Charles University in Prague

Faculty of Science

Study program: Molecular and cell biology, genetics and virology

Elena Tulupova, MSc

MODULATION OF SPORADIC COLORECTAL CANCER RISK BY POLYMORPHISMS AND HAPLOTYPES OF MISMATCH REPAIR GENES

PhD Thesis

Supervisor: Radim J. Šrám, MD, DSc
Institute of Experimental Medicine AS CR

Prague, 2016

Statutory Declaration

I hereby formally declare that this work was prepared by myself and that I cited all information sources and literature used. I have not submitted this work or it‘s substantial part to obtain another or the same academic degree.

In Prague, 11.6.2016

Signature:

ACKNOWLEDGMENT

I am grateful to all people who have been providing me with their help and who have supported me during the work on my PhD thesis. Many thanks belong especially to my supervisor Radim J. Šrám, MD, DSc, for his incredible support and patience; to all colleagues from the Department of Genetic Ecotoxicology, Institute of Experimental Medicine AS CR, as well as to colleagues from the Department of the Molecular Biology of Cancer leaded by Pavel E. Vodička, MD, PhD, and to my family and friends for their generous encouragement. My special acknowledgement belongs to my son Georgy Ryabchenko for improving my motivation and concentration; to Nicolas Wind for taking care of me; to Tatyana Kobets, PhD, for her invaluable consultations, and to Damien Smethurst, Irina Kim, Tomáš Tetiva and Zbyšek Mošna for their kind help with final proofreading and language corrections.

I apologize to the authors of many scientific works that due to space limitations I was not able to cite all their original publications, so in many instances I referred to corresponding review articles.

This work was kindly supported by grants IGA MH CR NR8563-5/2005, EU Diephy FOOD-CT-2003-505609; Grant Agency of the Czech Republic GACR 310/05/2626, GACR 310/07/1430; by AVOZ 50390512, AVOZ 50390703 and 50390512 of the Czech Republic and by BMBF for NGFN+ (publications 1-3), as well as by grants from Ministry of Environment of the Czech Republic VaV/340/2/00 and VaV/740/5/03, AS CR AV0Z50390512, and the Commission of the European Communities QLK4-CT-2000-00091 EXPAH (publications 4-5).

ABSTRACT

Sporadic colorectal cancer (CRC) is a common disease with complex aetiology and diverse molecular phenotypes. Failure of DNA repair systems is one of the leading determinants of cancer onset and development. The efficiency of these systems and susceptibility to cancer can be affected by genotype variations, including common single nucleotide polymorphisms (SNPs).

In this work, an association between SNPs and haplotypes of DNA mismatch repair (MMR) genes, SNPs and their combinations in other DNA repair genes, and a risk of sporadic CRC was investigated in a hospital-based case-control study. As result of our study, certain MMR SNPs and haplotypes altered CRC risk, as demonstrated for the first time in the Czech population. Individual SNPs in DNA repair genes seem to have a limited effect on CRC risk, with possible modification by age or smoking. Several of the associations observed were site-specific, confirming the molecular heterogeneity of CRC.

DNA repair capacity varies significantly between individuals, between different tissues of the same organism, and also between malignant and normal cells. To assess the background level of this variability, the association between SNPs in DNA repair genes and the individual DNA repair capacity in healthy individuals was investigated. Several polymorphisms in base-excision repair genes and their binary combinations affected either irradiation-specific DNA repair or oxidative DNA repair rates; smoking and occupational status play an important role.

Air pollution negatively affects acute and chronic morbidity, including cancer. The analysis of the polymorphisms in metabolizing gene *EPHX1* was a part of the research on relationships between occupational exposure to carcinogenic polycyclic aromatic hydrocarbons, chromosomal aberrations (CA), DNA adducts, and DNA polymorphisms. The *EPHX1* diplotype affected the frequency of CA, suggesting a protective role in metabolism of environmental carcinogens.

ABSTRAKT

Sporadický kolorektální karcinom (colorectal cancer, CRC) je časté nádorové onemocnění s komplexní etiologií a variabilitou molekulárních fenotypů. Poruchy reparativních systémů DNA jsou jedním z hlavních faktorů ovlivňujících vznik a vývoj rakoviny. Účinnost těchto systémů a náchylnost k rakovině může být ovlivněna individuální genetickou variabilitou, včetně jednonukleotidových polymorfismů (SNP).

Tato doktorská práce zkoumá souvislost mezi SNP a haplotypy genů DNA mismatch reparace (MMR) a SNP v dalších genech DNA reparace, a rizikem sporadického CRC v rámci asociační case-control studie typu. Výsledky ukazují, že některé MMR SNP a haplotypy mohou měnit riziko CRC, jak bylo prokázáno poprvé u české populace. Jednotlivé SNP v genech DNA reparace mají omezený vliv na riziko CRC, s možnou interakci s věkem nebo faktory životního stylu. Některé z pozorovaných asociací byly specifické pro určitou lokalizaci nádorů, což potvrzuje molekulární heterogenitu CRC.

Kapacita DNA reparačních systémů se liší mezi jednotlivci, v různých tkáních stejného organismu a také mezi maligními a normálními buňkami. Pro posouzení bazální úrovně této variability se zkoumal vztah mezi SNP v genech DNA reparace a individuální DNA reparační kapacitou u zdravých jedinců. Několik polymorfismů v genech bázové excisní reparace a jejich binárních kombinací ovlivňovaly účinnost opravy poškození DNA vyvolaného γ-ozařováním nebo oxidačního poškození DNA; kouření a pracovní expozice hrají důležitou roli.

Znečištění ovzduší negativně ovlivňuje akutní a chronickou nemocnost. Analýza polymorfismů v metabolickém genu *EPHX1* byla součástí výzkumu o vztazích mezi expozicí karcinogenními polycyklickými aromatickými uhlovodíky, chromozomálními aberacemi, DNA adukty a DNA polymorfismy. *EPHX1* diplotyp má vliv na frekvenci chromosomálních aberací, což naznačuje jeho ochrannou roli v metabolismu environmentálních karcinogenů.

CONTENT

Appendix 1:

Tulupova E, Kumar R, Hanova M, Slyskova J, Pardini B, Polakova V, Naccarati A, Vodickova L, Novotny J, Halamkova J, Hemminki K, Vodicka P. Do polymorphisms and haplotypes of mismatch repair genes modulate risk of sporadic colorectal cancer? *Mutation Research.* 2008, 648(1-2):40-5.

Appendix 2:

Pardini B, Naccarati A, Novotny J, Smerhovsky Z, Vodickova L, Polakova V, Hanova M, Slyskova J, **Tulupova E**, Kumar R, Bortlik M, Barale R, Hemminki K, Vodicka P. DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic. *Mutation Research.* 2008, 638(1-2):146-53.

Appendix 3:

Vodicka P, Stetina R, Polakova V, **Tulupova E**, Naccarati A, Vodickova L, Kumar R, Hanova M, Pardini B, Slyskova J, Musak L, De Palma G, Soucek P, Hemminki K. Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects. *Carcinogenesis.* 2007, 28(3):657-64.

Appendix 4:

Sram RJ, Beskid O, Binkova B, Chvatalova I, Lnenickova Z, Milcova A, Solansky I, **Tulupova E**, Bavorova H, Ocadlikova D, Farmer PB. Chromosomal aberrations in environmentally exposed population in relation to metabolic and DNA repair genes polymorphisms. *Mutation Research.* 2007, 620(1-2):22-33.

Appendix 5:

Binkova B, Chvatalova I, Lnenickova Z, Milcova A, **Tulupova E**, Farmer PB, Sram RJ. PAH-DNA adducts in environmentally exposed population in relation to metabolic and DNA repair gene polymorphisms. *Mutation Research.* 2007, 620(1-2):49-61.

LIST OF ABBREVIATIONS

5-FU – 5-fluorouracil
ADP – adenosine diphosphate
AP site - apurinic/apyrimidinic site
APC - adenomatous polyposis coli
APE1 - apurinic/apyrimidinic endodeoxyribonuclease 1
ATP – adenosine triphosphate
B[a]P - benzo[a]pyrene
BAX - Bcl-2 associated X protein
BER - base excision repair
BMI - body mass index
BRAF - v-Raf murine sarcoma viral oncogene homolog B
CA – chromosomal aberrations
CCA – conventional cytogenetic analysis
ChIP-seq - chromatin immunoprecipitation sequencing
CIMP – CpG-island methylator phenotype
CIN – chromosomal instability
c-PAH – carcinogenic polycyclic aromatic hydrocarbons
CRC – colorectal cancer
DHSseq - DNase I hypersensitive sites sequencing
DSB – double strand break
EPHX1 – epoxide hydrolase 1
eQTLs - quantitative trait loci
EXO1 - exonuclease 1
FISH - fluorescent in situ hybridization
FOBT - faecal occult blood test
GST – glutathione-S-transferase
GWAS – genome-wide association study
HDL cholesterol - high-density lipoprotein cholesterol
HNPCC - hereditary nonpolyposis colon cancer
hOGG1 - 8-Oxoguanine DNA glycosylase
IDL - insertion/deletion loops
IGFIIR - insulin-like growth factor II receptor
KRAS - Kirsten rat sarcoma viral oncogene homolog

LDL cholesterol - low-density lipoprotein cholesterol

LOH – loss of heterozygosity

MED1 - Mediator complex subunit 1

MGMT - O-6-methylguanine-DNA methyltransferase

miRNA – micro-ribonucleic acid

MLH – human MutL homolog

MMR – mismatch repair

MN - micronuclei

Mre – metal response element

MSH - human MutS homolog

MSI – microsatellite instability

MSI-H – high frequency microsatellite instability

MSI-L – low frequency microsatellite instability

MSS – microsatellite stable

MutL – mutatotr L

MutS – mutator S

MUTYH - MutY homolog

Myc – also c-Myc, avian myelocytomatosis viral oncogene homolog

NER - nucleotide excision repair

NQO1 - NAD(P)H quinone dehydrogenase 1

PAH – polycyclic aromatic hydrocarbons

PARE - personalized analysis of rearranged ends

PBL - peripheral blood lymphocytes

PCNA - proliferating cell nuclear antigen

PCR – polymerase chain reaction

PM – particulate matterRFLP - restriction fragment length polymorphism

rSNP - regulatory single nucleotide polymorphism

SNP – single nucleotide polymorphism

TCR - transcription-coupled DNA repair

TGFβ – tumor growth factor beta

TGFβIIR - transforming growth factor beta receptor 2

Twist - transcriptionally induced by activation of STAT3

VAPS - versatile air pollution sampler

XPD - xeroderma pigmentosum group G-complementing protein

XRCC1 - X-ray repair cross-complementing protein 1

INTRODUCTION

The determination of phenotypic differences in cancer on the grounds of an individual's genotype is currently a paradigm and, simultaneously, a major challenge in the modern molecular biology of cancer. Until now, phenotypic variability of cancer initiation and development often precludes a clear identification of a given tumor type. The presence of millions of polymorphic gene variants in the human genome provides extensive variations affecting physiological and pathogenic mechanisms (15 million single nucleotide polymorphisms (SNPs), 1 million short insertions and deletions, and 20,000 structural variants) (The 1000 Genomes Project Consortium, 2010). It refers, in particular, to polymorphisms in genes participating in DNA damage response and repair. As accepted, the higher levels of DNA damage and insufficiency of protective mechanisms may predispose individuals to cancer (Win et al, 2013b; Shilpa and Lakshmi, 2014; Iyama and Wilson, 2013). Thus, any polymorphism that affects cellular response to DNA damage may modulate tumor occurrence, growth, histological features, metastatic spread, and response to therapeutic interventions (Huhn et al, 2014; Joost et al, 2014; Stigliano et al, 2014; Zhang et al, 2014a). Haploinsufficiency of those genes and intense physical and functional interaction between components of DNA repair systems allow distinct polymorphisms to affect phenotypic outcomes through multiple pathways (Fridley and Biernacka, 2011; Mooney et al, 2014; Huang, 2015).

Initially considered as very promising biomarkers, coding and non-coding SNPs in candidate genes then outlasted the sharp recession of scientific enthusiasm due to a vast amount of false positives and false negatives and the low statistical power of results. However, genome-wide association studies (GWAS) revived the interest of these variants in the human genome as potential cancer modulators (Manolio, 2010; Win et al, 2013a; Mooney et al, 2014; Zhang et al,

2014a; Huang, 2015).

Some systems of the human organism are more prone to external and internal DNA-damaging insults, and in some cases, consequently, to cancer. The gastrointestinal tract is one of the major natural areas affected by such insults. Colorectal cancer (CRC) represents a complex disease with strong evidence of the involvement of both macro- and micro-environmental and genetic / epigenetic factors into the etiology of the disease (Suchanek et al, 2011; Boleij and Tjalsma, 2012). Diet, smoking, and drinking habits are among the most relevant environmental factors affecting the CRC risk (Ezzati et al, 2012; Alexandrova et al, 2014; Barrow and Michels, 2014). For many years, the Czech Republic has had one of the highest incidences of CRC in the world (Ferlay et al, 2013).

It appeared that CRC, considered for a long time as a homogeneous entity, consists in fact of several molecular subtypes. These subtypes give rise to CRC with different locations, histology behaviour, and pharmacogenomic aspects etc. Consequently, environmental and genetic factors will be involved in a different extent (Esteban-Jurado et al, 2014; Valle, 2014).

In the current work, we would like to consider how SNPs, representatives of individual genetic background, can affect individual CRC risk and some corresponding phenotypes (for example, tumor localization). Another important aspect necessary for the understanding of basic pathophysiological mechanisms would be assessment of several gene-gene and gene-environment interactions in CRC patients, as well as in healthy individuals.

1. BACKGROUND

1.1. Human DNA mismatch repair system

1.1.1. DNA mismatch repair: main functions and mechanisms

The maintenance of the integrity of the information in DNA molecules is essential to the survival of a particular organism, as well as to the survival of the species (Lindahl and Wood, 1999). The mismatch repair (MMR), a highly conserved DNA repair system, greatly determines the maintenance of genomic stability and possesses the widest spectrum of function. The MMR system is involved in the correction of errors introduced by replicative DNA polymerases δ and ε during DNA replication, including DNA base-base mispairing and insertion/deletion loops (IDLs), and in editing heteroduplexes occurring during genetic recombination. IDLs arise through slippage of the polymerase during the replication of microsatellites, resulting in a change in the length of oligonucleotide repeats. Different DNA-derived aberrant base pairs, such as those containing O6-methylguanine, 8-oxoguanine, carcinogen adducts, and UV photoproducts damage are also substrates of MMR machinery. MMR improves the fidelity of DNA biosynthesis by 100-1000-fold, complementing the intrinsic error-free and proofreading properties of replicative DNA polymerases and lowering the overall mutation rate to one error per 10^{10} nucleotides synthesized (Fukui, 2010; Jiricny, 2013).

At least six different MMR proteins are required to correct the mismatched DNA (Fig. 1). MMR system response is considered to be comprised of two major components: protein complexes Mutator S (MutS) and Mutator L (MutL). In eukaryotes, the MMR system includes two functional equivalents of *E. coli* MutS. The main MMR pathway in humans is initiated by the recognition of a mismatch by the heterodimer consisting of the human MutS homolog 2

(hMSH2) and human MutS homolog 6 (hMSH6) or human MutS homolog 3 (hMSH3[1]) proteins (MutS, α and β respectively) depending on the type of lesion to be repaired. These complexes are responsible for the recognition of base mismatches and IDLs in mono- to tetranucleotide repeats (Hsieh and Yamane, 2008; Iyama and Wilson, 2014). A complex of hMSH6 and hMSH2 performs the correction of single-base mispairs, whereas both hMSH3 and hMSH2 may contribute to the correction of IDLs of two or more bases (Iyama and Wilson, 2014). The MutS dimer has a general shape of two "opposing commas" or "praying hands" joined by interactions at a single, composite ATP-binding site on one side, and at the bound mispared/unpaired DNA on the opposite side. The dimer encircles two channels, one of which is occupied by the mismatched DNA (Shilpa and Lakshmi, 2014).

After the substrate recognition by one of the MutS complexes, the eukaryotic MutL homologs are required to help other proteins to organize. The MutL equivalent in humans exists in three heterodimeric forms: MutLα (human MutL homolog 1 (hMLH1) and human Post Meiotic Segregation (hPMS2)), MutLβ (hMLH1 and human MutL homolog 3 hMLH3), and MutLɣ (hMLH1 and hPMS1) (Iyama and Wilson, 2014). The hPMS2 is required for the correction of single-base mismatches; hPMS2 and hMLH3 both contribute to the correction of IDLs, while the role of hPMS1 in MMR awaits further research (Shilpa and Lakshmi, 2014). Upon DNA mismatch detection, the repair process proceeds with the participation of the MutLα, the major MutL homolog participating in MMR. It acts as an endonuclease and coordinates the interplay between the mismatch recognition complex and other proteins necessary for MMR: exonuclease 1 (hEXO1), a helicase(s), proliferating cell nuclear antigen (PCNA), single-strand DNA-binding protein (RPA) and DNA polymerases δ and ε (Iyama and Wilson, 2014; Shilpa and Lakshmi, 2014).

[1] In this work, we use a following formate: genes names are written in Italic font and protein names in normal font.

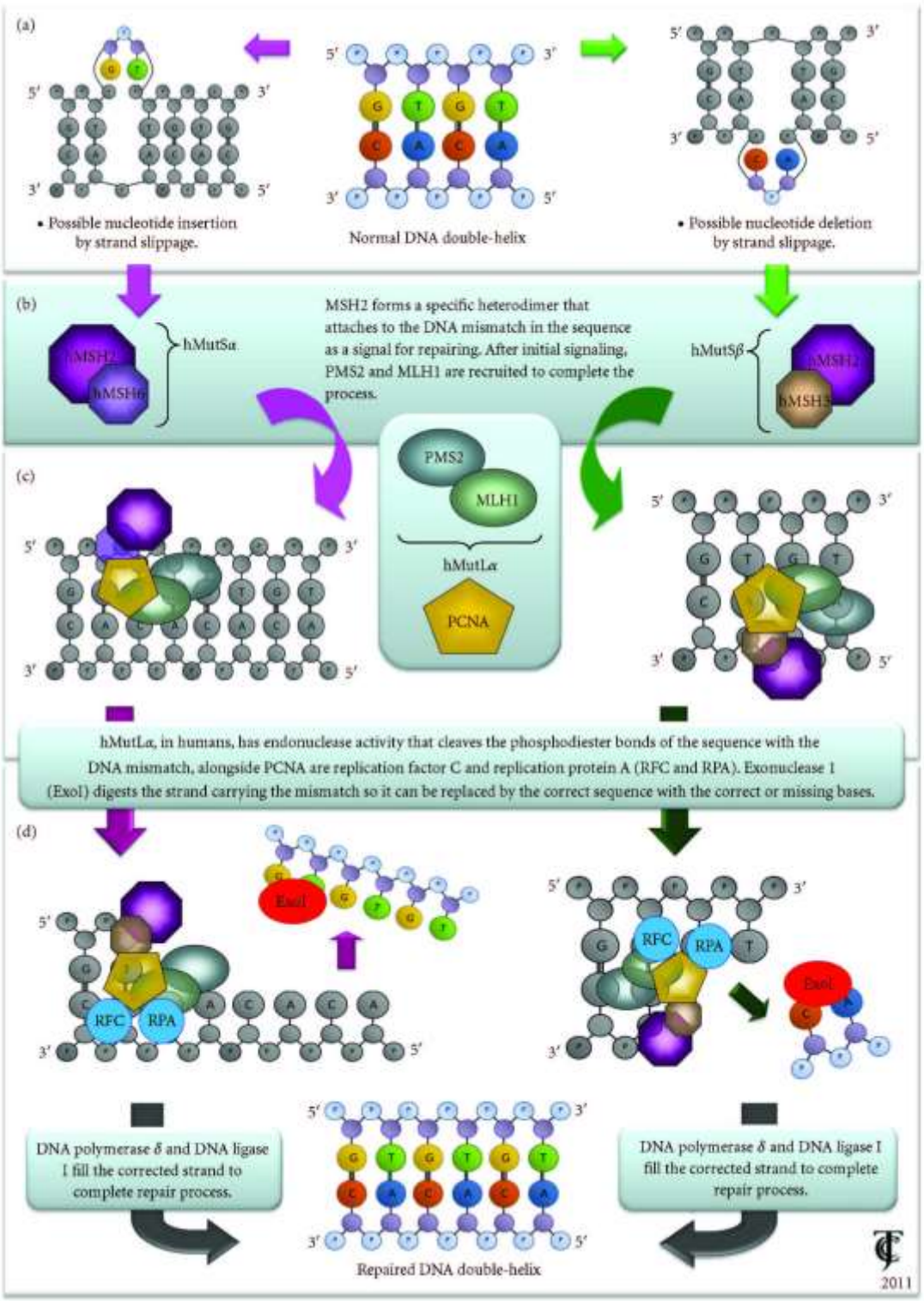

Fig. 1. General MMR mechanisms (Conde-Perezprina et al, 2012).

Upon DNA mismatch detection, the repair process proceeds with the participation of the MutLα, the major MutL homolog participating in MMR. It acts as an endonuclease and coordinates the interplay between the mismatch recognition complex and other proteins necessary for MMR: exonuclease 1 (EXO1), a helicase(s), proliferating cell nuclear antigen (PCNA), single-strand DNA-binding protein (RPA) and DNA polymerases δ and ε (Iyama and Wilson, 2014; Shilpa and Lakshmi, 2014). MutL has a conserved C-terminal region that mediates constitutive dimerization, and the N-terminal region undergoes ATP/ADP-regulated dimerization during MMR (Shilpa and Lakshmi, 2014). For leading strand synthesis, bound PCNA determines the orientation of MutLα incision while enhancing its endonuclease activity. PCNA coordinates with exonuclease EXO1, carrying out subsequent mismatch-provoked DNA excision directed by strand breaks located either 5' or 3' to the mispair (Hsieh and Yamane, 2008; Iyama and Wilson, 2014). Both MutS and MutL have an ATP-driven verification mechanism, by which the ATP-bound forms coordinate the downstream events in MMR, like EXO1 binding and nuclease activity (Jiricny, 2013).

1.1.2. Other functions of MMR system

The components of DNA MMR machinery are also involved in numerous cell processes:

- cell cycle arrest and induction of apoptosis in response to different types of DNA damage (O'Brien and Brown, 2006; Noonan et al, 2012, Lin et al, 2014);
- anti-recombination function: suppression of homologous recombination (Schofield and Hsieh, 2003; Tay et al, 2010);
- DNA-damage signaling (Iyer et al, 2006; Li, 2008; Edelbrock et al, 2013)
- meiotic recombination (as confirmed in studies on yeasts and mice) (Martini et al, 2011; Jiricny, 2013);

- mitotic recombination and maintenance of mitotic stability (Harfe and Jinks-Robertson, 2000; Schofield and Hsieh, 2003; Bak et al, 2014);
- somatic hypermutation (Iyer et al, 2006; Bak et al, 2014; Chen and Furano, 2015);
- cell aging (Conde-Perezprina et al, 2012).

1.1.3. Interaction of MMR system with other DNA repair mechanisms

Except for a direct role in MMR, mammalian MMR proteins can be involved in transcription-coupled repair (TCR) – as one of the nucleotide excision repair (NER) pathways, base excision repair (BER), and also recombination and meiosis. NER is responsible for the removal of UV light-induced pyrimidine dimers in adducts generated by benzo[a]pyrene and cisplatin. TCR is a NER pathway that preferentially corrects lesions in the transaction template strand of genes transcribed by RNA polymerase II. As detected in yeasts, MMR protein MSH2 can interact with NER proteins during the process of the recognition of a stalled RNA polymerase on the transcription template strand at sites of DNA damage or in the form of an MSH2-MSH6 heterodimer to bind to particular lesions repaired by NER (Shilpa and Lakshmi, 2014).

BER repairs oxidative lesions caused by reactive oxygen species, and also lesions, such as alkylation of purines and formation of adducts. The MutY homolog (MUTYH) protein is a DNA glycosylase involved in the excision of bases from the DNA, removing A from 8-oxo-G:A base pairs that generate apurinic/apyrimidinic (AP) sites in the DNA. It has been recently demonstrated that the MUTYH protein interacts with the MMR protein MSH6. Apparently, MutSα stimulates the DNA binding and glycosylase activities of MUTYH, showing a possible connection between the MMR system and the BER pathway (Niessen et al, 2006; David et al, 2007; Khan, 2015). Another link between MMR and BER is the MLH1 protein interacting with endonuclease Mediator complex subunit 1 **(**MED1) (Shilpa and Lakshmi, 2014).

The MMR system is also suggested to participate in protection against mammalian DNA damage induced by DNA alkylating agents, in this case interacting with O^6-methylguanine DNA methyltransferase (MGMT). The exact mechanism is not yet known; however, accumulation of O^6-methylguanine lesions could lead to interrupted transcription, recognition by the MMR system, and apoptosis (Iyama and Wilson, 2013).

1.1.4. Implication of MMR in carcinogenesis

MMR deficiency causes a reduction of DNA damage-induced apoptosis, an increase in mitotic recombination frequency, and an inefficient recombination, as well as increase in cell survival, which can consequently lead to an increase of the cancer susceptibility (Modrich and Lahue, 1996; Schofield and Hsieh, 2003; Shilpa and Lakshmi, 2014).

Genes involved in the DNA MMR system are considered as "high penetrance" genes of significant importance in hereditary nonpolyposis colon cancer (HNPCC) or Lynch syndrome and other forms of familial cancer in various organs (Martín-López and Fishel, 2013; Guillotin and Martin, 2014). Lynch syndrome patients have an increased risk of developing cancer, including an 80% risk of developing CRC. Inactivation of the MMR genes seems to be important in 15-20% (or even more) incidence of some sporadic cancers, especially concerning cancer in different organs of the "HNPCC spectrum": colon and rectum, uterine endometrium, stomach, and ovaries (Peltomäki, 2003).

There are two major mechanisms of MMR inactivation in the process of carcinogenesis:

1) epigenetic gene silencing by promoter hypermethylation,

2) germline and/or somatic mutation in MMR gene (Guillotin and Martin, 2014).

Methylation is a central epigenetic mechanism involved in the regulation of gene expression. Epigenetic inactivation, mainly due to methylation of the promoter of DNA repair

genes, has been reported in several DNA repair pathways related to cancer (Shilpa and Lakshmi, 2014). A range of studies showed that in the majority of microsatellite unstable sporadic CRC, and in sporadic endometrial carcinomas, MMR appears to be inactivated via hypermethylation of the *hMLH1* promoter (Cunningham et al, 1998; Esteller et al, 1998; Kuismanen et al, 2000; Kamory, 2003; Valo et al, 2015).

Other MMR genes are also controlled by promoter methylation. Aberrations of *hMSH6* and *hMSH3* methylation were found in sporadic CRC (Shilpa and Lakshmi, 2014). Very high levels of *hMSH3* promoter methylation were observed in oesophageal cancer tissues (Vogelsang et al, 2014). Germline 3'-end deletions affecting the epithelial cell adhesion molecule (EPCAM) gene located upstream of *hMSH2* were identified as a novel mechanism causing Lynch syndrome by epigenetic inactivation of the respective *hMSH2* allele (Ligtenberg et al, 2009). An explanation for the epigenetic transgenerational inheritance is that the loss of polyadenylation signals via deletions could result in an aberrant promoter methylation of neighbouring tumor suppressor genes (Yamamoto and Imai, 2015).

A range of studies showed that germline mutations in the MMR genes, in particular *hMLH1, hMSH2, hMSH6* or *hPMS2*, lead to predisposition to the autosomal dominant condition in Lynch syndrome (Martín-López and Fishel, 2013). In this case, only one mutated allele of an MMR gene is inherited. Loss of the second allele occurs somatically, either by loss of heterozygosity (LOH), mutation, or methylation. The rare case where both inherited alleles are mutated is called the constitutional MMR deficiency syndrome, and this leads to cancer during childhood (Guillotin and Martin, 2014; Yamamoto and Imai, 2015). One-allelic inherited germline mutation in *hMSH1* and *hMLH1* cause Muir-Torre syndrome, a rare autosomal dominant disorder characterized by the predisposition to numerous malignancies (Conde-Perezprina et al, 2012).

Less is known about the effect of mutations in *hMSH3*. It is suggested, however, that they can determine the onset of the tumor in association with the weak mutations in *hMLH1* and *hMSH2* due to a synergistic effect (Duraturo, 2011).

MMR deficiency leads to a mutator phenotype, which is characterized by accumulation of mutations in the DNA, primarily due to microsatellite instability (MSI) – a hallmark of abnormal MMR functioning. Microsatellites are short tandem (1-6 base pairs) repeats present through the whole genome with high susceptibility for replication errors. DNA polymerases slip over the tandem repeats, leading to replication length errors. These errors are normally repaired by the MMR pathway, therefore they remain fixed in deficient MMR cells. A deficient DNA MMR results in the widespread mutations to nucleotide repeats, some of which occur within the coding regions of cancer-related genes such as transforming growth factor beta receptor (*TGFβR2*), insulin-like growth factor II receptor (*IGFIIR*), Bcl-2 associated X protein (*BAX*) and DNA double-strand break (DSB) repair genes, including metal response element II (*Mre11*) and RAD50 gene homolog (*Rad50*) (Iyama and Wilson, 2014; Guillotin and Martin, 2014).

The MSI status of CRC is considered an extremely useful marker for population-based screening programs that aim to identify individuals and families with the Lynch syndrome (Iacopetta et al, 2010). MSI is commonly tested at five loci (BAT 25, BAT 26, D2S123, D5S346 and D17S250). Tumors are classified as having high frequency microsatellite instability (MSI-H) in case if >30% of the markers show instability of markers, low-frequency microsatellite instability (MSI-L) - <30% of the instability of markers, or as microsatellite stable (MSS) if there is no apparent instability (Boland et al, 1998).

Obviously, the above mentioned mechanisms are not the only ones causing MMR deficiency in cancer: for instance, an important MMR protein *hMSH3* can be physically removed from the nucleus as a result of inflammation and oxidative stress, probably mediated by cytokines

like interleukin 6, thus falling out of the MMR process (Tseng-Rogenski et al, 2015).

Functional gene variants can also be involved in predisposition to the development of MMR deficiency. For example, the recently detected unique single nucleotide variant at the codon 204, within the exon 2 of *hMLH1*, was predicted to decrease the activity of the MLH1 protein (Vodicka et al, 2015a). MicroRNAs (miRNAs) are also playing an important role in de- and reactivation of MMR. The molecule miR-155, which is frequently upregulated in MSI CRCs, is responsible for the silencing of *hMLH1, hMSH2* and *hMSH6*. Inhibition of miR-155 could thus be a method to reactivate MMR and to reverse the MSI phenotype (Valeri et al, 2010; Guillotin and Martin, 2014; Vymetalkova et al, 2014a). Downregulation of miR-422a and miR-16 expression in CRC tissues or serum as a possible consequence of reduced MutLα expression was demonstrated. In humans, miR-21 is also significantly overexpressed in CRC tissues with MSI compared to non-CRC mucosa, and this overexpression is associated with reduction of MSH2 and MSH6 protein levels (Li and Martin, 2016).

It is necessary to mention that, except for the role in cancer occurrence and different stages of development, MMR is also of importance in the organism response to cancer treatment. The MMR system takes part in the response to many drugs used in anti-cancer treatment. In particular, adducts of platinum compounds, alkylating and fluoropyrimidine agents are processed by the MMR system (Iyer et al, 2006). MMR deficient tumors can evolve resistance to these chemotherapeutic substances. The majority of studies suggest that, for example, MMR deficient colorectal tumors are more resistant to 5-FU than MMR proficient ones. The role of the MMR pathway in ionizing radiation response is not completely understood, and this issue is particularly important for the treatment of MMR deficient endometrial and rectal cancers. It was observed that MMR deficient cells become resistant after a high-dose of ionizing radiation compared to MMR proficient cells. These differences could potentially be explained by the role of the MMR

pathway in the regulation of cell cycle and homologous recombination pathway (Guillotin and Martin, 2014).

Cells with suppressed MLH1 protein function are more sensitive to ganciclovir, thus targeting MMR-deficient tumors may, for example, increase efficacy of a suicide gene therapy approach with the herpes simplex virus thymidine kinase and ganciclovir to cancer treatment (O'Konek et al, 2009). These findings confirm that MMR status is an important prognostic factor for colorectal tumors, and could help guide selection of an appropriate anti-cancer therapy.

1.2. Sporadic colorectal cancer: molecular phenotypes

Colorectal cancer (CRC) is one of the most frequently diagnosed cancers worldwide (Fig. 2, Fig. 3), with over one million diagnosed cases (9.8% of cancer diagnoses, cumulative lifetime risk of 2%) and ~600 000 deaths (8.1% of all cancer deaths). The CRC incidence rose in parallel with economic development, and the majority of cases occur in industrialized countries (Teixeira et al, 2014). For a long time, Czech Republic has been keeping one of the highest positions in CRC incidence and mortality (third place in the world in total according to Globocan 2012 study). At that rate, even if mortality has been relatively stabilized, the CRC incidence and prevalence are steadily growing in the Czech Republic during recent years (Dušek et al, 2014).

Initiation and progression of CRC is determined by complex interactions between a range of macro- and microenvironmental and genetic / epigenetic factors. Smoking, overweight, lack of physical activity, higher alcohol consumption, and diet poor in fiber and vitamins (elements of the so-called Western lifestyle), especially in combination, play an important role in different stages of CRC occurrence and development (Ezzati et al, 2005; Huxley et al, 2009). Microenvironmental agents have recently attracted more attention as important etiological factors in CRC, since the human intestine permits growth of over 500 different species of bacteria, as

well as the range of pathogenic and non-pathogenic viruses (JC virus, human papillomavirus), with the highest concentration of bacteria found in the colon.

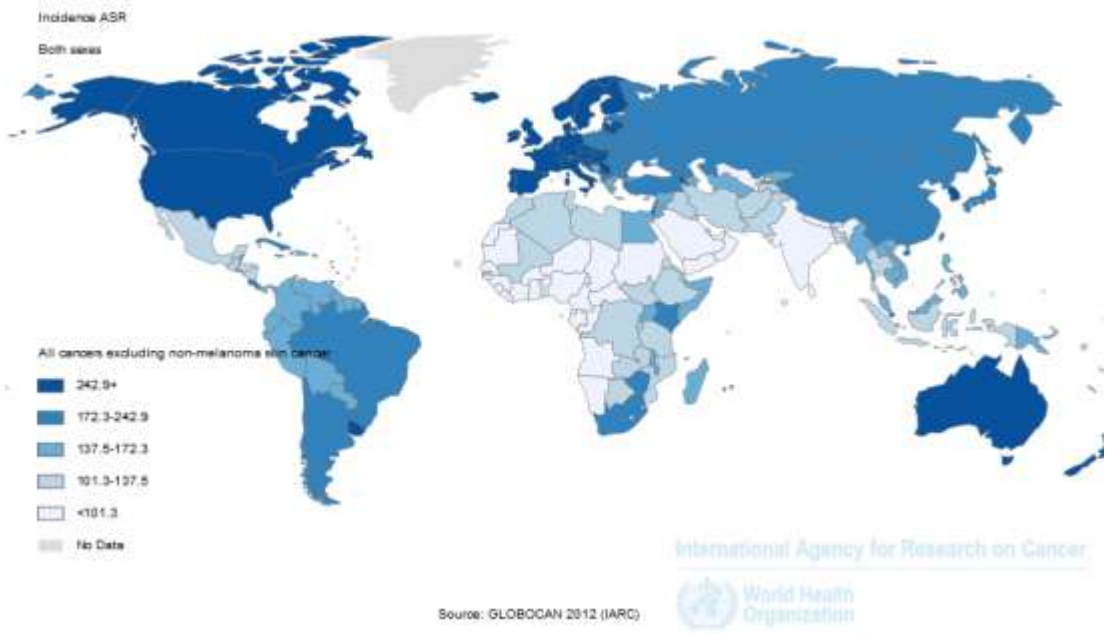

Fig. 2. Distribution of CRC incidence in the world (http://globocan.iarc.fr).

Fig. 3. Distribution of CRC incidence in Europe (http://globocan.iarc.fr).

Among them, *Helicobacter pylori, Bacteroides fragilis* and *Streptococcus bovis* may have association with CRC (Burnett-Hartman et al, 2008; Wu et al, 2009). Mechanical damage and inflammation of gut tissues are of importance in transformation of normal epithelium into malignancy as well (Desprat et al, 2008). All these factors are interrelated: inflammation in the colon can be triggered by the gut microbiota and their metabolites or byproducts. In general, modern research confirms a link between microRNAs, inflammation, the gut microbiota and tumorigenesis (Li and Martin, 2016). Mechanical pressure can alter gene expression, and in particular activate the avian myelocytomatosis viral oncogene homolog (*Myc*) and transcriptionally induced by activation of *STAT3* (*Twist*) oncogenes, which are implicated in the early stages of colon cancer (Desprat et al, 2008).

Depending on the stage and progression state of the disease, treatment regimens for CRCs include (Gustavsson et al, 2015):

1. Colectomy, surgical resection (stages 0, I and early stage II), up to date the most effective treatment.

2. Postoperative adjuvant chemotherapy (stage III and some stage II).

3. Chemotherapy, cytotoxic treatment with multi-drug therapy including 5-FU and leucovorin and CapeOx (capecitabine and oxaliplatin) (stage III).

4. Radiation therapy (recurrent or advanced disease).

Over past decades, significant progress has been achieved in the cytotoxic treatment of CRC by the use of fluoropyrimidines, irinotecan and oxaliplatin (Gustavsson et al, 2015). Another chemotherapeutic agent, imnotecan (CPT-11) has been shown to improve efficacy in CRCs (Mundade et al, 2014). Currently the evaluation of new methods of treatment using concurrent technologies, such as genetically engineered monoclonal antibodies and recombinant vaccines, is ongoing (Mundade et al, 2014).

Approximately 75% of colorectal cancer cases are sporadic, and the remaining are familial diseases. Hereditary syndromes are mostly transmitted as autosomal dominant traits, and are usually caused by highly penetrant mutations in single genes (Abdel-Rahman et al, 2006; Stigliano et al, 2014).

Currently, it is suggested that the following genetic pathways can lead to consequent transformation of normal gut mucosa to cancer through adenoma step (Fig. 4):

1) Chromosomal instability (CIN)

2) Microsatellite instability (MSI)

3) Serrated pathway

1) CIN pathway is observed in 65–70% of sporadic colorectal cancers. The complete mechanism of CIN still remains unclear, but it is characterized by defects in chromosomal segregation, telomere stability and the DNA damage response (Pino and Chung, 2011). CIN suggests that CRC appears as a result of stepwise mutational activation of oncogenes (e.g., Kirsten rat sarcoma viral oncogene homolog (*KRAS*), *c-Src,* avian myelocytomatosis viral oncogene homolog (*c-Myc*)) and the inactivation of tumor suppressor genes (Adenomatous polyposis coli (*APC*) gene, *TP53*) as well as heterozygosity for the long arm of the chromosome 18 (18q LOH**)** (Mundade et al, 2014)**.**

2) MSI pathway (10-15% of sporadic CRC). Overall, MMR deficiency is associated with 10-15% of CRC tumors either due to a germline mutation in one of the mismatch repair genes (*hMLH1, hMSH2, hMSH6* and *hPMS2*), especially in younger patients, or due to epigenetic silencing of *hMLH1* – in elder patients (Pino and Chung, 2011). According to this parameter, CRCs can be clinically classified as MSI-H, MSI-L or MSS as mentioned above**.**

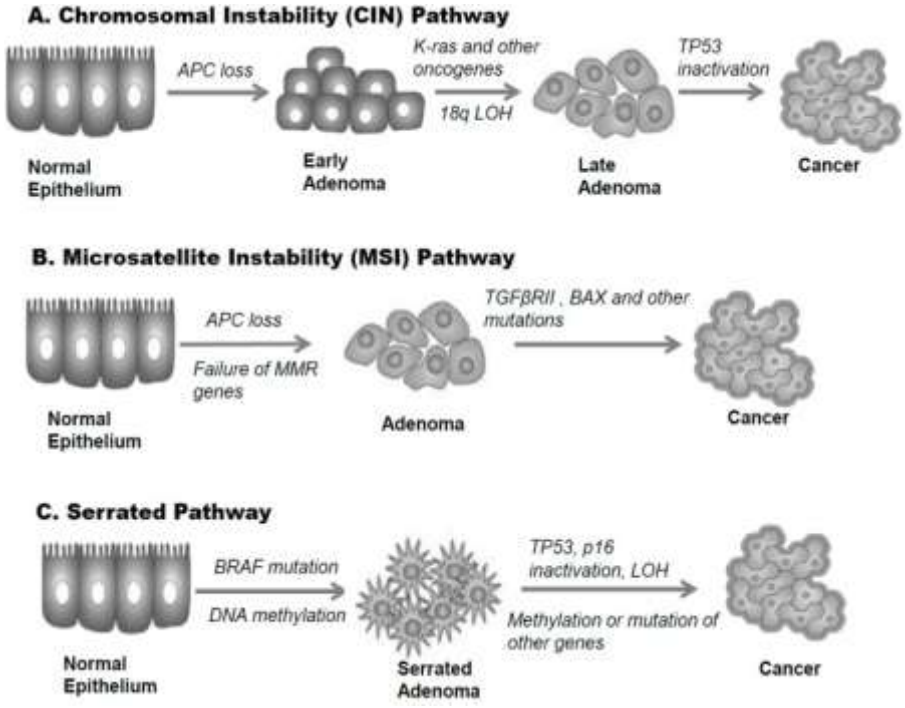

Fig. 4. Genetic pathways in CRC pathogenesis (Mundade et al, 2014).

Loss of *hMLH1* is identified in 60% of sporadic MMR deficient tumors, mainly due to hypermethylation of the promoter (Guillotin and Martin, 2014). Somatic *hMSH2* hypermethylation was present in 24% of Lynch syndrome CRCs with *hMSH2* deficiency (Nagasaka et al, 2010; Yamamoto and Imai, 2015). Finally, immunohistochemical studies demonstrated a rare phenomenon of heterogenous MMR status in CRC due to heterogenous expression of MMR molecules within a tumor (Joost et al, 2014). Loss of the *APC* product also happens on early in the MSI-dependent CRC progression (Mundade et al, 2014).

3) Serrated pathway (around 30% of CRC; named due to the morphologically serrated appearance of the precursor lesions)**,** characterized by the presence of *BRAF* (protein kinase B-raf, v-Raf murine sarcoma viral oncogene homolog B) mutation and epigenetic silencing of genes

involved in cell differentiation, DNA repair and cell-cycle control. The *APC* loss is not typical for this phenotype. A hallmark of serrated way is inactivation of tumor suppressor gene *p16* (cyclin-dependent kinase inhibitor 2A, multiple tumor suppressor 1) through promoter hypermethylation – the so-called CpG island methylator phenotype (CIMP) **(**O'Brien et al, 2015**)**.

A global DNA or LINE-1 hypomethylation (global loss of methylated cytosines) should also be mentioned as another genome-wide epigenetic mechanism leading to increase of mutability, aneuploidy and CIN, and, as result, to colon cancer (Figueiredo et al, 2009). The LINE-1 hypomethylation inversely correlates with MSI and CIMP status in CRC suggesting that this phenotype represents separate pathway in CRC (Ogino et al, 2008).

Some phenotypical difference as well as differences in chemotherapy response and prognosis can be observed between tumors developing by different pathways. The CIN phenotype was associated with a less favourable outcome for patients, compared with tumors with MSI: patients with CIN tumors had a decreased overall and progression-free survival, irrespective of ethnic background, tumor localization and adjuvant treatment with 5-FU (Walther et al, 2008).

A deficient MMR system has been detected in about 20% of right-sided colon cancers and 5% of left-sided colon and rectal cancers. Features of MSI-H CRC include a tendency to arise in the proximal colon, poor differentiation, lymphocytic infiltration, fast adenoma-adenocarcinoma progression and mucinous or signet-ring histology. Patients with MSI-H tumors appear to have a better prognosis than those with MSS tumors and can possibly have better response to irinotecan chemotherapy and poorer responses to 5-FU (Pino and Chung, 2011).

CIMP is independently associated with significantly worse prognosis of the disease-free survival and overall survival in CRC patients. The data on CIMP's value as a predictive factor for survival benefit due to adjuvant 5-FU therapy are still unclear (Juo et al, 2014).

In recent studies on the association of phenotypical differences and molecular characteristics of CRC, the following classification was also based on the CIMP and MSI status (Iacopetta, 2010):

I. CIMP – high, MSI-H, *BRAF* mutation

II. CIMP – high, MSI-L or MSS, *BRAF* mutation

III. CIMP – low, MSS or MSI-L, *KRAS* mutation

IV. CIMP – negative, MSS

V. CIMP – negative, MSI-H (Lynch syndrome).

Compared to patients with type IV tumors (the most predominant), the patients with type II tumors had the highest overall and disease-specific mortality; subjects with type III tumors also have higher disease-specific mortality. The lowest disease-specific mortality was observed in patients with type V tumors (Phipps et al, 2015).

Discovery of heterogeneous molecular pathways in CRC pathogenesis lead to the advent of "molecular pathologic epidemiology" as a multidisciplinary research direction investigating the interplay between exogenous and endogenous factors, molecular phenotypes, and cancer initiation and progression (Ogino et al, 2011). New molecular classifiers and biomarkers are being discovered continuously - for instance, numerous miRNA which can act to promote or inhibit tumorigenesis, thus affecting cell growth, proliferation, invasion and metastasis in CRC as well as outcomes of CRC chemotherapy and prognosis (Schetter et al, 2012; Aherne et al, 2015; Pardini et al, 2015).

1.3. Single nucleotide polymorphisms (SNPs) as modulators of individual cancer susceptibility

1.3.1. SNP analysis and its application for association studies

Given the substantial inter-individual differences in CRC initiation and consequent neoplastic events, researchers have been looking for the more abundant sources of genetic variations in the human genome for a long time (Mitchell et al, 2005). Common SNPs - single base pair polymorphisms with an abundance of the least frequent allele of 1% or greater in the given population - fall into this group.

Approximately 99.5% of the genome DNA sequence is identical among humans. The 0.5% sequence (15 million bps) accounts for all individual differences, including predisposition to diseases. SNPs are the most simple form and most common source of genetic polymorphism in the human genome. It was estimated that SNPs account for over 80% of all human individual phenotypic differences, as well as population diversity and disease predisposition (Roberts et al, 2010).

SNPs altering amino acids of protein-coding genes can drastically affect protein function and play an important role in molecular pathogenesis. By contrast, regulatory SNPs (rSNPs) were supposed to show modest effects that might modify gene function more subtly. Functions of rSNPs can involve gene expression regulation through the effect on RNA spicing, transcription factor binding, DNA methylation and miRNA recruitment. Different from simple Mendelian diseases that result from an individual variant, primarily in coding regions, the vast majority of SNPs that have been identified for common complex diseases map to non-coding intergenic and intronic regulatory regions, are enriched for expression of quantitative trait loci (eQTLs), DNase I hypersensitive sites sequencing (DHSseq) peaks, and chromatin immunoprecipitation sequencing (ChIP-seq) peaks (Huang, 2015). A number of computational tools, as well as *in vitro* assays, are

developed nowadays to assess possible functional effects of different types of SNPs (Table 1).

To explore mechanisms underlying common diseases, two main strategies have been applied: family-based linkage studies across the entire genome, and population-based association studies of individual candidate genes. Similar strategies were applied in discovering the role of SNPs in pathogenesis. Using this approach, substantial progress has been achieved; however, due to limitations of the methods, the progress was slower to some extent. Linkage analysis has low power, except when a single locus explains a substantial fraction of disease, and association studies of one or a few candidate genes allow evaluating only subtle part out of immense individual genotype variation (The International HapMap Consortium, 2005).

Development of SNP-detecting technologies and decreasing costs of analysis enabled wide SNP studies. SNP analysis largely contributed to various GWAS, in which large numbers of SNPs are tested for association with a disease in hundreds or thousands of persons. GWAS allowed us to identify thousands of genes and genetic variants involved in the complex diseases in humans: 11 912 SNPs associations with $P < 10^{-5}$ have been reported by 2013 (Welter et al, 2014).

Functional category	Computational tool	Experimental approach
RNA splicing	ESE finder, RESCUE-ESE	RNA-seq
Transcription factor binding	TRANSFAC, JASPER, MAPPER2	EMSAs ChIP, eQTL, reporter assays
DNA methylation	MethDB, EpiGraph, ENCODE	MeDIP-seq, MRE-seq, methylation array, bisulphate sequencing
miRNAs-mRNA interaction	miRNASNP, mrSNP, SNPinfo, MirSNP, miRdSNP	Reporter assays, Western blot, RNAi
LncRNA expression or structure	LincSNP, LncRNASNP 3dRNA	RNA-seq, RIP, microarray

Table 1. Functional classification of SNPs (Huang, 2015).

Genes with polymorphisms that have been previously associated with a particular disease can be grouped to form a disease biomarker (Mooney et al, 2014).

It stands to reason that the greatest challenge is to understand the functional consequences of detected SNPs and to accurately elucidate the biological mechanisms, by which they act (Huang, 2015). Most cancer-associated SNPs identified in GWAS are located outside protein-coding regions, their isolated effects are weak and can explain only a small proportion of the familial clustering of a particular disease trait. However, such SNPs can mark a region that is very important in cancer, for instance, containing a really high penetrance mutation (Manolio, 2010). Secondary analysis of GWAS data can be done by using a gene set analysis approach as well as whole-genome sequencing and association studies to clarify "missing" genetic heritability (Fridley and Biernacka, 2011). Modern sequencing technique allow detailed following of all stages of cancer. Some of the most important findings including fusions of essential genes were done using whole genome sequencing. Massively parallel sequencing using the PARE (personalized analysis of rearranged ends) is another powerful approach which facilitate the development of personalized biomarkers for clinical practice (Mundade et al, 2014).

Concerning the results of the modern association studies, which remain a necessary tool in cancer research, the accumulation of small effects of many genetic variants into a single analysis is expected to be more powerful than tests of individual SNP effects (Fridley and Biernacka, 2011). Multiple SNPs within the same gene can be used to calculate a gene-level effect by statistical combining of individual SNP *P*-values or multi-SNP modelling (Mooney et al, 2014).

Haplotype studies represent another powerful approach in the assessment of disease-related risk and mechanisms**.** Each allele is initially associated with the other alleles present on the particular chromosomal background on which it arose. The specific set of alleles observed on a single chromosome, or part of a chromosome, is called a haplotype (The International HapMap

Consortium, 2005). It is usually inherited as a single unit from parents, ultimately covering the variability within a gene. Haplotype inference is an important part of human disease studies. Sequencing-based discovery and direct or inferential phasing of alleles in a subset of tested patients enables us to fill in missing genotypes of the remaining patients on the basis of haplotype-block sharing, thereby increasing the power of association studies. If SNPs are close together in the genome, alleles will tend to be inherited as haplotypes more often than SNPs, which are further apart. Haplotype analysis can be especially useful for studies of populations, in which an individual founder mutation accounts for a substantial percentage of all cases, or for any population with a susceptibility allele in a region of very low recombination (Naccarati et al, 2007). Haplotype information substantially improves the power of GWAS with the support of reference panels of phased haplotypes. Haplotype data are commonly applied to studies of biological mechanisms. In application to cancer research, haplotypes can help in linkage evaluation while using the phasing of distant somatic variants from the same haplotype during tumor development (Snyder et al, 2015).

1.3.2. SNPs affecting susceptibility to cancer: an example of SNPs in MMR genes

Amino acid changes in genes coding components of MutS or MutL may affect mismatch recognition, matchmaking, ATPase cycling, and EXO1 binding and nuclease activity (Jiricny, 2013). The location in a regulatory element can allow even weak SNPs to play a decisive role in tumorigenesis; in some instances, an effect of such SNPs can be elevated due to epistatic interaction with other gene variants (Sur et al, 2013).

Here we mention some mechanisms, which show how SNPs located within MMR genes can affect carcinogenesis processes:

1) Regulation of the MMR gene transcription.

Epigenetic silencing of *hMLH1* has been found to underlie most MMR defective sporadic cancer cases (Gazzoli and Kolodner, 2003; Valle, 2014), some of these processes can be affected by polymorphisms in regulatory regions. SNP in the *hMLH1* core promoter region -93G>A was shown to have an association with sporadic CRC (Nizam et al, 2013), and also with an increased risk of lung cancer (Park et al, 2004), breast (Lee et al, 2005), and endometrial cancer (Beiner et al, 2006). Since this polymorphism is located 93 nucleotides upstream of the transcription start site in the putative consensus sequence for the binding of transcription factor AP-4, it possibly influences the activity of the *hMLH1* promoter (Shin et al, 2002). The -93G>A polymorphism allele was associated with a protective effect and reported as a determinant of CRC risk (Muniz-Mendosa et al, 2012; Nassiri et al, 2013).

A C to G nucleotide substitution at position -107 adjacent to the *hMLH1* gene translation initiation site reduced transcriptional activity and impeded the promoter-binding capacity of nuclear proteins. The effect of -107G on *hMLH1* gene transcription and nuclear protein binding to the promoter sequence implicates the site, including -107C, as a crucial element interacting with the activator that maintains *hMLH1* gene expression (Zhong et al, 2007).

The *hMSH2* -118T>C polymorphism is located in the core promoter region, 118 nucleotides upstream of the transcription start site in a potential transcription binding site. Strong association was observed between *hMSH2* -118T>C and family history of CRC, especially among female patients. It is possible that this SNP plays a role in the *MSH2* response to sex hormones (Mrkonjic et al, 2007; Nassiri et al, 2013).

The SNP does not have to be physically located in the regulatory region to affect its function. For example, the *hMSH3* polymorphism Ex23+3G>A is in a haplotype block with a distal enhancer of this gene which was found to have aberrant methylation in the tumor tissues of

patients with oesophageal cancer (Vogelsang et al, 2014). Some researchers suggest that methylation of distal DNA regulatory regions (e.g., enhancers) has more reliable correlation with expression levels of the corresponding genes than promoter methylation (Aran and Hellman 2013; Aran et al. 2013).

2) Synthesis of functionally different splicing variant.

It has been estimated that nearly 90% of protein coding genes are subject to alternative splicing in humans. In general, splicing of mRNA can be regulated by rSNPs occurring within branch sites, 5' and 3' splice sites, and intronic and exonic splicing enhancers and silencers. Disruption of normal splicing has been implicated in disease pathophysiology (Huang, 2015). Synthesis of functionally different splicing variants can be a reason for insufficient activity of MMR genes carrying intronic SNPs, for example, *hMSH2* IVS12-6T>C, which has been found to be associated with an increased risk of non-Hodgkin's lymphoma and sporadic CRC (Goessl et al, 1997; Paz-y-Mino et al, 2002). Intronic SNPs include IVS14-19A>G located 19 nucleotides upstream from the exon 15 splice acceptor site can potentially affect MMR function (Nassiri et al, 2013).

The unclassified gene variants discovered in *hMLH1* and *hMSH2* genes can cause defects in pre-mRNA splicing, either by altering degenerate positions of splice site sequences or by affecting intronic or exonic splicing regulatory sequences (Tournier et al, 2008). It is currently suggested that noncoding variants affecting RNA splicing of corresponding genes through eQTLs on a post-transcriptional level may be about the same importance in disease traits as variants modulating the gene expression (Li et al, 2016).

3) Modulation of interaction between elements of MMR machinery

SNPs located in the regions encoding interaction domains of MMR proteins can have a significant effect on the functioning of the MMR system (Fig. 6). SNPs with a similar potential

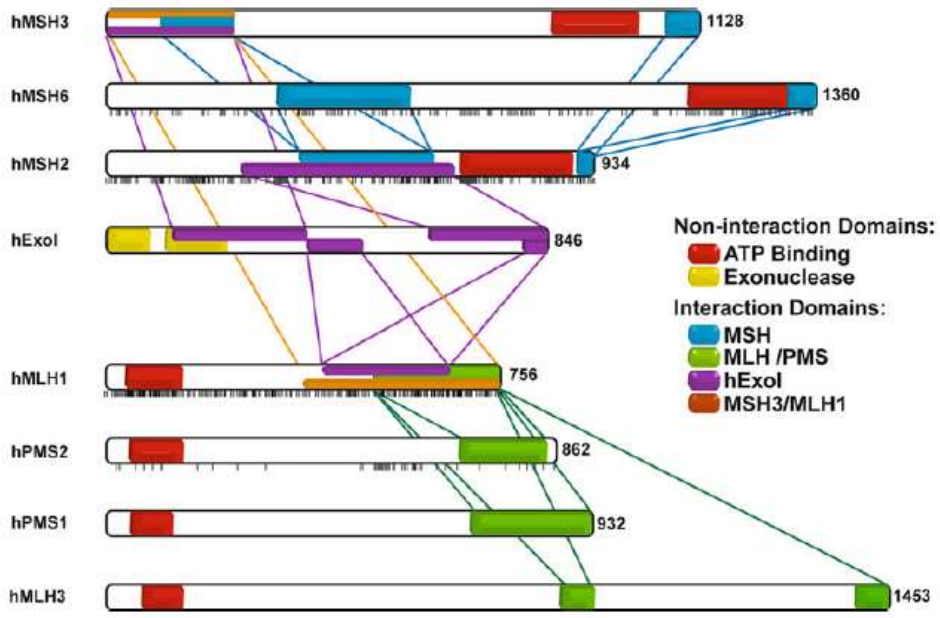

Fig. 6. Regions of interaction between MMR proteins (Martín-López and Fishel, 2013)

and association with the risk of different types of cancer have been found in several MMR genes. Two polymorphisms in the *hMSH6* exon 4, codon 389 and codon 396, are assumed to affect MMR efficiency, since these SNPs fall in the contact domain of *hMSH6* and *hMSH2* genes (Parc et al, 2000). A functional effect of the *hMSH3* Ex23 + 3G>A polymorphism is determined by Ala1045Thr amino acid change in ATP-binding domain. The polymorphism Ex12 +49C>T in *hEXO1* gene coding Thr439Met substitution may influence the interaction between hEXO1 and hMLH1 proteins (Schmutte et al, 2001). The coding polymorphism Ex6+23G>A in the *hMSH2* gene may have a putative influence on MMR function due to its location in a region stabilizing this gene's interaction with the hEXO1 protein (Schmutte et al, 2001). This was considered as affecting breast cancer susceptibility (Poplawski et al, 2005), with no significant association with CRC risk (Schafmayer et al, 2007). Once again, polymorphisms affecting interaction of MMR proteins are not always located in the corresponding gene domains. For instance, it is the case of

three missense alterations in *hPMS2* gene, which cause defective protein-protein interactions with *hMLH1* (Yuan et al, 2002).

4) Tissue-specific mechanisms

MMR genes are widely expressed in different types of cells. There is extensive natural heritable variation in levels of gene expression. Genetic linkage and association mapping have identified *cis*- and *trans*-acting DNA variants that influence expression levels of human genes in an organ-specific or tissue-specific manner (Sanli and Feil, 2015). DNA repair can be induced or inhibited by external or internal factors in various manners in exposed tissues. For example, -93G-A polymorphism in the *hMLH1* promoter was found to be associated with lung squamous cell carcinoma, but not with adenocarcinoma or small cell carcinoma (Park et al, 2004). The same polymorphism was associated with colon but not rectal localization in CRC (Campbell et al, 2009). Further investigations are necessary to recognize additional mechanisms responsible for the effect of SNPs of MMR genes on pro-carcinogenic biological processes in different tissues.

In this subchapter, we mentioned only a few putative mechanisms of the regulation of MMR activity by SNPs located in MMR genes. Obviously, there is an abundance of targets and points in a range of other cancer-related mechanisms (acting through numerous metabolic and regulatory pathways, gene-gene and gene-environment interactions etc.), where common polymorphisms can contribute to tumorigenesis or to affect cancer progression, metastasis, treatment outcome and prognosis. These mechanisms still need to be explored.

2. AIMS OF THE STUDY

The work was focused on realization of the following tasks:

1. to investigate a tentative association of polymorphisms in DNA MMR genes with the risk of sporadic CRC by means of a hospital-based case-control study on an ethnically homogeneous population from the Czech Republic;

2. to evaluate a possible influence of haplotypes of MMR genes constructed using tagging SNPs in MMR genes on sporadic CRC risk;

3. to assess an association of SNPs and their binary combinations in genes involved in DNA BER, NER and DSB repair and sporadic CRC risk;

4. to determine the differential distribution of colorectal adenocarcinomas by tumor localization in association with individual's genotype;

5. to assess the association between selected SNPs in DNA BER and DSB repair genes and the capacity to repair DNA damage induced by γ-irradiation and by base oxidation in a healthy population;

6. to assess an association of a diplotype in epoxide hydrolase 1 (*EPHX1*) gene and a level of chromosomal aberrations and polycyclic aromatic hydrocarbons (PAH)-DNA adducts in occupationally exposed policemen in Prague.

3. METHODS USED

SNP selection and genotyping

Selection of SNPs for genotyping and haplotype reconstruction (Hapmap database www.hapmap.org; SNP500 Cancer database http://variantgps.nci.nih.gov/cgfseq/pages/snp500.do);

Peripheral blood lymphocytes (PBL) DNA isolation using the phenol/chloroform extraction method (Vodicka et al, 2001);

SNP genotyping by real-time PCR (Taqman assays for allelic discrimination (Assay-on-demand, Applied Biosystems, Foster City, USA: numbers of assays are listed in Table 2);

SNP genotyping using PCR-RFLP (primer sequences and restriction enzymes are listed in Table 2);

Primer design for genotyping the *hMSH6* Ex1-145 G>A polymorphism (Primer3 software v.0.4.0 http://frodo.wi.mit.edu).

Statistical analyses (genotypes distribution - SPSS 13.0, Statgraphics 7; haplotype distribution - SAS/Genetics; linkage disequilibrium – Haploview (www.broad.mit.edu/mpg/haploview/documentation.php).

Other methods

Individual DNA repair capacity - γ-irradiation DNA repair test (Alapetite et al, 1999; Collins et al, 2001; Vodicka et al, 2003).

Analysis of repair capacity of PBL extracts towards repairing 8-oxoguanine (Collins et al, 1996)

Personal exposure monitoring using personal samplers for collection of PM2.5 particles (Binkova et al, 1998).

Lymphocyte cultures (PBL) for conventional cytogenetic analysis (CCA) (Carrano and Natarajan, 1988; Rossner et al, 2002)

Chromosomal aberrations - CCA (Carrano and Natarajan, 1988; Rossner et al, 2002)

Chromosomal aberrations - fluorescent in situ hybridization (FISH) analysis for chromosomes #1 and #4 (commercial whole chromosome painting probes (Cambio, UK) according to the manufacturer's protocol (Rubes et al, 1998).

Bulky-aromatic DNA adducts in lymphocytes - a 32P-postlabeling assay (Philips and Castegnaro, 1999; Binkova et al, 2003).

Urinary cotinine level by radioimmunoassay (Langone and van Vunakis, 1982)

Plasma level of vitamin C (Kiyoh and Megumi, 1993)

Plasma levels of vitamin E (alpha-tocopherol) and vitamin A (Driskell et al, 1982)

Folates level in plasma (CEDIA folate kit, Roche Diagnostics)

Plasma levels of triglycerides, total cholesterol, LDL and HDL cholesterol (spectrophotometry, Sigma diagnostic kits)

Table 2. Nomenclature of the SNPs studied, Taqman assays and PCR-RFLP conditions

Gene, position of SNP and nucleotide change	db number	Amino acid change	Taqman assay ID	Primers sequence for PCR-RFLP (5'→3')	Annealing temperature ,ºC	Restriction enzyme
Mismatch repair						
hMLH1 -93G>A	rs1800734	-	C_7535141_1_	-	60	-
hMLH1 IVS9 C>T	rs4647269	-	C_29968609_10	-	60	-
hMSH2 IVS12-6T>C	rs2303428		C_11804019_1	-	60	-
hMSH2 Ex6+23G>A	rs4987188	Gly322Asp	-	F: GTTTTCACTAATGAGCT R: AGTGGTATAATCATGTGGGT	Touch-down	HinfI
hMSH3 Ex4-100G>A	rs1805355	Pro231Pro	C_11434406_10	-	60	-
hMSH3 Ex23+3G>A	rs26279	Ala1045Thr	C_800002_1_	-	60	-
hMSH6 -556G>T	rs3136228	-	C_28985526_10	-	60	-
hMSH6 Ex1-145G>A	rs1042821	Gly39Glu	-	F: AGATGCGGTGCTTTTAGGAG R: CCCTCCGTTGAGGTTCTTC	Touch-down	SmaI
hMSH6 IVS4 G>C	rs2072447	-	C_22273199_10	-	60	-
EXO1 Ex12+49C>T	rs4149963	Thr439Met	C_25762095_10	-	60	-
Base-excision repair						
XRCC1 Ex6-22C>T	rs1799782	Arg194Trp	-	F GCC CCG TCC CAG GTA R AGC CCC AAG ACC CTT TCA CT	63	MspI
XRCC1 Ex9+16G>A	rs25489	Arg280His	-	F TTG ACC CCC AGT GGT GCT R CCC TGA AGG ATC TTC CCC AGC	57	RsaI
XRCC1 Ex10-4A>G	rs25487	Arg399Gln	-	F GCC CCT CAG ATC ACA CCT AAC R CAT TGC CCA GCA CAG GAT AA	65	MspI
hOGG1 Ex6-315C>G	rs1052133	Ser326Cys	-	F AGT GGA TTC TCA TTG CCT TCG R GGT GCT TGG GGA ATT TCT TT	59	Fnu4HI
APE1 Ex5+5T>G	rs1130409	Asn148Glu	C_8921503_10	F CTG TTT CAT TTC TAT AGG CTA R AGG AAC TTG CGA AAG GCT TC	59	BfaI
Nucleotide-excision repair						
XPD Ex23+61A>C	rs13181	Lys751Gln	C_3145033_10	F CCC CTC TCC CTT TCC TCT GTT	60	PstI

				R GCT GCC TTC TCC TGC GAT TA		
XPG Ex15-344G>C	rs17655	Asn1104His	C_1891743_10	F TGG ATT TTT GGG GGA GAC CT R CGG GAG CTT CCT TCA CTG AGT	56	Hsp92II
XPC Ex16+211C>A	rs2228001	Lys939Gln	C_234281_1_	F GAT GCA GGA GGT GGA CTC TCT R GTA GTG GGG CAG CAG CAA CT	61	PvuII
Double-strand break repair						
NBS1 Ex6-32G>C	rs1805794	Glu185Gln	C_26470398_10	F GGA TGT AAA CAG CCT CTT G R CAC AGC AAC TAT TAC ATC CT	59	HinfI
Metabolism						
EPHX1 Ex3+337T>C	rs1051740	Tyr113His	C_14938_30	-	60	-

4. RESULTS

In this chapter we describe findings which were obtained as results of experimental work carried out preferentially or partially by the author of this PhD thesis. Results on the association of SNPs in DNA BER, NER and double-strand break repair genes and risk of sporadic CRC, on association of SNPs in DNA repair genes, and on DNA repair capacity in healthy individuals were part of the research carried out by the Department of the Molecular Biology of Cancer, Institute of Experimental Medicine AS CR (head: Pavel E. Vodička, MD, PhD). The results on the association of DNA polymorphisms and a level of chromosomal aberrations and PAH-DNA adducts in an environmentally exposed population were a part of the research of the Department of Genetic Ecotoxicology of the same Institute (head: Radim J. Šrám, MD, DSc).

Association of SNPs in DNA MMR genes and risk of sporadic CRC (Appendix 1)

A modulating role of ten polymorphisms in MMR genes (*hMLH1, hMSH2, hMSH3, hMSH6* and *hEXO1*) and their haplotypes on a risk of CRC (as well as colon and rectal cancers separately) was investigated using a case-control approach. The study population consisted of 614 CRC patients and 614 controls, of Czech origin, matched for sex and age.

Ten common SNPs in genes *hMLH1*, *hMSH2*, *hMSH3*, *hMSH6* and *hEXO1* were genotyped in DNA from peripheral blood lymphocytes of all sample subjects using Pre-designed Taqman assays for allelic discrimination and a PCR-RFLP method. The selected polymorphisms were located in coding and non-coding regions of the genes and exerted a possible functional effect according to results of association and/or *in vitro* studies. Two tagging SNPs in the *hMLH1* gene and the *hMSH6* gene were selected based on Hapmap data (www.hapmap.org) for corresponding haplotypes reconstruction. The author of this PhD thesis carried out the vast majority of genotyping experiments and became a first author of the corresponding publication.

RESULTS SUMMARY

1. Two polymorphisms in the *hMSH6* gene showed an association with the altered risk of CRC. The carriers of variant allele for the promoter -556G>T polymorphism were at the increased risk of CRC (OR 1.29; 95% CI 1.02–1.62; P = 0.04), more pronounced for rectal cancer (OR 1.42; 95% CI 1.03–1.95; P = 0.03); whereas the carriers of the variant A allele for the -145G>A polymorphism in the *hMSH6* exon 1 (Gly39Glu) were at the decreased risk of CRC (OR 0.76; 95% CI 0.60–0.98; P = 0.03).

2. In the *hMSH6* gene, the variant allele of the intronic IVS4-101G>C SNP was associated with an increased risk of colon cancer (OR 1.34; 95% CI 1.03–1.74; P = 0.03). On the other hand, carriers of the T allele for the IVS9-1406C>T polymorphic variant *hMLH1* exhibited the decreased risk of rectal cancer (OR 0.71; 95% CI 0.51–0.98; P = 0.04). None of the above associations was significant after correcting for multiple hypotheses testing.

3. No association with either smoking habit or family history of CRC was shown for any of the polymorphisms in the MMR genes.

4. The TAG haplotype in the *hMSH6* gene (-556G>T - Ex1-145G>A - IVS4-101G>C) was associated with the significantly decreased risk in both colon and rectal cancer patients (OR 0.74; 95% CI 0.59–0.92; global P = 0.02). The GGG haplotype was found to be exclusively associated with an increased risk of rectal cancer (OR 1.32; 95% CI 1.05–1.65). None of the MMR haplotypes was significantly associated with the familial aggregation of the disease.

Association of SNPs in DNA BER, NER and DSB repair genes and risk of sporadic CRC (Appendix 2)

The study was carried out using the same sample[2], and the study design was in general the

[2] This analysis actually precluded the study of DNA MMR polymorphisms, thus the sample size was lower in that period: 532 CRC cases and 532 sex- and age-matched controls.

same as for the above-mentioned MMR polymorphisms. The SNPs in DNA BER (*XRCC1, hOGG1* and *APE1*), NER (*XPD, XPG and XPC*) and DSB repair (*NBS1*) genes and interactions in binary genotype combinations in genes involved in the same DNA repair pathway were tested for association with CRC risk. Particular attention was paid to the modifying effect of smoking and age on the association between DNA repair polymorphisms and CRC risk. The author of the thesis contributed into genotyping of above-mentioned SNPs.

RESULTS SUMMARY

1. The increased risk of CRC was shown in individuals carrying simultaneously variant alleles (homozygous genotypes) for *APE1* Ex5+5T>G (Asn148Glu) and *hOGG1* Ex6-315C>G (Ser326Cys)[3] polymorphisms (OR: 6.37; 95% CI: 1.40–29.02; P = 0.02). The same binary genotype combination also showed the increased risk for colon cancer.

2. Smoking modifies the extent of the association between individual CRC risk and the *hOGG1* Ex6-315C>G polymorphism. Smokers with the variant GG genotype for the polymorphism showed the increased risk of CRC (OR: 4.17; 95% CI: 1.17–15.54; P = 0.03).

3. In the group of oldest individuals (64-86 years), the association of some SNPs with CRC risk became more pronounced than in the whole studied population. It was valid for *APE1* Ex5+5T>G polymorphism - heterozygous and homozygous genotypes (OR: 1.79; 95% CI: 1.04–3.07; P = 0.04 and OR: 2.57; 95% CI: 1.30–5.06; P = 0.007, respectively), which increased CRC risk, and for homozygous variant CC genotype *XPG* Ex15-344G>C (Asn1104His), which was associated with a higher risk of rectal cancer.

[3] Here and in the following chapter we use unified SNPs nomenclature according to the SNP500 database http://variantgps.nci.nih.gov/cgfseq/pages/snp500.do as well as nomenclature based on the corresponding amino acid changes, as it was used in the original manuscript.

Association of SNPs in DNA BER and DSB repair genes and DNA repair capacity in healthy individuals (Appendix 3)

In this study, the association between SNPs in DNA BER (*XRCC1, hOGG1* and *APE1*) and DSB repair (*NBS1*) genes, and the capacity to repair DNA damage induced by *γ*-irradiation (using a comet assay) and DNA oxidative damage (induced by visible light in the presence of a photosensitizer) was investigated in a healthy population consisting of 244 healthy individuals (183 men and 61 women, mean age 41.3 ± 11.3 years, 90 individuals were smokers and 154 non-smokers) employed in local administration, medical centres, and various branches of the plastics industry. The study population was recruited in the regions of Western Slovakia and Eastern Bohemia with common socioeconomic conditions. Single SNPs and combinations of different polymorphisms in BER genes were analysed in relation to irradiation-specific DNA repair rates. The author of the thesis contributed into genotyping of above-mentioned SNPs within this study using a PCR-RFLP method.

RESULTS SUMMARY

1. Irradiation-specific DNA repair rates were mainly affected by *XRCC1* Ex10-4A>G (Arg399Gln) polymorphism and smoking. Irradiation-specific DNA repair rates were significantly decreased in individuals with the homozygous variant in *XRCC1* Ex10-4A>G (0.45 ± 0.47 SSB/109 Da) than those with the wild-type (1.10 ± 0.70 SSB/109 Da, P = 0.0006, Mann–Witney U-test).

2. The capacities to repair oxidative DNA damage were significantly decreased in individuals with the homozygous variant genotype in *hOGG1* Ex6-315C>G (Ser326Cys) (0.37 ± 0.28 SSB/109 Da) when compared to those with wild-type and heterozygous genotypes (0.83 ± 0.79 SSB/109 Da, P = 0.008, Mann–Witney U-test).

3. A significant decrease in the capacity to repair DNA oxidative damage was also associated with a combination of variant alleles in *hOGG1* Ex6-315C>G and *APE1* Ex5+5T>G (Asn148Glu), predominantly due to the variant G allele in *hOGG1* Ex6-315C>G (P = 0.018, Kruskal–Wallis test). The trend was also observed in binary combinations of polymorphisms *XRCC1* Ex10-4A>G with either *XRCC1* Ex6-22C>T (Arg194Trp) or *XRCC1* Ex9+16G>A (Arg280Hys) (P = 0.002 and P = 0.005, respectively; Kruskal–Wallis test).

Association of *EPHX1* polymorphism and a level of chromosomal aberrations and PAH-DNA adducts in an environmentally exposed population (Appendices 4 and 5)

The analysis of polymorphism in xenobiotic metabolizing gene *EPHX1* was a part of the extensive study, which addressed relationships between exposure to carcinogenic polycyclic aromatic hydrocarbons (c-PAHs), chromosomal aberrations (CA), DNA adducts, and polymorphisms in DNA repair and metabolizing genes in a group of occupationally exposed policemen (N= 53, males, aged 22–50 years) working outdoors in the downtown area of Prague and in matched "unexposed" volunteers (N= 52). The author of the thesis contributed by genotyping of polymorphism in exon 3 of *EPHX1* (Tyr113His) by TaqMan allelic discrimination assay.

RESULTS SUMMARY

1. The lower frequency of genomic translocations as identified by the FISH method was observed in individuals with high activity genotype of *EPHX1* gene in comparison to low and medium activity genotypes.

2. No significant association was shown between *EPHX1* polymorphism and a level of total and B[α]P-"like" DNA adducts.

5. DISCUSSION

Using a case-control approach, we tested the hypotheses that polymorphic variants in the genes encoding proteins involved in DNA MMR, BER, NER and DSB repair can modulate susceptibility to sporadic CRC. This study was part of the extensive research of the association of SNPs in genes involved in different molecular pathways with CRC risk in an ethnically homogeneous population from the Czech Republic. Tentative association with sporadic CRC has been investigated, in particular, for SNPs in metabolizing and DNA repair genes (Naccarati et al, 2006; Naccarati et al, 2007), cell cycle related genes (Polakova et al, 2009; Hemminki et al, 2014), miRNA-encoding genes (Pardini et al, 2015) and many others. In this chapter, we discuss the results of our research, with a particular emphasis on SNPs in MMR genes. While speaking about possible functional effects of SNPs in BER, NER and DSB repair genes, we illustrate it by the results of *in vitro* study on the association between these polymorphisms and the capacity to repair DNA damage induced by γ-irradiation and by base oxidation in healthy individuals. Taking into account the importance of environmental factors in DNA damage and cancer, we regard polymorphisms in xenobiotic-metabolizing gene *EPHX1* under the aspect of gene-environment interaction. Perspectives of further implications of SNPs in modern cancer research and clinical practice will be outlined briefly.

5.1. Individual SNPs in DNA repair genes and CRC risk

In the earlier stages of our research on SNPs in DNA repair genes as potential biomarkers of CRC risk, the candidate SNPs located in protein-encoding areas of DNA repair genes were selected, as it seemed the most meaningful way to test polymorphisms causing amino acid changes and subsequent steric and functional effects in the corresponding proteins. These SNPs

were also common ones, in accordance with the "common disease – common variant" model (Tomlinson et al, 2008). The strengths of our study included age- and sex- matching of cases and controls; a sufficient sample size; representative character of the study population (nine hospital centres around the whole country included into the study); inclusion of only 'colonoscopically negative' individuals to ensure participation of CRC-free control individuals (Brenner et al, 2010). The selection of SNPs in genes participating in DNA BER, NER and DSB (recombination) repair was solidly supported by data obtained in earlier *in vitro* and *in vivo* studies (considered in more details in appendices 2-3).

However, none of these SNPs *per se* showed significant association with overall risk of sporadic CRC in our hospital based case-control association study. Indeed, results of numerous recent GWAS suggest that only a relatively minor part of SNPs with possible disease association are located in, or occur in tight linkage disequilibrium, with protein-coding regions of genes: around 40% in intergenic regions, and another 40% in introns (Manolio, 2010). Thus, while designing the following study of SNPs in MMR genes, we dedicated more attention to regulatory, non-coding regions – in the firstinstance, to SNPs located in promoter and intronic regions.

Our results suggested a possible association between the increased risk of CRC and a variant T allele of the **-556G>T** polymorphism in the ***hMSH6* promoter**, probably due to its modulating effect on transcription. Results of an *in vitro* study on CHO cells indicated that this sequence variation results in the loss of a Sp1 binding site involved in the gene transcription regulation. In addition, -556T allele in combination with two other polymorphisms in *hMSH6* (-448G>A and -159C>T), apparently affects gene expression by the promoter methylation (Gazzoli and Kolodner, 2003). However, no association of the -159C>T polymorphism with CRC risk was reported in an independent study (Mrkonjic et al, 2007).

We failed to find the expected association between the -93G>A SNP located in the core

promoter region of the *hMLH1* gene and CRC risk in the studied population. The -93G>A polymorphism is located in a CpG island, and the A-allele is presumably involved in the epigenetic silencing of *hMLH1*, via promoter methylation (Chen et al, 2007). One of the plausible reasons for such a result could be that our case subjects were not classified according to MSI status. This explanation can be supported by findings within studies carried out by Raptis et al. (2007) and Campbell et al. (2009) on larger samples: the *hMLH1* -93A variant allele didn't show the association with overall CRC risk or risk of MSS CRC; however, it was significantly associated with the risk of MSI-H CRC. Further meta-analysis of results of six studies (17,791 cases and 13,782 controls in total), confirmed the association of the *hMLH1* -93G>A polymorphism with higher risk of MSI-H CRC (Wang et al, 2012b), and meta-analysis of 33 studies supported the hypothesis of the association of the presence of variant A allele for all types of CRC, especially in non-Asian populations (Xu et al, 2012). Indeed, this SNP seems to have quite wide variability in frequency in populations of different ethnic origin (Zhi et al, 2011). Other molecular signatures of CRC phenotypes could be also useful in stratification of CRC risk associated with MMR polymorphisms in our study: -93G>A polymorphism has been shown to be associated with *hMLH1* methylation, CIMP phenotype, and *BRAF* Val600Glu mutation in MSI-H CRC (Samowitz et al, 2008). Its impact into MSI in gastric cancer was recently reported, supporting the assumption on the modulating effect of this *MLH1* promoter polymorphism into the processes of MMR inactivation during tumorigenesis (Zhu et al, 2015).

Three intronic SNPs in *hMLH1, hMLH2* and *hMSH6* genes MMR genes were included into our research, but none of them was found to be associated with CRC risk. One of the studied polymorphisms, the intronic substitution IVS12-6T>C in the *hMSH2* gene, was associated with gastric cancer in a Chinese population (Wang et al, 2012a). However, our data did not show any association for this polymorphism with the risk of the disease; no increased CRC risk was also

observed in studies carried out on a Korean population (Kim et al, 2004) and on a population with mixed ethnicity from Canada (Raptis et al, 2007).

While analysing SNPs located in protein-encoding regions of MMR genes, we found that the carriers of the variant allele for the **Ex1-145G>A** polymorphism in the ***hMSH6*** coding for Gly39Glu amino acid change exhibited a significantly lower risk of CRC than non-carriers, both for colon and rectal cancers. Previous studies did not reveal any association of the same polymorphism with CRC risk (Berndt et al, 2007), and with the risk of adenomatous polyposis in populations of Caucasian origin (Yu et al, 2006). Glycine-to-glutamic acid substitution can determine the formation of sterically different helical structures, polypeptide folding, and intrinsic aggregation due to a hydrophilic side chain of glutamic acid, thus affecting the protein function (Branden et al, 1999; Chiti et al, 2003).

We found no significant associations of SNPs in *hMSH3* and *hEXO1* genes with CRC risk. An association with increased CRC risk has been shown for Ex23+3G>A polymorphism in *hMSH3* in patients of European descent (Mrkonjic et al, 2007; Koessler et al, 2008). The GG genotype was also positively associated with oesophageal cancer risk in the group of South Africans with mixed ethnic ancestry, especially in smokers (Vogelsang et al, 2012). According to the computational tool GeneVar (GENe Expression VARiation) platform, this polymorphism correlated with an increased expression of the *hMSH3* gene (Stranger et al, 2012; Yang et al, 2010).

Age is an important confounder in the assessment of CRC risk since the incidence of this cancer increases significantly in individuals older than 50 years, and therefore it is often considered as a disease that affects elderly people (Stigliano, 2014). There is also a strong association between age and the rate of transition from advanced adenoma to CRC; at the same time, this rate doesn't associate with gender (Brenner et al, 2007). In our association study we did

not observe any significant interaction between CRC risk, MMR polymorphisms, and age. A role of MMR in aging is still not sufficiently established; however, there are some indications of increasing MSI caused by MMR impairment in human organisms due to age (Neri et al, 2005). MMR can also be involved in other processes playing a significant role in aging: the repair of oxidative damage (*hMSH2, hMLH1*), signalling of apoptosis (*hMSH2*), and senescence (Conde-Perezprina et al, 2012). The **gender** didn't modulate CRC susceptibility according to results of our study; it seems to be in line with results of other research in the field.

Regarding age and SNPs in other DNA repair pathways under the study, the significant association was observed between ***APE1*** **Ex5+5T>G** (coding for Asn148Glu amino acid change) polymorphism and an increased CRC risk in individuals **between 64 and 86 years old**. The apurinic/apyrimidinic endodeoxyribonuclease 1 (APE1) is the essential enzyme in the **BER** pathway, playing a key role in repair of oxidative damage (Hoeijmakers, 2001), as well as a transcriptional co-activator for numerous transcription factors (AP-1, NF-κB, Myb, HIF-1α, HLF, PAX, and p53) involved in cancer promotion and progression by regulating the expression of their target genes (Ando et al, 2008). The *APE1* Ex5+5T>G polymorphism was recently associated with an increased CRC risk in Chinese population, however, without modulation by age (Zhang et al, 2014b).

Among SNPs in **NER** genes, only the **Ex15-344G>C** (Asn1104His) polymorphism in the ***XPG*** (xeroderma pigmentosum group G) gene exhibited an association with rectal cancer risk in the group of **older individuals** included in our study. A homozygous genotype of this polymorphism has been shown to be associated with higher level of DNA damage, strand breaks in particular, in healthy individuals (Slyskova et al, 2014). As aging is characterized by higher levels of DNA damage (Moller, 2006; Schumacher et al, 2008), SNPs promoting accumulation of DNA damage can introduce an additional load to the organ systems prone to carcinogenic

processes. Recently, a significantly lower level of DNA NER capacity and the increased level endogenous DNA damage have been demonstrated in blood cells of newly diagnosed CRC patients in comparison to age-matched healthy controls using a comet assay; however, no correlation with age was observed (Slyskova et al, 2012b).

In general, it seems to be accepted that it is often not feasible to recognize a distinct role of a single common polymorphism in carcinogenesis (Naccarati et al, 2007; Picelli et al, 2013). However, indications obtained in such association studies can be a valuable source for evaluation on a larger scale and this data can be also used in pooled analyses. For example, in a systematic review and meta-analysis based on the "Venice criteria", out of 241 associations investigated, only three SNP demonstrated a strong grade of cumulative evidence (among them the Ex6-32G>C polymorphism in the *NBS1* gene involved into DNA DSB repair associated with higher risk of bladder cancer; in our study we didn't observe any association of this SNP with CRC risk) (Ricceri et al, 2012). Sometimes, this findings are also attributed to discrepancies in the literature associated with different sample size, ethnicity, gene-environment and gene-lifestyle interactions, study design and experimental protocols etc. Therefore, investigation of the combined effect of multiple SNPs in several genes and in one or more relevant DNA repair pathways could be more relevant in finding cancer associated traits.

The main limitation of our study was a lack of additional data on molecular phenotype of CRC cases in our study (for example, MSI, CIMP, *KRAS* and *BRAF* mutations). More detailed data on life style could also help in more interesting stratification of our results; on the other side, it would decrease the power of our findings due to sample size limitations.

5.2. Binary combinations of SNPs and haplotypes in DNA repair genes and CRC risk

Our data confirms the assumption that common gene variants are less likely to be associated with "single-gene" effects. The analysis of SNP combinations and haplotypes seems to be more promising and to help to define the more complex patterns of reactions.

We found that the TAG haplotype, based on three SNPs of the ***hMSH6*** gene **(-556G>T - Ex1-145G>A - IVS4-101G>C)**, was associated with a decreased overall risk of CRC. However, we recorded a significant association with sporadic CRC risk for none of haplotypes of *hMLH1*, *hMSH2*, or *hMSH3* genes. In the currently available scientific literature, significant association with cancer risk is mostly reported on haplotypes in regulatory regions of some of these genes. Three-way gene-gene interactions were detected within *hMSH2* gene among intronic polymorphisms IVS11+107A>G, IVS11+183A>G and IVS8+719T>C in association with CRC risk (Li, 2015). Furthermore, the combined *IVS12−6CC* and *IVS10+12AA* genotypes of *hMSH2* also significantly increased the risk of gastric cancer (Wang et al, 2012a). The increased CRC risk was proposed for a rare *hMLH1* haplotype (Koessler et al, 2008) and for *hMSH3* haplotype constructed using two SNPs in the gene including Ex23+3G>A which we also analysed (Berndt et al, 2007).

As the result of our case-control association study of SNPs in BER, NER and DSB repair genes, the increased risk of CRC was shown for individuals carrying variant alleles homozygous genotypes for both ***APE1*** **Ex5+5T>G** (Asn148Glu) and ***hOGG1*** **Ex6-315C>G** (Ser326Cys) polymorphisms. The association with CRC risk was recently reported for both polymorphism in Turkish population, with additional association of *hOGG1* Ex6-315C>G with higher CRC grade and liver metastasis; however, the sample size was very low (Canbay et al, 2011). Variant alleles of two above mentioned polymorphisms in combination with *XRCC1* (X-ray repair cross-complementing protein 1) Ex10-4A>G variant G allele also significantly reduced lung cancer risk

(Li et al, 2011), earlier this polymorphism was associated with many types of cancer (Naccarati et al, 2007). Controversial results have been published on the association of the *hOGG1* Ex6-315C>G polymorphism with cancer risk; however, the recent meta-analysis confirmed only association of this SNP with head-and-neck cancer and only in Asian populations (Wang et al, 2012c).

Concerning the latter three polymorphisms and their combinations, the approval of their functional effect was obtained in our ***in vitro*** **study on association of DNA repair polymorphisms with DNA repair capacity in healthy individuals**. Deficient DNA repair capacity and the accumulation of different types of DNA damage can represent both early and intermediate markers in carcinogenesis (Schumacher et al, 2008). DNA repair capacity vary significantly among individuals (Collins et al, 2001), with age and sex as important parameters (Slyskova et al, 2014), reflecting types of exposure specific for the individual's environment (e.g., Vodicka et al, 2004; Cebulska-Wasilewska et al, 2005; Taioli et al, 2007 etc.), and also among different tissues of the same organism due to their specific functional characteristics (Azqueta et al, 2013). To understand the relationship between genotype and phenotype, it is important to be able to differentiate between this natural variation and the effects determined by the misdirected metabolism of malignant cells.

To assess the background level of the potential modification of individual DNA efficiency by SNPs in DNA repair genes in the PBL of healthy individuals, the capacity to repair DNA damage induced by γ-irradiation and by base oxidation was analysed in our *in vitro* study on a healthy population from Western Slovakia and Eastern Bohemia. Within this study, a significant decrease in the capacity to repair DNA oxidative damage was associated with the same combination of variant alleles in ***APE1*** **Ex5+5T>G** and ***hOGG1*** **Ex6-315C>G,** which we considered above in relation to CRC risk in our case-control study. This effect was predominantly

due to the variant G allele in *hOGG1* Ex6-315C>G (as approved in multifactorial analysis). In healthy individuals, the *hOGG1* wild type C allele was associated with 2-fold higher capacity to repair oxidative DNA damage (CC vs. GG genotype) independently or in combinations with *XRCC1* Ex6-22C>T (Arg194Trp) or *XRCC1* Ex9+16G>A (Arg280His) in our study. Many researchers reported an impaired DNA repair under the oxidative stress associated with the *hOGG1* G allele, with suggestion that the effect can be underlid by the protein localization and post-transcriptional regulation (e.g., Luna et al, 2005; Kershaw et al, 2012). Chronic inflammation due to oxidative stress imposed by reactive oxygen species, is considered one of the main etiological factors in CRC and continues playing an important role on all further steps of CRC pathogenesis (Perše et al, 2013). Slightly decreased expression levels of the *hOGG1* gene was found in tumor tissues of CRC patients in comparison to their normal mucosa (Slyskova et al, 2012a). On the other side, the increased expression levels of *APE1* and *hOGG1* in PBL of patients with CRC and benign adenoma in comparison to healthy controls were reported by other researchers, suggesting elevated oxidative stress for CRC and adenoma (Obtułowicz et al, 2010).

Naturally, *in vitro* studies have not a few limitations and their results cannot be directly extrapolated into *in vivo* conditions. It is especially valid when we consider processes taking place in normal and tumor cells, particularly in human populations differed in age (participants of our study on healthy individuals were medium 10 years younger than in our CRC study).

In conclusion, the study of multi-gene SNP interactions and haplotypes of DNA repair genes seems to be a relevant approach in association studies, yet the lack of appropriate research and publications is still noted in this field (Ricceri et al, 2012). Our study provides a range of important indications in this aspect to be developed in further studies using larger samples and supported by data on individual cancer molecular phenotype.

5.3. SNPs and haplotypes in DNA repair genes and tumor localization in CRC

Every region of a large intestine has distinct features from point of view of anatomy, mucosa morphology, biochemical and immunological parameters that gradually change along the gut (Iacopetta, 2002). It is assumed nowadays that the oncogenesis of left- and right-sided CRC may involve, at least partially, differing mechanisms. Earlier it was suggested that some of these mechanisms can be mediated by nuclear β-catenin either within the APC/β-catenin pathway and/or independently of APC (Kapiteijn et al, 2001). The gradual changes in a CRC stage, grade, serration and mucin production as well as predominant sex, MSI, CIMP and other histological and molecular parameters were discovered in CRC with different bowel localization (Bae et al, 2013). The extent of LINE-1 hypomethylation also differed between tumor sites, so this mechanism can also be responsible for specific CRC phenotypes (Figueiredo et al, 2009).

For some polymorphisms under the study, we observed differential distribution according to cancer localization. In the ***hMSH6*** gene, a variant allele of the intronic **IVS4-101G>C** SNP was associated with increased risk of **colon cancer** but not overall CRC risk. On the other hand, carriers of the T allele for another intronic polymorphic variant the **IVS9-1406C>T** in ***hMLH1*** exhibited a decreased risk of **rectal cancer**. Though after correcting for multiple hypotheses testing, none of the above associations was convincingly significant, it can be important indication for future studies, confirming the role of SNPs located in regulatory regions of MMR genes.

We also observed that the most frequent GGG haplotype in ***hMSH6*** gene **(-556G>T - Ex1-145G>A - IVS4-101G>C)** was associated with an increased risk of **rectal cancer**. One of these polymorphisms, the non-synonymous *hMSH6* Ex1-145G>A polymorphism was found to be significantly associated with the risk of MSI-H colon cancer in Caucasian, African-American or Hispanic subjects (Campbell et al, 2009). It has been shown recently that

expression of MMR genes vary between tissues of tumors with different localization. The *hEXO1, hMSH2, hMSH3, hMSH6,* and *hPMS*2 genes were up-regulated in colon tumors in comparison to rectal tumors (Vymetalkova et al, 2014b).

As for SNPs in other DNA repair pathways, the variant allele of **Ex5+5T>G** polymorphism of the ***APE1*** gene showed significant association with **colon cancer** in our study, while the overall CRC risk was significant only in the **eldest individuals**. It was suggested that increased expression of DNA repair genes in colon can be a manifestation of protective mechanism, as colon is exposed to higher amounts of carcinogens than other parts of a gut due to significantly longer time of a bolus standing in colon (Vymetalkova et al, 2014b).

The **Ex15-344G>C** (Asn1104His) polymorphism in the NER gene ***XPG*** exhibited an association with **rectal cancer** risk but only in the group of **older individuals** in our study. It seems much more likely that these phenotypic differences can be mediated rather by rSNPs than by polymorphisms in protein-encoding areas of NER genes: for instance, two SNPs in miRNA-binding sites of other NER genes were associated with increased risk of rectal cancer in the Czech population (Naccarati et al, 2012).

Additional data about molecular characteristics of tumors in our sample (MSI status, *KRAS* and *BRAF* mutations, CIMP and CIN status) could, apparently, clarify underlying molecular mechanisms of such differentiation and/or to refer tumors under consideration to particular molecular phenotype.

5.4. Gene-environment interactions in CRC risk: SNPs in DNA repair genes and diet

The complex role of environmental/life-style factors in cancer etiology and pathogenesis remains the most fascinating and the least explored area of cancer research nowadays. Among life-style factors, the most acute attention should be dedicated to food quality and dietary habits,

as results of various studies demonstrate the importance of **diet-genomic interactions** in cancer development (Ogino et al, 2011). Sporadic CRC is among the types of cancer in which nutrition has an extreme significance. Factors such as obesity, low physical activity, high red meat consumption, tobacco smoking, and alcohol abuse bring definite negative effects into occurrence and development of CRC (Alexandrova et al, 2014).

The situation with regards the quality of the nutrition of the Czech population is quite alarming nowadays. According to results of the HAPIEE study on dietary habits in Central and Eastern European countries, even though some progress has been achieved concerning the intake of fresh seasonal vegetables and fruits, Healthy Diet Indicator in the country appeared to be more than two or three times less than the average in Europe and the USA (Boylan et al, 2009), with a forecasted epidemic of obesity in the proximal years (Statistics, F.A.P., 2013). As regarding obesity, no significant interaction with BMI was observed in our study. It can be mentioned, though, that according to the extensive analysis carried out by Harding et al. (2015) in Australia and New Zealand, waist circumference seems to be a better predictive factor for CRC risk than BMI. In our study, an analysis of questionnaire-based lifestyle data revealed a significant difference in dietary habits in cases with more vegetarians (12%) compared to controls (6%). However, hardly any far-reaching conclusions can be drawn based on this observation due to the subjectivity of questionnaires, which decreases the reliability of the given information. Moreover, questionnaire-based observations can introduce profound bias due to dietary recall and the influence of confounders, including diet modulation during the lifetime and inflammatory bowel diseases (Martínez et al, 2008).

Components of food can have a substantial effect on the process of tumorigenesis, either in protecting against cancer or promoting the malignant transformation. Dietary factors may act as genotoxicants, through epigenetic mechanisms, or to modulate DNA repair capacity. Examples

include aflatoxin B1, ochratoxin A, ptaquiloside, various pyrrolizidine alkaloids, heterocyclic amines, and PAH, such as benzo[a]pyrene (B[a]P). Mutations and/or different types of DNA damage (e.g., CA and copy number variants) may be induced by endogenously or exogenously formed reactive species, inhibitors of topoisomerase II enzymes (e.g., flavonoids), DNA repair (e.g. caffeine), or of the mitotic spindle (acrylamide) (Ferguson and Philpott, 2008; Dahm et al, 2010). Some micronutrients such as zinc, magnesium, selenium, folic acid, and vitamin B12, are involved in DNA methylation, synthesis, and repair, mainly as co-factors of important proteins (e.g., p53), thus their inadequate intake may mimic important mutations that promote genomic instability and cancer initiation (Fenech, 2002).

On the other hand, the individual cancer risk can be decreased by nutrients acting as antimutagens (e.g. vitamin C, carotenoids, chlorophyllin, dietary fibers, and plant polyphenols) (Ferguson and Philpott, 2008). These substances are considered as possible agents of chemoprevention - the use of natural or synthetic chemical agents to reverse, suppress or prevent the carcinogenic progression (Teixeira, 2014; Castells, 2015).

A collection of the most detailed information about dietary habits, covering the precise individual intake of different nutrients, can give researchers different insights into the influence of food components on CRC risk. For example, in the study of a polymorphism in the *ApoE* gene involved in lipid metabolism, no significant differences in genotype frequencies were observed between CRC cases and healthy controls. However, among high red meat consumers, *ApoE* isoforms modulated the risk of MSI-H and MSS/MSI-L CRCs. In non-ApoE4 (Epsilon 4) carriers, 2–4 and more red meat servings/week were associated with developing MSS/MSI-L CRC, whereas among ApoE4 allele carriers, four or more red meat servings/week were associated with MSI-H CRC when compared with the controls (Mrkonjic et al, 2010). An interaction was observed between the intake of processed meat and the MMR gene *hMSH3*

Ex23+3107G>A in CRC (Berndt, 2007), and also between *h*MSH2 *IVS12−6T>C* polymorphism and high pickled food intake or fried food intake in relation to the risk of gastric cancer in the Chinese population (Wang et al, 2012a). Significant gene-environment interaction was observed between intronic polymorphisms in *hMSH2* gene IVS15-214T>C and IVS11+107A>G and cereal consumption (Li, 2015).

Interaction between genotype and nutrition can be substantially affected by the mucosa-associated bacteria that have a primary impact on the colonic environment. Enteropathogenic *Escherichia coli* is able to down-regulate MMR proteins such as *hMSH2* and *hMLH1* (Maddocks et al, 2009), so the presence of these bacteria can theoretically further attenuate deficient expression of MMR genes caused by rSNPs.

5.5. Gene-environment interactions in CRC risk: SNPs in DNA repair genes and smoking

Smoking was another important environmental/lifestyle factor we attempted to address. Smoking is considered a main preventable reason of cancer (Fisher et al, 2011). In the Czech Republic, 24.3% of adults (19% women, 30% men) are daily smokers (OECD, 2010). It would also be extremely interesting to assess the role of **alcohol drinking** in CRC susceptibility within our population, since higher alcohol consumption is an established risk factor in CRC (Fedirko et al, 2011; Alexandrova et al, 2014). Czech Republic is well known for long-term high alcohol consumption, with a threatening trend in the last two decades (OECD, 2012; WHO, 2014). According to results of the questionnaire-based study in the Czech Republic, Poland, and Russia, the Czech men had the highest intake of alcohol among men and women from these countries (Boylan et al, 2009). Regretfully, alcohol consumption was not included as covariate into our analysis, even though the data on alcohol consumption have been collected for our study

population. However, we found that it was important to touch upon the issue of alcohol drinking in this discussion.

We didn't discover any modulation of CRC risk in relation to any of the polymorphisms in the MMR genes caused by smoking. A variety of results were observed for the effect modification by smoking or drinking in the MMR-deficient background, which might be due to the ethnic difference, the different definitions for smoking and drinking, the heterogeneity in history of smoking and drinking, and the potential confounding effect of other known exposures. It can perhaps also be partially explained by a lack of data on molecular signatures of CRC tumors in our study: for instance, Campbell et al. (2009) observed that smoking (but not drinking) had a significant interaction with *hMLH1* −93G>A variant in MSS colon cancers. This polymorphism was also significantly associated with the risk of lung squamous cell carcinoma with a gene-smoking interaction (Park et al, 2004). The effect of smoking on CRC risk can be dose-dependent: Pande et al. (2010) discovered that individuals with MMR-deficient background may be at increased risk of CRC if they smoke regularly, while former smokers, short-term smokers, and light smokers were at decreased CRC risk.

Smoking affected the CRC risk in carriers of SNPs in **BER** genes in our case-control study: smokers with variant allele homozygous genotype for the ***hOGG1* Ex6-315C>G** polymorphism were at increased risk of CRC compared to control smokers. Conversely, in upper aerodigestive tract cancer, this polymorphism (CG or GG vs. CC genotypes) was interpreted as a protective factor in moderate smokers (Marques et al, 2014). In association studies on CRC risk, this polymorphism is often reported with smoking and drinking, especially in combination with ***XRCC1* Ex10-4A>G** polymorphism. The *XRCC1* Ex10-4A>G genotype was associated with a decreased risk of CRC among smokers and drinkers (Zhang et al, 2014b). In other studies performed in Chinese population, individuals carrying G allele of the *XRCC1* Ex10-4A>G, and

having a smoking or drinking habit, had an increased risk of CRC (Huang et al, 2015).

The *XRCC1* Ex10-4A>G combined with smoking status seems to predict progression-free survival to radiotherapy patients with nasopharyngeal carcinoma, with possible correlation between these genetic polymorphism and p53 protein status (Jin et al, 2014). In the Thai population, polymorphisms in *XRCC1* and *hOGG1* genes, particularly in combination, were associated with increased susceptibility to cholangiocarcinoma, modified by smoking and alcohol consumption (Sonqserm et al, 2014). Interaction between tobacco and polymorphisms of *XRCC1* Ex10-4A>G increases the risk of head and neck squamous cell carcinoma in Indian population (Choudhury et al, 2014). The risk of gastric cancer was significantly elevated in individuals with *XRCC1* Ex10-4A>G AG and GG genotypes (Ghosh et al, 2016).

In our study of DNA repair capacity on healthy individuals, irradiation-specific DNA repair rates, attributable mostly to BER repair, were significantly decreased in subjects with the homozygous variant in ***XRCC1* Ex10-4A>G** than those with the wild-type. The trend was also observed in binary combination of this SNP with either *XRCC1* Ex6-22C>T or *XRCC1* Ex9+16G>A polymorphisms.

In general, irradiation-specific DNA repair rates in PBL of healthy individuals (mostly attributed to the functioning of BER system) were affected by smoking in our study, while no such effect was observed for oxidative damage-induced DNA repair rates (those were modulated by occupational exposure status of the subjects); age and sex didn't modulate this effect as well. In the recent study, though, sex but not smoking appeared to predetermine changes in DNA damage or DNA BER and NER in healthy individuals (Slyskova et al, 2014).

5.6. SNPs as potential biomarkers for clinical practice: perspectives

Modern schemes of CRC treatment result in high tumor response rates and relatively long overall patient survival; at that, the outcome is greatly influenced by tumor stage and grade at diagnosis. Tumor, node and metastasis (TNM) classification is used nowadays to plan CRC treatment and to predict outcomes of the therapy (Greene, 2003). Tumor response to a therapy is modulated by specific clinico-pathological features of the patient, environmental factors, ethnic origins and the individual's genetic and epigenetic status (Mundade, 2014). Tumor resistance and the adverse drug reactions (in particular, myelo-, gastro- and neurotoxic reactions) represent the main problems of CRC management. Since at present CRC is considered not one entity but rather a heterogeneous group of diseases with variant profiles of genetic and epigenetic alterations, reliable predictive and prognostic biomarkers are essential to develop a targeted and personalized approach to decision-making in CRC treatment (Mohelnikova-Duchonova et al, 2014).

Up to date the majority of molecular markers demonstrate quite divergent and inconsistent results on their prognostic and/or predictive value, except for MSI, *KRAS* and *BRAF* mutations (Schmoll et al, 2012; Duffy et al, 2014). However, a clinical use of these markers is still limited: *KRAS* Gly13Asp mutation for identification of resistance to anti-EGFR monoclonal antibodies in metastatic CRC patients; *BRAF* Val600Glu mutation for identification of resistance to anti-EGFR monoclonal antibodies in metastatic CRC patients and for exclusion of Lynch syndrome; MSI for confirmation of Lynch syndrome (Reimers et al, 2013). In addition, the Oncotype DX colon cancer assay (the 12-gene colon cancer Recurrence Score) from Genomic Health Inc. (Redwood City, CA, USA) has been validated for clinical planning of CRC treatment for patients with stage II and stage III (Reimers et al, 2013)

A potential of germline mutations and SNPs as CRC markers is extensively studied during last decades and there are several important reasons for that. SNPs are highly abundant in the

genome, the mutation rate in germline is quite low, they can be detected in blood, feces or other tissues; genotyping is technically easy and reliable nowadays and it can be automated (Strimpakos et al, 2009). All these features make SNPs a very attractive goal for biomarker search, since testing for tumor markers is highly dependent on sampling techniques, preservation, laboratory conditions and methodology, as well as on population characteristics (ethnicity, macro-and microenvironment, tumor localization, stage and grade, age etc.) Moreover, results can be changed under influence of some therapy making outcomes discrepant or unreliable (Strimpakos et al, 2009).

There are three main potential applications of SNPs for clinical use: as biomarkers of susceptibility, predictive and prognostic biomarkers (Reimers et al, 2013). In our case-control study, we addressed SNPs in DNA repair genes as potential biomarkers of susceptibility – predisposition to CRC in the Czech population. In line with other research in the field, we demonstrated that the effect of SNPs on modulation of CRC risk is quite limited; our results also indicate the role of tumor localization, age and environmental factors in cancer predisposition. Individual SNPs rarely have a pronounced effect on disease risk and phenotype but their synergetic effect can change DNA repair activity; they can be studied as a part of multi-marker panels for assessment CRC risk (Picelli et al, 2013; Reimers et al, 2013). To verify results of GWAS, PARE sequencing and other modern high-throughput methods, other association and observational studies in large cohorts with sufficient environmental and pathological data (tumor markers, lifestyle, occupational exposures etc.) should be applied.

Approved functional effect of SNPs doesn't automatically mean that their effect on cancer risk can be detected (Huhn et al, 2014). Since many chemotherapeutic substances used in cancer treatment are processes by DNA MMR, BER and other repair systems (Iyer et al, 2006), SNPs in DNA repair genes can participate in modulating of response to these drugs and can be studied as

tentative biomarkers of prediction and prognosis for different types of cancer (Reimers et al, 2013). For instance, one of MMR polymorphisms we analysed (*hMSH2* IVS12-6T>C) has been found by other researchers to be associated with an adverse response to chemotherapy with alkylating agents: the variant C allele was associated with the increased risk of developing of therapy-related acute myeloid leukaemia following (Worrillow et al, 2003). The association with survival rate and with the risk of relapse or metastasis for CRC patients was found for polymorphisms in *hMSH3*, and the association with recurrence rates for polymorphisms *hMSH6* (Vymetalkova et al, 2014a). SNPs in another MMR gene *hEXO1* were also associated with CRC survival (Yoshiva et al, 2008). The *XRCC1* Ex10-4A>G polymorphism was used as a predictor of response to radiotherapy and progression-free survival in patients with nasopharyngeal carcinoma (Jin et al, 2014); it was also associated with better prognosis in CRC (Moreno et al, 2006).

Undoubtedly, we should always keep in mind that craving new fundamental biological knowledge is not the only and primary goal of cancer research from point of view of the public benefit. Integrated efforts of different stakeholders including scientists, clinical and public health professionals and educational institutions, are necessary for proper dissemination of results of cancer epidemiology and biology as well as for implementation of this data into practice. The risk of CRC is influenced by both environmental and genetic factors. However, according to results of twin studies, effects of heritable factors in sporadic CRC do not exceed 13-35% (Lichtenstein et al, 2000; Czene et al, 2002). The rest of the CRC risk is determined by life style and other environmental factors mediated, primarily, via epigenetic mechanisms (Li and Martin, 2016). Unsurprisingly, according to some estimations, up to 50% of CRC incidence can be a subject of primary prevention based on dietary and lifestyle modifications including chemoprevention (Teixeira et al, 2014). Thus, CRC prevention is the issue of first-rate importance.

Nowadays a CRC screening using faecal occult blood test (FOBT) and structural

examination - sigmoidoscopy/colonoscopy stays the most important strategy in early detection of CRC. Since CRC and most of its precursors – adenomas and serrated polyps - are usually asymptomatic during the early stages, screening makes possible to detect them and to eliminate by endoscopy or surgery, as result substantially reducing CRC morbidity and mortality (Castells, 2015). To decrease the risk of CRC in individuals/families with inherited MMR mutations, a germline mutation screening is also indicated, followed by a recommended screening colonoscopy, prophylactic hysterectomy and bilateral salpingo-oophorectomy for MMR mutation carriers and their family members (Win et al, 2013b). A National program of CRC screening in the Czech Republic has been started in the year 2000. Currently, the annual CRC screening is intended for asymptomatic 50 to 54 years old individuals and includes FOBT, with a subsequent colonoscopy in case of the positive FOBT result. From 55 years of age it is possible to choose either a FOBT repeated once per two years, or primary screening colonoscopy as an alternative, to be possibly repeated within 10 years (Suchánek et al, 2011). However, the situation with CRC prevention in the Czech Republic can be considered satisfactory only with a very long stretch of imagination: the screening remains the main and basically only national-wide program on CRC, being officially presented as a prevention, while in reality it constitutes mainly an early diagnostic measure. Without massive campaigns directed on promotion of healthy lifestyle and health literacy, no dramatic decrease of CRC incidence in the country can be foreseen in the long term perspective.

In particular, the issues of healthy nutrition, tobacco smoking and alcohol consumption seem to be drastically underestimated in the Czech Republic. Even though some progress in a quality of diet of Czech inhabitants took place since the beginning of democratization processes in the country in 1990's (Boylan et al, 2009), tobacco and alcohol are not considered as „harmful" drugs by significant part of the Czech population, and number of smokers has not been

decreasing substantially during two decades (Sovinová et al, 2010). Results of cancer epidemiological and genetic (genomic, epigenomic etc.) research, especially those concerning molecular mechanisms of gene-environment interactions, are awaiting for careful consideration and implementation into evidence-based national cancer prevention strategy.

5.7. Gene-environment interaction: polimorphisms in *EPHX1* gene and air pollution

Air pollution is a negative factor of urban development, with a growing global concern (Lave and Seskin, 2013). It is associated with acute and chronic morbidity and mortality, including cardiovascular-specific, pulmonary diseases, immune syndromes and certain types of cancer (Beelen et al, 2014).

A level of air pollution in Prague in the period of our study was higher than official standard for the Czech Republic, mostly due to intense traffic (Sram et al, 2007). Among components of urban pollution, particulate matter (PM) of aerodynamic equivalent diameter <2.5μm (PM2.5) belongs to the most harmful air pollutants, in particular, since it carries many dangerous chemicals adsorbed on their surface (Anderson et al, 2012). The diverse PAH, including B[a]P, represent a group of highly hazardous pollution-derived carcinogens causing high levels of DNA damage (Topinka et al, 2011; Rossner and Sram, 2014). The effect of exposure to c-PAH adsorbed on respirable air particles was extensively studied in several groups of occupationally exposed healthy men working in the downtown area in comparison to volunteers spending working hours indoors in suburban areas of Prague. A personal exposure to eight c-PAH and PM2.5 during a work-shift, supported by daily air pollution data from stationary monitoring, was assessed in association with DNA adducts and CA. DNA adducts (in our study measured using ^{32}P-postlabelling assay) are considered a biomarker of personal exposure, while CA serve as validated complex biomarkers of early effect. Using both CCA and FISH methods in

our study allowed to cover a range of unstable and stable CA, respectively (Sram et al, 2011). As main results of this study, the CCA didn't show any difference in CA levels between exposed policemen and controls, while all CA endpoints detected by FISH (concerning stable CA mostly), were elevated in the exposed group compared to controls (Sram et al, 2007). As for DNA adducts, the level of the B[a]P-"like" adducts was significantly higher in exposed individuals and the total DNA adduct level was similar in the exposed and control groups (Binkova et al, 2007).

Within this study, it was hypothesized that polymorphisms in DNA repair and metabolic genes can modulate susceptibility to formation of DNA adducts and CA in PBL of the studied subjects. Our part of the study concerned an SNP in the exon 3 (T>C, Tyr113His), in combination with an SNP in the exon 4 (A>G, His139Arg) of the same gene coding for the microsomal epoxide hydrolase (EPHX1) enzyme.

The EPHX1 plays a distinct role in the organism's response to xenobiotics and in maintenance of organ- and tissue-specific homeostasis. Substrates of this enzyme are, the most typically, of exogenous origin and include highly reactive oxidative metabolites - epoxides (Ginsberg et al, 2010). Conversely, EPHX1 can be induced or deactivated by certain exogenous or endogenous compounds. For instance, several PAH have been demonstrated to induce *in vitro* EPHX1 activity in human liver tissues (Pushparajah et al, 2008). The EPHX1 is suggested to be involved into processes of neuro-, mielo-, and hematotoxicity under exposure to different xenobiotics (Vaclavíková et al, 2015). Two above mentioned *EPHX1* polymorphisms in the exon 3 and 4 were suggested to affect the enzyme activity. In the exon 3, C allele shows reduced enzyme activity by at least 50% (slow allele), while exon 4 G allele leads to 25% activity increase, with adverse influence on the protein level (Hassett et al, 1994). Thus, these polymorphisms have been widely used as markers of EPHX1 activity in molecular epidemiology studies, with a specific classification on low activity, medium activity and high activity for

particular allele combinations of both polymorphisms (diplotype)[4]. However, there is still no general agreement on the role and extent of the effect of these SNPs especially due to their contra-directional impacts on the enzyme activity, and a fact that most of the corresponding data were obtained within *in vitro* studies and their results were not completely reproducible *in vivo* (Ginsberg et al, 2010). A few recent studies also suggest that other SNPs, including those located in promoter regions, can participate in regulation of the *EPHX1* gene expression and the corresponding enzyme activity (Ginsberg et al, 2010). Modulation of toxicity of EPHX1 substrates by the *EPHX1* diplotype was investigated in variety of studies, mostly regarding occupational pollutants including vinyl chloride, 1,3-butadiene, styrene (Ginsberg et al, 2010)

In our study, no association was observed between separate polymorphisms or diplotypes of *EPHX1* and chromosomal damage according to results of the CCA. However, the FISH, as more sensitive method, allowed to detect a decreased level on translocations in individuals with high activity *EPHX1* genotype in comparison to low and medium activity genotype. This trend was observed in the exposed group as well as in controls. A protective effect of high activity *EPHX1* genotype against chromosomal damage was shown in the literature on other studies on healthy individuals. A lower frequency of aberrant cells (as determined by CCA) in association with high activity *EPHX1* genotype was reported earlier in occupationally exposed individuals (Vodicka et al, 2004), as well as confirmed in a recent study on larger sample which included subjects with a wide range of lifestyle and occupational exposure variables (Hemminki et al, 2015). However, in human PBL *in vitro, EPHX1* genotypes of higher activity were associated with increased level of B[a]P-induced CA (Salama et al, 2001).

Among obvious strengths of our study, we can mention abundant data on lifestyle and biochemical parameters (level of vitamins, cotinine, lipids etc.); no gender as confounder since

[4] See the classication in the Materials and Methods section of the Appendix IV (Sram et al, 2007).

only men were participating in the study, the possibility to use both CCA and FISH methods for CA measurement. However, our work has certain limitations. First, the sample size was quite low, so our results can be considered as a sort of indication of the effect. Within such sample, possibilities for stratification by important confounders are limited. At that, smoking and age can substantially affect the association between CA and *EPHX1* genotypes: in the study on Norwegian men, the *EPHX1* high activity genotype was associated with lower level of chromatid gaps mostly in older smokers (Skjelbred et al, 2011). Besides, in CCA within our study the chromatid gaps were recorded, but they were not scored as aberrations. In our study, number of smokers in exposed group was about two-fold higher than in controls, thus causing potential bias.

Obviously, the *EPHX1* genotypes can be associated at different extent with different types on CA. In the study of Skjelbred et al. (2011), the *EPHX1* genotypes were mostly attributed to association with chromatid-type CA; Hemminki et al. (2015) report association of the *EPHX1* high activity genotype with lower level of total CA (individually and in combination with polymorphism in another metabolizing gene *GSTP1*). However, in the same study the EPHX1 diplotype in combination with polymorphism in the *NQO1* gene modulates a level of chromatid type CA in healthy individuals (Hemminki et al, 2015). These finding seem to be logically correct: as shown by *in vitro* experiments, the chromatid-type CA are mostly formed as result of chemical exposure, while chromosome-type CA are supposed to be an outcome of direct DNA damage, in particular, caused by ionizing radiation (Albertini et al, 2000).

For instance, when the formation of micronuclei (MN) has been used as a biomarker of chromosomal damage under exposure to c-PAH, confirmation of the role of the *EPHX1* diplotype in modulation of MN frequency is mostly related to occupationally exposed population, with less effect observed in control individuals in different gender and age groups (reviewed by Iarmarcovai et al, 2008; Rossnerova et al, 2011; Merlo et al, 2014). An extent and a vector of the

effect can also be exposure-specific: for instance, lower frequency of MN was observed for individuals with high activity *EPHX1* genotype at individuals occupationally exposed to 1,3-butadiene (Tan et al, 2010). Since 1,3-butadiene is detoxified by the EPHX1 enzyme, the effect of genotypes determining different enzyme activity can be more straightforward than in case of c-PAH, that are activated by EPHX1. Results of further studies in this population assume that *EPHX1* polymorphisms can affect c-PAH-induced chromosomal damage through more complicated metabolic steps, probably involving oxidative stress mechanisms (Novotna et al, 2007; Rossner et al, 2011; Ghosha et al, 2013).

No association between *EPHX1* polymorphisms and a level of total and B[α]P-"like" DNA adducts was observed in our study. As well as in the case of *EPHX1* polymorphisms and CA, quite diverse results are presented in the available literature. It can be partially explained by controversial role of EPHX1 in genotoxicity and carcinogenesis: on the one hand, it detoxify a range of environmental and industrial carcinogens (e.g., 1,3-butadiene, styrene, benzene), on the other, it leads to activation of other hazardous xenobiotics as c-PAH (Vaclavíková et al, 2015).

In the study carried out in five regions of Spain in both men and women, the significantly increased level of bulky DNA adducts was found in relation to the high activity genotype *EPHX1*; it was concluded that the metabolic pathways of oxidation, hydrolysis and acetylation are relevant for the formation of bulky DNA adducts (Agudo et al, 2009).

Smoking could also play a confounding role in interaction between formation of DNA adducts and *EPHX1* since tobacco smoke also contain B[α]P and other c-PAH: smokers carrying the variant G allele of EPHX1 exon 4 polymorphism were significantly more likely to have higher DNA adduct levels in blood (Peluso et al, 2013). From other side, slow *EPHX1* alleles were associated with lower level of B[α]P-protein- and DNA adducts in lung cancer patients (Pastorelli et al, 1998). A significant increase in tobacco-related lung cancer risk was observed

for high activity *EPHX1* genotypes as compared to low activity genotypes (Park et al, 2005).

The association with B[α]P-DNA adducts was observed for several other metabolizing genes, including those coding glutathione-S-transferases (GST) (Lodovici et al, 2004). These enzymes can also use epoxides as substrate; therefore, the effect of *EPHX1* diplotypes can be masked or, reversely, to exhibit a synergy in combination with different polymorphisms of these genes (Ginsberg et al, 2010). In our study, an increased level of total DNA adducts was observed in subjects with null *GSTM1* genotype, while the *EPHX1* genotype didn't predict a level of either total or B[α]P-"like" DNA adducts. As mentioned above, it seems that the risk or, alternatively, the protective effects of *EPHX1* genotypes can be increased in combination with other polymorphisms: for instance, carriers of five "at risk" alleles in *EPHX1, NQO1* (NAD(P)H quinone dehydrogenase 1), *APE1* and other metabolizing and DNA repair genes had significantly higher DNA adduct levels in conditions of urban environmental exposure and/or smokers than subjects carrying fewer than two "at risk" alleles (Peluso et al, 2013).

As mentioned above, the EPHX1 can be induced or inhibited by a range of exogenous and endogenous compounds. Glucocorticoids inhibit *EPHX1* expression through interaction with the gene promoter (Ginsberg et al, 2010). It is widely known that elevated cortisol levels are related to higher work stress (job strain, work overload, over-commitment to work) as well as to general life stress (Chida and Steptoe, 2009). It was shown that traffic policemen of both sexes had a significantly higher levels of plasma adrenocorticotropic hormone compared to control subjects with indoor activity (Tomei et al, 2003). Therefore, we can speculate that in our study, the policemen working in a downtown area of heavy traffic and crowd can potentially have higher work stress than control indoor workers, and that higher glucocorticoid level can be one of the factors supressing c-PAH-activating function of the EPHX1 in their blood cells.

As a final note, we can mention that *EPHX1* genotype has been studied in relation to lung

and oesophageal cancer, leukaemia and lymphomas, and also in relation to susceptibility to CRC risk. Results of these studies vary significantly, however, there can be a higher risk of CRC and polyps in relation to smoking and red meat consumption (reviewed by Ginsberg et al, 2010). Newly emerging studies indicate that epigenetic mechanisms including methylation of different genes and miRNA expression may play a significant role in interaction between environmental insults, *EPHX1* genotype, and subsequent DNA damage and cancer (Xing et al, 2013; Ravengini et al, 2015).

In conclusion, our study provides important indication on potential role of *EPHX1* genotypes in modulation of susceptibility to environmental genotoxic insults. An increased frequency of CA is believed to predict cancer risk (Bonassi et al, 2008; Vodicka et al, 2010). Thus genotypes of *EPHX1* which increased susceptibility of CA can modulate risk of carcinogen-induced cancer in environmentally exposed individuals.

6. CONSLUSIONS

The main conclusions of the PhD thesis are as follows:

1. The modulation of susceptibility to sporadic CRC by SNPs and haplotypes in genes involved in DNA MMR was assessed for the first time in the Czech population within a hospital-based case-control study. Our results demonstrate a limited role of MMR polymorphisms; however, haplotypes and SNPs in regulatory gene regions can be involved in the modulation of CRC risk in this population.

2. Gene-environment interactions and age can be among the most important factors modulating CRC risk associated with polymorphisms in DNA repair genes. Complex assessment of these interactions is of primary importance and should be supported by studies on molecular mechanisms underlying these interactions.

3. Some SNPs and haplotypes modulate the risk of either colon or rectal cancer, without significant association with the overall CRC risk. These findings confirm the heterogeneous nature of CRC and the importance of tumor molecular profiling for the proper interpretation of the results of association studies.

4. SNPs within the same DNA repair pathway can interact with each other, resulting in a synergistic effect on the CRC risk. Some proportion of this influence can be mediated by the functional effect of these SNPs and their combinations on individual DNA repair capacity, as shown in the frame of the *in vitro* study on healthy individuals.

5. The SNPs coding for amino acid changes in DNA repair proteins can modulate DNA repair capacities in healthy individuals. These results are important for individualizing a natural genotype-phenotype variation to be used as a baseline to study in conditions of malignancy.

6. Individuals carrying higher activity genotypes of the metabolizing gene *EPHX1* can be more resistant to stable chromosomal damage under exposure to c-PAH, i.e., they can potentially have a lower risk of cancers induced by environmental carcinogens.

7. REFERENCES

Abdel-Rahman WM, Mecklin J-P, Peltomaki P: The genetics of HNPCC: application to diagnosis and screening. Crit Rev Oncol Hematol 2006, 58:208-220.

Agudo A, Peluso M, Sala N, et al: Aromatic DNA adducts and polymorphisms in metabolic genes in healthy adults: findings from the EPIC-Spain cohort. Carcinogenesis 2009, 30: 968–976.

Aherne ST, Madden SF, Hughes DJ, et al: Circulating miRNAs miR-34a and miR-150 associated with colorectal cancer progression. BMC Cancer 2015, 15: 1.

Alapetite C, Thirion P, De La Rochefordiere A, et al: Analysis by alkaline comet assay of cancer patients with severe reactions to radiotherapy: defective rejoining of radioinduced DNA strand breaks in lymphocytes of breast cancer patients. Int J Cancer 1999; 83: 83–90.

Albertini RJ, Anderson D, Douglas GR, et al: IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. International Programme on Chemical Safety. Mutat Res 2000, 463: 111–172.

Aleksandrova K, Pischon T, Jenab M, et al: Combined impact of healthy lifestyle factors on colorectal cancer: a large European cohort study. BMC Med 2014, 12: 1.

Anderson JO, Thundiyil JG, Stolbach A: Clearing the air: a review of the effects of particulate matter air pollution on human health. J Med Toxicol 2012, 8: 166-175.

Ando K, Hirao S, Kabe Y, et al. A new APE1/Ref-1-dependent pathway leading to reduction of NF-kappaB and AP-1, and activation of their DNA-binding activity. Nucleic Acids Res 2008, 36: 4327–4336.

Aran D, Hellman A: DNA methylation of transcriptional enhancers and cancer predisposition. Cell 2013, 154: 11-13.

Azqueta A, Langie SA, Slyskova J, Collins AR: Measurement of DNA base and nucleotide excision repair activities in mammalian cells and tissues using the comet assay–a methodological overview. DNA repair 2013, 12: 1007-1010.

Bae JM, Kim JH, Cho NY, et al: Prognostic implication of the CpG island methylator phenotype in colorectal cancers depends on tumour location. Brit J Cancer 2013, 109: 1004-1012.

Bak ST, Sakellariou D, Pena-Diaz J: The dual nature of mismatch repair as antimutator and mutator: for better or for worse. Front Genet 2014, 5: 287.

Barrow TM, Michels KB: Epigenetic epidemiology of cancer. Biochem Biophys Res Comm 2014, 455: 70-83.

Beelen R, Raaschou-Nielsen O, Stafoggia M, et al: Effects of long-term exposure to air pollution on natural-cause mortality: an analysis of 22 European cohorts within the multicentre ESCAPE project. The Lancet 2014, 383: 785-795.

Beiner ME, Rosen B, Fyles A, et al: Endometrial cancer risk is associated with variants of the mismatch repair genes *MLH1* and *MSH2*. Cancer Epidemiol Biomarkers Prev 2006, 15: 1636–1640.

Berndt SI, Platz EA, Fallin MD, et al: Mismatch repair polymorphisms and the risk of colorectal cancer. Int J Cancer 2007, 120: 1548-1554.

Binkova B, Cerna M, Pastorkova A, et al: Biological activities of organic compounds adsorbed onto ambient air particles: comparison between the cities of Teplice and Prague during the summer and winter seasons 2000–2001. Mutat Res 2003, 525: 43–59.

Binkova B, Chvatalova I, Lnenickova Z, et al: PAH–DNA adducts in environmentally exposed population in relation to metabolic and DNA repair gene polymorphisms. Mutat Res 2007, 620: 49-61.

Binkova B, Topinka J, Mrackova G, et al: Coke oven workers study: the effect of exposure and GSTM1 and NAT2 genotypes on DNA adduct levels in white blood cells and lymphocytes as determined by32P-postlabelling. Mutat Res 1998, 416: 67–84.

Boland CR, Thibodeau SN, Hamilton SR, et al: A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer. Cancer Res 1998, 58: 5248-5257.

Boleij A, Tjalsma H: Gut bacteria in health and disease: a survey on the interface between intestinal microbiology and colorectal cancer. Biological Reviews 2012, 87: 701-730.

Bonassi S, Norppa H, Ceppi M, et al: Chromosomal aberration frequency in lymphocytes predicts the risk of cancer: Results from a pooled cohort study of 22 358 subjects in 11 countries. Carcinogenesis 2008, 29: 1178–1183.

Boylan S, Welch A, Pikhart H, et al: Dietary habits in three Central and Eastern European countries: the HAPIEE study. BMC Public Health 2009, 9:439.

Branden C, Tooze J: Introduction to Protein Structure. Garland Publishing, 2nd edition, Taylor and Francis Group, New York, 1999.

Brenner H, Haug U, Arndt V, et al: Low risk of colorectal cancer and advanced adenomas more than 10 years after negative colonoscopy. Gastroenterology 2010, 138: 870-876.

Brenner H, Hoffmeister M, Stegmaier C, et al: Risk of progression of advanced adenomas to colorectal cancer by age and sex: estimates based on 840 149 screening colonoscopies. Gut 2007, 56:1585-1589.

Burnett-Hartman AN, Newcomb PA, Potter JD: Infectious agents and colorectal cancer: a review of Helicobacter pylori, Streptococcus bovis, JC virus, and human papillomavirus. Cancer Epidem Biomar 2008, 17: 2970-2979.

Campbell PT, Curtin K, Ulrich CM, et al: Mismatch repair polymorphisms and risk of colon cancer, tumour microsatellite instability and interactions with lifestyle factors. Gut 2009, 58: 661-667.

Canbay E, Cakmakoglub B, Zeybekb U, et al: Association of APE1 and hOGG1 polymorphisms with colorectal cancer risk in a Turkish population. Curr Med Res Opin 2011, 27: 1295-1302.

Carrano A, Natarajan AT: Considerations for population monitoring using cytogenetic techniques. Mutat Res 1988, 204: 379–406.

Castells A: Choosing the optimal method in programmatic colorectal cancer screening: current evidence and controversies. Therap Adv Gastroenterol 2015, 8: 221-233.

Cebulska-Wasilewska A, Wiecheć A, Panek A, et al: Influence of environmental exposure to PAHs on the susceptibility of lymphocytes to DNA-damage induction and on their repair capacity. Mutat Res 2005, 588: 73-81.

Chen H, Taylor NP, Sotamaa KM, et al: Evidence for heritable predisposition to epigenetic silencing of MLH1. Int J Cancer 2007, 120: 1684–1688.

Chen J, Furano AV: Breaking bad: The mutagenic effect of DNA repair. DNA repair 2015, 32: 43-51.

Chida Y, Steptoe A: Cortisol awakening response and psychosocial factors: a systematic review and meta-analysis. Biol Psychol 2009, 80: 265-278.

Chiti F, Stefani M, Taddei N, et al: Rationalization of the effects of mutations on peptide and protein aggregation rates. Nature 2003, 424: 805–808.

Choudhury JH, Choudhury B, Kundu S, Ghosh S, et al: Combined effect of tobacco and DNA repair genes polymorphisms of XRCC1 and XRCC2 influence high risk of head and neck squamous cell carcinoma in northeast Indian population. Med Oncol 2014, 31: 1-9.

Collins AR, Dusinska M, Gedik C, Stetina R: Oxidative damage to DNA: do we have a reliable biomarker? Environ Health Perspect 1996, 104: 465–469.

Collins AR, Dušinská M, Horváthová E et al: Inter-individual differences in repair of DNA base oxidation, measured in vitro with the comet assay. Mutagenesis 2001, 16: 297-301.

Conde-Perezprina JC, Leon-Galvan MA, Konigsberg M: DNA mismatch repair system: repercussions in cellular homeostasis and relationship with aging. Oxid Med Cell Longev 2012, 2012: 728430.

Cunningham JM, Christensen ER, Tester DJ, et al: Hypermethylation of the hMLH1 promoter in colon cancer with microsatellite instability. Cancer Res 1998, 58: 3455-3460.

Czene K, Lichtenstein P, Hemminki K: Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish family-cancer database. Int J Cancer 2002, 99: 260-266.

Dahm CC, Keogh RH, Spencer EA, et al: Dietary Fiber and Colorectal Cancer Risk: A Nested Case–Control Study Using Food Diaries. JNCI J Natl Cancer Inst 2010.

David SS, O'Shea VL, Kundu S: Base-excision repair of oxidative DNA damage. Nature 2007, 447: 941-950.

Desprat N, Supatto W, Pouille P-A, et al: Tissue Deformation Modulates Twist Expression to Determine Anterior Midgut Differentiation in Drosophila Embryos. Develop Cell 2008, 15: 470–477.

Driskell JW, Neese CC, Bryant MM, Bashor C: Measurement of vitamin A and vitamin E in human serum by highperformance liquid chromatography. J Chromatogr 1982, 231: 439–444.

Duffy MJ, Lamerz R, Haglund C et al: Tumor markers in colorectal cancer, gastric cancer and gastrointestinal stromal cancers: European group on tumor markers 2014 guidelines update. Int J Cancer 2014, 134: 2513–2522.

Duraturo F, Liccardo R, Cavallo A, et al: Association of low-risk MSH3 and MSH2 variant alleles with Lynch syndrome: probability of synergistic effects. Int J Cancer 2011, 129: 1643-1650.

Dušek L, Mužík J, Májek O: Nové trendy v epidemiologii kolorektálního karcinomu v ČR. Časopis lékařů českých 2014, 153: 100-102.

Edelbrock MA, Kaliyaperumal S, Williams KJ: Structural, molecular and cellular functions of MSH2 and MSH6 during DNA mismatch repair, damage signaling and other noncanonical activities. Mutat Res 2013, 743-744: 53-66.

Esteban-Jurado C, Garre P, Vila M, et al: New genes emerging for colorectal cancer predisposition. World J Gastroenterol 2014, 20: 1961-1971.

Esteller M, Levine R, Baylin SB, et al: MLH1 promoter hypermethylation is associated with the microsatellite instability phenotype in sporadic endometrial carcinomas. Oncogene 1998, 17(18).

Ezzati M, Henley SJ, Lopez AD, Thun MJ: Role of smoking in global and regional cancer epidemiology: current patterns and data needs. Int J Cancer 2005, 116: 963-971.

Fedirko V, Tramacere I, Bagnardi V, et al: Alcohol drinking and colorectal cancer risk: an overall and dose–response meta-analysis of published studies. Ann Oncol 2011, 22: 1958-1972.

Fenech M: Chromosomal biomarkers of genomic instability relevant to cancer. Drug discovery today 2002, 7: 1128-1137.

Ferguson LR, Philpott M: Nutrition and mutagenesis. Annu Rev Nutr 2008, 28: 313-329.

Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, et al: Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. Eur J Cancer 2013, 49: 1374-1403.

Figueiredo JC, Maria V. Grau MV, Kristin Wallace K, et al: Global DNA Hypomethylation (LINE-1) in the Normal Colon and Lifestyle Characteristics and Dietary and Genetic Factors. Cancer Epidemiol Biomarkers Prev 2009, 18: 1041.

Fisher EB, Fitzgibbon ML, Glasgow RE, et al: Behavior matters. Am J Prev Med 2011, 40: e15-30.

Fridley BL, Biernacka JM: Gene set analysis of SNP data: benefits, challenges, and future directions. Eur J Hum Gen 2011, 19: 837-843.

Fukui K: DNA mismatch repair in eukaryotes and bacteria. J Nucl Acids 2010.

Gazzoli I, Kolodner RD: Regulation of the human *MSH6* gene by the Sp1 transcription factor and alteration of promoter activity and expression by polymorphisms. Mol Cell Biol 2003, 23: 7992–8007.

Ghosha R, Topinka J, Joad JP: Air pollutants, genes and early childhood acute bronchitis. Mutat Res 2013, 749: 80– 86.

Ginsberg G, Guyton K, Johns D, et al: Genetic polymorphism in metabolism and host defense enzymes: Implications for human health risk assessment, Crit Rev Toxicol 2010, 40: 575-619.

Goessl C, Plaschke J, Pistorius S, et al: An intronic germline transition in the HNPCC gene

hMSH2 is associated with sporadic colorectal cancer. Eur J Cancer 1997, 33: 1869–1874.

Ghosh S, Ghosh S, Bankura B, et al: Association of DNA repair and xenobiotic pathway gene polymorphisms with genetic susceptibility to gastric cancer patients in West Bengal, India. Tumor Biology 2016, 1: 11.

Greene FL: TNM staging for malignancies of the digestive tract: 2003 changes and beyond. Semin Surg Oncol 2003, 21: 23–29.

Guillotin D, Martin SA: Exploiting DNA mismatch repair deficiency as a therapeutic strategy. Exp Cell Res 2014, 329: 110-115.

Gustavsson B, Carlsson G, Machover D, et al: A review of the evolution of systemic chemotherapy in the management of colorectal cancer. Clin Colorectal Cancer 2015, 14: 1-10.

Harding JL, Shaw JE, Anstey KJ, et al: Comparison of anthropometric measures as predictors of cancer incidence: A pooled collaborative analysis of 11 Australian cohorts. Int J Cancer 2015, 137: 1699–1708.

Harfe BD, Jinks-Robertson S. Mismatch repair proteins and mitotic genome stability. Mutat Res. 2000, 451: 151-167.

Harrison DJ, Hubbard AL, MacMillan J, et al: Microsomal epoxide hydrolase gene polymorphism and susceptibility to colon cancer. Br J Cancer 1999, 79: 168–171.

Hassett C, Aicher L, Sidhu LA, Omiecinski CJ: Human microsomal epoxide hydrolase: genetic polymorphism and functional expression in vitro of amino acid variants. Hum Mol Genet 1994, 3: 421–428.

Hemminki K, Frank C, Försti A, et al: Metabolic gene variants associated with chromosomal aberrations in healthy humans. Genes, Chromosomes and Cancer 2015, 54: 260–266.

Hemminki K, Mušák L, Vymetálková V, et al: Cyclin D1 splice site variant triggers chromosomal aberrations in healthy humans. Leukemia 2014, 28: 721-2.

Hoeijmakers JH: Genome maintenance mechanisms for preventing cancer. Nature 2001, 411: 366–374.

Hsieh P, Yamane K: DNA mismatch repair: molecular mechanism, cancer, and ageing. Mech Ageing Dev 2008, 129: 391-407.

Hu JJ, Smith TR, Miller MS, et al: Genetic regulation of ionizing radiation sensitivity and breast cancer risk. Environ Mol Mutagen 2002, 39: 208–215.

Huang Q: Genetic study of complex diseases in the post-GWAS era. J Genet Genomics 2015, 42: 87-98.

Huang Y, Li X, He J, et al: Genetic polymorphisms in XRCC1 genes and colorectal cancer susceptibility.World J Surg Oncol 2015, 13: 1.

Huhn S, Bevier M, Pardini B, et al: Colorectal cancer risk and patients' survival: influence of polymorphisms in genes somatically mutated in colorectal tumors. Cancer Causes & Control 2014, 25: 759-769.

Huxley RR, Ansary-Moghaddam A, Clifton P et al: The impact of dietary and lifestyle risk factors on risk of colorectal cancer: a quantitative overview of the epidemiological evidence. Int J Cancer 2009, 125: 171–180.

Iacopetta B, Grieu F, Amanuel B: Microsatellite instability in colorectal cancer. Asia-Pacific Journal of Clinical Oncology 2010, 6: 260-269.

Iacopetta B: Are there two sides to colorectal cancer? Int J Cancer 2002, 101: 403–408.

Iarmarcovai G, Bonassi S, Botta A, et al: Genetic polymorphisms and micronucleus formation: a review of the literature. Mutat Res 2008, 658: 215-233.

Iyama T, Wilson DM, 3rd: DNA repair mechanisms in dividing and non-dividing cells. DNA Repair 2013, 12: 620-636.

Iyer RR, Pluciennik A, Burdett V, Modrich PL: DNA mismatch repair: functions and mechanisms. Chem Reviews 2006, 106: 302-323.

Jain M, Tilak AR, Upadhyay R et al: Microsomal epoxide hydrolase (EPHX1), slow (exon 3, 113His) and fast (exon 4, 139Arg) alleles confer susceptibility to squamous cell esophageal cancer. Toxicol App Pharmacol 2008, 230: 247–251.

Jin H, Xie X, Wang H, et al: ERCC1 Cys8092Ala and XRCC1 Arg399Gln polymorphisms predict progression-free survival after curative radiotherapy for nasopharyngeal carcinoma. PloS one 2014,9: e101256.

Jiricny J: Postreplicative mismatch repair. Cold Spring Harbor perspectives in biology 2013, 5: a012633.

Joost P, Veurink N, Holck S, et al: Heterogenous mismatch-repair status in colorectal cancer. Diagn Pathol 2014, 9: 126.

Juo YY, Johnston FM, Zhang DY, et al: Prognostic value of CpG island methylator phenotype

among colorectal cancer patients: a systematic review and meta-analysis. Annals Oncol 2014, 25: 2314-2327.

Kámory E, Kolacsek O, Ottó S, Csuka O: hMLHl and hMSH2 somatic inactivation mechanisms in sporadic colorectal cancer patients. Pathol Oncol Res 2003, 9: 236-241.

Kapiteijn E, Liefers GJ, Los LC, et al: Mechanisms of oncogenesis in colon versus rectal cancer. J Pathol 2001, 195: 171–178.

Kershaw RM, Hodges NJ: Repair of oxidative DNA damage is delayed in the Ser326Cys polymorphic variant of the base excision repair protein OGG1. Mutagenesis 2012, 27: 501–510.

Khan S: Potential role of Escherichia coli DNA mismatch repair proteins in colon cancer. Crit Rev Oncol Hematol 2015, 96: 475-482.

Kim JC, Roh SA, Koo KH, et al: Genotyping possible polymorphic variants of human mismatch repair genes in healthy Korean individuals and sporadic colorectal cancer patients. Fam Cancer 2004, 3: 129–137.

Kiyoh T, Megumi K: High-performance liquid chromatographic determination of plasma ascorbic acid in relationship to health care. J Chromatogr 1993, 613: 275–280.

Koessler T, Oestergaard MZ, Song H, et al: Common variants in mismatch repair genes and risk of colorectal cancer. Gut 2008, 57: 1097–1101.

Kuismanen SA, Holmberg MT, Salovaara R, et al: Genetic and epigenetic modification of MLH1 accounts for a major share of microsatellite-unstable colorectal cancers. Am J Pathol 2000, 156: 1773-1779.

Langone JJ, van Vunakis H: Radioimmunoassay of nicotine, cotinine and (3-pyridyl)-oxo-*N*-methylbutyramine. Methods Enzymol 1982, 84: 628–640.

Lave LB, Seskin EP (2013). Air pollution and human health (Vol. 6). Routledge.

Lee KM, Choi JY, Kang C, et al: Genetic polymorphisms of selected DNA repair genes, estrogen and progesterone receptor status, and breast cancer risk. Clin Cancer Res 2005, 11: 4620–4626.

Li G, Hu F, Yuan F, et al: Intronic and promoter polymorphisms of hMLH1/hMSH2 and colorectal cancer risk in Heilongjiang Province of China. J Cancer Res Clin Oncol 2015, 141: 1393-1404.

Li GM: Mechanisms and functions of DNA mismatch repair. Cell Res. 2008, 18: 85-98.

Li SK, Martin A: Mismatch Repair and Colon Cancer: Mechanisms and Therapies Explored. Trends Molec Med 2016, 22: 274-289.

Li YI, van de Geijn B, Raj A, et al: RNA splicing is a primary link between genetic variation and disease. Science 2016, 352: 600-604.

Li Z, Guana W, Li M-x et al: Genetic polymorphism of DNA base-excision repair genes (APE1, OGG1 and XRCC1) and their correlation with risk of lung cancer in a Chinese population. Arch Med Res 2011, 42: 226–234.

Lichtenstein P, Holm NV, Verkasalo PK, et al: Environmental and heritable factors in the causation of cancer—analyses of cohorts of twins from Sweden, Denmark, and Finland. New Eng J Med 2000, 343: 78-85.

Ligtenberg MJ, Kuiper RP, Chan TL, et al: Heritable somatic methylation and inactivation of MSH2 in families with Lynch syndrome due to deletion of the 3' exons of TACSTD1. Nature Gen 2009, 41: 112-117.

Lin B, Gupta D, Heinen CD: Human pluripotent stem cells have a novel mismatch repair-dependent damage response. J Biol Chem 2014, 289: 24314-24324.

Lindahl T, Wood RD: Quality control by DNA repair. Science 1999, 286: 1897-1905.

Lodovici M, Luceri C, Guglielmi F, et al: Benzo(a)pyrene diolepoxide (BPDE)-DNA adduct levels in leukocytesof smokers in relation to polymorphism of CYP1A1, GSTM1, GSTP1, GSTT1, andmEH. Cancer Epidemiol Biomarkers Prev 2004, 13: 1342–1348.

Luna L, Rolseth V, Hildrestrand GA, et al: Dynamic relocalization of hOGG1 during the cell cycle is disrupted in cells harbouring the hOGG1-Cys326 polymorphic variant. Nucleic Acids Res 2005, 33: 1813–1824.

Maddocks ODK, Short AJ, Donnenberg MS, et al: Attaching and effacing Escherichia coli downregulate DNA mismatch repair protein in vitro and are associated with colorectal adenocarcinomas in humans. PLoS ONE 2009, 4: e5517.

Manolio TA: Genome-wide association studies and assessment of the risk of disease. N Engl J Med 2010, 363: 166-176.

Marques CR, Da Silva TM, De Albuquerque DM, et al: NAT2, XRCC1 and hOGG1 Polymorphisms, Cigarette Smoking, Alcohol Consumption and Risk of Upper Aerodigestive Tract Cancer. Anticancer Res 2014, 34: 3217-3224.

Martínez, ME, Marshall JR, Giovannucci E: Diet and cancer prevention: the roles of observation and experimentation. Nature Rev Cancer 2008, 8: 694-703.

Martini E, Borde V, Legendre M, et al: Genome-wide analysis of heteroduplex DNA in mismatch repair-deficient yeast cells reveals novel properties of meiotic recombination pathways. PLoS Genet 2011, 7: e1002305.

Martín-López JV, Fishel R: The mechanism of mismatch repair and the functional analysis of mismatch repair defects in Lynch syndrome. Familial cancer 2013, 12: 159-168.

Merlo DF, Agramunt S, Anna L, et al: Micronuclei in cord blood lymphocytes and associations with biomarkers of exposure to carcinogens and hormonally active factors, gene polymorphisms, and gene expression: the NewGeneris cohort. Environ Health Perspectives 2014, 122: 193-200.

Mitchell RJ, Campbell H, Farrington SM, et al: Prevalence of family history of colorectal cancer in the general population. British J Surgery 2005, 92: 1161-1164.

Modrich P, Lahue R: Mismatch repair in replication fidelity, genetic recombination, and cancer biology. Ann Rev Biochem 1996, 65: 101-133.

Mohelnikova-Duchonova B, Melichar B, Soucek P: FOLFOX/FOLFIRI pharmacogenetics: The call for a personalized approach in colorectal cancer therapy. World J Gastroenterol 2014, 20: 10316–10330.

Moller P: Assessment of reference values for DNA damage detected by the comet assay in human blood cell DNA. Mutat Res 2006, 612: 84–104.

Mooney MA, Nigg JT, McWeeney SK, Wilmot B: Functional and genomic context in pathway analysis of GWAS data. Trends Genet 2014, 30: 390-400.

Moreno V, Gemignani F, Landi S, et al: Polymorphisms in Genes of Nucleotide and Base Excision Repair: Risk and Prognosis of Colorectal Cancer. Clin Cancer Res 2006, 12: 2101.

Mrkonjic M, Chappell E, Pethe VV, et al: Association of apolipoprotein E polymorphisms and dietary factors in colorectal cancer. British J Cancer 2009, 100: 1966 – 1974.

Mrkonjic M, Raptis S, Green RC, et al: MSH2 118T>C and MSH6 159C>T promoter polymorphisms and the risk of colorectal cancer. Carcinogenesis 2007, 28: 2575-2580.

Mrkonjic M, Roslin NM, Greenwood CM, et al. Specific variants in the MLH1 gene region may drive DNA methylation, loss of protein expression, and MSI-H colorectal cancer. PloS one 2010, 5: e13314.

Mundade R, Imperiale TF, Prabhu L, et al: Genetic pathways, prevention, and treatment of sporadic colorectal cancer. Oncoscience 2014, 1: 400-406.

Muniz-Mendoza R, Ayala-Madrigal ML, Partida-Perez M, et al: MLH1 and XRCC1

polymorphisms in Mexican patients with colorectal cancer. Genet Mol Res 2012, 11: 2315-2320.

Naccarati A, Pardini B, Hemminki K, Vodicka P: Sporadic colorectal cancer and individual susceptibility: a review of the association studies investigating the role of DNA repair genetic polymorphisms. Mutat Res 2007, 635: 118-45.

Naccarati A, Pardini B, Landi S, et al: Polymorphisms in miRNA-binding sites of nucleotide excision repair genes and colorectal cancer risk. Carcinogenesis 2012, 33: 1346–1351.

Naccarati A, Soucek P, Stetina R et al: Genetic polymorphisms and possible gene–gene interactions in metabolic and DNA repair genes: Effects on DNA damage. Mutat Res 2006, 593: 22–31.

Nagasaka T, Rhees J, Kloor M, et al. Somatic hypermethylation of MSH2 is a frequent event in Lynch Syndrome colorectal cancers. Cancer Res 2010, 70: 3098-3108.

Nassiri M, Kooshyar MM, Roudbar Z, et al: Genes and SNPs associated with non-hereditary and hereditary colorectal cancer. Asian Pac J Cancer Prev 2013, 14: 5609-5614.

Neri S, Gardini A, Facchini A, et al: Mismatch repair system and aging: microsatellite instability in peripheral blood cells from differently aged participants. J Gerontol Series A: Biol Sci Med Sci 2005, 60: 285-292.

Niessen RC, Sijmons RH, Ou J, et al: MUTYH and the mismatch repair system: partners in crime? Hum Gen 2006, 119: 206-211.

Nizam ZM, Abdul Aziz AA, Kaur G, et al: Contribution of the MLH1 -93G>a promoter polymorphism in modulating susceptibility risk in Malaysian colorectal cancer patients. Asian Pac J Cancer Prev 2013, 14: 619-624.

Noonan EM, Shah D, Yaffe MB, et al: O6-Methylguanine DNA lesions induce an intra-S-phase arrest from which cells exit into apoptosis governed by early and late multi-pathway signaling network activation. Integr Biol (Camb) 2012, 4: 1237-1255.

Novotna B, Topinka J, Solansky I, et al: Impact of air pollution and genotype variability on DNA damage in Prague policemen. Toxicol Lett 2007, 172: 37–47.

O'Brien MJ, Zhao Q, Yang S: Colorectal serrated pathway cancers and precursors. Histopahol 2015; 66: 49-65.

O'Brien V, Brown R: Signalling cell cycle arrest and cell death through the MMR system. Carcinogenesis 2006, 27: 682-692.

Obtułowicz T, Swoboda M, Speina E, et al: Oxidative stress and 8-oxoguanine repair are enhanced in colon adenoma and carcinoma patients. Mutagenesis 2010, 25: 463-471.

Ogino S, Chan AT, Fuchs CS, Giovannucci E: Molecular pathological epidemiology of colorectal neoplasia: an emerging transdisciplinary and interdisciplinary field. Gut 2011, 60: 397e411.

Ogino S, Kawasaki T, Nosho K, et al: LINE-1 hypomethylation is inversely associated with microsatellite instability and CpG island methylator phenotype in colorectal cancer. Int J Cancer 2008, 122: 2767–2773.

O'Konek JJ, Boucher PD, Iacco AA, et al: MLH1 deficiency enhances tumor cell sensitivity to ganciclovir. Cancer gene therapy 2009, 16: 683-692.

Organisation for Economic Co-operation and Development (OECD): Health at a glance: Europe 2010. Paris: OECD, 2010.

Organisation for Economic Co-operation and Development (OECD): Tackling Harmful Alcohol Use Country note - Czech Republic 2012 [Internet]. OECD [cited 2016 Feb. 12]. Available from: http://www.oecd.org/czech/Tackling-Harmful-Alcohol-Use-Czech-Republic-en.pdf.

Pande M, Lynch PM, Hopper JL, et al: Smoking and colorectal cancer in Lynch syndrome: results from the Colon Cancer Family Registry and the University of Texas M.D. Anderson Cancer Center. Clin Cancer Res 2010, 16: 1331–1339.

Parc YR, Halling KC, Wang L, et al: hMSH6 Alterations in Patients with Microsatellite Instability-Low Colorectal Cancer. Cancer Res 2000, 60: 2225.

Pardini B, Fabio Rosa F, Alessio Naccarati A, et al: Polymorphisms in microRNA genes as predictors of clinical outcomes in colorectal cancer patients. Carcinogenesis 2015, 36: 82–86.

Park JY, Chen L, Elahi A, et al: Genetic analysis of microsomal epoxide hydrolase gene and its association with lung cancer risk.Eur J Cancer Prev 2005, 14: 223–230.

Park SH, Lee GY, Jeon HS, et al: -93G-A polymorphism of hMLH1 and risk of primary lung cancer. Int J Cancer 2004, 112: 678–682.

Pastorelli R, Guanci M, Cerri A, et al: Impact of inherited polymorphisms in glutathione S-transferase M1, microsomal epoxide hydrolase, cytochrome P450 enzymes on DNA, and blood protein adducts of benzo(a)pyrene-diolepoxide. Cancer Epidemiol Biomarkers Prev 1998, 7: 703–709.

Paz-y-Mino C, Perez JC, Fiallo BF, Leone PE: A polymorphism in the hMSH2 gene (gIVS12-6T>C) associated with non-Hodgkin lymphomas. Cancer Genet Cytogenet 2002,133: 29–33.

Peltomäki P: Role of DNA mismatch repair defects in the pathogenesis of human cancer. J Clin Oncol 2003, 21: 1174-1179.

Peluso MEM, Munnia A, Srivatanakul P, et al: DNA Adducts and Combinations of Multiple Lung Cancer At-Risk Alleles in Environmentally Exposed and Smoking Subjects. Environ Molec Mut 2013, 54: 375-383.

Perše M: Oxidative stress in the pathogenesis of colorectal cancer: cause or consequence? BioMed Res Int, 2013.

Phillips DH, Castegnaro M: Standardization and validation of DNA adduct postlabelling methods: report of interlaboratory trials trials and production of recommended protocols. Mutagenesis 1999, 14: 301–315.

Phipps AI, Limburg PJ, Baron JA, et al: Association between molecular subtypes of colorectal cancer and patient survival. Gastroenterology 2015, 148: 77–87.

Picelli S, Zajac P, Zhou XL, et al: Common variants in human CRC genes as low-risk alleles. Eur J Cancer 2010, 46: 1041-1048.

Pino MS, Chung DC: Microsatellite instability in the management of colorectal cancer. Exp Review Gastroenterol Hepatol 2011, 5: 385-399.

Polakova V, Pardini B, Naccarati A, et al: Genotype and haplotype analysis of cell cycle genes in sporadic colorectal cancer in the Czech Republic. Human Mutat 2009, 30: 661–668.

Poplawski T, Zadrozny M, Kolacinska A, et al: Polymorphisms of the DNA mismatch repair gene HMSH2 in breast cancer occurence and progression, Breast Cancer Res Treat 2005, 94: 199–204.

Pushparajah DS, Umachandran M, Plant KE, et al: Differential response of human and rat epoxide hydrolase to polycyclic aromatic hydrocarbon exposure: studies using precision-cut tissue slices. Mutat Res 2008, 640: 153–161.

Raptis S, Mrkonjic M, Green RC, et al: MLH1-93G >A promoter polymorphism and the risk of microsatellite-unstable colorectal cancer. J Natl Cancer Inst 2007, 99: 463–474.

Ravegnini G, Sammarini G, Hrelia P, Angelini S: Key Genetic and Epigenetic Mechanisms in Chemical Carcinogenesis. Toxicol Sci 2015, 148: 2-13.

Reimers MS, Zeestraten ECM, Kuppen PJK, et al: Biomarkers in precision therapy in colorectal cancer. Gastroenterology Rep 2013, 1: 166–183.

Ricceri F, Matullo G, Vineis P: Is there evidence of involvement of DNA repair polymorphisms

in human cancer? Mut Res 2012, 736: 117–121.

Roberts R, Wells GA, Stewart AF, et al: The genome-wide association study - a new era for common polygenic disorders. J Cardiovasc Transl Res 2010, 3: 173-182.

Rossner Jr P, Sram RJ (2014). Environmental Pollution and Health Consequences. In Studies on Pediatric Disorders 2014 (pp. 283-299). Springer New York.

Rossner P, H. Bavorova H, Ocadlikova D, et al: Chromosomal aberrations in peripheral lymphocytes of children as biomarkers of environmental exposure and life style. Toxicol Lett 2002, 134: 79–85.

Rossner P, Rossnerova A, Sram RJ: Oxidative stress and chromosomal aberrations in an environmentally exposed population. Mutat Res 2011, 707: 34-41.

Rossnerova A, Spatova M, Rossner P, et al: Factors affecting the frequency of micronuclei in asthmatic and healthy children from Ostrava. Mutat Res 2011, 708: 44-49.

Rubes J, Kucharova S, Vozdova M, et al: Cytogenetic analysis of peripheral lymphocytes in medical personnel by means of FISH. Mutat Res 1998, 412: 293–298.

Salama SA, Sierra-Torres CH, Oh HY, et al: Variant metabolizing gene alleles determine the genotoxicity of benzo[a]pyrene. Environ Mol Mutagen 2001, 37: 17–26.

Samowitz WS, Curtin K, Wolff RK, et al: The MLH1 -93 G>A promoter polymorphism and genetic and epigenetic alterations in colon cancer. Genes Chromosomes Cancer 2008, 47: 835-844.

Sanli I, Feil R: Chromatin mechanisms in the developmental control of imprinted gene expression. Int J Biochem Cell Biol 2015.

Schafmayer C, Buch S, Egberts JH, et al: Genetic investigation of DNA-repair pathway genes *PMS2*, *MLH1*, *MSH2*, *MSH6*, *MUTYH*, *OGG1* and *MTH1* in sporadic colon cancer. Int J Cancer 2007, 121: 555–558.

Schetter AJ, Okayama H, Harris CC: The role of microRNAs in colorectal cancer. Cancer Journal (Sudbury, Mass.) 2012, 18: 244.

Schmoll HJ, Van Cutsem E, Stein A, et al: ESMO Consensus Guidelines for management of patients with colon and rectal cancer. A personalized approach to clinical decision making. Ann Oncol 2012, 23: 2479–2516.

Schmutte C, Sadoff MM, Shim K-S, et al: The interaction of DNA mismatch repair proteins with human exonuclease I. J Biol Chem 2001, 276: 33011–33018.

Schofield MJ, Hsieh P: DNA mismatch repair: molecular mechanisms and biological function. Annu Rev Microbiol 2003, 57: 579-608.

Schumacher B, Garinis GA, Hoeijmakers JHJ: Age to survive: DNA damage and aging. Trends Gen 2008, 24: 77–85.

Shilpa V, Lakshmi K: Mechanisms of mismatch repair genes in cancer – a brief review. J Proteom Genom 2014, 1: 101.

Shin KH, Shin JH, Kim JH, Park JG, et al: Mutational analysis of promoters of mismatch repair genes hMSH2 and hMLH1 in hereditary nonpolyposis colorectal cancer and early onset colorectal cancer patients: identification of three novel germ-line mutations in promoter of the hMSH2 gene. Cancer Res 2002, 62: 38–42.

Skjelbred CF, Svendsen M, Haugan V, et al: Influence of GSTM1, GSTT1, GSTP1, NAT1,NAT2, EPHX1, MTR and MTHFR polymorphism on chromosomal aberration frequencies in human lymphocytes. Carcinogenesis 2011, 32: 399-405.

Slyskova J, Korenkova V, Collins AR, et al: Functional, Genetic, and Epigenetic Aspects of Base and Nucleotide Excision Repair in Colorectal Carcinomas. Clin Cancer Res 2012a; 18: 5878–5887.

Slyskova J, Lorenzo Y, Karlsen A, et al: Both genetic and dietary factors underlie individual differences in DNA damage levels and DNA repair capacity. DNA Repair 2014, 16: 66-73.

Slyskova J, Naccarati N, Pardini B, et al: Differences in nucleotide excision repair capacity between newly diagnosed colorectal cancer patients and healthy controls. Mutagenesis 2012b, 27: 225–232.

Snyder MW, Adey A, Kitzman JO, Shendure J: Haplotype-resolved genome sequencing: experimental methods and applications. Nat Rev Genet 2015, 16: 344-358.

Songserm N, Promthet S, Pientong C, et al: Gene–environment interaction involved in cholangiocarcinoma in the Thai population: polymorphisms of DNA repair genes, smoking and use of alcohol. BMJ Open 2014, 4: e005447.

Sovinová, H., Sadílek, P., & Csémy, L. (2010). Vývoj prevalence kuřáctví v dospělé populaci ČR. Názory a postoje občanů ČR k problematice kouření.(Období 1997–2009): výzkumná zpráva [Prevalence of smoking in the adult population of the Czech Republic. Opinions and attitudes of Czech citizens on the issue of smoking.(1997–2009): Research Report]. Státní zdravotní ústav 2010. *Státní zdravotní ústav, Prague*.

Sram RJ, Beskid O, Binkova B, et al: Chromosomal aberrations in environmentally exposed population in relation to metabolic and DNA repair genes polymorphisms. Mutat Res 2007, 620: 22-33.

Sram RJ, Binkova B, Beskid O, et al: Biomarkers of exposure and effect--interpretation in human risk assessment. Air Quality, Atmosphere, & Health 2011, 4: 161-167.

Statistics, F. A. P. (2013). Food and Agriculture Organization of the United Nations. 2013. Cyprinus carpio.

Stigliano V, Sanchez-Mete L, Martayan A, Anti M: Early-onset colorectal cancer: a sporadic or inherited disease? World J Gastroenterol 2014, 20: 12420-12430.

Stranger BE, Montgomery SB, Dimas AS, et al: Patterns of cis regulatory variation in diverse human populations. PLoS Genet 2012, 8: e1002639.

Strimpakos AS, Syrigos KN, Saif MW: Pharmacogenetics and biomarkers in colorectal cancer. Pharmacogenom J 2009, 9: 147-160.

Suchánek Š, Vepřeková G, Májek O, et al: Epidemiologie, etiologie, screening a diagnostika kolorektálního karcinomu, včetně diagnosticko-terapeutických zákroků na tlustém střevě. Onkologie 2011, 5: 261-265.

Sur I, Tuupanen S, Whitington T, et al: Lessons from functional analysis of genome-wide association studies. Cancer Res 2013, 73: 4180-4184.

Taioli E, Sram RJ, Binkova B, et al: Biomarkers of exposure to carcinogenic PAHs and their relationship with environmental factors. Mutat Res 2007, 620: 16-21.

Tan H, Wang Q, Wang A, et al: Influence of GSTs, CYP2E1 and mEH polymorphisms on 1, 3-butadiene-induced micronucleus frequency in Chinese workers. Toxic App Pharmacol 2010, 247: 198–203.

Tay YD, Sidebotham JM, Wu L: Mph1 requires mismatch repair-independent and -dependent functions of MutS alpha to regulate crossover formation during homologous recombination repair. Nucleic Acids Res 2010, 38: 1889-1901.

Teixeira MC, Braghiroli MI, Sabbaga J, Hoff PM. Primary prevention of colorectal cancer: myth or reality? World J Gastroenterol 2014, 20: 15060-15069.

The 1000 Genomes Project Consortium: A map of human genome variation from population-scale sequencing. Nature 2010, 467: 1061–1073.

The International HapMap Consortium A haplotype map of the human genome. *Nature* 2005, 437: 1299-1320.

Tomei F, Rosati MV, Baccolo TP, et al: Plasma concentration of adrenocorticotropic hormone in

traffic policemen. J Occupat Health 2003, 45: 242-247.

Tomlinson I, Webb E, Carvajal-Carmona L, et al: A genome-wide association study identifies colorectal cancer susceptibility loci on chromosomes 10p14, 8q23.3. Nat Genet 2008, 40: 623–630 (The CORGI Consortium).

Topinka J, Rossner Jr. P, Milcova A, et al: DNA adducts and oxidative DNA damage induced by organic extracts from PM2.5 in an acellular assay. Toxicol Lett 2011, 202: 186–192.

Tournier I, Vezain M, Martins A, et al: A large fraction of unclassified variants of the mismatch repair genes MLH1 and MSH2 is associated with splicing defects. Human Mut 2008, 29: 1412-1424.

Tseng-Rogenski SS, Hamaya Y, Choi DY, Carethers JM: Interleukin 6 Alters Localization of hMSH3, Leading to DNA Mismatch Repair Defects in Colorectal Cancer Cells. Gastroenterology 2015, 148: 579–589.

Václavíková R, Hughes DJ, Souček P: Microsomal epoxide hydrolase 1 (EPHX1): Gene, structure, function, and role in human disease. Gene 2015, 571: 1–8.

Valeri N, Gasparini P, Fabbri M, et al. Modulation of mismatch repair and genomic stability by miR-155. Proc Natl Acad Sci USA 2010, 107: 6982-6987.

Valle L: Genetic predisposition to colorectal cancer: where we stand and future perspectives. World J Gastroenterol 2014, 20: 9828-9849.

Valo S, Kaur S, Ristimaki A, et al: DNA hypermethylation appears early and shows increased frequency with dysplasia in Lynch syndrome-associated colorectal adenomas and carcinomas. Clin Epigenetics 2015, 7:71.

Vodicka P, Caja F, Vymetalkova V, et al: A novel c. 204 Ile68Met germline variant in exon 2 of the mutL homolog 1 gene in a colorectal cancer patient. Oncol Lett 2015a, 9: 183-186.

Vodicka P, Koskinen M, Stetina R, et al: The role of various biomarkers in the evaluation of styrene genotoxicity. Cancer Detect Prevent 2003, 27: 275–284.

Vodicka P, Kumar R, Stetina R, et al: Markers of individual susceptibility and DNA repair rate in workers exposed to xenobiotics in a tire plant. Environ Molec Mutagenesis 2004, 44: 283-292.

Vodicka P, Musak L, Frank C, et al: Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects. Carcinogenesis 2015b, 36: 1299-1306.

Vodicka P, Polivkova Z, Sytarova S, et al: Chromosomal damage in peripheral blood

lymphocytes of newly diagnosed cancer patients and healthy controls. Carcinogenesis 2010, 31: 1238–1241.

Vodicka P, Soucek P, Tates AD, et al: Association between genetic polymorphisms and biomarkers in styrene-exposed workers. Mutat Res 2001; 482: 89–103.

Vodicka P, Stetina R, Polakova V, et al: Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects. Carcinogenesis 2007, 28: 657-64.

Vogelsang M, Paccez JD, Schafer G et al: Aberrant methylation of the *MSH3* promoter and distal enhancer in esophageal cancer patients exposed to first‐hand tobacco smoke. J Cancer Res Clin Oncol 2014, 140: 1825–1833.

Vogelsang M, Wang Y, Veber Net al: The Cumulative Effects of Polymorphisms in the DNA Mismatch Repair Genes and Tobacco Smoking in Oesophageal Cancer Risk. PLoS ONE 2012, 7: e36962.

Vymetalkova V, Pardini B, Rosa F, et al: Variations in mismatch repair genes and colorectal cancer risk and clinical outcome. Mutagenesis 2014a, 29: 259-265.

Vymetalkova VP, Slyskova J, Korenkova V, et al: Molecular characteristics of mismatch repair genes in sporadic colorectal tumors in Czech patients. BMC Med Genet 2014b, 15: 1.

Walther A, Houlston R, Tomlinson I: Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis. Gut 2008; 57:941–950.

Wang D, Zhou J, Wang T et al: Polymorphisms in *MSH2* gene and risk of gastric cancer, and interactions with lifestyle factors in a Chinese population. Cancer Epidemiology 2012a, 36: e171–e176.

Wang T, Liu Y, Sima L, et al: Association between MLH1 -93G.A Polymorphism and Risk of Colorectal Cancer. PLoS ONE 2012b, 7: e50449.

Wang W, Wang M, Chen Y, et al: The hOGG1 Ser326Cys polymorphism contributes to cancer susceptibility: evidence from 83 case–control studies. Mutagenesis 2012c, 27: 329-336.

Welter D, MacArthur J, Morales J, et al: The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucl Acid Res 2014, 42: D1001-D1006.

Win AK, Hopper JL, Buchanan DD, et al: Are the common genetic variants associated with colorectal cancer risk for DNA mismatch repair gene mutation carriers? Eur J Cancer 2013a, 49: 1578-1587.

Win AK, Macinnis RJ, Dowty JG, Jenkins MA: Criteria and prediction models for mismatch repair gene mutations: a review. J Med Genet 2013b, 50: 785-793.

World Health Organization (WHO): Global status report on alcohol and health 2014 [Internet]. WHO [cited 2015 Oct. 12]. Available from: http://www.who.int/substance_abuse/publications/global_alcohol_report/en/

Worrillow LJ, Travis LB, Smith AG, et al: An intron splice acceptor polymorphism in hMSH2 and risk of leukemia after treatment with chemotherapeutic alkylating agents. Clin Cancer Res 2003, 9: 3012-3020.

Wu S, Rhee KJ, Albesiano E, et al: A human colonic commensal promotes colon tumorigenesis via activation of T helper type 17 T cell responses. Nature Med 2009, 15: 1016-1022.

Xing C, Chen Q, Li G, et al: Microsomal Epoxide Hydrolase (EPHX1) Polymorphisms are Associated with Aberrant Promoter Methylation of ERCC3 and Hematotoxicity in Benzene-Exposed Workers. Environ Molec Mutagenesis 2013, 54: 397-405.

Xu J-L, Yin Z-Q, Ming-De Huang M-D, et al: MLH1 polymorphisms and cancer risk: a meta-analysis based on 33 Case-control studies. Asian Pacific J Cancer Prev 2012, 13: 901-907.

Yamamoto H, Imai K: Microsatellite instability: an update. Arch Toxicol 2015, 89: 899-921.

Yang TP, Beazley C, Montgomery SB, et al. (2010) Genevar: a database and Java application for the analysis and visualization of SNP-gene associations in eQTL studies. Bioinformatics 26: 2474-2476.

Yu JH, Bigler J, Whitton J, Potter JD, Ulrich CM: Mismatch repair polymorphisms and colorectal polyps: hMLH1-93G >A variant modifies risk associated with smoking. Am J Gastroenterol 2006, 101: 1313–1319.

Yuan ZQ, Gottlieb B, Beitel LK, et al: Polymorphisms and HNPCC: PMS2–MLH1 protein interactions diminished by single nucleotide polymorphisms. Hum Mutat 2002, 19: 108–113.

Zhang K, Civan J, Mukherjee S, et al: Genetic variations in colorectal cancer risk and clinical outcome. World J Gastroenterol 2014a, 20: 4167-4177.

Zhang S-H, Wang L-A, Li Z, et al: APE1 polymorphisms are associated with colorectal cancer susceptibility in Chinese Hans. World J Gastroenterol 2014b; 20: 8700-8708.

Zhi W, Xue B, Wang L, et al: The MLH1 2101C>A (Q701K) variant increases the risk of gastric cancer in Chinese males. BMC Gastroenterology 2011, 11: 133.

Zhong X, Arita M, Yamada K, et al A single nucleotide substitution (-107C-->G) in the hMLH1 promoter found in colorectal cancer population reduces transcriptional activity. Biochem Genet 2007, 45: 671-81.

Zhu H, Li X, Zhang X, et al: Polymorphisms in mismatch repair genes are associated with risk and microsatellite instability of gastric cancer, and interact with life exposures. Gene 2015, 579:1.

8. APPENDICES 1-5

Appendix 1:

Tulupova E, Kumar R, Hanova M, Slyskova J, Pardini B, Polakova V, Naccarati A, Vodickova L, Novotny J, Halamkova J, Hemminki K, Vodicka P. Do polymorphisms and haplotypes of mismatch repair genes modulate risk of sporadic colorectal cancer? *Mutation Research.* 2008, 648(1-2):40-5.

Appendix 2:

Pardini B, Naccarati A, Novotny J, Smerhovsky Z, Vodickova L, Polakova V, Hanova M, Slyskova J, **Tulupova E**, Kumar R, Bortlik M, Barale R, Hemminki K, Vodicka P. DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic. *Mutation Research.* 2008, 638(1-2):146-53.

Appendix 3:

Vodicka P, Stetina R, Polakova V, **Tulupova E**, Naccarati A, Vodickova L, Kumar R, Hanova M, Pardini B, Slyskova J, Musak L, De Palma G, Soucek P, Hemminki K. Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects. *Carcinogenesis.* 2007, 28(3):657-64.

Appendix 4:

Sram RJ, Beskid O, Binkova B, Chvatalova I, Lnenickova Z, Milcova A, Solansky I, **Tulupova E**, Bavorova H, Ocadlikova D, Farmer PB. Chromosomal aberrations in environmentally exposed population in relation to metabolic and DNA repair genes polymorphisms. *Mutation Research.* 2007, 620(1-2):22-33.

Appendix 5:

Binkova B, Chvatalova I, Lnenickova Z, Milcova A, **Tulupova E**, Farmer PB, Sram RJ. PAH-DNA adducts in environmentally exposed population in relation to metabolic and DNA repair gene polymorphisms. *Mutation Research.* 2007, 620(1-2):49-61.