Differences in nucleotide excision repair capacity between newly diagnosed colorectal cancer patients and healthy controls

Jana Slyska\textsuperscript{1,2,3}, Alessio Naccarati\textsuperscript{1}, Barbara Pardini\textsuperscript{1}, Veronika Polakova\textsuperscript{1,3}, Ludmila Vodickova\textsuperscript{1,3,4}, Zdenek Smerhovsky\textsuperscript{1,4,5}, Miroslav Levy\textsuperscript{6}, Ludmila Lipska\textsuperscript{6}, Vaclav Liska\textsuperscript{7} and Pavel Vodicka\textsuperscript{1,3,8}

\textsuperscript{1}Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Academy of Science of the Czech Republic, Videnska 1083, 14220 Prague, Czech Republic, \textsuperscript{2}Department of Genetics and Microbiology, Faculty of Science, Charles University, Albertov 6, 120 00 Prague 2, Czech Republic, \textsuperscript{3}Institute of Biology and Medical Genetics, \textsuperscript{4}Department of Genetics, Institute of Experimental Toxicogenomics and Occupational Health, National Institute of Public Health, Srborava 48, 100 00 Prague 10, Czech Republic, \textsuperscript{5}Institute of Epidemiology, \textsuperscript{6}Department of Surgery, Teaching Hospital and Medical School Pilsen, University, Alej Svobody 80, 304 00 Pilsen, Czech Republic.

\textsuperscript{7}2nd Medical Faculty, Charles University, V uvalu 58, 160 00 Prague 6, Czech Republic, \textsuperscript{8}Department of Surgery, 1st Medical Faculty of Charles University and Thomayer University Hospital, Videnska 800, 142 00 Prague 4, Czech Republic and \textsuperscript{1}Department of Surgery, Teaching Hospital and Medical School Pilsen, University, Alej Svobody 80, 304 00 Pilsen, Czech Republic.

\textsuperscript{*}To whom correspondence should be addressed. Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Academy of Science of the Czech Republic, Videnska 1083, 14220 Prague, Czech Republic. Tel: +420 241 062 694; Fax: +420 241 062 782; Email: pvodicka@biomed.cas.cz

\textsuperscript{1}Received on August 5, 2011; revised on November 3, 2011; accepted on November 14, 2011.

Alteration of DNA integrity is a potential cause of cancer and it is assumed that reduced DNA repair capacity and accumulation of DNA damage may represent intermediate markers in carcinogenesis. In this case-control study, DNA damage and nucleotide excision repair capacity (NER-DRC) were assessed in association with sporadic colorectal cancer (CRC). Both parameters were quantified by comet assay in blood cells of 70 untreated incident patients and 70 age-matched healthy controls. mRNA expression and polymorphisms in relevant NER genes were concurrently analyzed. The aim of this study was to characterize incident CRC patients for NER-DRC and to clarify possible relations between investigated variables. Comet assay and mRNA expression analysis showed that CRC patients differ in repair capacity as compared to controls. Patients had a lower NER-DRC and simultaneously they exhibited higher endogenous DNA damage (for both $P < 0.001$). Accumulation of DNA damage and decreasing NER-DRC behaved as independent modulating parameters strongly associated with CRC. Expression levels of 6 out of 9 studied genes differed between groups ($P \leq 0.001$), but none of them was related to CRC or to any of the studied NER polymorphisms. However, in patients only, $XPC$ Ala499Val modulated expression levels of $XPC$, $XPB$ and $XPD$ gene, whereas $XPC$ Lys939Gln was associated with $XPA$ expression level in controls (for all $P < 0.05$). This study provides evidence on altered DRC and DNA damage levels in sporadic CRC and proposes the relevance of the NER pathway in this malignancy. Further, alterations in a complex multigene process like DNA repair may be better characterized by functional quantification of repair capacity than by quantification of individual genes transcripts or gene variants alone.

Introduction

Colorectal neoplasia is the third most common cancer worldwide and the fourth leading cause of cancer death (1). Recently, decreasing colorectal cancer (CRC) mortality rates have been recorded in developed countries, most likely due to larger efforts in cancer screening and/or improved treatment (2). Since an early diagnosis of CRC is associated with substantially better prognosis and curability, early biomarkers and intermediary end points are called for to dissect the complex process of colorectal carcinogenesis (3).

While in the etiopathogenesis of inherited CRC forms the influence of particular genetic features prevails, sporadic forms (around 80% of all cases) show a multifactorial pattern and are determined by an interplay of multiple genetic and environmental/lifestyle factors (4). Well-established risk factors for sporadic CRC are obesity, diet low in vegetables/fruits and rich in meat, physical inactivity and smoking (5,6). Concurrently, the specific nature of colon epithelium, which is a dynamically changing system, makes it prone to genotoxic attacks of external source and/or spontaneous genetic changes. The above reasons underline the importance of efficient DNA repair machinery to maintain the cellular genomic integrity. Indeed, well-supported associations exist between DNA repair insufficiency and inherited forms of CRC: hereditary nonpolyposis CRC is caused by deficient mismatch repair and MUTYH-associated polyposis is associated with deficient mutation in the base excision repair (BER) gene MutY (7).

In sporadic CRC, there is no single germ-line mutation causing a strong deficiency in DNA repair activity. However, alterations of the individual DNA repair capacity (DRC) may significantly modulate the susceptibility to this cancer, especially in the context of gene–environment interactions. Family-based studies have suggested that DRC is a phenotype with a strong genetic basis, estimating heritability in the range of 48–75% (8). DNA repair may be further modulated by environmental/lifestyle factors via several possible mechanisms, such as activation/inhibition of repair enzymes, different provision of building blocks (nucleotides) for the repair machinery or regulation of repair gene expression (9). Thus, DRC, comprising the effect of both hereditary components and variable environmental factors, characterizes the actual phenotype of the cells. Due to the above aspects, estimation of DRC is naturally becoming a representative biomarker and an integral part of modern molecular epidemiological studies on cancer.

Nucleotide excision DNA repair (NER), one of the major players in maintaining genomic integrity, may modulate predisposition to various cancers as well as response to subsequent treatment (10). So far, reduced NER-DRC was...
associated with several types of sporadic malignancies (summarized in Table 1). However, there is a lack of data on NER-DRC in association with sporadic CRC. Additionally, the characterization of individual DNA repair profiles may be of further importance for the implementation of targeted therapies. For instance, platinum-based compounds are first-line cytostatic agents whose rate is measured by our functional assay. Further, we have genotyped common single nucleotide polymorphisms (SNPs) in four NER genes. Additionally, the above variables were also analyzed in the context of individual biological and lifestyle factors and principal clinical characteristics.

The strength of our approach was that we included in the study only patients with newly diagnosed, histologically confirmed sporadic CRC. The homogeneity of the study group was further potentiated by the fact that all patients were sampled prior to any surgical intervention or therapy, thus eliminating the effects of these factors on the individual level of investigated parameters and minimizing the bias on the results.

### Materials and methods

#### Study population

The study population comprised 70 patients with sporadic CRC and 70 healthy controls frequency matched for age. All participants were of Caucasian origin. Incident cases were recruited at the time of the diagnosis among patients visiting two surgical departments (Thomayer Teaching Hospital in Prague and Teaching Hospital and Medical School in Pilsen, Czech Republic) between 2008 and 2010. Only new histologically confirmed CRC cases who did not receive any surgery or other specific treatment prior to sampling were included into the study. Collaborating clinicians provided clinical and pathological characteristics for each CRC patient: presence of inflammatory processes, number of polyps at colonoscopy, tumor localization, stage of the malignancy [tumor, nodes, metastases staging system (TNM) according to Union of International Cancer Control] and microsatellite instability (MSI) status. Controls have not been exposed to any potentially harmful chemicals except for those from environmental sources. No other selection criteria have been applied.

Participating subjects were properly informed about the aim of the research; they signed a written consent and the approval for genetic analysis, in accord with the Helsinki declaration. The Ethics Committees of the Thomayer Teaching Hospital in Prague and Teaching Hospital and Medical School in Pilsen (Czech Republic) approved the design of the study. Trained personnel

#### Materials and methods

**Table I. Reduced NER capacity detected in blood cells of patients with various types of sporadic cancer—an overview of the literature**

<table>
<thead>
<tr>
<th>Cancer</th>
<th>Cases</th>
<th>Controls</th>
<th>Damage-inducing agent</th>
<th>Assay</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder</td>
<td>106 incident cancer, previously untreated</td>
<td>137 matched by age, sex and ethnicity</td>
<td>BPDE</td>
<td>Comet assay</td>
<td>Schabath et al. (11)</td>
</tr>
<tr>
<td>Breast</td>
<td>33 breast carcinoma patients 69 newly diagnosed, previously untreated</td>
<td>47 matched by age 79 matched by age and ethnicity</td>
<td>UVC BPDE</td>
<td>HCR</td>
<td>Ramos et al. (12)</td>
</tr>
<tr>
<td>Colorectal</td>
<td>40 sporadic CRC 28 hereditary predisposition to CRC</td>
<td>39 not specified</td>
<td>N-AcO-2-FAA UDS</td>
<td>HCR</td>
<td>Peto et al. (14)</td>
</tr>
<tr>
<td>Cutaneous melanoma</td>
<td>312 newly diagnosed, previously untreated</td>
<td>324 matched by age, sex and ethnicity</td>
<td>UV HCR</td>
<td>Wei et al. (15)</td>
<td></td>
</tr>
<tr>
<td>Head and neck</td>
<td>55 newly diagnosed, previously untreated 123 newly diagnosed, previously untreated 744 newly diagnosed, previously untreated</td>
<td>61 matched by age, sex, and smoking status 136 matched by age and sex 753 matched by age, sex and ethnicity</td>
<td>BPDE HCR</td>
<td>Cheng et al. (16)</td>
<td></td>
</tr>
<tr>
<td>Non-melanoma skin</td>
<td>88 primary cancer 255 newly diagnosed, previously untreated</td>
<td>135 matched by sex and age 333 unmatched</td>
<td>UV HCR</td>
<td>Wei et al. (19)</td>
<td></td>
</tr>
<tr>
<td>Non-small cell lung</td>
<td>316 newly diagnosed 467 newly diagnosed, previously untreated 108 with second primary cancer 271 lung carcinoma patients</td>
<td>316 matched by age, sex, and smoking status 488 matched by age, sex, ethnicity and smoking status 99 matched by age and sex 271 matched by age and smoking</td>
<td>BPDE HCR</td>
<td>Wei et al. (21)</td>
<td></td>
</tr>
<tr>
<td>Prostate</td>
<td>75 newly diagnosed, previously untreated 65 prevalent cancer</td>
<td>96 matched by age and race</td>
<td>UV HCR</td>
<td>Hu et al. (25)</td>
<td></td>
</tr>
</tbody>
</table>

UV, ultraviolet light; HCR, host-cell reactivation assay; N-AcO-2-FAA, N-acetoxy-N-2-fluorenylacetamide; UDS, unscheduled DNA synthesis.
ner and colorectal cancer

Table II. Characteristics of the study population

<table>
<thead>
<tr>
<th>Characteristics Category</th>
<th>Controls*</th>
<th>CRC patients</th>
<th>P-valueb</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td>n = 70</td>
<td>n = 70</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>36</td>
<td>24</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>34</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>n = 70</td>
<td>n = 70</td>
<td></td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>62.1 ± 12.7</td>
<td>65.4 ± 10.1</td>
<td></td>
</tr>
<tr>
<td>Median (quartiles)</td>
<td>60 (52–74)</td>
<td>66 (58–74)</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>39–86</td>
<td>39–84</td>
<td></td>
</tr>
<tr>
<td>Diagnosis</td>
<td>n = 70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td></td>
<td>38 (54.3%)</td>
<td></td>
</tr>
<tr>
<td>Rectum</td>
<td></td>
<td>32 (45.7%)</td>
<td></td>
</tr>
<tr>
<td>TNM</td>
<td>n = 69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td></td>
<td>7 (10.2%)</td>
<td></td>
</tr>
<tr>
<td>II</td>
<td></td>
<td>25 (36.2%)</td>
<td></td>
</tr>
<tr>
<td>III</td>
<td></td>
<td>16 (23.2%)</td>
<td></td>
</tr>
<tr>
<td>IV</td>
<td></td>
<td>21 (30.4%)</td>
<td></td>
</tr>
<tr>
<td>Smoking status</td>
<td>n = 69</td>
<td>n = 66</td>
<td></td>
</tr>
<tr>
<td>Non-smokers</td>
<td>54 (78.3%)</td>
<td>50 (75.8%)</td>
<td></td>
</tr>
<tr>
<td>Smokers</td>
<td>15 (21.7%)</td>
<td>16 (24.2%)</td>
<td></td>
</tr>
<tr>
<td>Number of cigarettes/day</td>
<td>n = 14</td>
<td>n = 14</td>
<td></td>
</tr>
<tr>
<td>≤10</td>
<td>9 (64.3%)</td>
<td>6 (42.9%)</td>
<td></td>
</tr>
<tr>
<td>&gt;10</td>
<td>5 (35.7%)</td>
<td>8 (57.1%)</td>
<td></td>
</tr>
<tr>
<td>Alcohol consumption</td>
<td>n = 69</td>
<td>n = 65</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>26 (37.7%)</td>
<td>18 (27.7%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>43 (62.3%)</td>
<td>47 (72.3%)</td>
<td></td>
</tr>
<tr>
<td>Alcohol intake (grams/day)</td>
<td>n = 43</td>
<td>n = 43</td>
<td></td>
</tr>
<tr>
<td>&lt;25</td>
<td>24 (55.8%)</td>
<td>28 (65.1%)</td>
<td></td>
</tr>
<tr>
<td>25–49.9</td>
<td>12 (27.9%)</td>
<td>80 (18.6%)</td>
<td></td>
</tr>
<tr>
<td>50–74.9</td>
<td>40 (9.3%)</td>
<td>50 (11.6%)</td>
<td></td>
</tr>
<tr>
<td>75–100</td>
<td>3 (7%)</td>
<td>2 (4.7%)</td>
<td></td>
</tr>
<tr>
<td>Body mass index</td>
<td>n = 61</td>
<td>n = 54</td>
<td></td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>26.1 ± 4.2</td>
<td>27.5 ± 4.5</td>
<td></td>
</tr>
<tr>
<td>&lt;18.5</td>
<td>1 (1.7%)</td>
<td>2 (3.7%)</td>
<td></td>
</tr>
<tr>
<td>18.5–24.9</td>
<td>29 (47.5%)</td>
<td>14 (25.9%)</td>
<td></td>
</tr>
<tr>
<td>25.0–29.9</td>
<td>24 (39.3%)</td>
<td>23 (42.6%)</td>
<td></td>
</tr>
<tr>
<td>30.0–40.0</td>
<td>7 (11.5%)</td>
<td>15 (27.8%)</td>
<td></td>
</tr>
<tr>
<td>Family history of cancer</td>
<td>n = 68</td>
<td>n = 66</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>41 (60.3%)</td>
<td>35 (53%)</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>27 (39.7%)</td>
<td>31 (47%)</td>
<td></td>
</tr>
<tr>
<td>CRC in family</td>
<td>n = 61</td>
<td>n = 63</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>53 (86.9%)</td>
<td>56 (88.9%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>8 (13.1%)</td>
<td>7 (11.1%)</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>n = 66</td>
<td>n = 65</td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>59 (89.4%)</td>
<td>53 (81.5%)</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>7 (10.6%)</td>
<td>12 (18.5%)</td>
<td></td>
</tr>
</tbody>
</table>

*Data were not available for all study participants for some parameters.

bOne-way ANOVA for categorical variables, Kruskal–Wallis test for continuous variables. Significant P-values shown in bold.

Interviewed patients and controls, using a structured questionnaire. Study subjects provided information on their lifestyle habits, body mass index (BMI), diabetes and family/personal history of cancer (Table II). Lifelong or long-term (at least six consecutive months) drug use or exposures to genotoxins were also investigated by the questionnaire.

Isolation of PBMC

Eight milliliters of peripheral venous blood were drawn from each subject into heparinized tubes, mixed 1:1 with RPMI 1640 medium (HEPES modification, containing 2 mM L-glutamine, 10% heat-inactivated fetal bovine serum, 1.5% phytohemagglutinin and 0.2% penicillin/streptomycin, Sigma-Aldrich), layered over Histopaque-1077 (Sigma-Aldrich) and centrifuged at 320 g for 40 min at room temperature (RT). Isolated PBMC were counted and their viability was checked by trypan blue exclusion. When viability was higher than 95%, cells were aliquoted into cultivation tubes with medium (~10^5 cells per 5-ml medium). Tubes were incubated at 37°C. After a mitogen-stimulation period of 20 h, PBMC were further processed for the challenge assay.

Challenge assay for evaluating NER capacity

NER-DRC was analyzed as a level of intermediate single-strand breaks (SSBs) in DNA of (+)-anti-Benz[a]pyrene-7,8-dihydrodiol-9,10-epoxide (BPDE)-treated cells, originated during the incision of BPDE-DNA adducts by NER pathway. The increase in DNA breaks reflects the ability of NER machinery to recognize and remove corresponding adducts from DNA.

The methodology is described in detail in Slyskova et al. (29). Briefly, BPDE was added into the medium with mitogen-stimulated PBMC at a concentration of 1 μM for 30 min at 37°C. After challenge, old medium was replaced by a new one to remove BPDE excess. PBMC were harvested immediately after the treatment (time 0) or further cultured and harvested at 1, 2 and 4 h. Separated from medium by centrifugation, cells from each experimental point were rewashed with PBS and further processed using a standard comet assay protocol.

Comet assay

SSBs in DNA were analyzed by the alkaline comet assay based on a routinely used protocol (30). Experimental conditions for lysis, alkali treatment, electrophoresis, neutralization and scoring are presented in (29). Data are reported as tail DNA%, determined in 50 randomly selected cells from two parallel slides per experimental point.

DRC and endogenous DNA damage calculation

For each experimental point, the net DNA damage value was calculated by subtracting the basal control tail DNA% of untreated cells from the tail DNA% of treated cells. Final and reported DRC value was obtained as a difference between the net level of tail DNA% measured immediately after the treatment
with BPDE at time 0, and the net value at the maximal increase of tail DNA% detected up to 4 h of culturing.

The parameter reported as ‘endogenous DNA damage’ represents the mean value of all independent measurements of tail DNA% of the untreated control PBMCs.

SSBs in DNA per 10^6 dalton can be derived from tail DNA% by multiplying by a conversion factor of 0.042 based on a calibration curve (over the range of damage detected in the current study, the calibration curve is linear) as reported in (31).

Expression analysis
RNA isolation and quality control. Total RNA from 2 ml of fresh peripheral blood sampled into EDTA vacutainers was isolated using TRizol according to the procedure supplied by manufacturer (Invitrogen, Paisley, UK) and was kept at −80°C. RNA integrity (RIN) was measured using capillary electrophoresis performed on Agilent Bioanalyzer 2100, with RNA 6000 Nano Assay (Agilent Technologies, Palo Alto, CA). RIN of all samples were in the range between 8.0 and 10.0. RNA quantity and purity was measured using ASP-3700 Micro-volume UV/Vis Spectrophotometer (Avans-Biotechnology, Taiwan). OD_{260/280} ratios for all samples were between 1.8 and 2.0. Inhibition testing was performed for all samples by adding internal control template DNA (spike DNA), using Internal DNA extraction control kit (Primer Design, Southampton, UK) and following manufacturer’s instructions.

cDNA synthesis and RT-qPCR. Complementary DNA (cDNA) was synthesized from 1 μg of total RNA by using a ReverTaq™ First strand cDNA synthesis kit (MBI Fermentas, Vilnius, Lithuania) with random hexamer primers in a final volume of 40 μl following manufacturer’s instructions. cDNA was stored at −20°C. qPCR was performed on 7500 Real Time PCR system (Applied Biosystems, Foster City, CA) using chemicals produced by Primer Design Ltd. Precision™ 2 × qPCR Mastermix and custom designed real-time PCR assays with PerfectProbe™ were used. All target genes (ERCC1, RAD23B, RPA1, XPA, XPD, XPC, XPD, XPF, XPG) assays were individually designed and were fully validated, with guaranteed priming specificity (BLAST screening) and > 90% of efficiency. Primer sequences are shown in Supplementary Table I, available at Mutagenesis Online. The PCR reactions were performed in a volume of 20 μl, containing 1 μg of cDNA for each sample. Cycling program was set at initial hold at 95°C for 10 min, followed by 50 cycles of denaturation at 95°C for 15 sec, annealing and extension at 60°C for 32 sec and 72°C for 15 sec. Each run contained positive (interplate calibrator, 25 ng of human cDNA) and negative (no template) control. Results were analyzed using integrated 7500 System SDS Software version 1.3.1 (Applied Biosystems).

Reference genes were selected from a geNorm™ housekeeping gene selection kit of 12 genes with PerfectProbe™ and analyzed by both Genorm and NormFinder algorithms (GenEx Professional, MultiD Analyses AB, Göteborg, Sweden). Two combinations of selected reference genes (TOP1, EIF4A2 and B2M, CYCl) were tested for stability in all study samples, but none of them proved to be a reliable normalization factor, and the same was observed for total RNA amount. Therefore, Cq values of target genes were normalized to mean expression of all genes, as it was shown to be the best normalization factor, applying both Genorm and Normfinder algorithms. M-value for mean expression of all genes was 0.1, when ignoring groups, or 0.02, when stratified for groups (patients versus controls). Data are expressed as relative to maximum quantities (lowest expression was considered as 1).

Expression analyses were performed following MIQE guidelines (11).

Genotyping analysis
Considering the size of our study population, SNPs were chosen according to the minor allele frequency (MAF > 0.25) and according to the expected effect on DRC phenotype based on (12). All subjects were genotyped for five polymorphisms in four NER genes: XPA, XPC, XPD and XPG. For XPD Lys751Gln (rs28365048), XPG Asn104His (rs17655) and XPC Lys939Gln (rs2228001) genotyping a PCR-RFLP procedure was carried out using primers and conditions previously described (13). For XPC Ala499Val (rs2228000), primers and conditions of reaction have been described in (15). XPA G23A (rs1800975) has been analyzed with TaqMan allelic discrimination assay (Applied Biosystems; Assay-on-demand, SNP genotyping products: C_482935_1). The results were regularly confirmed by random regenotyping of >10% of the samples for each polymorphism and showed concordant results. The genotypes with ambiguous and/or no results were excluded from the data set. Distribution of genotypes in the study group is shown in Supplementary Table II, available at Mutagenesis Online.

Statistical analysis
Investigated parameters were normalized by logarithmic transformation due to their asymmetric distribution in the study population. The relationships between variables of interest at the bivariate level were studied by means of T-test, ANOVA and Pearson correlation. The strength of associations between CRC occurrence and categorical variables at binary level were tested by Chi-square test. The binary logistic regression was employed to study the simultaneous association of the DNA damage, NER-DRC and gene expression with CRC, adjusted for age, sex, smoking habit, alcohol consumption, BMI and family history of cancer. Genotype frequencies for each polymorphism were tested for compliance with Hardy–Weinberg equilibrium. All statistical tests were performed at 5% level of statistical significance; for expression data, correction for multiple testing analyses (significant P-value after correction being 0.005) was applied. The SPSS analytical package version 16.0 (Chicago, IL, USA) was employed for all statistical analyses.

Results
Study population
The study was carried out on 70 incident CRC patients and 70 healthy controls (mean age ± SD 65.4 ± 10.1 and 62.1 ± 12.7 years, respectively). In patients, malignancy in colon accounted for 54.3% of cases, whereas the rest were diagnosed for rectal cancer. The TNM staging was available for 69 patients: 7 were classified as stage I, 25 belonged to the stage II, 16 to stage III and stage IV was assigned for 21 individuals. In one patient, the pathologist failed to determine TNM. MSI status was available for 41 cases and 7 out of them (17.1%) were MSI unstable. The distribution of all clinical, biological or lifestyle characteristics is reported in Table II. No differences were observed between patients and controls, except for sex distribution (males prevailed among patients, $P = 0.042$) and for BMI (lower in controls, $P = 0.039$).

Endogenous DNA damage
Significantly higher endogenous DNA damage was observed in CRC patients, with median 25.9 (interquartile range 4.0–43.0) tail DNA% as compared to controls, median 9.3 (interquartile range 2.4–21.5) tail DNA%, ($P < 0.001$; Figure 1). After categorizing DNA damage into quartiles, we observed that incident CRC patients were over-represented in the fourth quartile category with the highest DNA damage (odds ratio [OR] 11.49, 95% confidence interval [CI] 3.36–39.34, $P < 0.001$), while the controls prevailed among those with the lowest level of DNA damage (first quartile). Investigated biological or lifestyle factors (age, sex, smoking status, alcohol consumption, BMI, family history of cancer and diabetes),
when included into the binary logistic regression model, did not significantly affect the value of regression coefficients.

There was no association between DNA damage and TNM stage or localization of the tumor.

NER-DRC capacity

The CRC patients, with median 8.7 (interquartile range 2.7–14.3) tail DNA%, exhibited significantly lower NER-DRC than the controls with median 12.9 (interquartile range 7.4–20.6) tail DNA% (P < 0.001; Figure 2a). After categorizing NER-DRC into quartiles, the subjects with the lowest NER-DRC (the first quartile) comprised mainly CRC patients, whereas control subjects were the majority in the fourth quartile with the highest NER-DRC (OR 0.1, CI 0.03–0.32, P < 0.001). Similar to DNA damage, none of the investigated biological or lifestyle factors were associated with NER-DRC and their inclusion into the binary logistic model did not affect the value of regression coefficient. DNA damage and NER-DRC did not significantly correlate, either in the pooled study population or according to the diagnosis. Despite only a moderate decrease of NER-DRC in patients with TNM from I to III, a most pronounced reduction in NER-DRC was observed in patients with stage IV (P = 0.036, Figure 2b).

Expression and genotyping analyses

Expression profile was analyzed in a subgroup of 66 patients and 42 controls, for which RNA material was available and which pass the selection criteria for RNA purity and quality and control of PCR inhibition (as described in Materials and methods). Out of nine studied genes, mRNA levels of six of them significantly differed between patients and controls, also after applying correction for multiple testing analysis (significant P-value after correction being 0.005). Expression levels of XPB and XPF genes were higher in control group, while higher expression levels of XPA, XPG, ERCC1 and RAD23B were detected in patients (for all P ≤ 0.001, Figure 3). A strong relationship between expression levels of RAD23B, XPG and ERCC1 genes was observed in the whole study group (R = 0.98, P < 0.001) and after stratification for patients and controls. Expression levels of any studied repair gene did not correlate with DNA damage or DRC and were not modulated by any of clinical, biological or lifestyle factors.

Distribution of the analyzed genotypes was in agreement with the Hardy–Weinberg equilibrium. None of the studied

Fig. 2. (a) NER-DRC in CRC patients and healthy controls. (b) NER-DRC in CRC patients stratified according to TNM staging.

Fig. 3. mRNA expression levels of nine NER genes in CRC patients and healthy controls, expressed as quantities relative to the lowest detected expression assigned as value 1.
SNPs were significantly associated with DNA damage levels or DRC, either in patients or in controls. Expression levels were only moderately modulated by some SNPs. In cases only, variant allele of *XPC* Ala499Val was associated with lower *XPC* and higher *XPB* and *XPD* expression levels (for all *P* < 0.05). In controls, variant allele of *XPC* Lys939Gln was associated with higher *XPA* expression level (*P* < 0.05; data not shown).

### Discussion

DRC reflects the actual capacity of the organism to maintain DNA integrity and constitutes an informative biomarker of intermediate cancer phenotype (19). Our study represents an investigation on basal DNA damage and DRC in relation to sporadic CRC. In particular, we have focused our interest on evaluation of individual DRC characterizing NER activity. To approach this, we have challenged PBMC of study subjects by BPDE and quantified the removal of BPDE adducts from DNA (which reflects rate-limiting incision step of NER pathway) by a modified version of comet assay. BPDE was chosen as a model compound for two main reasons. First, it is a metabolite of carcinogen benzo(a)pyrene (BaP), to which an organism is commonly exposed from various sources. Along with environmental or occupational pollution, BaP is generated also by pyrolysis in burning tobacco or in meat prepared at high temperatures. Second, BPDE binds to DNA, forming predominantly N2-deoxyguanosine bulky adducts, which have been detected in colonic mucosa (20) and are specifically removed by NER (16). Generally, NER recognizes a wide spectrum of bulky DNA lesions induced by UV light and a variety of helix-distorting agents (17). A reduced NER capacity may consequently enhance the CRC risk due to a diminished protection of intestinal epithelium against genotoxic compounds, present in the lumen or transported by the blood.

In our study group, newly diagnosed CRC patients had significantly lower NER-DRC and higher levels of SSBs in DNA as compared to the age-matched healthy subjects. This was clearly documented after categorization of the above parameters into quartiles, where accumulation of SSBs and decreasing DRC were characteristic for CRC patients. Interestingly, the two parameters behaved as factors independently associated with sporadic CRC. Furthermore, when the cases were stratified according to the TNM staging, significantly lower NER-DRC was observed in patients with the stage IV in comparison to those with less severe/invasive stages of the disease. A gradual decrease of DRC with increasing TNM stage in a surrogate tissue possibly reflects important biological phenomenon in relation to progression of the disease. Deficient or reduced NER-DRC was already reported as a risk factor for several different cancers, including bladder (18), breast (21,22), skin (23–25), head and neck (14,32,33), lung (34–37) and prostate cancer (38), as summarized in Table I. Our results contribute to the list of evidences on the importance of NER-DRC in carcinogenesis, showing the same relevance also for sporadic CRC. To our knowledge, the only study investigating specifically NER capacity in sporadic CRC patients was performed almost 30 years ago by unscheduled DNA synthesis in smaller study group (39).

Simultaneously with reduced NER-DRC, CRC patients exhibited elevated endogenous DNA damage. The arbitrary unit of tail DNA% may be converted to SSBs/10^6 nucleotides (31,40). By expressing our data in this latter unit, we obtain 0.18 versus 0.37 SSBs/10^6 nucleotides in studied controls and patients, respectively which means ~540 versus 1113 breaks/cell. Pooling together data from 119 publications reporting endogenous DNA damage extent in cancer-free populations (41), a clear positive correlation was observed between level of SSBs and age. For individuals belonging to the 50- to 64-year-old group, an average SSB level of 510 breaks/cell was calculated, which is in full agreement with our data. This demonstrates that CRC patients bear >2-fold higher level of strand breaks in DNA than the reference value for healthy population of the same age. The enhanced levels of DNA damage represent an additional suggestion for a generally altered status of the DNA repair machinery among cancer patients. In agreement with our findings, several studies have shown that basal DNA damage is indeed increased in leukocytes of patients suffering from a variety of different forms of cancer, as reviewed by (42). On the other hand, we cannot rule out that the level of endogenous DNA damage may comprise various alkali-labile DNA lesions, converted during comet assay into SSBs. These lesions are a target for BER pathway, which was also reported to be deficient in cancer patients (43,44).

Additionally, mRNA quantity of nine genes involved in the recognition/incision step of NER was studied. Expression levels of individual genes were not significantly related to either NER-DRC or the extent of DNA damage. However, a coordinated expression of *RAD23B*, *XPG* and *ERCCI1* genes was observed. This is an interesting finding considering that XPC-RAD23B complex acts in DNA damage recognition and subsequently recruits the XPG and ERCC1/XPF to the site of damage directly or via strong interactions with TFIIH nine subunits complex (45,46). Six NER genes were found to be differently transcribed between patients and controls. *XPB* and *XPF* had higher expression in controls, while *XPA*, *XPG*, *ERCC1* and *RAD23B* were more expressed in patients. Such a finding is somehow surprising, as total DRC was shown to be lower in patients. Observed expression profiles and the lack of correlation between quantity of mRNA and DRC may be in concordance with previous observations showing that mRNA quantity does not necessarily reflect the activity of protein (47,48) or overall repair capacity (49,50). In some cases, it is the absence rather than the relative content of a functional protein that is important for cellular activity. Moreover, genes act in interactive networks and alterations of each of them might have different impact on the overall cell function (51). Communication between the NER system and DNA damage signaling may also play a critical role (52).

In the present study, we have also attempted to relate outcomes of functional tests and expression levels to relevant variation in some NER genes. The individual differences in DNA damage levels and DRC as well as expression levels have been hypothesized for many years to be associated to individual genetic background in DNA repair genes (13,29). In our hands, investigated SNPs were not significantly associated with either DNA damage or DRC after stratification for health status. However, there were some associations with NER gene expression, the most interesting being a modulation of expression levels of *XPC*, *XPB* and *XPD* by *XPC* Ala499Val variant allele in CRC patients. A potential functional effect of this SNP is supported by previous studies on susceptibility to cancer (53,54). A role of SNPs in NER pathway on sporadic CRC risk has been recently postulated (55,56), but the limited size of our currently studied population precludes evaluation of the association of individual SNPs with CRC risk. The actual
association of variation in DNA repair genes and CRC risk has not been clearly disclosed either in the context of recent genome-wide association studies (57).

In conclusion, our study provides evidence on alterations of cellular DRC among sporadic CRC patients and suggests the role of NER in its etiology. However, one of the main future challenges with intermediate biomarkers, like DNA damage and DRC, is to understand whether they belong to the causal pathway of a disease, whether they are simply a side effect of a disease or whether their measurement may be confounded by some other factors.

Genes involved in DNA damage recognition/incision phase of NER act in an interplay and in a synchronized way (45,58). Based on our findings, this process is more comprehensively characterized by functional quantification of repair capacity than by quantification of individual gene transcripts or gene variants. Overall, our observation points out the usefulness of DRC analyses, which measure the real outcome of a complex multigene process, as also recently concluded by Collins and Azqueta (59). Comet assay is a convenient methodology for DRC evaluation and high-throughput versions of it are currently under development. This will allow the simultaneous determination of DRC in multiple samples and thus making this assay suitable for large population screening minimizing inter-experimental variation (60).

Supplementary data
Supplementary Tables I and II are available at Mutagenesis Online.

Funding

Acknowledgements
The authors are thankful to Jitka Bila for her excellent technical help and Vendula Rusnakova for helping with expression data analysis. We appreciate the kind expertise and advice of Prof. Rudolf Stetina in finalizing the manuscript. We are also grateful to medical staff involved in the biological samples and data collection and to all individuals that participated to the study.

Conflict of interest statement: None declared.

References


