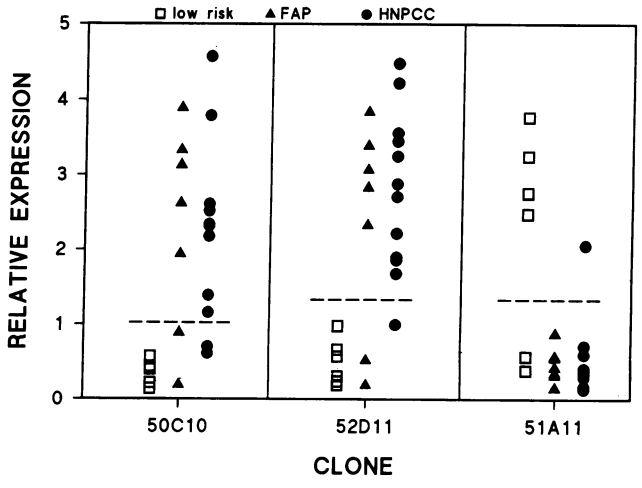

FIG. 3. Relative expression of individual clones in each biopsy sample. Individual data points for three of the selected clones of Figs. 1 and 2 are shown for each biopsy sample, rather than the mean level of expression. □, Low-risk mucosa; ▲, FAP high-risk mucosa; ●, HNPCC high-risk mucosa. A value (illustrated by broken line) can be defined for each clone that distinguishes most of the low-risk biopsy samples from most of the high-risk samples.

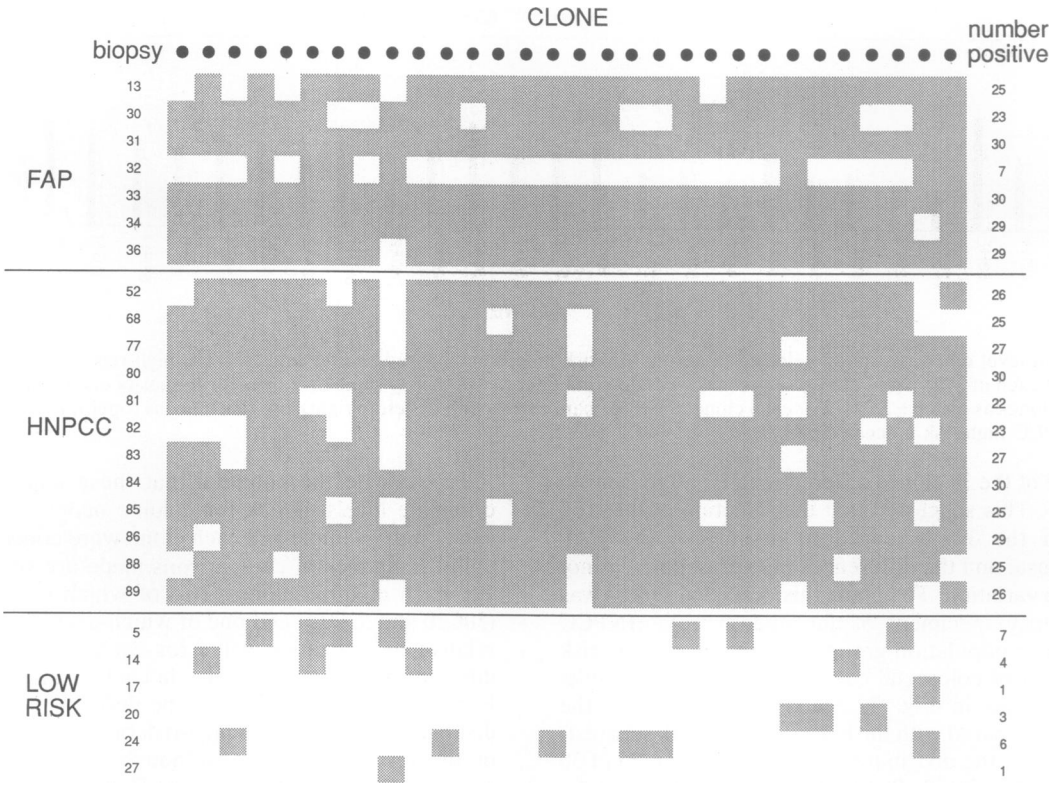

FIG. 4. Pattern of expression of selected clones in the low- and high-risk biopsy samples. Data for each low- and high-risk biopsy sample are summarized for each of the selected clones (as shown in the examples of Fig. 3). Each biopsy sample is identified by its number at left; each of the 30 selected clones, in the same order as that shown in Fig. 1, is represented by a solid circle at the top. For each biopsy sample, a shaded area was placed in the clone column if the clone categorized the sample as high-risk, as illustrated for the examples in Fig. 3. The numbers at right are the number of clones of the 30 selected for which the biopsy sample was so categorized as high-risk.

10 of the 30 clones. Therefore, although for any individual clone there are false negatives among the data for the high-risk tissues, and false positives among the low-risk tissues, the pattern of expression of these 30 clones clearly distinguishes ≈95% of the high-risk samples from all of the low-risk samples. The sole exception is FAP patient 32, whom only 7 of 30 clones classify as high-risk. Since the pattern for this patient is not intermediate, but so similar to the low-risk biopsy samples, this may represent either a misclassification of the sample or a sampling variable in taking the biopsy specimen that is not understood. However, the data for this sample reinforce the conclusion that the expression of these 30 clones is coordinately regulated in the colonic mucosa.

DISCUSSION

These experiments were designed to provide an analysis of a cross section of the many changes in gene expression that accompany both neoplastic transformation and risk for colonic tumor development, and to identify panels of sequences that could be used to characterize tissue phenotype (16, 20). While only sequences that are relatively abundant can be quantitated by this method, and the source of the reference cDNA library used, in this case, the colon carcinoma cell line HT29, further limits the analysis to sequences expressed in that cell line (16, 20), the method has fulfilled these original objectives. It has been suggested that the complex pattern of deletions in colorectal tumors detected by analysis of restriction fragment length polymorphism, termed the tumor allelotype, may similarly be useful as "molecular correlates of tumor behavior" (11, 22).

The data presented here extend to an HNPCC high-risk group our previous finding of extensive modulations in gene expression in the FAP high-risk mucosa (16). It is not surprising that the inheritance of an abnormal gene in FAP results in a large perturbation of gene expression in the mucosa, since it can be postulated that the inherited defect involves a gene of importance in normal lineages of colonic cell differentiation (16); thus, abnormalities that develop may affect a wide variety of cellular functions and interactions in this complex tissue.

It is more difficult to understand why changes in normal-appearing flat mucosa of FAP patients are more extensive than changes occurring in either the adenomas or carcinomas that subsequently progress from the mucosa at risk (16). The reason may be related to the recent clarification that there are many pathways along which colonic cell transformation may proceed (11, 12, 16, 20, 22), a concept that was clearly enunciated over 30 years ago for mammary tumors by Foulds (23, 24). The high-risk flat mucosa that exhibits the wide range of pleiotropic effects arising from the inherited defect may be initiated along many of these paths, with only one selected in each clonal outgrowth that leads first to adenoma and then to carcinoma (8, 15, 22). Thus, as progression takes place, many of the large number of changes detected in the initiated mucosa may be distributed in different subsets of the heterogeneous population of tumors that arise. The alterations in gene expression, similar to the alterations in genetic deletions (11), therefore become variable among tumors that have progressed along different paths, and hence the total number of statistically significant changes between the population of tumors and the normal flat mucosa are reduced relative to the number of such alterations detected in the early initiated mucosa.

An important hypothesis evolves from this view of inherited colonic cancer: the proximal cause of increased frequency of tumors and earlier age of onset in individuals with increased genetic susceptibility is the large number of mod-

ulations of gene expression that prime the cells along many abnormal developmental pathways, any one of which may progress under the influence of other factors (e.g., dietary mutagens or tumor promoters). The role of the inherited defect is to establish the abnormal pleiotropic effects and consequently an elevated probability of subsequent progression to tumor formation, rather than the initiation of a specific or single biochemical chain of events that inexorably leads to transformation. This is consistent with the variability in numbers of polyps that can be seen in individuals in the same pedigree, in which presumably the same defect is inherited and in which similar diets are consumed (5, 25).

One of the clones in this panel (50F1, Fig. 1) has been identified as the mitochondrial gene for subunit III of cytochrome oxidase (21). Alterations in expression of this sequence in colonic tumors and transformed cell lines (21) may underlie the reported alterations in structure and function of mitochondria in colorectal cancer (26, 27). This report of changes in the flat mucosa of genetically susceptible individuals is consistent with data indicating that shifts in energy source for human colonic epithelial cells is a determining factor in their normal differentiation both *in vitro* (28, 29) and *in vivo* (30, 31). Further sequence analysis of the 30 selected cDNA clones may provide additional insight into the complex mechanisms that place the mucosa at increased risk for development of colon cancer.

This work was supported in part by Grants CA41372, CA40558, and P30-CA1330 from the National Cancer Institute and SIG-7 from the American Cancer Society.

1. Bodmer, W. F., Bailey, C. J., Bodmer, J., Bussey, H. J. R., Ellis, A., Gorman, P., Lucibello, F. C., Murday, V. A., Rider, S. H., Scambler, P., Sheer, D., Solomon, E. & Spurr, N. K. (1987) *Nature (London)* **328,** 614–616.
2. Leppert, M., Dobbs, M., Scambler, P., O'Connell, P., Nakamura, Y., Stauffer, D., Woodward, S., Burt, R., Hughes, J., Gardner, E., Lathrop, M., Wasmuth, J., Lalouel, J.-M. & White, R. (1987) *Science* **238,** 1411–1413.
3. Solomon, E., Voss, R., Hall, V., Bodmer, W. F., Jass, J. R., Jeffreys, A. J., Lucibello, F. C., Patel, I. & Rider, S. H. (1987) *Nature (London)* **328,** 616–619.
4. Ashton-Rickardt, P. G., Dunlop, M. G., Nakamura, Y., Morris, R. G., Purdie, C. A., Steel, C. M., Evans, H. J., Bird, C. C. & Wyllie, A. H. (1989) *Oncogene* **4,** 100–105.
5. Leppert, M., Burt, R., Hughes, J. P., Samowitz, W., Nakamura, Y., Woodward, S., Gardner, E., Lalouel, J.-M. & White, R. (1990) *N. Engl. J. Med.* **322,** 904–908.
6. Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlaan-de Vries, M., van Boom, J. H., van der Eb, A. J. & Vogelstein, B. (1987) *Nature (London)* **327,** 293–297.
7. Forrester, K., Almoguera, C., Han, K., Grizzle, W. E. & Perucho, M. (1987) *Nature (London)* **327,** 298–303.
8. Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Preisinger, A. C., Leppert, M., Nakamura, Y., White, R., Smits, A. M. M. & Bos, J. L. (1988) *N. Engl. J. Med.* **319,** 525–532.
9. Baker, S. J., Fearon, E. R., Nigro, J. M., Hamilton, S. R., Preisinger, A. C., Jessup, J. M., vanTuinen, P., Ledbetter, D. H., Barker, D. F., Nakamura, Y., White, R. & Vogelstein, B. (1989) *Science* **244,** 217–221.
10. Nigro, J. M., Baker, S. J., Preisinger, A. C., Jessup, J. M., Hostetter, R., Cleary, K., Bigner, S. H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F. S., Weston, A., Modali, R., Harris, C. C. & Vogelstein, B. (1989) *Nature (London)* **342,** 705–708.
11. Vogelstein, B., Fearon, E. R., Kern, S. E., Hamilton, S. R., Preisinger, A. C., Nakamura, Y. & White, R. (1989) *Science* **244,** 207–211.
12. Law, D. J., Olschwang, S., Monpezat, J. P., Lefrancois, D., Jagelman, D., Petrelli, N. J., Thomas, G. & Feinberg, A. (1988) *Science* **241,** 961–964.
13. Fearon, E. R., Cho, K. R., Nigro, J. M., Kern, S. E., Simons, J. W., Ruppert, J. M., Hamilton, S. R., Preisinger, A. C., Thomas, G., Kinzler, K. W. & Vogelstein, B. (1990) *Science* **247,** 49–56.
14. Heerdt, B. G., Molinas, S., Deitch, D. & Augenlicht, L. H. (1991) *Oncogene* **6,** 125–129.
15. Fearon, E. R., Hamilton, S. R. & Vogelstein, B. (1987) *Science* **238,** 193–196.
16. Augenlicht, L. H., Wahrman, M. Z., Halsey, H., Anderson, L., Taylor, J. & Lipkin, M. (1987) *Cancer Res.* **47,** 6017–6021.
17. Lipkin, M., Winawer, S. J. & Sherlock, P. (1981) *Clin. Bull.* **11,** 13–21.
18. Lipkin, M., Blattner, W. E., Fraumeni, J. F., Lynch, H. T., Deschner, E. & Winawer, S. (1983) *Cancer Res.* **43,** 1899–1904.
19. Lipkin M., Blattner, W. E., Gardner, E. J., Burt, R. W., Lynch, H., Deschner, E., Winawer, S. & Fraumeni, J. F. (1984) *Cancer Res.* **44,** 4201–4207.
20. Augenlicht, L. H. (1989) in *Cell and Molecular Biology of Colon Cancer*, ed. Augenlicht, L. H. (CRC, Boca Raton, FL), pp. 165–186.
21. Heerdt, B. G., Halsey, H. K., Lipkin, M. & Augenlicht, L. H. (1990) *Cancer Res.* **50,** 1596–1600.
22. Fearon, E. R. & Vogelstein, B. (1990) *Cell* **61,** 759–767.
23. Foulds, L. (1954) *Cancer Res.* **14,** 327–339.
24. Foulds, L. (1956) *J. Natl. Cancer Inst.* **17,** 701–711.
25. Bussey, H. J. R. (1975) *Familial Polyposis Coli* (Johns Hopkins Univ. Press, Baltimore).
26. Summerhayes, I. C., Lampidis, T. J., Bernal, S. D., Nadakavukaren, J. J., Nadakavukaren, K. K., Shepherd, E. L. & Chen, L. B. (1982) *Proc. Natl. Acad. Sci. USA* **79,** 5292–5296.
27. Modica-Napolitano, J. S., Steele, G. D. & Chen, L. B. (1989) *Cancer Res.* **49,** 3369–3373.
28. Pinto, M., Appay, M.-D., Simon-Assmann, P., Chevalier, G., Dracopoli, N., Fogh, J. & Zweibaum, A. (1982) *Biol. Cell* **44,** 193–196.
29. Zweibaum, A., Pinto, M., Chevalier, G., Dussaulx, E., Triadou, N., Lacroix, B., Haffen, K., Brun, J.-L. & Rousset, M. (1985) *J. Cell. Physiol.* **122,** 21–29.
30. Roediger, W. E. (1980) *Gut* **21,** 793–798.
31. Harig, J. M., Soergel, K. H., Komorowski, R. A. & Wood, C. M. (1989) *N. Engl. J. Med.* **320,** 23–28.