Gastroenterology 2015;149:1872–1883

Genetic Variants That Predispose to DNA Double-Strand Breaks in Lymphocytes From a Subset of Patients With Familial Colorectal Carcinomas

Sanjeevani Arora,[1,2] Hong Yan,[1] Iltaeg Cho,[3] Hua-Ying Fan,[3] Biao Luo,[4] Xiaowu Gai,[5] Dale L. Bodian,[6] Joseph G. Vockley,[6] Yan Zhou,[7] Elizabeth A. Handorf,[7] Brian L. Egleston,[2,7] Mark D. Andrake,[2] Emmanuelle Nicolas,[8] Ilya G. Serebriiskii,[2,9] Timothy J. Yen,[1] Michael J. Hall,[10,11] Erica A. Golemis,[2] and Greg H. Enders[1,12]

[1]*Cancer Biology Program,* [2]*Molecular Therapeutics Program,* [4]*Institute for Personalized Medicine,* [7]*Biostatistics,* [8]*Genomics,* [10]*Medical Oncology,* [11]*Clinical Genetics,* [12]*Medicine, Fox Chase Cancer Center, Philadelphia, Pennsylvania;* [3]*Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania;* [5]*Massachusetts Eye and Ear Infirmary, Harvard Medical School, Boston, Massachusetts;* [6]*Inova Translational Medicine Institute, Inova Health System, Falls Church, Virginia;* [9]*Kazan Federal University, Kazan, Russia*

BACKGROUND & AIMS: DNA structural lesions are prevalent in sporadic colorectal cancer. Therefore, we proposed that gene variants that predispose to DNA double-strand breaks (DSBs) would be found in patients with familial colorectal carcinomas of an undefined genetic basis (UFCRC). **METHODS:** We collected primary T cells from 25 patients with UFCRC and matched patients without colorectal cancer (controls) and assayed for DSBs. We performed exome sequence analyses of germline DNA from 20 patients with UFCRC and 5 undiagnosed patients with polyposis. The prevalence of identified variants in genes linked to DNA integrity was compared with that of individuals without a family history of cancer. The effects of representative variants found to be associated with UFCRC was confirmed in functional assays with HCT116 cells. **RESULTS:** Primary T cells from most patients with UFCRC had increased levels of the DSB marker γ(phosphorylated)histone2AX (γH2AX) after treatment with DNA damaging agents, compared with T cells from controls ($P < .001$). Exome sequence analysis identified a mean 1.4 rare variants per patient that were predicted to disrupt functions of genes relevant to DSBs. Controls (from public databases) had a much lower frequency of variants in the same genes ($P < .001$). Knockdown of representative variant genes in HCT116 CRC cells increased γH2AX. A detailed analysis of immortalized patient-derived B cells that contained variants in the Werner syndrome, RecQ helicase-like gene (*WRN*, encoding T705I), and excision repair cross-complementation group 6 (*ERCC6*, encoding N180Y) showed reduced levels of these proteins and increased DSBs, compared with B cells from controls. This phenotype was rescued by exogenous expression of WRN or ERCC6. Direct analysis of the recombinant variant proteins confirmed defective enzymatic activities. **CONCLUSIONS:** These results provide evidence that defects in suppression of DSBs underlie some cases of UFCRC; these can be identified by assays of circulating lymphocytes. We specifically associated UFCRC with variants in *WRN* and *ERCC6* that reduce the capacity for repair of DNA DSBs. These observations could lead to a simple screening strategy for UFCRC, and provide insight into the pathogenic mechanisms of colorectal carcinogenesis.

Keywords: Colon Cancer; Hereditary Cancer; Genomic Instability; Tumorigenesis.

BASIC AND TRANSLATIONAL AT

Familial colorectal carcinoma (FCRC) is characterized by early disease onset and/or occurrence of CRC in multiple family members. Several FCRC syndromes have been linked with specific germline defects: familial adenomatous polyposis coli with the Wnt pathway gene *adenomatous polyposis coli*, Lynch syndrome with a group of mismatch repair genes (most commonly *MLH1*, *MSH2*, *MSH6*, and *PMS2*), and MutY-H polyposis with the eponymous base excision repair gene.[1] However, most FCRC remains genetically undefined (UFCRC), accounting for approximately 20% of CRC in the United States.

Clinical guidelines advise starting CRC screening in UFCRC families at earlier ages and, depending on family history, more frequent intervals.[2] Although beneficial, this strategy is inefficient. Family members who are not predisposed genetically are subjected to unnecessary costs and morbidity, although some of those actually at risk may be underscreened. Intensive genome-wide association studies have sought to identify additional genes underlying UFCRC. These studies have yielded only moderate associations at multiple genome locations, implying dauntingly complex genetics.[3–8] No common molecular defect has been recognized.

Abbreviations used in this paper: **BMP, bone morphogenetic protein; DSB, double-strand break; EVS, Exome Variant Server; ExAC, Exome Aggregation Consortium; FA, Fanconi's anemia; FCRC, familial colorectal cancer; HQV, high-quality variant; NER, nucleotide excision repair; PBL, peripheral blood lymphocytes; PP2, PolyPhen2; Pt1, patient 1; ROC, receiver operator characteristic; sCRC, sporadic colorectal cancer; siRNA, small interfering RNA; UFCRC, undefined familial colorectal cancer.**

Most current article

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2015.08.052

Most sporadic cases of CRC (sCRCs) show chromosomal instability, however, its molecular basis has remained ill defined.[9,10] In a few instances, somatic mutations have been found in genes involved in mitosis or mitotic checkpoints.[11] Recent studies have suggested that replicative stress, rather than mitotic defects, may underlie chromosomal instability in many sCRCs.[12] One FCRC family was described in which a germline BUBR1 variant perturbed genome stability in peripheral blood lymphocytes (PBLs),[13] suggesting that PBLs, the cells most readily obtained from patients, might show defects in other UFCRC patients. We hypothesized that genetic defects causing constitutional genome instability underlie a major fraction of UFCRCs and can be detected by biological assays in PBLs. Validation of this hypothesis would suggest strategies to improve screening for CRC.

Methods

Patients and Controls

The Institutional Review Board approved all work. Patients were seen in the Fox Chase Cancer Center Familial Risk Assessment Program. They were diagnosed with CRC before age 50 and/or had another family member with CRC and tested negative for known FCRC syndromes (see the Supplementary Materials and Methods section). sCRCs were defined as individuals with CRCs without a family history of cancer. Our cohort included patients ages 40–82 years. Control samples for biological studies for both UFCRC and CRC were collected from individuals who denied a personal or family history of CRC, and were age- and sex-matched to the patients. For the large-scale exome sequencing comparison, controls were drawn from a population-based study in Virginia[14] and similarly denied a personal or family history of cancer. The Exome Variant Server (EVS) (http://evs.gs.washington.edu/EVS/) and the Exome Aggregation Consortium (ExAC) web site (Cambridge, MA) (version 0.3) (http://exac.broadinstitute.org/) were used to assess the frequency of the selected variants in the general population or in a particular ethnic group. The ExAC data set contains information on 60,706 unrelated individuals. In each patient group, European-descent Caucasians were the dominant racial group.

Lymphocyte Cell Preservation, Culture, Drug Treatments, Flow Cytometry, and Metaphase Spreads

PBLs were collected from patients and controls in an identical fashion and preserved by standard methods and stimulated with phytohemaglutinin and interleukin 2. At 72 hours, cells were left untreated or treated under the following conditions: 20 μmol/L aphidicolin, 100 μmol/L etoposide, or 25 μmol/L camptothecin and fixed in paraformaldehyde 2 hours later, or 5 J/m^2 UV and fixed 5 hours later. For flow cytometry, cells were fixed in ethanol and stained with propidium iodide. Metaphase spreads were generated by classic methods (see the Supplementary Materials and Methods section).

Exome Sequencing and Variant Calling

DNA libraries were prepared from 4 μg genomic DNA using NEBNext Ultra DNA library prep Kit for Illumina (New England Biolabs, Beverly, MA) and sequenced on a HiSeq2500 (Illumina, San Diego, CA). All variants were analyzed on the current PolyPhen2 (PP2) website (VAR version, http://genetics.bwh.harvard.edu/pph2/).[15] To be classified as a high-quality variant (HQV), a patient variant had to receive a score of at least 0.95 (probably damaging) and at least two thirds of the following: a scale-invariant feature transform (SIFT) score of damaging, a Provean score of deleterious (<-2.5), and a MutationAssessor score of at least moderately damaging. HQVs were required to map to the major transcript in the Uniprot or ENTREZ gene databases. Exome sequencing controls were seen at Inova Fairfax Hospital (see the Supplementary Materials and Methods section). Deleteriousness was predicted as described earlier except that any PP2 score greater than 0.85 was included, regardless of its presence in a dominant transcript.

Immunofluorescence and Biochemistry

For scoring in primary lymphocytes, cells were allowed to attach to poly-d-lysine–coated slides or 96-well plates and stained with anti-γ(phosphorylated)histone2AX (γH2AX) antibody (05-636; Millipore, Temecula, CA). Cells on slides were photographed with a Nikon Eclipse E800 microscope (Tokyo, Japan) and the number of bright foci per nucleus was scored. Cells in 96-well plates were imaged on the ImageXpress Micro automated microscope (Molecular Devices, Sunnyvale, CA) driven by MetaXpress software. Images were analyzed in the Multiwavelength Scoring module of MetaXpress and results were displayed and exported using the AcuityXpress software package (Molecular Devices). Confocal microscopy was performed on a BioRad Radiance2000 confocal microscope (Hercules, CA). For immunoblotting for γH2AX, chromatin extracts were prepared from cell nuclei that were disrupted by sonication.[16]

WRN Helicase Assays

The helicase domain (amino acids 467–1031) of Xenopus WRN (Werner syndrome, RecQ helicase-like gene) and the T646I variant (homologous to the human WRN T705I variant in patient 1 [Pt1]) were subcloned into a glutathione-S-transferase fusion expression (pGEX) vector in frame with the glutathione-S-transferase open reading frame. The DNA unwinding assay to detect the helicase activity was performed as previously described[17] (see the Supplementary Materials and Methods section).

ERCC6 Chromatin Remodeling Assay

Constructs encoding the Pt1 excision repair cross complementation group 6 (ERCC6) variant were generated by site-directed mutagenesis (Stratagene/Agilent, Santa Clara, CA).[16] ERCC6 and the ERCC6 N180Y variant were C-terminally tagged with the Flag epitope, expressed using the insect SF9 culture system, purified by affinity chromatography, and assayed for chromatin remodeling as described[18] (see the Supplementary Materials and Methods section).

Comet Assays

B-cell lines cells either were left untreated or treated with aphidicolin 20 μmol/L, camptothecin 25 μmol/L, or UV at 8 J/m^2 where indicated. The presence of DNA double-strand breaks (DSBs) was assessed by neutral comet assay[19] (see the Supplementary Materials and Methods section).