Am. J. Hum. Genet. 43:638–644, 1988

Localization of the Genetic Defect in Familial Adenomatous Polyposis within a Small Region of Chromosome 5

Yusuke Nakamura,* Mark Lathrop,* Mark Leppert,* Marc Dobbs,§ John Wasmuth,§ Erica Wolff,* Mary Carlson,* Esther Fujimoto,* Karen Krapcho,* Tena Sears,* Scott Woodward,* J. Hughes,‡ Randy Burt,† Eldon Gardner,‖ Jean-Marc Lalouel,* and Ray White*

*Howard Hughes Medical Institute and Department of Human Genetics, and †Division of Gastroenterology, University of Utah Health Sciences Center, and ‡St. Mark's Hospital, Salt Lake City; §Department of Biological Chemistry, California College of Medicine, Irvine; and ‖Department of Biology, Utah State University, Logan

Summary

Familial adenomatous polyposis (FAP), a Mendelian disorder that includes familial polyposis coli (FPC) and Gardner syndrome (GS), has an autosomal dominant mode of inheritance. It is characterized by hundreds to thousands of adenomatous polyps that can progress to carcinoma of the colon, suggesting that the gene that harbors the FAP germ-line mutation may play an important role in the somatic genetic pathway to colon cancer. The defect responsible for FAP was recently mapped to the long arm of chromosome 5 by linkage between the FPC phenotype and a locus defined by DNA probe pC11p11 (D5S71), located at 5q21–22. Because an important next step in the paradigm for identification of a disease gene is to obtain a more precise localization, we isolated and mapped by linkage six additional polymorphic DNA markers in the FAP region. Subsequent linkage analysis in six pedigrees, three having the FPC phenotype and three segregating GS, placed the FAP locus very close to a new marker, YN5.48 (D5S81), that is approximately 17 centimorgans distal to C11p11 on the genetic map. The analysis revealed no evidence of genetic heterogeneity between the two phenotypes, a question that had not been clearly resolved by the earlier studies. The new set of markers in the near vicinity of the FAP locus represents a further step toward isolation of the genetic defect and provides the opportunity for preclinical diagnosis of risk status for colon cancer among individuals in families that are segregating adenomatous polyposis.

Introduction

Colon cancer is a good model for investigations of oncogenesis. It is one of the most common cancers in North America (140,000 new patients/year), and a precancerous stage, the adenomatous polyp, is present for study. Moreover, a familial syndrome of colon cancer exists: familial adenomatous polyposis (FAP), a highly penetrant, autosomal dominant genetic defect. FAP is characterized by hundreds to thousands of polyps in the colon; the chance is nearly 100% that one or more of the polyps will become malignant. By analogy to the paradigm of carcinogenesis in retinoblastoma, the mutant gene in the germ-line cells of affected individuals may be the same as the gene that converts a normal cell to an adenoma or cancer cell following somatic mutation. For all these reasons, isolation of the gene responsible for FAP should be a significant step toward an understanding of the mechanism of carcinogenesis.

In 1986, a partial deletion of the long arm of chromosome 5 (q13–15 or q15–22) in lymphocytes of a patient with Gardner syndrome (GS) and mental retardation was reported (Herrera et al. 1986). Following this finding, two groups (Bodmer et al. 1987; Leppert et al. 1987) mapped the gene responsible for FAP to chromosome 5 when significant linkage was observed be-

Received June 2, 1988; revision received July 11, 1988.

Address for correspondence and reprints: Yusuke Nakamura, M.D., Howard Hughes Medical Institute, 603 Wintrobe Building, University of Utah Health Sciences Center, Salt Lake City, UT 84132.

0002-9297/88/4305-0010$02.00

tween the familial polyposis coli (FPC) phenotype and a DNA marker, pC11p11 (D5S71), that had been mapped by physical methods to chromosome 5q21–22. A number of genetic diseases have now been mapped by linkage analysis, and for two of them—Duchenne muscular dystrophy and chronic granulomatous disease—the genes harboring the mutation have subsequently been isolated (Monaco et al. 1986; Royer-Pokora et al. 1986). Identifying markers closer to a mutant locus after the first evidence of linkage has been detected is a laborious but vital step toward identification of the gene.

FAP is divided on clinical grounds into two types, FPC and GS. GS exhibits extracolonic manifestations in addition to polyposis in the colon: benign tumors such as multiple osteomas, often in mandibular bone or skull; fibrous desmoid or epidermoid cysts; and congenital hypertrophy of retinal pigment epithelium. The initial linkage reports cited above yielded no definitive information about possible genetic heterogeneity between FPC and GS.

To establish a map with relatively high resolution for locating the FAP gene(s) more precisely and to address the question of heterogeneity, we isolated six new DNA probes that define loci close to C11p11. These markers were ordered into a genetic map of a small region of chromosome 5q, by linkage analysis in a panel of 59 3-generation pedigrees. By further linkage analysis in six families segregating FAP, we found that the FAP locus was tightly linked to a new DNA marker approximately 17 centimorgans (cM) distal to C11p11, with no evidence of genetic heterogeneity between FPC and GS.

Material and Methods

Cosmid Library

High-molecular-weight DNA was isolated from somatic hybrid cell line HHW141, in which the human component consists of approximately 40% of the long arm of chromosome 5 (Cirullo et al. 1983), according to methods published elsewhere (Barker et al. 1984). DNA was partially digested with *Sau*3AI and fractionated by centrifugation in a 10%–38% sucrose density gradient at 22,000 rpm for 16 h, with a SW50 rotor (Beckman Co., Ltd.). DNA fractions of 35–45 kb mean size were collected. Cosmid vector pWE15 (Stratagene) was digested with *Bam*HI and treated with calf intestinal phosphatase. Genomic DNA (1 μg) and vector DNA (600 ng) were ligated at room temperature for 2 h (Hayashi et al. 1986). One-fourth of the ligation mixture was packaged using Gigapak gold™ (Stratagene), and after transformation 20 genome equivalents (1,500,000 kb) of independent clones were obtained.

Screening for RFLPs

Cosmid clones with human DNA inserts were selected by colony hybridization with ^{32}P-labeled human lymphocyte DNA. Cosmid DNAs were isolated from 5 ml of overnight culture at 37 C. Bacteria containing cosmid were precipitated and resuspended in 500 μl of 8% sucrose, 5% Triton-X, 50 mM EDTA, and 50 mM Tris-HCl (pH 8.0). After addition of 50 μl of lysozyme (10 mg/ml), samples were put into boiled water for 3 min and centrifuged at 15,000 rpm for 10 min. Supernatants were treated twice with phenol-chloroform-isoamyl alcohol, and DNA was recovered by ethanol precipitation. Under these conditions we were usually able to prepare 3–10 μg of cosmid DNA from 5 ml of culture.

For screening RFLPs, we digested DNAs from six unrelated individuals with *Msp*I, *Taq*I, *Rsa*I, *Bgl*II, *Pvu*II, or *Pst*I. DNA probes were labeled with [^{32}P]-dCTP (Amersham) by random hexanucleotide priming (Feinberg and Vogelstein 1984). To eliminate the background of repetitive sequences, we added 125 μg of human placental DNA to prehybridization solution (50% formamide, 50 mM $NaPO_4$ at pH 6.8, 10 × Denhardt's solution, 5 × NaCl/citrate [1 × = 0.15 M NaCl/15 mM sodium citrate] and 0.1% SDS) and hybridization solution (50% formamide, 50 mM $NaPO_4$ at pH 6.8, 1 × Denhardt's solution, 5 × NaCl/citrate, 0.1% SDS, and 10% dextran sulfate). Additionally, before adding membranes we preincubated [^{32}P]-labeled cosmid DNA in hybridization solution at 42 C for 1 h.

Genotyping

For genotyping with DNA markers, 5 μg of each DNA was digested. Southern blottings were done according to a method described elsewhere (Barker et al. 1984). Prehybridization and hybridization were the same as above. The filters were washed once at room temperature with 2 × NaCl/citrate, 0.5% SDS and twice for 15 min with 0.1 × NaCl/citrate, 0.5% SDS at 65 C. Autoradiography was done at −80 C, using Kodak XAR-5 film and intensifying screens, for 24–72 h.

Linkage Analysis

Linkage analysis was performed with the LINKAGE program package (Lathrop et al. 1985). The genetic map was constructed from genotype information obtained in 59 3-generation reference families (40 of which

belong to the Centre d'Étude du Polymorphisme Humain panel), by using the GMS (Gene Mapping System) algorithm, as described elsewhere (Lathrop et al. 1988). Sex-specific differences in recombination estimates were investigated by comparing the likelihoods of two different models: (1) a model assuming equal male and female recombination estimates and (2) a model assuming a constant ratio of female/male genetic distance throughout the chromosome region spanned by the marker loci.

Results

New RFLP Markers

When we tested 120 cosmid clones containing human DNA inserts for RFLPs, 48 clones revealed polymorphism with one or more enzymes; two RFLPs were of the type that reflect a variable number of tandem repeats (VNTRs) (data not shown). Of 30 polymorphic clones tested in 59 3-generation families, six showed close linkage to locus C11p11 (recombination fraction <.20).

Four of the six new cosmids showed RFLPs with two enzymes (table 1). Table 2 summaries the results of pairwise tests between C11p11 and each of the new loci in the reference families. From these data we constructed a genetic linkage map in the vicinity of C11p11 (fig. 1) with the GMS program (Lathrop et al. 1988). The best-supported gene order was YN5.120 (D5S83)-C11p11-KK5.33(D5S85)-YN5.64(D5S82)-YN5.48(D5S81)-MC5.61(D5S84)-EW5.5(D5S86), although the odds against two alternative orders for the first four markers are <50:1. All other tested alternatives could be rejected with odds of at least 100:1. As no evidence of differences in recombination rates between males and females was found within this portion of chromosome 5 (χ_1^2 = 0.67), equal recombination distances in the two sexes were assumed in subsequent analyses.

Linkage of New Markers to FAP in Affected Families (Table 3)

Affected individuals in K109, K1498, and K1507 exhibit typical GS, with extracolonic tumors such as jaw osteoma, desmoid tumors, and/or epidermoid cysts. K415, K1441, and K1452 show classical familial polyposis without extracolonal disorders. For the haplotypes of *MspI* and *TaqI* polymorphisms from MC5.61 the combined lod score in all six families was 8.05 at

Table 1

New RFLPs Mapped near the FAP Locus

Probe (Locus)	Restriction Enzyme	Allele Size (kb)	Allele Frequency
cYN5.48 (D5S81)	*TaqI*	6.0	.50
		3.6	.50
cYN5.64 (D5S82)	*PstI*	5.2	.61
		5.1	.39
	TaqI	8.0	.15
		6.0	.85
cYN5.120 (D5S83)	*MspI*	9.0	.40
		8.0	.60
cMC5.61 (D5S84)	*MspI*	5.5	.49
		5.0	.51
	TaqI	20.0	.54
		18.0	.46
cKK5.33 (D5S85)	*TaqI*	9.0	.43
		8.0	.57
	PstI	6.0	.23
		8.0	.77
cEW5.5 (D5S86)	*MspI*	5.0	.28
		4.5	.72
	TaqI	6.0	.30
		8.0	.70

Table 2

Two-Locus Lod Scores (below Diagonal) and Recombination Estimates (above Diagonal) between Marker Loci Typed in Reference Families

	C11p11	YN5.48	YN5.64	YN5.120	MC5.61	KK5.33	EW5.5
C11p11		.08	.08	.04	.07	.05	.19
YN5.48	9.7		.11	.03	.04	.09	.05
YN5.64	11.7	12.2		.10	.13	.06	.11
YN5.120	12.5	11.2	7.8		.06	.07	.19
MC5.61	12.0	18.8	9.9	7.4		.10	.02
KK5.33	13.3	11.5	34.1	8.2	12.8		.20
EW5.5	5.9	15.5	16.2	2.8	19.9	5.2	

θ = .04; the haplotypes of YN5.64 at θ = .06. The haplotypes of two enzymes from KK5.33, which is about 3.0 cM from C11p11, showed linkage to FAP with 7% recombination (maximum lod score 7.80 at θ = .07); these recombinants were not informative with C11p11. EW5.5 and YN5.120, at the end of the map in figure 1, show greater than 10% recombination with FAP (lod scores <3).

MCS.61 and KK5.33 showed significant lod scores for FPC and GS separately, with some recombinants. YN5.48, which is flanked by YN5.64 and MC5.61, showed no recombination with FAP (total lod score 8.25). Table 4 lists lod scores for YN5.48 with each pedigree listed separately; combined lod scores in the three FPC families (3.50) or in the three GS families (4.75) indicated significant linkage to this locus for both forms of FAP.

Location Score of the Gene for FAP

The location-score method was used to estimate the location of the gene responsible for FAP with respect to the six new markers and C11p11. To reduce the computation times required for multilocus analysis in such extended pedigrees, we initially considered a reduced map that included only C11p11-KK5.33-YN5.64-YN5.48. As figure 2 shows, FAP has a maximum-likelihood position at the location of YN5.48. Although no recombinants are observed between FAP and C11p11 in our data, the latter is a relatively uninformative locus with low heterozygosity; the odds against the placement of FAP at the C11p11 locus in the reduced map are 43,000:1. With the addition of YN5.120 to the map, the odds against the placement of FAP at C11p11 versus its placement at YN5.48 are 10^5:1. Additional calculations were made with another reduced map, YN5.64-YN5.48-MC5.61-EW5.5; this also gave a maximum location score for FAP at YN5.48 (figure 2, dashed curve). In both reduced maps, all locations of FAP outside the interval spanned by YN5.64-YN5.48-MC5.61 are rejected with odds >10,000:1. The 1-lod-unit confidence interval calculated from the second reduced map gives an upper bound of .03 for recombination between FAP and YN5.48, with FAP either in the interval YN5.64-YN5.48 or in the interval YN5.48-MC5.61.

Discussion

We have constructed a high-resolution map of DNA markers in the vicinity of the gene for FAP on chromosome 5, to define the location of the deleterious muta-

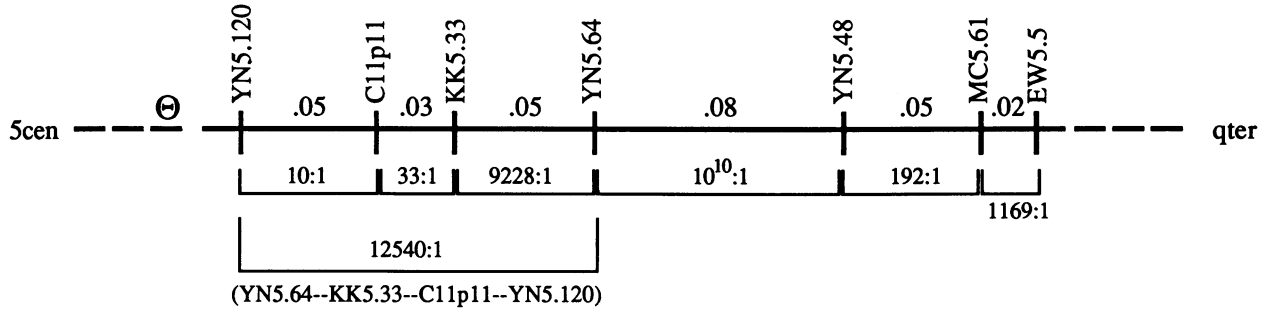

Figure 1 Map of markers on a short region of chromosome 5. Odds against inversion of adjacent loci are indicated in brackets. Odds are 12,540:1 against reversing the orientation of the subgroup containing the proximal four loci.

Table 3

Lod-Score Tables of Markers with FAP

Locus	Recombination Fraction							θ	Z_{max}
	.00	.05	.10	.15	.20	.30	.40		
C11p11	3.32	2.99	2.64	2.27	1.88	1.10	.37	.00	3.32
YN5.48	8.25	7.41	6.55	5.67	4.75	2.83	1.03	.00	8.25
YN5.64	–	5.61	5.53	5.13	4.57	3.15	1.45	.06	5.64
YN5.120	–	1.69	2.00	1.96	1.74	1.08	.40	.11	2.01
MC5.61	–	8.03	7.59	6.83	5.91	3.79	1.55	.04	8.05
KK5.33	–	7.66	7.69	7.16	6.36	4.26	1.84	.07	7.80
EW5.5	–	1.92	2.27	2.21	1.98	1.25	.47	.11	2.28

Note.—A dash (–) = significantly negative.

tion FAP more precisely: seven markers have been mapped within a region about 29 cM long. Although the order of three of the marker loci (YN5.120, C11p11, and KK5.33) is not yet certain, multipoint analysis in FAP-segregating families has strongly supported localization of the gene involved in FAP near the marker YN5.48, which is likely to be 17 cM distal to C11P11.

This result is consistent with linkage data on different families, presented by Bodmer et al. (1987), in which a maximum lod score of 3.36 between FAP and C11p11 was found at $\theta = 0$, with 1-lod-unit upper bound on recombination of about .15. Although our data also show no recombination with C11p11, several FAP recombinants are found for a new marker (KK5.33) that is approximately 3 cM from C11p11, in families that are not informative for the latter marker. Moreover, both YN5.48 and MC5.61, which we estimate to be >10 cM distal to C11p11, have shown strong evidence of linkage to FAP. The odds are >40,000:1 in favor of the region near YN5.48, rather than the region near C11p11, being the most likely location of the genetic defect responsible for FAP.

We can indirectly estimate a physical distance of ≤3,000 kb between the closest DNA marker (YN5.48) and the FAP locus. The map of new markers presented here therefore delimits and significantly narrows the target region for efforts to isolate the gene or genes involved in FAP. For some diseases that have been linked to genetic markers, identification of closer markers has been the labor of several years; in other cases no other markers have been identified. The present study has demonstrated the efficiency that can be achieved in high-resolution mapping when a hybrid cell line with an appropriate human component is available.

Table 4

Lod Scores in Families Informative for Locus YN5.48

Family	Recombination Fraction						
	.00	.05	.10	.15	.20	.30	.40
FPC:							
415	.54	.50	.46	.42	.37	.27	.15
1452	1.29	1.14	.97	.81	.64	.32	.08
1441	1.67	1.48	1.31	1.14	.95	.56	.22
Total	3.50	3.12	2.74	2.37	1.96	1.15	.45
GS:							
1498	1.11	1.00	.89	.77	.65	.38	.13
1507	.90	.81	.72	.62	.51	.30	.09
109	2.74	2.47	2.20	1.92	1.62	.99	.36
Total	4.75	4.28	3.81	3.31	2.78	1.67	.58
Combined total	8.25	7.40	6.55	5.68	4.74	2.82	1.03

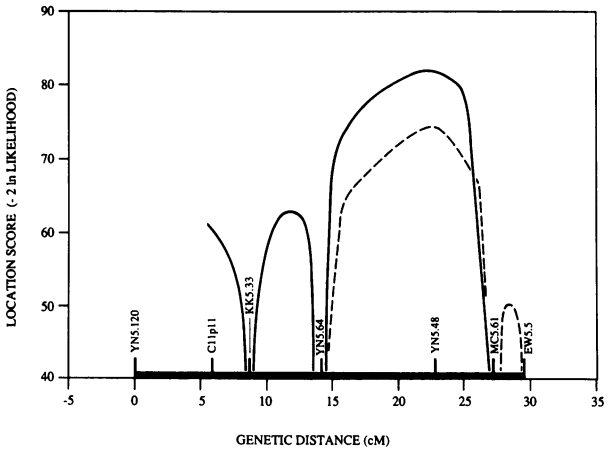

Figure 2 Location score for the FAP gene on the map in fig. 1. The solid line indicates location score with respect to the subset of loci C11p11, KK5.33, YN5.64, and YN5.48; the dashed line indicates location score calculated with respect to the subset YN5.64, YN5.48, MC5.61, and EW5.5. Both curves indicate that the most likely location of FAP is very close to the YN5.48 locus.

The two reports on linkage of FPC and C11p11 cited above (Bodmer et al. 1987; Leppert et al. 1987) did not resolve the question of possible genetic heterogeneity between FPC and GS. The present study showed a significant lod score (>3.0 at $\theta = 0$) between YN5.48 and either FPC or GS. Although we cannot eliminate the possibility that two different defective genes lie very close together, the simplest interpretation is that separate mutations within the same gene are generating the differences between the phenotypes.

Several genetic alterations associated with colon cancer have been reported recently. Frequent mutation in the K-*ras* gene has been identified by two groups (Bos et al. 1987; Forrester et al. 1987). Loss of heterozygosity has been observed on chromosomes 5, 17p, and 22 (Fearon et al. 1987; Solomon et al. 1987; Okamoto et al. 1988). Carcinogenesis of colon cancer therefore seems to require multiple mutations. A genetic defect that confers predisposition to FAP probably operates at the stage of conversion of normal mucosal cells to adenomatous polyps, because the presence of hundreds to thousands of colonic polyps is the most characteristic symptom in FAP patients; later development from adenomatous polyp to colon cancer is seen not only in FAP patients but also in patients with no family history of polyposis. In sporadic cases, somatic mutation at the FAP locus could play an important role in generating adenomatous polyps. The elevated risk for colon cancer in FAP patients may simply reflect the extremely high number of adenomatous polyps; in such a scenario the risk in FAP patients would be up to several thousand times that in a patient with only one adenomatous polyp. Additional events, such as K-*ras* point mutations and deletions in some chromosomes following somatic mutations, may change an adenoma cell to a cancer cell. Isolation of the gene for FAP and elucidation of its function could yield many answers to questions about oncogenesis.

In the meantime, because a set of DNA markers is now available for preclinical diagnosis of risk status for individuals in FAP families, unaffected family members with no elevated risk for colon cancer can be reassured, while affected members can be identified for medical intervention.

Acknowledgments

We are grateful to Dora Stauffer for expert technical assistance and to Ruth Foltz for editorial assistance in the preparation of the manuscript. J.-M. Lalouel and R. White are investigators for the Howard Hughes Medical Institute.

References

Barker, D., M. Schafer, and R. White. 1984. Restriction sites containing CpG show a higher frequency of polymorphism in human DNA. Cell 36:131–138.

Bodmer, W., C. Bailey, J. Bodmer, H. Bussey, A. Ellis, P. Gorman, F. Lucibello, V. Murday, S. Rider, P. Scambler, D. Sheer, E. Solomon, and N. Spurr. 1987. Localization of the gene for familial adenomatous polyposis on chromosome 5. Nature 328:614–616.

Bos, J., E. Fearon, S. Hamilton, M. Verlan-de Vries, J. van Boom, A. van der Eb, and B. Vogelstein. 1987. Prevalence of *ras* gene mutations in human colorectal cancers. Nature 327:293–297.

Cirullo, R., S. Dana, and J. Wasmuth. 1983. Efficient procedure for transferring specific human genes into Chinese hamster cell mutants: interspecific transfer of the human genes encoding leucyl- and asparaginyl-tRNA synthetases. Mol. Cell. Biol. 3:892–902.

Fearon, E., S. Hamilton, and B. Vogelstein. 1987. Clonal analysis of human colorector tumors. Science 238:193–196.

Feinberg, A., and B. Vogelstein. 1984. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. 137:P266–267.

Forrester, K., C. Almoguera, K. Han, W. Grizzle, and M. Perucho. 1987. Detection of high incidence of K-*ras* oncogenes during human colon tumorigenesis. Nature 327: 298–303.

Hayashi, K., M. Nakazawa, Y. Ishizaki, N. Hiraoka, and A. Obayashi. 1986. Regulation of inter- and intramolecu-

lar ligation with T4 DNA ligase in the presence of polyethylene glycol. Nucleic Acids Res. **14**:7617–7630.

Herrera, L., S. Kakati, L. Gibas, E. Pietrzak, and A. Sandberg. 1986. Brief clinical report: Gardner syndrome in a man with an interstitial deletion of 5q. Am. J. Med. Genet. **25**:473–476.

Lathrop, G. M., J.-M. Lalouel, C. Julier, and J. Ott. 1985. Multilocus linkage analysis in humans: detection of linkage and estimation of recombination. Am. J. Hum. Genet. **37**:482–498.

Lathrop, M., Y. Nakamura, P. Cartwright, P. O'Connell, M. Leppert, C. Jones, H. Tateishi, T. Bragg, J.-M. Lalouel, and R. White. 1988. A primary genetic map of markers for human chromosome 10. Genomics **2**:157–164.

Leppert, M., M. Dobbs, P. Scambler, P. O'Connell, Y. Nakamura, D. Stauffer, S. Woodward, R. Burt, J. Hughes, E. Gardner, M. Lathrop, J. Wasmuth, J.-M. Lalouel, and R. White. 1987. The gene for familial polyposis coli maps to the long arm of chromosome 5. Science **238**:1411–1413.

Monaco, A., R. Neve, C. Colletti-Feener, C. Bertelson, D. Kurnit, and L. Kunkel. 1986. Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. Nature **323**:646–650.

Okamoto, M., M. Sasaki, K. Sugio, C. Sato, T. Iwama, T. Ikeuchi, A. Tonomura, T. Sasazuki, and M. Miyaki. 1988. Loss of constitutional heterozygosity in colon carcinoma from patients with familial polyposis coli. Nature **331**: 273–277.

Royer-Pokora, B., L. Kunkel, A. Monaco, S. Goff, P. Newburger, R. Baehner, F. Cole, J. Curnutte, and S. Orkin. 1986. Cloning the gene for an inherited human disorder—chronic granulomatous disease—on the basis of its chromosomal location. Nature **322**:32–38.

Solomon, E., R. Voss, V. Hall, W. Bodmer, J. Jass, A. Jeffreys, F. Lucibello, I. Patel, and S. Rider. 1987. Chromosome 5 allele loss in human colorectal carcinomas. Nature **328**: 616–619.