GASTROENTEROLOGY 2005;129:537–549

BASIC–ALIMENTARY TRACT

Functional Significance and Clinical Phenotype of Nontruncating Mismatch Repair Variants of *MLH1*

TIINA E. RAEVAARA,* MARI K. KORHONEN,* HANNES LOHI,† HEATHER HAMPEL,§ ELLY LYNCH,¶ KARIN E. LÖNNQVIST,* ELKE HOLINSKI–FEDER,‖ CHRISTIAN SUTTER,# WENDY McKINNON,** SEKHAR DURAISAMY,** ANNE–MARIE GERDES,†† PÄIVI PELTOMÄKI,† MAIJA KOHONEN–CORISH,§§ ELISABETH MANGOLD,¶¶ FINLAY MACRAE,¶ MARC GREENBLATT,** ALBERT DE LA CHAPELLE,§ and MINNA NYSTRÖM*

Departments of *Biological and Environmental Sciences and †Medical Genetics, University of Helsinki, Helsinki, Finland; §Human Cancer Genetics Program, Comprehensive Cancer Center, The Ohio State University, Columbus, Ohio; ¶Familial Bowel Cancer Clinic, The Royal Melbourne Hospital, Melbourne, Australia; ‖Institute of Human Genetics, University of Munich, Munich, Germany; #Department of Molecular Pathology, University of Heidelberg, Heidelberg, Germany; **Vermont Cancer Center, University of Vermont, Burlington, Vermont; ††Clinical Genetics Department of KKA, Odense University Hospital, Odense, Denmark; §§Garvan Institute of Medical Research, Sydney, Australia; and ¶¶Institute of Human Genetics, University of Bonn, Bonn, Germany

See editorial on page 751.

Background & Aims: **Germline mutations in mismatch repair genes are associated with hereditary nonpolyposis colorectal cancer. A significant proportion of mutations are nontruncating and associated with a variability of clinical phenotype and microsatellite instability and with occasional presence of residual protein in tumor tissue that suggests impaired functional activity but not total lack of mismatch repair. To address pathogenic significance and mechanism of pathogenicity, we studied the functionality of 31 nontruncating *MLH1* mutations found in clinically characterized colorectal cancer families and 3 other variations listed in a mutation database.** ***Methods:*** **Mutations constructed by site-directed mutagenesis were studied for protein expression/stability, subcellular localization, protein-protein interaction, and repair efficiency. The genetic and biochemical data were correlated with clinical data. Finally, comparative sequence analysis was performed to assess the value of sequence homology as a tool for predicting functional results.** ***Results:*** **Altogether, 22 mutations were pathogenic in more than one assay, 2 variants were impaired in one assay, and 10 variants acted like wild-type protein. Twenty of 34 mutations affected the quantity of MLH1 protein, whereas only 15 mainly amino-terminal mutations were defective in an in vitro repair assay. Comparative sequence analysis correctly predicted functional studies for 82% of variants.** ***Conclusions:*** **Pathogenic nontruncating alterations in *MLH1* may interfere with different biochemical mechanisms but generally more than one. The severe biochemical defects are mirrored by phenotypic characteristics such as early age at onset and high microsatellite instability, whereas variants with no or mild defects in functionality are associated with variable clinical phenotypes.**

Hereditary nonpolyposis colorectal cancer (HNPCC) is associated with dominantly inherited mutations in mismatch repair (MMR) genes. Even in the presence of uniform genetic predisposition, the clinical phenotypes, age at onset, and tumor spectrum vary significantly among families and even among patients from the same family. Furthermore, most but not all tumors show lack of the mutated protein and high microsatellite instability (MSI) in the genome, the hallmarks of HNPCC tumors. The first international criteria (Amsterdam I) for the diagnosis of HNPCC were developed in 1991 and required (1) histologically verified colorectal cancer in 3 or more relatives, one of whom is a first-degree relative of the other two; (2) colorectal cancer involving at least 2 generations; and (3) one or more colorectal cancer cases diagnosed before the age of 50 years.[1] However, many putative HNPCC families do not

Abbreviations used in this paper: EGFP, enhanced green fluorescent protein; HNPCC, hereditary nonpolyposis colorectal cancer; MMR, mismatch repair; MSI, microsatellite instability; NPV, negative predictive value; OPV, overall predictive value; PCR, polymerase chain reaction; PPV, positive predictive value; Sf 9, *Spodoptera frugiperda* 9; TE, total protein extract; WT, wild-type.

0016-5085/05/$30.00
doi:10.1053/j.gastro.2005.06.005

fulfill the original criteria; new, less stringent criteria were developed that incorporate extracolonic cancers, notably endometrial carcinomas in addition to colorectal carcinomas (modified Amsterdam criteria, Amsterdam criteria II).[2] Sometimes HNPCC should be considered even in the absence of a family history of cancer if the patient is young and has high MSI in tumor tissue (Bethesda guidelines).[3,4] Remarkably, even susceptibility to sporadic, microsatellite-stable colorectal carcinoma may be linked to germline variants of MMR genes.[5] Such a wide variety of clinical phenotypes complicates diagnostics, counseling, and design of appropriate follow-up and treatment strategies for individuals carrying the predisposing mutations. Biologic tools to predict the pathogenicity of the different mutations would be of prime clinical importance.

To date, the database of sequence variations found in HNPCC or putative HNPCC kindreds (http://www.insight-group.org/) contains information on 448 different germline alterations that are likely to be pathogenic and 108 nonpathogenic alterations, affecting mostly the MMR genes *MLH1* (50%), *MSH2* (39%), and *MSH6* (7%).[6] Germline mutations in the 2 major HNPCC-associated genes, *MLH1* and *MSH2*, are found in two thirds of classic HNPCC that fulfill the stringent Amsterdam criteria and display MSI in tumor tissue,[7] compared with only 8%–47% of the less typical colon cancer families with late age at onset, atypical tumor spectrum, and/or lack of MSI.[8,9] Despite certain "typical" features, even the former group does not show a uniform clinical picture. Instead, the gene involved and the site and type of mutation seem to have an effect on the phenotype.[10–12] In particular, missense mutations leading to single amino acid substitutions appear to be associated with a wide range of clinical phenotypes,[13] and in many cases their pathogenicity is difficult to interpret.

To address the pathogenic significance and mechanism of pathogenicity, we studied 34 nontruncating *MLH1* mutations for expression/stability, nuclear localization, and functionality of the respective mutated protein. We also assessed how well comparative sequence analysis using alignments of eukaryotic MLH1 sequences predicts the results of functional assays. The genetic and biochemical data were then correlated with clinical data. Characterization of the biochemical defects facilitated the definition of functional domains of the MLH1 protein and revealed different mechanisms through which the pathogenic effects of the mutations were mediated. Our classification of the investigated variants as either pathogenic or nonpathogenic based on our in vitro results was supported by the clinical associations.

Materials and Methods

MLH1 Mutations and HNPCC Families

The present study comprised 27 *MLH1* missense mutations and 4 amino acid deletions found in 52 putative HNPCC families that have been subjected to molecular genetic and clinical studies suitable for prospective genotype-phenotype analysis. Table 1 shows the mutations and clinical characteristics of the index patients and their families collected through an international HNPCC collaboration. Because the HNPCC families included in the study were found and analyzed by many research groups, different methods were used for mutation detection (Table 1). The alterations are scattered throughout the MLH1 polypeptide but are mainly clustered in its amino-terminus, which is responsible for adenosine triphosphate binding and hydrolysis,[14–16] and in the carboxyl-terminus, which contains the region where MLH1 interacts with its counterparts, PMS2, MLH3, and PMS1 (Figure 1).[17–19]

Half of the families did not fulfill the Amsterdam criteria. Mean age at onset of HNPCC-related tumors within the families varied from 27 to 77 years, and age at onset of index patients varied from 19 to 76 years. Altogether, 69 verified mutation carriers have had only colorectal carcinoma(s), 8 had colorectal carcinoma and endometrial carcinoma, and 5 had only endometrial carcinoma (data not shown).

MSI phenotypes were high in 38, low in 4, and stable in 2 tumors originating from mutation carriers in different families (Table 1). Immunohistochemical staining was performed in tumors from 38 families, and MLH1 protein was lost in 24, decreased in 3, and present in 11 tumors (Table 1). In 4 tumors associated with mutations V213M, K618A, and V716M, MLH1 staining was positive and MSI status was low or stable.

In addition to 31 mutations found in HNPCC families, 3 MLH1 variations listed in the international HNPCC mutation database (http://www.insight-group.org/) were studied: (1) I219V, which has been shown to be nonpathogenic in previous functional assays[20–23] and used here as a functional control; (2) K618T, whose pathogenicity according to previous functional assays has remained partly unsettled[17,20,22,23]; and (3) an in-frame deletion of codons 633–663, comprising exon 17, which we previously found to be pathogenic[12] and used here as a nonfunctional control. All studies were approved by the institutional review boards of the collaborating universities or local ethical committees.

Functional Assays

Site-directed mutagenesis and production of different expression vectors. The MLH1 variants were created using site-directed mutagenesis as described previously.[12,24] All of the primer sequences, polymerase chain reaction (PCR) parameters, and restriction sites are available from the authors upon request. In the first PCR, wild-type (WT) *MLH1* complementary DNA (cDNA) cloned into pFastBac1 plasmid (Invitrogen, Carlsbad, CA) was used as template; the fragment A was created using primer pair forward A (fA) and reverse A (rA) and the fragment B was with primers fB and rB. In the