3′-UTR Poly(T/U) Tract Deletions and Altered Expression of EWSR1 Are a Hallmark of Mismatch Repair–Deficient Cancers

Shivendra Kishore1, Salvatore Piscuoglio2, Michal B. Kovac2, Annette Gylling7, Friedel Wenzel2, Francesca Trapani1,4,8, Hans Joerg Altermann5, Valentina Mele3,4, Giancarlo Marra6, Päivi Peltonäki2, Luigi Terracciano4, Mihaela Zavolan1, and Karl Heinimann2

Abstract
The genome-wide accumulation of DNA replication errors known as microsatellite instability (MSI) is the hallmark lesion of DNA mismatch repair (MMR)–deficient cancers. Although testing for MSI is widely used to guide clinical management, the contribution of MSI at distinct genic loci to the phenotype remains largely unexplored. Here, we report that a mononucleotide (T/U)16 tract located in the 3′ untranslated region (3′-UTR) of the Ewing sarcoma breakpoint region 1 (EWSR1) gene is a novel MSI target locus that shows perfect sensitivity and specificity in detecting mismatch repair–deficient cancers in two independent populations. We further found a striking relocation of the EWSR1 protein from nucleus to cytoplasm in MMR-deficient cancers and that the nonprotein-coding MSI target locus itself has a modulatory effect on EWSR1 gene expression through alternative 3′ end processing of the EWSR1 gene. Our results point to a MSI target gene–specific effect in MMR-deficient cancers. Cancer Res; 74(1); 224–34. ©2013 AACR.

Introduction
With an estimated population incidence of about 1 in 370, Lynch syndrome (formerly known as hereditary nonpolyposis colorectal cancer) represents the most common, autosomal dominantly inherited cancer predisposition worldwide (1), caused by germline mutations in DNA mismatch-repair (MMR) genes (MLH1, MSH2, MSH6, and PMS2). Mutation carriers are at an increased risk to develop colorectal cancer as well as a distinct spectrum of extracolonic cancers (predominantly of the endometrium, ovaries, and stomach) at a young age (2).

At the cellular level, biallelic inactivation of the MMR system leads to genome-wide accumulation of DNA replication errors at specific repetitive nucleotide sequences, a condition termed microsatellite instability (MSI). MSI is the hallmark lesion of MMR-deficient cancers (3), but is also observed in 10% to 20% of the sporadic colorectal, gastric, and endometrial cancers. The tumor’s MSI status is increasingly used to guide clinical management (3, 4), because genome-wide gene expression data from MMR-deficient cancers (3), but is also observed in 10% to 20% of the sporadic colorectal, gastric, and endometrial cancers. The tumor’s MSI status is increasingly used to guide clinical management (3, 4), because genome-wide gene expression data from MMR-deficient cancers follows distinct pathogenetic paths (5).

A large number of distinct genic loci affected by MSI have been described, consisting mainly of mono- and dinucleotide repeats within the 5′ untranslated regions (5′-UTR; e.g., NR-27), introns (BAT-26), and the 3′-UTR (CAT25) of specific genes (6). The functional significance of MSI at these noncoding repeat loci and how they may contribute to the pathogenic process is, however, largely unknown (7). Here, we report a novel target gene locus, EWS16T, consisting of a monomorphic polythymine (16T) tract within the 3′-UTR of the Ewing sarcoma gene (EWSR1). We assessed this locus in two independent populations and found that EWS16T discriminates MMR-deficient from MMR-proficient cancers, be they sporadic or hereditary, with high sensitivity and specificity. Despite its location in a noncoding region of the EWSR1 gene, EWS16T contractions are associated with changes in EWSR1 expression and subcellular localization. Our findings thus directly implicate the RNA/DNA-binding Ewing sarcoma protein, better known for its fused variants in Ewing sarcoma, in MSI-associated colorectal tumorigenesis.

Materials and Methods
DNA and RNA Isolation
To isolate genomic DNA and total RNA from cell lines and fresh/frozen tumor tissue samples, Qiagen QIAamp DNA/RNA
Mini Kit (Qiagen) and for formalin-fixed paraffin-embedded tumor samples RecoverAll Total Nucleic Acid Isolation Kit (Ambion, Invitrogen) were used according to the manufacturers’ guidelines.

Analysis of MSI

MSI was assessed on two independent cohorts of patients encompassing 85 Lynch syndrome–related cancers (78 colorectal cancers and 7 endometrial carcinomas), 113 sporadic colorectal cancers, including 14 cases with MLH1 promoter hypermethylation and 12 MMR-proficient, MSI-low cancers, and a Finnish cohort, including 8 Lynch syndrome gastric-related cancers, 50 sporadic colorectal cancers, and 64 sporadic gastric cancers, all of which were compared with matched tumor-free mucosa. On the basis of the recommendations of the National Cancer Institute workshop on MSI, a panel of microsatellite loci (BAT25, BAT26, D2S123, D5S346, and D17S250) and two additional microsatellite markers (BAT40 and MYCL1) were used to determine MSI status (8).

Relative expression of EWSR1 by quantitative PCR

Relative expression of EWSR1 by quantitative PCR was calculated using the standard curve method (10, 11), and the fold-changes in gene expression were normalized using the HPRT1 probes Hs01003270_g1 (Applied Biosystems) and the EWSR1 probe validated by ATCC by short tandem repeat profiling, and used for functional studies within 6 months after thawing from liquid nitrogen tank. All cell lines were maintained at 37°C with 5% CO₂.

Five colorectal cancer cell lines were used for this study: three MMR-deficient cell lines (HCT116, LoVo, and HCT15) and two MMR-proficient ones (SW480 and HT29). HCT116 and HCT15 cell lines were cultured in RPMI 1640 (Invitrogen) supplemented with 10% FBS, 1% kanamycin sulfate, 1% GlutaMAX-I, 1% sodium pyruvate, 1% nonessential amino acids (NEAA), 1% HEPES (all from Invitrogen), and 0.1% 2-mercapto-ethanol (Sigma-Aldrich). LoVo cells were cultured in Ham’s F-12 media supplemented with 10% FBS. HT29 cells were grown in McCoy’s 5A Medium (Invitrogen) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I, all from Invitrogen. SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldrich) with 10% FBS, 1% kanamycin sulfate, and GlutaMAX-I (all from Invitrogen). SW480 cells were cultured in L-15 Medium (Sigma-Aldr...
Mutant constructs with U deletions (EWSR1-2/3/4/5/6 del-3′-UTR/psiCHECK-2-SPAm) and insertions (EWSR1-2/4/ins-3′-UTR/psiCHECK-2-SPAm) were introduced using standard overlap extension PCRs.

Antibodies for Western blot analyses against NF90, SFPQ, and human antigen R (HuR) were obtained from Santa Cruz Biotech, and antibody against EWSR1 for Western blot analyses and Immunoprecipitation were obtained from Abcam.

In vitro transcription

In vitro transcription for pull-down assay using S1 aptamer was performed using T7 RiboMax Express Large Scale RNA Production System (Promega) as per manufacturer’s instructions. Region flanking the wild-type (wt)- and deleted poly T/U tracks in the 3′-UTR of EWSR1 constructs was amplified using primers T7_EWSR1_IVT_CE 5′-GCTTCTAATACGACTCAGTCCATAGGGAGAAATGGGAAACCCTTTGTGAG-3′ and EWSR1_IVT_Cr: 5′-GGGAAGGGCCGAGAAGGAT-3′ to introduce T7 promoter sequences at the 5′ end of the amplicons. S1 aptamer sequence was introduced at the 3′ end using another round of PCR using T7_EWSR1_IVT_CE 5′-GCTTCTAATACGACTCAGTCCATAGGGAGAAATGGGAAACCCTTTGTGAG-3′ and EWSR1_S1apt_IVT_Cr: 5′-CATGGCGCGCCGCGACTATCTTACGCACTTG CATGATTC TGGTCGG-0

Protein pull-down assay

For each sample, 100 µL of the Dynabeads MyOne Streptavidin C1 (Invitrogen) was washed twice with one bead-volume of solution A [diethyl pyrocarbonate (DEPC)-treated 0.05 mol/L NaOH, DEPC-treated 0.1 mol/L NaCl], once with one bead-volume of solution B (DEPC-treated 0.1 mol/L NaCl), and once with RNA binding buffer. The beads were resuspended in one bead-volume of RNA binding buffer (100 mmol/L NaCl, 50 mmol/L HEPES 7.5, 0.5% NP-40, and 10 mmol/L MgCl2) with 100 µg of in vitro transcribed RNA with S1 aptamer sequence and incubated at 10°C for 40 minutes in a thermomixer with intermittent shaking. The beads were washed twice with one volume of RNA wash buffer before incubation with the lysate. HEK293 cell pellet from a 15-cm2 dish was lysed in 3 mL native lysis buffer (25 mmol/L HEPES-KOH pH 7.5, 100 mmol/L KCl, 0.5% NP-40, 5 mmol/L MgCl2, 0.5 mmol/L dithiothreitol, protease inhibitor cocktail, 1 mmol/L NaF, 1 mmol/L NaVO4, and 300 U of RNasin) for 15 minutes on ice. The lysate was subsequently gently sonicated and centrifuged to remove any cell debris. 200 µg of E. coli tRNA was additionally added to prevent nonspecific binding of proteins to the beads. One milliliter of the lysate was added to the beads coupled to S1 aptamer RNA and also to the beads alone for no RNA control. The mixture was incubated at 4°C on rotation wheel. After 1 hour of incubation, the beads were washed three times with native lysis buffer. The bound proteins were eluted with 100 µL native lysis buffer supplemented with 25 mmol/L LiCl to 30-45 minutes at 10°C on a thermomixer with intermittent shaking. Of note, 900 µL of 100% ethanol was added to the eluate and incubated at −80°C for 2 hours followed by centrifugation to precipitate the eluted proteins. The pellet was air-dried and dissolved in 35 µL of (SDS loading dye). Before loading on the NuPAGE Novex 4–12% Bis-Tris gels (Invitrogen), the samples were heated at 90°C for 5 minutes. After SDS-PAGE electrophoresis, the gel was stained with colloidal blue and bands of interest were excised and sent for mass spectrometry.

Poly(A) site selection assay

HeLa cells were transfected with psiCHECK-2-SPAm constructs for 24 hours. Total RNA was isolated from the HeLa cells using Tri Reagent (Sigma) followed by DNase I (Promega) treatment according to the manufacturer’s protocol. Reverse transcription was done with oligo d (T)16 primers for 1 hour. For poly(A) site selection assay, multiplexed PCR was set up using a single forward primer specific to the psiCHECK-2 vector (psiCHECK-2-SeqFor: 5′-ATGAAATGGTGAAGTACA-3′) and two 4 nucleotide terminal anchored oligo d (T)16 reverse primers specific to the two isoforms variants of EWSR1 (EWSpolyAproxR: 5′-TTTCTTTTTTTTTTTTACCA-3′ and EWSpolyAdistR: 5′-TTTTTTTTTTTTTTTTGACT-3′ respectively) to detect only the isoforms generated specifically from the psiCHECK-2-SPAm constructs. PCR was run for 28 cycles and the products were separated on a 2% agarose gel. The bands were quantified using the ImageJ software (http://rsweb.nih.gov/ij/).

Luciferase assays

HeLa cells were seeded in a 48-well plate one day before transfection. Of note, 0.2 µg of plasmids (psiCHECK-2-SPAm constructs) were transfected with Lipofectamine 2000 (Invitrogen) for 24 hours. Luciferase assays were done on the transfected cells using The Dual-Luciferase Reporter Assay System (Promega). Both transfections and luciferase assays were done according to the manufacturer’s protocol.

siRNA transfections

Control-siRNAs and siRNAs against hnRNPC, HuR, NF90, and SFPQ were and were obtained from Santa Cruz Biotechnology. HeLa cells were reverse transfected with siRNA oligos using RNAiMAX (Invitrogen). After 48 hours, the cells were transfected with EWSR1-wt-3′-UTR- and EWSR1-6del-3′-UTR-psiCHECK-2-SPAm-constructs for another 24 hours. The cells were subsequently harvested and split into two aliquots. One aliquot was used to assess the knockdown efficiency of siRNA using Western blot analysis, whereas the other was used for RNA isolation and subsequent poly(A) site selection assay.

Statistical analyses

For statistical analysis, the χ2 test and Fisher exact test for nonparametric variables and Student t test for parametric variables were used, with all probabilities reported as two-tailed, considering a P value of less than 0.05 to be statistically significant. Calculations were performed using the software program SPSS 17.0 (IBM Corporation).

Ethical approval

The study is part of the so-called “Basler Studie über familiare Tumorkrankheiten,” reference number EK258/05 and has been approved by the “Ethikkommission beider Basel.” Furthermore, written informed consent was obtained from all
patients with Lynch syndrome as well as from the sporadic patients.

Results

The initial investigation of a primary (vimentin- and CD99-positive) extra-osseous Ewing sarcoma from a MSH6 mutation carrier (c.3696dupT) previously affected by colon cancer revealed that the tumor displayed the typical hallmarks of Lynch syndrome, i.e., MSI at mononucleotide markers and specific loss of MSH6 expression (Supplementary Fig. S1), but none of the molecular features commonly associated with Ewing sarcoma, i.e., chromosomal translocations involving the EWSR1 gene locus on 22q12 (Supplementary Fig. S2). Reasoning that the EWSR1 gene is likely involved in the development of this Ewing sarcoma tumor and in view of the tumor’s MSI-high phenotype, we focused on a mononucleotide tract in the 3'-UTR of the EWSR1 gene consisting of 16 thymines (EWS16T; c.318_333), which we thought would be a likely MSI target. Strikingly, both of the patient’s tumors, the Ewing sarcoma and the colon cancer, were found to carry somatic contractions/deletions of four and five thymines, respectively. This prompted us to assess the prevalence of somatic alterations at EWS16T in MMR-deficient cancers in general. In contrast with Wheeler and colleagues (16), who reported a deletion variant (c.331_c.333delTTT; dbSNP: rs76631619) in James D. Watson’s genome, we found the EWS16T locus to be quasi-monomorphic in 370 (99.7%) out of 371 constitutional (i.e., tumor-free) DNA samples tested.

EWS16T, a novel target gene locus in Lynch syndrome-related cancers

To determine whether EWS16T is indeed a novel target gene locus in Lynch syndrome, we analyzed 85 matched tumor-free tissue and cancer samples (78 colorectal and 7 endometrial) from 78 Swiss patients with Lynch syndrome with confirmed MMR germline mutation (58 MLH1, 20 MSH2). We further investigated 14 sporadic MMR-deficient colorectal cancers with MLH1 promoter hypermethylation as well as 86 sporadic MMR-proficient colorectal cancers. Assessment of EWS16T tract length by capillary electrophoresis of fluorescently labeled PCR products revealed that all MMR-deficient cancers but none of the MMR-proficient ones displayed novel alleles, i.e., contractions or expansions at the EWS16T tract (Fig. 1). These findings were further confirmed independently in a Finnish cohort of 122 patients. In this patient cohort as well all of the 29 MMR-deficient (12 colorectal cancer, 17 gastric cancers) but none of the 93 MMR-proficient (38 colorectal cancer and 55 gastric) cancers showed EWS16T tract instability. The majority (72.7%) of somatic alterations consisted of contractions/deletions of four or more base pairs (Supplementary Fig. S3). Consistently, we found that all MMR-deficient cell lines investigated (LoVo, HCT15, and HCT116) carried solely mutated EWS16T alleles (contractions) without evidence for a wt tract allele. In contrast, MMR-proficient cell lines (HT29 and SW480) as well as 37 microsatellite-stable colon adenomas and 12 MMR-proficient, MSI-low colorectal cancers from Swiss patients had a stable EWS16T locus (Table 1). Thus, the EWS16T tract represents a novel, quasi-monomorphic MSI target locus identifying hereditary and sporadic MMR-deficient cancers with 100% sensitivity [95% confidence interval (CI), 97-100] and specificity (95% CI, 98–100).

EWSR1 poly(T/U) tract deletions promote distal poly(A) site usage

Fusion of EWSR1 with ETS family transcription factors as a result of chromosomal translocations has long been implicated in the development of Ewing sarcoma (9). It is still debated to what extent is tumorigenesis the consequence of the fusion protein as opposed to the loss-of-function of the normal EWSR1 protein. Because somatic EWS16T tract alterations were exclusively present in MMR-deficient cancers, whether hereditary or sporadic, and occurred in all types of cancer investigated (colorectal, gastric, and endometrial), we wondered about their possible functional role(s) in MMR-related carcinogenesis. The poly(T/U) tract deletions occur in a region that encodes the 3'-UTR of EWSR1 and therefore do not alter the coding sequence of the EWSR1 gene. Nonetheless, 3'-UTRs contain sequence elements that are important for the posttranscriptional regulation of protein levels. Furthermore, it has been recently demonstrated that changes in 3'-UTR length through alternative polyadenylation activates oncogenes (17). We thus set out to characterize the effect of EWS16T tract deletions on EWSR1 expression levels. A recently generated catalog of polyadenylation sites identified through 3' end sequencing from
HEK293 cells (18) shows that EWSRI undergoes alternative polyadenylation, generating two transcript forms that differ in the length of their 3′-UTRs (Supplementary Fig. S4). The EWS16T tract deletions occur very close to the distal poly(A) site and may thus result in changes in 3′ end processing factor assembly, thereby altering the poly(A) site selection. To investigate this possibility, we cloned the 3′-UTR of EWSR1 downstream of the Renilla luciferase-coding region in a psiCHECK-2–mutant vector in which the synthetic poly(A) signal and may thus result in changes in 3′-UTR tract deletions occur very close to the distal poly(A) site (Fig. 2a-2b). We further investigated the MMR-proficient colon cancer samples

Table 1. EWS16T tract instability in 34 colon adenoma and 319 cancer samples

<table>
<thead>
<tr>
<th>Samples analyzed</th>
<th>N</th>
<th>MSI status</th>
<th>EWS16Tr tract status</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>wt n (%)</td>
</tr>
<tr>
<td>MMR-deficient cancers</td>
<td>128</td>
<td>MSI-high</td>
<td>0</td>
</tr>
<tr>
<td>Colorectal</td>
<td>104</td>
<td></td>
<td>126 (98.4)</td>
</tr>
<tr>
<td>MLH1 deficient</td>
<td>76</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MSH2 deficient</td>
<td>28</td>
<td></td>
<td>26 (92.9)</td>
</tr>
<tr>
<td>Gastric</td>
<td>17</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MLH1 deficient</td>
<td>16</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MSH2 deficient</td>
<td>1</td>
<td></td>
<td>11 (100)</td>
</tr>
<tr>
<td>Endometrial</td>
<td>7</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MLH1 deficient</td>
<td>4</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MSH2 deficient</td>
<td>3</td>
<td></td>
<td>0</td>
</tr>
<tr>
<td>MMR-proficient cancers</td>
<td>191</td>
<td>MSS/MSI-low</td>
<td>0</td>
</tr>
<tr>
<td>Colorectal</td>
<td>124</td>
<td></td>
<td>124 (100)</td>
</tr>
<tr>
<td>Colorectal</td>
<td>12</td>
<td>MSI-Low</td>
<td>12 (100)</td>
</tr>
<tr>
<td>Gastric</td>
<td>55</td>
<td>MSS</td>
<td>55 (100)</td>
</tr>
<tr>
<td>MMR-proficient colon adenomas</td>
<td>37</td>
<td>MSS</td>
<td>37 (100)</td>
</tr>
<tr>
<td>Lynch syndrome-related</td>
<td>10</td>
<td></td>
<td>10 (100)</td>
</tr>
<tr>
<td>MLH1 deficient</td>
<td>8</td>
<td></td>
<td>8 (100)</td>
</tr>
<tr>
<td>MSH2 deficient</td>
<td>2</td>
<td></td>
<td>2 (100)</td>
</tr>
<tr>
<td>Sporadic</td>
<td>27</td>
<td></td>
<td>27 (100)</td>
</tr>
</tbody>
</table>

NOTE: The text in bold provides the overall number of cancers according to their respective MSI status. MSI was determined according to Umar and colleagues (8).

To determine the factors involved in EWSR1 pre-mRNAs with EWS16T tract deletions

SFPQ, but not HuR or NF90, mediates distal poly(A) site usage in EWSR1 pre-mRNAs with EWS16T tract deletions

To determine the factors involved in EWSR1 poly(A) site selection, we used S1 aptamer-tagged, in vitro transcribed wt 3′-UTR and a 3′-UTR variants with 6U deletions in the EWS16T region to pull down the proteins that associated with these RNAs (Supplementary Fig. S5). In three independent experiments, we reproducibly identified a set of A/U-rich element-binding proteins that associate with these constructs (Fig. 3A and 3B). Interestingly, we found that nuclear factor 45/90/110 (NF45/90/110), heterogeneous nuclear ribonucleoprotein C (hnRNPC) and HuR associate with the wt but not with the mutant 3′-UTR. NF45 and NF90 have been previously shown to regulate mRNA stability and redistribution of nuclear mRNAs in the cytoplasm (19). Conversely, the EWS16T-mutant preferentially associated with the SFQ/NON0 heterodimer, which is an essential pre-mRNA splicing factor that couples splicing with polyadenylation as a component of a snRNP-free complex with SNRPA/U1A (20). Strikingly, NON0 has recently been implicated in the cellular response to DNA double-strand breaks (PMID: 22941645).

To determine which of the RNA-binding proteins (RBP) identified earlier influenced the poly(A) site selection, we knocked down their expression individually with siRNAs (Fig. 4a) and assessed the poly(A) site usage in reporter constructs that had either the wt or the mutant EWS16T tract (6U deletions) cloned downstream of luciferase. None of the siRNAs influenced the poly(A) site usage in the wt constructs in which the shorter isoform was expressed (Fig. 4b). When the 6U deletion construct was used, however, the knockdown of SFQ and hnRNPC strongly influenced
poly(A) site selection. Specifically, knockdown of SFPQ promoted proximal site usage, whereas the knockdown of hnRNPC led to increased expression of the longer 3'UTR isoform (Fig. 4c). This result was consistent with the observed binding pattern of the RBPs. hnRNPC binding was most prominent immediately downstream of the proximal poly(A) site (our unpublished PAR-CLIP data on hnRNPC), whereas SFPQ most likely bound to the truncated U-tract immediately upstream of the distal poly(A) site, as inferred from the pull-down experiments (Fig. 3a). Interestingly, knockdown of SFPQ did not affect poly(A) site selection in wt constructs, consistent with our earlier findings that SFPQ specifically associated with the construct carrying the 6U deletion. Thus, our results indicate that hnRNPC and SFPQ have antagonistic activity on the processing of mutant EWSR1 pre-mRNA with hnRNPC promoting the generation of the shorter and SFPQ of the longer 3'UTR isoform.

Distal poly(A) site usage is associated with decreased EWSR1 expression

To determine if the choice in polyadenylation site may influence EWSR1 expression, we performed luciferase assays on the constructs that carried the wt or various EWS16T deletion variants. The results shown in Supplementary Fig. S6 indicate a significant downregulation (up to 30%) of protein levels associated with EWS16T tract deletions. We attempted to corroborate our observations in vivo, assessing EWSR1 mRNA expression in a set of eight Lynch syndrome–related MMR-deficient and 5 sporadic MMR-proficient colorectal cancers relative to matched, tumor-free mucosa by quantitative PCR (qPCR). Four (50%) out of the eight Lynch syndrome, but only one (20%) out of the five sporadic colorectal cancer showed significantly (>1.2-fold) increased expression of the longer 3'UTR isoform (Fig. 5A). The total EWSR1 mRNA levels, however, were significantly reduced (>1.2-fold) in both Lynch syndrome–related and sporadic colorectal cancer tissues compared with the matched, tumor-free mucosa (Fig. 5B), pointing to additional mechanisms that regulate EWSR1 expression in sporadic colorectal cancers. Alternatively, qPCR interpretation may have been compromised by the presence of noncancer cells, especially in MMR-deficient cancer specimens.

Figure 2. Poly(A) site selection assay through multiplexed PCR. A, schematic diagram of the multiplexed semi-quantitative protocol for poly(A) site selection assay. A single forward primer, which was either vector-specific [for analyzing poly(A) usage of reporter constructs] or EWSR1-3'UTR-specific [for analyzing endogenous EWSR1 poly(A) usage], was used along with two terminally anchored reverse primers to yield amplicons representing both long and short variants of EWSR1 3'UTR in a single PCR reaction. B, poly(A) site usage in EWSR1 3'UTR constructs with variable poly(T/U) tracts lengths cloned in psiCHECK-2-SPAm vector along with the quantification of distal poly(A) site usage (percentages). C, poly(A) site usage in endogenous EWSR1 gene across various MMR-deficient and -proficient cell lines.
Consistent with the data obtained at the mRNA level, the cancers displayed (on average) an approximately 30% reduction in EWS expression when compared with matched, tumor-free mucosa. Surprisingly, however, MMR-deficient and -proficient cancers significantly differed with regard to the subcellular localization of EWS \((P < 0.001)\). Tumor-free colon mucosa and adenomas from patients with Lynch syndrome as well as sporadic adenocarcinomas showed exclusive nuclear expression (Fig. 6). In contrast, Lynch syndrome–related colorectal cancers displayed diffuse cytoplasmic EWS expression (Fig. 6). These results were subsequently confirmed by IHC analysis of a TMA containing 64 sporadic and 94 Lynch syndrome–related colorectal cancers; we observed an approximately 30% reduction in EWS expression in both groups, but only the Lynch syndrome–related cancers showed diffuse cytoplasmic staining for EWS \((61\% \text{ vs. } 3\% \text{ of the sporadic cancers}; \ P < 0.001)\). Thus, MMR-deficient colorectal cancers, all carrying somatic EWS16T tract alterations, display a distinct subcellular EWS distribution pattern \textit{in vivo}. Further studies, ideally performed on...
EWS16T-mutated colorectal cancer cell lines, are needed to assess if this is directly related to EWS16T tract alterations or, rather, an indirect consequence of MSI-associated genetic instability affecting regulators of EWSR1 protein localization such as the methyl-transferases like PRMT1, which is known to regulate the localization of EWS (21) by methylating glycine/arginine-rich motifs located in the arginine-glycine-glycine domains of EWS.

Discussion

EWS16T, a poly(T/U) tract in the 3′-UTR of the EWSR1 gene, represents a novel, quasi-monomorphic MSI target gene locus that identifies both hereditary and sporadic MMR-deficient colorectal cancers with perfect (100%) diagnostic sensitivity and specificity. Given this high diagnostic accuracy and quasi-monomorphic nature ascertained in more than 300 cancers, EWS16T has the potential to substantially facilitate and improve the accuracy of MSI detection in routine daily practice and prospective studies are now needed to determine whether EWS16T alone could replace the set of multiple diagnostic markers currently used for MSI testing.

3′-UTRs play an essential role in regulating the stability, subcellular localization, and translation of corresponding mRNAs via sequence-specific interactions with trans-acting factors, including small RNAs and proteins (22). Several 3′-UTR point mutations have been linked to the risk of developing cancer in humans (23–28). Recent reports have shown that contractions at 3′-UTR mononucleotide repeats can exert deleterious effects by destabilizing the respective mRNAs associated with them and thereby be of direct pathophysiological relevance (29, 30). Characterization of 3′-UTR MSI target genes thus can provide important insights into the consequences of MMR for gene expression. Our investigations of the in vitro and in vivo effects of EWS16T alterations conclusively show that contractions at this locus indeed affect multiple regulatory mechanisms such as alternative polyadenylation, mRNA/protein expression, and possibly subcellular localization. Given that EWS is an RNA-/DNA–binding protein with many RNA/DNA targets, these alterations provide new leads into the mechanisms behind MSI-associated colorectal tumorigenesis.

The physiologic role of EWS is largely unknown, but based on its domain structure, the protein is thought to be involved in...
such diverse processes as gene expression, RNA processing/transport, and cell signaling. Knockout of EWS in mice results in high postnatal lethality, defects in pre-B cell development, meiotic arrest/germ cell apoptosis, premature cellular senescence, and hypersensitivity to ionizing radiation (31). These observations suggest roles for EWS in homologous recombination, DNA damage response, and maintenance of genome integrity (32). Indeed, we recently found that EWSR1 binds RNAs that originate at intrinsically unstable genomic loci and its knockdown increases the frequency of double-strand DNA breaks (Kishore and colleagues; data not shown).

With respect to tumorigenesis, genetic alterations in EWSR1 were first observed in Ewing sarcoma, the second most common malignant bone tumor in adolescents and young adults after osteosarcoma (16, 33, 34). In about 85% of cases, an EWSR1-FLI1 fusion protein is observed that has retained the N-terminal transcription activation domain but lost the RNA-binding domains, which are replaced with the DNA-binding domain of the fusion partner (35). The fusion protein is constitutively active and has been shown to alter the transcription of several downstream targets. Ewing sarcoma is thus largely thought of as a gain-of-function phenotype. Loss of the EWSR1 function has been largely overlooked, despite the fact that the protein has a canonical RNA-binding domain and has been shown to regulate several RNA-processing events in the nucleus (36, 37). As a member of the TET (TLS/FUS, EWS, and TAF15) family of RNA- and DNA-binding proteins, it has been involved in transcriptional regulation and RNA processing (38–42). Consistent with this potential dual role, EWS has been shown to regulate cyclin D1 transcripts both transcriptionally and at the level of splicing, with the oncogenic fusion protein EWS-FLI1, promoting the expression of the oncogenic cyclin D1b splice variant in Ewing sarcoma cells (43). More recently, EWS has been shown to regulate alternative splicing of the p53 repressor MDM2 and to be part of the microprocessor complex that mediates the genesis of microRNAs (32, 42, 44).

In summary, the EWS16T locus represents a novel, quasi-monomorphic MSI target locus to accurately identify both hereditary and sporadic MMR-deficient cancers. Constructions therein affect EWSR1 at multiple levels, including alternative polyadenylation, mRNA/protein expression, and possibly subcellular localization. Our results thereby implicate the
RNA-DNA–binding protein EWSR1, critical for the maintenance of genome integrity (32), in MSI-associated colorectal tumorigenesis.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: S. Kishore, S. Piscuoglio, M. Kovac, M. Zavolan, K. Heinimann
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): S. Kishore, S. Piscuoglio, M. Kovac, L. Terracciano, M. Zavolan, K. Heinimann
Analysis and interpretation of data (e.g., statistical analysis, bios-statistics, computational analysis): S. Kishore, S. Piscuoglio, M. Kovac, L. Terracciano, M. Zavolan, K. Heinimann
Writing, review, and/or revision of the manuscript: S. Kishore, S. Piscuoglio, M. Kovac, A. Gylling, F. Wenzel, F. Trapani, H.J. Altermatt, V. Mele, G. Marra, P. Peltomäki, I. Terracciano, M. Zavolan, K. Heinimann
European Research Council FP7-ERC-232635 (P. Peltomäki), the Janssen Research Foundation (P. Peltomäki), the Sigrid Juselius Foundation (P. Peltomäki), Biocentrum Helsinki (P. Peltomäki), the European Research Council FP7-ERC-232635 (P. Peltomäki), the University of Basel (M. Zavolan), the Krebsliga beider Basel (K. Heinimann), the Krebsliga Zentralschweiz (K. Heinimann), and Oncosuisse (K. Heinimann).

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.

Received July 23, 2013; revised September 16, 2013; accepted September 29, 2013; published OnlineFirst October 24, 2013.

EWS16T, a Novel Genetic Marker for Microsatellite Instability

References


Deficient cancers are a hallmark of mismatch repair–deficient cancers

Shivendra Kishore, Salvatore Piscuoglio, Michal B. Kovac, et al.


Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-13-2100

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2013/10/24/0008-5472.CAN-13-2100.DC1

Cited articles
This article cites 44 articles, 17 of which you can access for free at:
http://cancerres.aacrjournals.org/content/74/1/224.full.html#ref-list-1

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.