Prevalence of Alterations in DNA Mismatch Repair Genes in Patients With Young-Onset Colorectal Cancer

PAUL J. LIMBURG,* WILLIAM S. HARMSEN,‡ HELEN H. CHEN,‖ STEVEN GALLINGER,¶ ROBERT W. HAILE,# JOHN A. BARON,** GRAHAM CASEY,# MICHAEL O. WOODS,‡‡ STEPHEN N. THIBODEAU,§ and NORALANE M. LINDOR‖

*Division of Gastroenterology & Hepatology, ‡Division of Biomedical Statistics and Informatics, §Department of Laboratory Medicine, and ‖Department of Medical Genetics, Mayo Clinic, Rochester, Minnesota; ¶Cancer Care Ontario, Mount Sinai Hospital, Samuel Lunenfeld Research Institute, University of Toronto, Toronto, Ontario, Canada; #Department of Preventive Medicine, Keck School of Medicine, University of Southern California, Los Angeles, California; **Department of Community and Family Medicine, Dartmouth Medical School, Hanover, New Hampshire; and ‡‡Department of Genetics, Faculty of Medicine, Memorial University, St John's, Newfoundland, Canada

See related article, Mukherjee B et al, on page 1919 in *Gastroenterology*; see editorial on page 458.

BACKGROUND & AIMS: Direct germline analysis could be used to screen high-risk patients for mutations in DNA mismatch repair genes associated with Lynch Syndrome. We examined the prevalence of mutations in *MLH1*, *MSH2*, and *MSH6* in a population-based sample of patients with young-onset (age <50 years) colorectal cancer (CRC). **METHODS:** Young-onset CRC cases were randomly selected from 3 Colon Cancer Family Registry sites. DNA was extracted from peripheral blood leukocytes; *MLH1*, *MSH2*, and *MSH6* were sequenced, and duplication and deletion analyses was performed for *MLH1* and *MSH2*. Results were reported as deleterious or suspected deleterious, likely neutral, variant of uncertain significance, or no alteration detected. Germline data were compared to Amsterdam II criteria (ACII) and immunohistochemistry results in secondary analyses. **RESULTS:** Among 195 subjects, 11 had deleterious/suspected deleterious mutations (5.6%; 95% confidence interval [CI], 2.8%–9.9%), 12 had likely neutral alterations (6.2%; 95% CI, 3.2%–10.5%), 14 had variants of uncertain significance (7.2%; 95% CI, 4.0%–11.8%), 2 had a likely neutral alteration and a variant of uncertain significance (1.0%; 95% CI, 0.1%–3.7%), and 156 had no alteration detected (80.0%; 95% CI, 73.7%–85.4%). Sensitivity, specificity, and positive and negative predictive values for detecting deleterious/suspected deleterious mutations, based on ACII, were 36.4% (4/11), 96.7% (178/184), 40.0% (4/10), and 96.2% (178/185), respectively; based on immunohistochemistry these values were 85.7% (6/7), 91.9% (136/148), 33.3% (6/18), and 99.3% (136/137), respectively. **CONCLUSIONS: In a population-based sample of young-onset CRC cases, germline mutations in *MLH1*, *MSH*, and/or *MSH6* were more prevalent than reported for CRC patients overall. Because only about 5% of young-onset CRC cases had confirmed deleterious or suspected deleterious mutations, further comparative effectiveness research is needed to determine the most appropriate screening strategy for Lynch Syndrome in this high-risk group.**

Keywords: Colon Cancer; Genetics; Mutation Analysis; Cancer Genetics.

View this article's video abstract at www.cghjournal.org

In the United States, approximately 6.5% of all incident colorectal cancers (CRCs) are diagnosed before age 50 years.[1] Unfortunately, these young-onset CRC cases often present with advanced stage tumors[2] and contribute disproportionately to the estimated 800,000 person-years of life lost annually to CRC.[3] Recent data from the Surveillance, Epidemiology and End Results (SEER) program demonstrate that CRC incidence rates are gradually increasing in younger adults.[4] Improved clinical management strategies are promptly needed to address the effects of this discouraging trend.

Lynch syndrome is caused by germline defects in 1 or more DNA mismatch repair (MMR) genes and characterized by young-onset cancer(s) arising in the colorectum and other target organs.[5] Most Lynch syndrome–associated CRCs are reportedly due to *MLH1*, *MSH2,* or *MSH6* mutations.[6–8] Several clinical algorithms, molecular tests, and statistical models have been developed to predict the presence or absence of Lynch syndrome.[7,9–14] Despite these readily accessible tools, Lynch syndrome patients continue to be widely under-recognized in clinical practice.[15–17]

Direct germline mutation analysis, ie, without clinical or molecular triage, represents a straightforward, potentially applicable strategy for Lynch syndrome screening in young-onset CRC patients. However, the feasibility of this approach depends, in part, on the prevalence of MMR gene mutations in this defined patient group. To our knowledge, no studies have examined the prevalence of *MLH1*, *MSH2*, and *MSH6* mutations by using an unselected, population-based sample of young-onset CRC cases identified at multiple North American centers. In the present study, we used data and tissue resources from the Colon Cancer Family Registry (Colon CFR)[18] to address this knowledge gap.

Materials and Methods

Subject Population

Subjects were recruited through the Colon CFR, an international collaboration of 6 participating centers (University of Hawaii, Honolulu, HI; Fred Hutchinson Cancer Research Center, Seattle, WA; Mayo Clinic, Rochester, MN; University of Southern California Consortium, Los Angeles,

Abbreviations used in this paper: ACII, Amsterdam II criteria; CI, confidence interval; Colon CFR, Colon Cancer Family Registry; CRC, colorectal cancer; IHC, immunohistochemistry; MMR, mismatch repair; NPV, negative predictive value; PPV, positive predictive value.

1542-3565/$36.00
doi:10.1016/j.cgh.2010.10.021

CA; Cancer Care Ontario, Toronto, Canada; and University of Melbourne, Melbourne, Australia) organized in 1997 to create a comprehensive resource for genetic epidemiology studies. As described elsewhere,[18] recruitment strategies differed across centers with respect to family ascertainment (population-based vs clinic-based) and CRC subject enrollment (all incident cases vs oversampling by family history or early age of onset). For the current investigation, a random sample of exclusively population-based, young-onset (diagnosed before age 50 years) CRC cases who (1) were recruited during phase I of the Colon CFR collaboration (1997–2002) through Cancer Care Ontario (n = 67), Mayo Clinic Rochester (n = 67), and University of Southern California (n = 67) and (2) had extracted DNA available for germline *MLH1*, *MSH2,* and *MSH6* mutation analyses was included. Specifically, there was no preselection based on family history or availability/results of tumor testing. From the initial sample of 201 young-onset CRC cases, 6 subjects were excluded on the basis of tumor location in the appendix (n = 4) or anus (n = 2) rather than the colorectum. Demographic data, venipuncture samples, tumor blocks, and pathology reports were collected according to established protocols (available at http://epi.grants.cancer.gov/CFR/research_app-instructions.html). CRC anatomical subsites were ascertained from pathology reports. Proximal cancers were defined as tumors located in the cecum, ascending colon, hepatic flexure, transverse colon, and splenic flexure; distal cancers were defined as tumors located in the descending colon, sigmoid colon, rectosigmoid colon, and rectum, respectively.

Gene Mutation Analyses

Extracted DNA samples (from peripheral blood leukocytes) were shipped to Myriad Genetic Laboratories (Salt Lake City, UT) for full mutation analyses of *MLH1*, *MSH2,* and *MSH6*. The testing center was blinded to all clinical data associated with the specimens. Assay methods were identical to those used for clinical testing. On receipt, samples were assigned a unique bar code for robotic specimen tracking. DNA was amplified by polymerase chain reaction for each subject. The amplified products were each directly sequenced in forward and reverse directions by using fluorescent dye–labeled sequencing primers: *MLH1* (approximately 2300 base pairs comprising 19 exons and approximately 560 adjacent noncoding intronic base pairs), *MSH2* (approximately 2800 base pairs comprising 16 exons and approximately 470 adjacent noncoding intronic base pairs), and *MSH6* (approximately 4080 base pairs comprising 10 exons and approximately 290 adjacent noncoding intronic base pairs). The analyzed, noncoding intronic regions of *MLH1*, *MSH2,* and *MSH6* do not extend more than 20 base pairs proximal to the 5′ end and 10 base pairs distal to the 3′ end of each exon. Chromatographic tracings of each amplicon were analyzed by a proprietary computer-based review, followed by visual inspection and confirmation. Genetic variants were detected by comparison with a consensus wild-type sequence constructed for each gene. Sequence information of the coding region was derived from RefSeq NM_000249.3 (MLH1), NM_000251.1 (MSH2), and NM_000179.2 (MSH6). Intronic nucleotide information was derived from genomic sequences from NCBI — AC011816.17 (MLH1), AC079775.6 (MSH2), AC006509.15 (MSH6). All potential genetic variants were independently confirmed by repeated polymerase chain reaction amplification of the indicated gene region(s) and sequence determination as above. Large rearrangement testing for *MLH1* and *MSH2* was performed by Southern blot analysis in conjunction with multiplex ligation-dependent probe amplification (reagents from MRC-Holland, Amsterdam, the Netherlands). For Southern analysis of *MLH1* and *MSH2*, three sets of restriction digests were performed for each gene by using individual enzymes or a combination of enzymes to yield appropriate banding patterns. Digested DNA was electrophoresed on agarose gels, transferred to nylon membranes, and hybridized with a gene-specific probe radiolabeled with 32P. Autoradiographs and phosphorimages were analyzed for the presence of novel bands and for fragment dosage to assess the presence of deletions or duplications in exon regions. Southern blot and multiplex ligation-dependent probe amplification data were subjected to dual reviews involving technical personnel and at least 1 laboratory director for confirmation. Germline alterations were categorized as deleterious/suspected deleterious, likely neutral, or variant of uncertain significance, with specific sequence alterations recorded for each gene.

Immunohistochemistry Testing

Immunohistochemistry (IHC) for MLH1, MSH2, and MSH6 protein expression was performed as previously described[9,10] at Mayo Clinic Rochester (for Mayo Clinic Rochester and University of Southern California Consortium cases) and at Cancer Care Ontario (for Cancer Care Ontario cases), according to established protocols for clinical and research evaluation. In brief, 4- to 6-μm tissue sections were cut from formalin-fixed, paraffin-embedded tissue blocks and stained by using the avidin-biotin complex method of Ventana Medical Systems (Oro Valley, AZ) (Bio-Tek Solutions buffer kit and DAB detection kit) and the Tech Mate 500 (Ventana) automated immunohistochemical stainer. Antibodies to MLH1 (clone G168-728, 1/250; Pharmingen, San Diego, CA), MSH2 (clone FE11, 1/50; Oncogene Research Products, Cambridge, MA), and MSH6 (clone 44, 1/500; Transductions Laboratories, Lexington, KY) were used. MMR protein expression was reported as present, absent, or inconclusive for each immunostain. Because tumor phenotype was not considered a primary study end point, complete IHC data were only available for 155 of 195 (79%) subjects.

Statistical Analyses

All genotyping results were provided to the Colon CFR participating sites by Myriad Genetics. Data analyses were conducted by the Colon CFR investigators. Summary statistics were reported as mean (standard deviation) or frequency (percent), as appropriate. Baseline demographic characteristics were compared across the 3 Colon CFR participating sites by using the Student *t* test or χ^2 test, as appropriate. Prevalence estimates for germline MMR gene alterations, in aggregate (ie, finding present in 1 or more genes) and for *MLH1*, *MSH2,* and *MSH6* separately, were defined a priori as the primary study end point, with accompanying 95% exact binomial confidence intervals (CIs) reported. Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) for detecting deleterious/suspected deleterious MMR gene mutations were also estimated on the basis of the Amsterdam II criteria (ACII[12]; fulfilled vs not fulfilled) and IHC testing for MLH1, MSH2, and MSH6 protein expression (any absent vs all present) as secondary analyses in subjects for whom relevant data were available.