[CANCER RESEARCH 62, 4579–4582, August 15, 2002]

Advances in Brief

Allele Separation Facilitates Interpretation of Potential Splicing Alterations and Genomic Rearrangements[1]

Hidewaki Nakagawa, Hai Yan, Janet Lockman, Heather Hampel, Kenneth W. Kinzler, Bert Vogelstein, and Albert de la Chapelle[2]

Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio 43210 [H. N., J. L., H. H., A. d. l. C.]; and Howard Hughes Medical Institute, The Oncology Center, The Johns Hopkins University, Baltimore, Maryland 21231 [H. Y., K. W. K., B. V.]

Abstract

Mutations that alter normal splice patterns and genomic rearrangements are common causes of hereditary diseases including hereditary nonpolyposis colorectal cancer. However, abnormal transcripts can be difficult to detect and interpret because splicing patterns are often heterogeneous even in normal cells. Standard techniques including sequencing and Southern hybridization fail to detect some genomic rearrangements. We show here that separation of alleles in somatic cell hybrids, through "conversion" technology, considerably facilitates the interpretation of abnormal splicing patterns and the detection of genomic rearrangements. We detected novel mutations in *MLH1* in each of four hereditary nonpolyposis colorectal cancer patients. The genomic mutations were CAG>CAA predicting Q346Q; GAG>AAG predicting E102K; a>g at nucleotide 1559–2 at intron 13, and a tandem duplication involving exons 7–12. By separating the two alleles, we showed that one allele produced only abnormal transcript or no transcript whereas the other allele produced only normal transcript. These results allowed pathogenicity to be unambiguously assigned to the mutations and increased the sensitivity of genomic testing.

Introduction

Mutation detection is key to the diagnosis of inherited disorders. In some diseases, searching for one or a small number of recurrent or widespread mutations is sufficient for most diagnostic purposes. For example, in sickle cell anemia most patients have a missense founder mutation of the β-globin gene (1), and in achondroplasia almost all patients have a recurrent missense mutation of the fibroblast growth factor receptor-3 gene (2). By contrast, in many diseases the phenotype results from a variable number of different mutations. Moreover, some diseases are genetically heterogeneous in that mutations in more than one gene can cause the phenotype. For instance, in HNPCC[3], germ-line mutations of at least four mismatch repair genes have been implicated (*MLH1*, *MSH2*, *MSH6*, *PMS2*; Refs. 3 and 4). The total number of different mutations in these genes is presently greater than 300,[4] and the list is rapidly growing (3).

In 20–50% of putative HNPCC families and patients, no mutation is detectable by standard methods (5–7). Many methods have been devised to improve mutation detection (8, 9); however, one class of changes remains particularly difficult to detect and interpret (*i.e.*, rearrangements of the transcripts, many of which lead to the decrease or disappearance of the transcript through NMD; Ref. 10). These mutations mainly consist of sequence changes causing splicing errors or consist of large genomic rearrangements. The clinical significance of abnormal transcripts is particularly difficult to interpret in genes where alternatively spliced isoforms of the transcript occur naturally and frequently. One such gene is *MLH1* (11, 12), which shows several alternative splicing patterns. In this study, we show that the separation of alleles in somatic cell hybrids (here referred to as conversion; Ref. 13) considerably facilitates the detection and interpretation of abnormal transcripts.

Materials and Methods

Patients and Samples. Four unrelated patients belonging to HNPCC families fulfilling the Amsterdam criteria (4), each with a microsatellite unstable colorectal carcinoma, were studied. Three of the patients were previously mentioned in the initial report on the conversion technology (13) as: patient 1 (ML), patient 2 (CG), and patient 3 (GS). Patient 4 has not been published previously. For each patient, an EBV-transformed lymphoblastoid cell culture was available as a source of DNA and RNA. In patients 1–3 lymphoblastoid cells were used to produce the fusion clones converted to haploidy, whereas in patient 4 lymphocytes from a fresh blood sample were used for this purpose.

Diploid-to-Haploid Conversion. Haploid-converted clones of patients 1–3 were created by using the conversion technology of Yan *et al.* (13). Haploid-converted clones from the fourth patient were created by GMP Genetics, Inc. In brief, human fresh lymphocytes or established lymphoblastoid cells were electrofused with a specifically designed mouse cell line (E2). Unfused mouse parental cells were negatively selected by sodium hypoxanthine, aminopterin, and thymidine (HAT; Life Technologies, Inc.), and unfused human lymphocytes were negatively selected by Geneticin (Life Technologies, Inc.). Hybrid cells were maintained in DMEM (Life Technologies, Inc.) including 10% FBS, 0.5 mg/ml Geneticin, 1× HAT, and penicillin-streptomycin.

RT-PCR Analysis. RNA was extracted from original lymphoblastoids and converted cells by Trizol according to the manufacturer (Life Technologies, Inc.). Two micrograms of total RNA were treated with Superscript II (Life Technologies, Inc.) to produce cDNA, using random hexamers according to the manufacturer. We designed four sets of primers to amplify the entire coding sequence (*a*) in a single amplicon and (*b*) as three overlapping amplicons. These were designed so as not to amplify mouse *MLH1*. The primers for the amplification of the entire coding sequence were 5′-gcgccaaaatgtcgttcgtg-3′(Ex1F) and 5′-cccacagtgcataaataacc-3′(Ex19R). The primers for the first part of *MLH1* were Ex1F and 5′-ctgagtagtttgcattggatatgtaa-3′(Ex9R), for the second part 5′-gagacagtagctgatgttaggacacta-3′(Ex8F) and 5′-cagtcatttcctttcggg-gaat-3′(Ex13R), and for the third part 5′-agaggacctacttccagcaacc-3′(Ex12F) and Ex19R. Other primers used for amplification are available on request.

Received 4/18/02; accepted 6/24/02.

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

[1] Supported by Grants CA67941, CA16058, CA43460, and CA62924 from the National Institutes of Health. H.N. is a postdoctoral fellow from Japan Society for Promoting of Science and is supported by Sumitomo Life Social Welfare Service's Foundation, The Osaka Medical Research Foundation for Incurable Disease, and Yamanouchi Foundation for Research on Metabolic Disorders. Under agreements between The Johns Hopkins University and GMP Genetics, Inc., K. W. K. and B. V. are entitled to a share of the royalties received by the University on sales of products related to the use of Conversion technologies. K. W. K. and B. V. are consultants to GMP Genetics, Inc. The University, K. W. K., and B. V. also own stock in GMP Genetics, Inc., which is subject to certain restrictions under University policy. The terms of these arrangements are being managed by the University in accordance with its conflict of interest policies.

[2] To whom requests for reprints should be addressed, at Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Medical Research Facility, Room 646, 420 West 12th Avenue, Columbus, OH 43210. Phone: (614) 688-4781; Fax: (614) 688-4772; E-mail: delachapelle-1@medctr.osu.edu.

[3] The abbreviations used are: HNPCC, hereditary nonpolyposis colorectal cancer; NMD, nonsense-mediated decay; M, mutated; W, wild type.

[4] Internet address: www.nfdht.nl/.

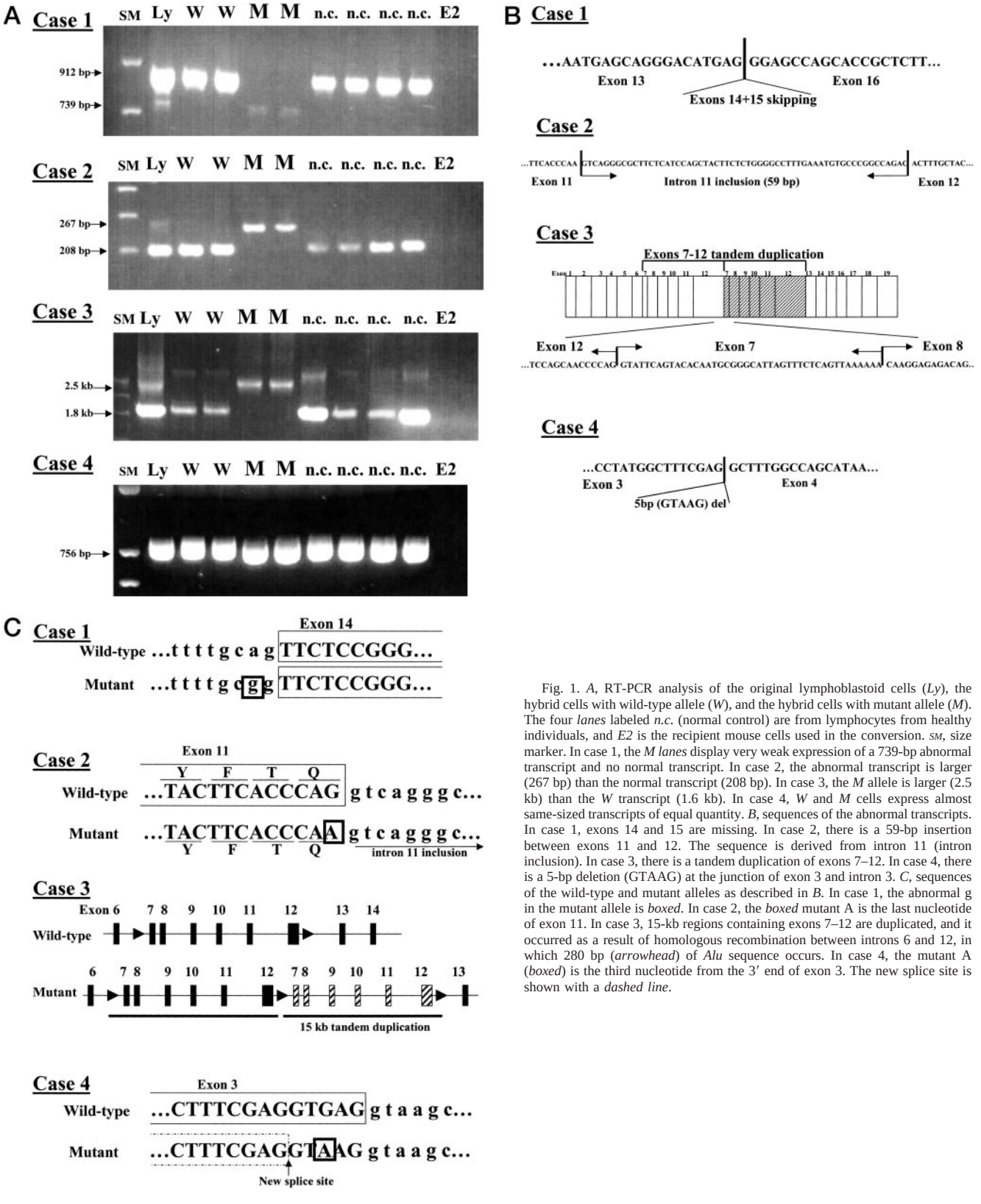

Fig. 1. *A*, RT-PCR analysis of the original lymphoblastoid cells (*Ly*), the hybrid cells with wild-type allele (*W*), and the hybrid cells with mutant allele (*M*). The four *lanes* labeled *n.c.* (normal control) are from lymphocytes from healthy individuals, and *E2* is the recipient mouse cells used in the conversion. *SM*, size marker. In case 1, the *M lanes* display very weak expression of a 739-bp abnormal transcript and no normal transcript. In case 2, the abnormal transcript is larger (267 bp) than the normal transcript (208 bp). In case 3, the *M* allele is larger (2.5 kb) than the *W* transcript (1.6 kb). In case 4, *W* and *M* cells express almost same-sized transcripts of equal quantity. *B*, sequences of the abnormal transcripts. In case 1, exons 14 and 15 are missing. In case 2, there is a 59-bp insertion between exons 11 and 12. The sequence is derived from intron 11 (intron inclusion). In case 3, there is a tandem duplication of exons 7–12. In case 4, there is a 5-bp deletion (GTAAG) at the junction of exon 3 and intron 3. *C*, sequences of the wild-type and mutant alleles as described in *B*. In case 1, the abnormal g in the mutant allele is *boxed*. In case 2, the *boxed* mutant A is the last nucleotide of exon 11. In case 3, 15-kb regions containing exons 7–12 are duplicated, and it occurred as a result of homologous recombination between introns 6 and 12, in which 280 bp (*arrowhead*) of *Alu* sequence occurs. In case 4, the mutant A (*boxed*) is the third nucleotide from the 3′ end of exon 3. The new splice site is shown with a *dashed line*.

Long-range RT-PCR was performed by using Expand Long Template PCR system (Roche), and PCR was carried out at 94°C (10 s), 60°C (30 s), and 68°C (2 min) for 10 cycles, and at 94°C (10 s), 60°C (30 s), and 68°C (2 min + cycle elongation for more yields of 20 s for each cycle) for 25 cycles. The PCR products were run on a 1.5% agarose gel to analyze their size differences.

DNA Sequencing. The RT-PCR products were isolated using QIAquick PCR product purification kit (Qiagen). The purified RT-PCR products were sequenced directly by the primer used for PCR, or cloned into TA vector pCR2.1 (Invitrogen). DNA from plasmid clones was extracted by QIAprep Spin Miniprep kit (Qiagen) and sequenced using the ABI sequencing system (Perkin-Elmer Applied Biosystems). Amplification of the exons and exon-intron junctions of *MLH1* and sequencing was described previously (14).

Table 1 *Summary of findings*

Case no.	Genomic change detected by sequencing of original cell DNA	RT-PCR change detected by analysis of original cell RNA	Transcript in mutated allele	Genomic change and consequence
1	a→g at nt[a] 1559-2	Faint band of reduced size	Deletion of exons 14 and 15	a→g at nt 1559-2 splicing error
2	None	Faint band of increased size	Addition of 59 bp from intron 11	CAG→CAA Q346Q splicing error
3	None	Faint band of increased size	Tandem duplication of exons 7–12	Tandem duplication of 15-kb comprising exons 7–12
4	GAG→AAG at codon 102, exon 3	No alteration	5-bp deletion in exon 3	GAG→AAG E102K splicing error

[a] nt, nucleotide.

Results and Discussion

Genomic Sequencing. Exon-by-exon genomic sequencing including the promoter regions of *MLH1* and *MSH2* revealed no changes in lymphoblast DNA from cases 2 and 3. In case 1, there was an a to g change at the conservative splice acceptor site of exon 14 of *MLH1* (a→g at nucleotide 1559-2). In case 4, there was a G to A change at codon 102 predicting a glutamic acid to lysine, E102K, amino acid substitution of unknown significance.

RT-PCR of Diploid Cell Transcripts. Using RNA from lymphoblastoid cells, in all four cases a normal-sized transcript was seen by RT-PCR; in addition, in cases 1–3, a very faint aberrant band was also seen (Fig. 1*A*). In case 4, only a normal-sized transcript was seen in the diploid cells (Fig. 1*A*).

Haploid Conversion Hybrids. In each case hybrid clones were studied for their content of alleles for markers on chromosome 3 (*D3S1263*, *D3S1569*, *D3S1300*, and *D3S1601*). On the basis of data not shown, we selected two clones having one copy of one of the chromosomes 3, and two clones having the other copy of chromosome 3. In Fig. 1, these clones are referred to as *W* for wild type and *M* for presumed mutated allele, respectively. RT-PCR and sequencing of cDNA and genomic DNA from these clones revealed the nature of the mutation in each case as:

In case 1 (Fig. 1), RT-PCR demonstrated only normal transcript in the W clones and only a smaller than normal transcript in the M clones; this transcript was of extremely weak abundance, apparently as a consequence of NMD. Sequencing revealed that this transcript lacked exons 14 and 15. The a→g change at nucleotide 1559-2 seen previously in heterozygous form in lymphoblastoid cell DNA occurred in the M clones but not in the W clones. Thus, this change apparently caused the splicing error.

In case 2 (Fig. 1), RT-PCR showed a normal product in W clones but a larger product in M clones. By sequencing, a 59-bp insertion emanating from the 5′ end of intron 11 was seen in this transcript. Sequencing of the M clones revealed a G→A change in the last nucleotide of exon 11 that had not been noticed in the sequencing of lymphoblastoid cell DNA. This change obviously abrogates the splicing that instead occurs at the next downstream splicing-donor consensus gt in intron 11, resulting in the addition of 59 bp to the transcript. As can be seen (Fig. 1*A*) in lymphoblastoid cells, the abnormal transcript is present in greatly reduced quantity compared with the wild-type transcript, hence it is subject to NMD.

In case 3 (Fig. 1), RT-PCR showed normal transcript in W clones but only an abnormally large transcript in M clones. By sequencing, this turned out to represent a tandem duplication of exons 7–12. By genomic sequencing of the M clones, a 15-kb tandem duplication comprising exons 7–12 was identified. It had arisen through homologous recombination between introns 6 and 12 in *Alu* regions showing 88% sequence identity. For unknown reasons, Southern blot analysis by using a cDNA probe of *MLH1* against *HindIII*- and *NsiI*-digested genomic DNA did not detect this genomic rearrangement of *MLH1* (data not shown).

In case 4 (Fig. 1), RT-PCR showed a normally sized product of *MLH1* in both W clones and M clones. By sequencing of the transcript from the M clones, a 5-bp deletion in the 3′ end of exon 3 was present as well as the minor population of the normally sized transcript with one nucleotide change. Sequencing of genomic DNA from the M clones revealed the missense mutation at codon 102 (GAG→AAG) that had been noticed in the sequencing of lymphocyte DNA. This nucleotide change abrogates the splicing pattern of exon 3, which instead occurs 5 bp upstream within exon 3, resulting in a 5-bp deletion in the transcript. This change is obviously pathogenic.

Table 1 summarizes the findings. Without conversion, a mutation was signaled in two of the cases by genomic sequencing and in three of the cases by RT-PCR. However, in all four cases it was the study of the haploid-converted clones that made it possible to delineate and assess the potential disease-causing role of the mutations. In particular, the presence of one or several minor species of transcript, as seen in cases 1–3, is hard to assess in the diploid context. This is the case even when the sequence of the minor species can be determined. In contrast, by demonstrating that the abnormal transcript emanates from one of the alleles as shown in all four cases here, and that it is of greatly diminished quantity, as shown in 2 cases here, a decisive, deleterious role is likely.

Exonic splicing enhancers and silencers have been proposed to explain how some missense or silent mutations can affect the splicing pattern and result in the inactivation of the protien (15, 16). In this study, we demonstrated two missense and one silent mutation of *MLH1* affecting its splicing pattern. This approach could be generally applied to the analysis of presumptive splicing mutations and complicated genomic rearrangements in any gene, as long as that gene continues to be expressed in somatic cell hybrids. Such studies usually allow pathogenicity to be assessed and, thus, increase the sensitivity of mutation testing for inherited diseases (17).

Acknowledgments

We thank Tamara Hemingway and Getachew Boru for assistance with cell culture and DNA preparation and Barbara Fersch for preparation of this manuscript.

References

1. Kan, Y. W., and Dozy, A. M. Polymorphism of DNA sequence adjacent to human β-globin structural gene: relationship to sickle mutation. Proc. Natl. Acad. Sci. USA, *75:* 5631–5635, 1978.
2. Rousseau, F., Bonaventure, J., Legeai-Mallet, L., Pelet, A., Rozet, J. M., Maroteaux, P., Le Merrer, M., and Munnich, A. Mutations in the gene encoding fibroblast growth factor receptor-3 in achondroplasia. Nature (Lond.), *15:* 252–254, 1994.
3. Peltomäki, P., and Vasen, H. F. Mutations predisposing to hereditary nonpolyposis colorectal cancer: database and results of a collaborative study. The International Collaborative Group on Hereditary Nonpolyposis Colorectal Cancer. Gastroenterology, *113:* 1146–1158, 1997.
4. Vasen, H. F., Watson, P., Mecklin, J. P., and Lynch, H. T. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative Group on HNPCC. Gastroenterology, *116:* 1453–1456, 1999.
5. Nystrom-Lahti, M., Wu, Y., Moisio, A. L., Hofstra, R. M., Osinga, J., Mecklin, J. P., Jarvinen, H. J., Leisti, J., Buys, C. H., de la Chapelle, A., and Peltomäki, P. DNA mismatch repair gene mutations in 55 kindreds with verified or putative hereditary non-polyposis colorectal cancer. Hum. Mol. Genet., *5:* 763–769, 1996.
6. Scott, R. J., McPhillips, M., Meldrum, C. J., Fitzgerald, P. E., Adams, K., Spigelman, A. D., du Sart, D., Tucker, K., and Kirk, J. Hereditary nonpolyposis colorectal cancer in 95 families: differences and similarities between mutation-positive and mutation-negative kindreds. Am. J. Hum. Genet., *68:* 118–127, 2001.

7. Papadopoulos, N., and Lindblom, A. Molecular basis of HNPCC: mutations of MMR genes. Hum. Mutat., *10:* 89–99, 1997.
8. Liu, W., Smith, D. I., Rechtzigel, K. J., Thibodeau, S. N., and James, C. D. Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. Nucleic Acids Res., *26:* 1396–1400, 1998.
9. Bateman, J. F., Freddi, S., Lamande, S. R., Byers, P., Nasioulas, S., Douglas, J., Otway, R., Kohonen-Corish, M., Edkins, E., and Forrest, S. Reliable and sensitive detection of premature termination mutations using a protein truncation test designed to overcome problems of nonsense-mediated mRNA instability. Hum. Mutat., *13:* 311–317, 1999.
10. Hentze, M. W., and Kulozik, A. E. A perfect message: RNA surveillance and nonsense-mediated decay. Cell, *96:* 307–310, 1999.
11. Kohonen-Corish, M., Ross, V. L., Doe, W. F., Kool, D. A., Edkins, E., Faragher, I., Wijnen, J., Khan, P. M., Macrae, F., and St. John, D. J. RNA-based mutation screening in hereditary nonpolyposis colorectal cancer. Am. J. Hum. Genet., *59:* 818–824, 1996.
12. Genuardi, M., Viel, A., Bonora, D., Capozzi, E., Bellacosa, A., Leonardi, F., Valle, R., Ventura, A., Pedroni, M., Boiocchi, M., and Neri, G. Characterization of *MLH1* and *MSH2* alternative splicing and its relevance to molecular testing of colorectal cancer susceptibility. Hum. Genet., *102:* 15–20, 1998.
13. Yan, H., Papadopoulos, N., Marra, G., Perrera, C., Jiricny, J., Boland, C. R., Lynch, H. T., Chadwick, R. B., de la Chapelle, A., Berg, K., Eshleman, J. R., Yuan, W., Markowitz, S., Laken, S. J., Lengauer, C., Kinzler, K. W., and Vogelstein, B. Conversion of diploidy to haploidy. Nature (Lond.), *403:* 723–724, 2000.
14. Aaltonen, L. A., Salovaara, R., Kristo, P., Canzian, F., Hemminki, A., Peltomäki, P., Chadwick, R. B., Kääriäinen, H., Percesepe, A., Ahtola, H., Härkönen, N., Julkunen, R., Kangas, E., Ojala, S., Tulikoura, J., Valkamo, E., Eskelinen, M., Järvinen, H., Mecklin, J. P., and de la Chapelle A. Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. N. Engl. J. Med., *338:* 1481–1487, 1998.
15. Liu, H. X., Cartegni, L., Zhang, M. Q., and Krainer, A. R. A mechanism for exon skipping caused by nonsense or missense mutations in *BRCA1* and other genes. Nat. Genet., *27:* 55–58, 2001.
16. Maquat, E. L. The power of point mutations. Nat. Genet., *27:* 5–6, 2001.
17. Yan, H., Kinzler, K. W., and Vogelstein, B. Genetic testing–present and future. Science (Washington DC), *289:* 1890–1892, 2000.

Allele Separation Facilitates Interpretation of Potential Splicing Alterations and Genomic Rearrangements

Hidewaki Nakagawa, Hai Yan, Janet Lockman, et al.

Cancer Res 2002;62:4579-4582.