BIOCHEMICAL EFFECTS OF INHERITED MMR GENE MUTATIONS AND DIET ON COLON CANCER RISK

DENIS ĐERMADI BEBEK

Division of Genetics
Department of Biosciences
Faculty of Biological and Environmental Sciences
Doctoral School in Health Sciences
Integrative Life Science Doctoral Program (ILS)
University of Helsinki

To be presented for public examination with the permission of the Faculty of Biological and Environmental Sciences of the University of Helsinki in the Main Building, Auditorium 14, Fabianinkatu 33, on the 31st of October 2014 at 12.00 p.m.

Supervisor

Professor Minna Nyström
Division of Genetics
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Thesis committee

Docent Pekka Heino
Division of Genetics
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Professor Juha Partanen
Division of Genetics
Department of Biosciences
Faculty of Biological and Environmental Sciences
University of Helsinki, Finland

Reviewers

Docent Anne-Maria Pajari
Department of Food and Environmental Sciences
Faculty of Agriculture and Forestry
University of Helsinki, Finland

Docent Jukka Partanen
Research Director
Finnish Red Cross Blod Service, Finland

Opponent

Professor Riccardo Fodde
Department of Pathology
Erasmus MC, Netherlands

Cover layout by Anita Tienhaara
Cover image: Mouse goes Pop Art by *Denis Đermadi Bebek*

ISBN 978-951-51-0222-5 (paperback)
ISBN 978-951-51-0223-2 (PDF)
ISSN 2342-3161 (paperback)
ISSN 2342-317X (PDF)
Electronic version: https://ethesis.helsinki.fi/en

Helsinki 2014, Hansaprint

To both of you who were stable pillars of my childhood and youth; both victims of cancers - to my grandpa and grandma

Za dedu i baku, koji ste bili stabilni oslonac u djetinjstvu i mladosti; oboje žrtve raka

“I am a deeply superficial person.”

— Andy Warhol

CONTENTS

LIST OF ORIGINAL PUBLICATIONS **i**

ABBREVIATIONS **ii**

ABSTRACT **9**

INTRODUCTION **11**

REVIEW OF THE LITERATURE **13**

CANCER 13

Oncogenes and tumor suppressor genes 13

Misdirected epigenetic modifications in cancer 14

Colorectal cancer (CRC) 15

RISK FACTORS IN CRC 18

Inherited predisposition to CRC 18

Lynch syndrome (LS) *19*

Mismatch repair mechanism *21*

Structure and function of MutSα 23

Functional analysis of non-truncating MMR variants *25*

Western diet as a CRC risk factor 25

METABOLIC STRESS AND CANCER 27

MOUSE MODELS IN CRC STUDIES 29

MMR deficient mouse models 29

Mouse as a model for diet effects in the intestine 31

AIMS OF THE PRESENT STUDY **32**

MATERIALS AND METHODS **33**

MSH2 VARIANTS (I-II) 33

PROTEIN PRODUCTION AND PURIFICATION (I-II) 35

MutSα production 35

Viral amplification *35*

Infection of the Sf9 cells for protein production *35*

Protein extraction *35*

Ni-NTA purification *35*

BANDSHIFT ASSAY (I-II) 36

Construction of labeled homo- and heteroduplex oligomers 36

Mismatch binding and release assays 37

MOUSE MODELS AND DIETS (III) 37

Diets 37

Feeding experiment 38

Histological studies 38

QUANTITATIVE ANALYSIS OF PROTEOMES (III) 38

Sample preparation from mouse colonic mucosa 38

Protein extraction from mucosa samples 39

2D Difference gel electrophoresis (DIGE) 39

Protein identification 40

WESTERN BLOT ANALYSIS (I-III) 41

Expression levels of purified MutSα variants (I-II) 41

Expression levels of MLH1, SELENB1 and IDH1 in mouse samples (III) 41

RT-QPCR (III) 41

STATISTICAL ANALYSES (I-III) 42

GENE ONTOLOGY AND NETWORK ANALYSIS (III) 43

RESULTS & DISCUSSION **44**

BIOCHEMICAL ANALYSIS OF INHERITED *MSH2* VARIATIONS (I-II) 44

Variants in the connector or lever domain cause protein degradation 44

Variants in the ATPase domain impair DNA binding or ATP-mediated release of protein (I-II) 45

MSH2 variants p.A636P and p.C697F do not bind to a mismatch *45*

MSH2 variants p.G674A and p.E749K impair release from a mismatch *46*

Inconclusive MSH2 variants 48

Nonpathogenic MSH2 variants 48

PROTEOMIC ANALYSIS OF THE EFFECTS OF DIET ON CRC DEVELOPMENT (III) 49

Mice fed with Western diet show increased weight gain and trend in development of neoplasias 49

Differences in protein expression were mainly found between the *Mlh1*$^{+/-}$ WD* and AIN mice 50

Western diet affects proteins involved in energy metabolism and cellular toxicity 51

Western diet disrupts cellular ROS homeostasis *53*

Neutralization of ROS by increased expression of enzymes that generate NADPH *54*

mRNA expression of energy metabolism genes correlates with their protein expression *55*

Increased expression of SELENBP1 and LGALS4 – cellular response to neoplastic processes *55*

Unfolded protein response may be disrupted in mice fed with Western diet *56*

Expression changes of detected proteins and CRC tumorigenesis *57*

CONCLUSIONS **58**

FUTURE PROSPECTS **59**

ACKNOWLEDGEMENTS **60**

REFERENCES **67**

LIST OF ORIGINAL PUBLICATIONS

This thesis is based on the following publications:

I. OLLILA, S., **DERMADI BEBEK, D.**, GREENBLATT, M. & NYSTRÖM, M. 2008. Uncertain pathogenicity of MSH2 variants N127S and G322D challenges their classification. *Int J Cancer,* 123**,** 720-4.

II. OLLILA, S., **DERMADI BEBEK, D.**, JIRICNY, J. & NYSTRÖM, M. 2008. Mechanisms of pathogenicity in human MSH2 missense mutants. *Hum Mutat,* 29**,** 1355-63.

III. **ĐERMADI BEBEK, D.**, VALO, S., PUSSILA, M., REYHANI, N., SARANTAUS, L., LALOWSKI, M., BAUMANN, M. & NYSTRÖM, M. 2014. Inherited cancer predisposition sensitizes colonic mucosa to address Western diet effects and putative cancer-predisposing changes on mouse proteome. *J Nutr Biochem,* 2014 Jul 17. pii: S0955-2863(14)00134-X. doi: 10.1016/j.jnutbio.2014.06.002.

In the publications I-II D.Đ.B. conducted:

- expression of his-tagged recombinant proteins using *Sf9* cells
- extraction of proteins (14 variants)
- checking of expression with Western blot
- purification of recombinant proteins in Ni-NTA system
- bandshift assay in order to study mismatch recognition and binding/releasing of recombinant proteins from DNA heteroduplex (mostly for variants I extracted).

D.Đ.B. helped in analysis of bandshift assays and with writing.

In the publication III D.Đ.B. conducted:

- design of the study
- all laboratory work (mice and diet work, selection of the main methodology used in this study 2D- difference gel electrophoresis (DIGE), set up of the 2D-DIGE, mass spectrometry MALDI for identification of the proteins, statistical and network analysis (DeCyder, SPSS, GeneMANIA, HIPPIE), additional RT-qPCR and Western blot analysis)

D.Đ.B. wrote the paper.

Chapter: *Unfolded protein response may be disrupted in mice fed with Western diet* is not included in any of the listed publications.

The publications are referred to in the text by their roman numerals.

ABBREVIATIONS

ADP – adenosine diphosphate
AIN – AIN-93G American Institute of Nutrition, purified diet for laboratory rodents
ATP – adenosine triphosphate
APC – adenomatous polyposis coli
CRC – colorectal cancer
C-terminus – Carboxy terminus
2D DIGE – 2 dimensional differential gel electrophoresis
E. coli – *Escherichia coli*
EXO1 – Exonuclease 1
FAO – fatty acid oxidation
FAP – familial adenomatous polyposis coli
GO – Gene Ontology
HNPCC – Hereditary nonpolyposis colorectal cancer
IDL – insertion / deletion loop
IHC – immunohistochemistry
InSiGHT – International Society for Gastrointestinal Hereditary Tumors
LOH – loss of heterozygosity
LS – Lynch syndrome
MLH1, 3 - MutL Homolog 1, 3
MMR – mismatch repair
MSH2, 3, 6 – MutS homolog 2, 3, 6
MSI – microsatellite instability
NADP – nicotinamide adenine dinucleotide phosphate
N-terminus – Amino terminus
OXPHOS – oxidative phosphorylation
PCNA – proliferating cell nuclear antigen
PMS1, 2 – human postmeiotic segregation increased homolog 1, 2
PMSF – phenylmethylsulfonyl fluoride
PPP – pentose phosphate pathway
RFC – replication factor C
ROS – reactive oxygen species
RPA – replication protein A
SDS-PAGE – sodium dodecyl sulphate – polyacrylamide gel electrophoresis
*Sf*9 – *Spodoptera frugiperda* 9
ssDNA – single stranded DNA
TCA – citric acid cycle
TSG – tumor suppressor gene
UPR – unfolded protein response
WD – Western-style diet
WD* – Western-style diet used in the study
Wnt – Wingless-Type MMTV Integration Site Family
WT – wild type

ABSTRACT

Colorectal cancer (CRC) is one of the leading causes of death in developed countries. Although, a small fraction of cancers are caused by inherited genetic predisposition most of the CRCs are sporadic. In CRC, cancer risk is associated with lifestyle factors and aging. Even in dominantly inherited CRC predisposition such as in Lynch syndrome (LS), which is linked to germline mutations in the mismatch repair (MMR) genes *MLH1, MSH2, MSH6* and *PMS2*, cancer develops as a result of accumulation of genetic and epigenetic changes. After diagnosing an LS family, to be able to offer contiguous pre-symptomatic surveillance and predictive gene counseling to mutation carriers in a family, the pathogenicity assessment of a mutation is needed. Dependent on the type and the site of a germline mutation, inherited cancer risk may vary from high to low and especially in the latter case cancer risk may be strongly affected by lifestyle factors such as diet. Epidemiological studies on humans and previous studies on mice have shown that especially a Western-style diet (WD) may predispose colon mucosa to CRC. However, the mechanisms, which mediate the effects of diet on tumorigenesis are largely unknown.

Since both genetic and lifestyle factors have been shown to predispose to cancer, this thesis analyzed biochemical defects caused by inherited MMR gene mutations and Western diet exposure. Different MMR gene mutations may compromise MMR function through various biochemical defects. Here, we studied 18 inherited non-truncating mutations in *MSH2*, the second most frequently mutated gene among Lynch syndrome patients. We assessed protein stability, DNA binding, and ATP mediated DNA release abilities of the MSH2 variants. The majority of variants in the amino terminal region including the connector and lever domains p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p.C333Y, p.D603N) affected protein stability. Variations in the ATPase domain (p.A636P, p.G674A, p.C697F, p.I745-I746del, p.E749K) totally abolished either mismatch binding or release. Four protein variants (p.T33P, p.A272 V, p.G322D, p.V923E) expressed slightly reduced mismatch binding and/or release efficiencies compared to wild-type (WT) MSH2 protein, while two variants (p.N127S, p.A834T) were indistinguishable from WT.

To define the effects of Western-style diet, we analyzed protein expression changes in histologically normal colon mucosa of wild type ($Mlh1^{+/+}$) and CRC predisposed mice ($Mlh1^{+/-}$) after a long-term feeding experiment with WD and AIN-93G control diet. Using network analysis and data mining we also determined which of the affected proteins might be putative players in early CRC development. Our results pinpoint changes in a complex protein interaction network involved in ATP synthesis coupled proton transport, oxidoreduction coenzyme and nicotinamide nucleotide metabolic processes, which are important in the generation of reactive oxygen species (ROS) and cellular protection against ROS toxicity. Additionally, we detected SELENBP1 and LGALS4, which are implied in neoplastic processes.

Our studies show that mutations in the MMR gene affect the biochemistry of MMR, can have an effect on the phenotype of the mutation carriers and in the latest study suggest that the high sensitivity to Western diet may be linked to haplo-insufficiency caused by a loss of function mutation in the *Mlh1*$^{+/-}$ mice.

INTRODUCTION

Cancer is one of the leading causes of deaths worldwide (www.who.int), and colorectal cancer (CRC), which has a five-year prevalence of 8.9% in the United States and 14.3% in Europe (Finland 10.7%) (http://globocan.iarc.fr) is the third most common diagnosed cancer (www.cancer.org). The lifetime risk of developing CRC is about 5% in the US and 6% in Europe, indicating the need to characterize the most notable risk factors.

One of the most common inherited CRCs is Lynch syndrome (LS), previously known as hereditary non-polyposis colorectal cancer (HNPCC) syndrome, which accounts for 2 - 4% of all CRC cases (Lynch & de la Chapelle, 2003). The susceptibility is caused by germline mutations in the genes *MLH1, MSH2, MSH6* and *PMS2*, which encode proteins responsible for DNA mismatch repair (MMR) (Peltomaki, 2005). LS related cancers (1) are inherited in an autosomal dominant manner, (2) have an earlier average age of onset of colorectal cancer than in the general population (45 years in Lynch syndrome v. 65 years in the general population), (3) are mostly found in the proximal colon, with an increased risk of malignant disease at certain extra-colonic sites and (4) have accelerated carcinogenesis. LS cancers express high microsatellite instability (MSI) due to MMR malfunction (Aaltonen *et al.*, 1993), which leads to an increased mutation rate. Moreover, MSI is found in 10 - 15% of sporadic CRCs. Hypermethylation of the promoter region and subsequent inactivation of *MLH1* leads to MMR malfunction and causes the MSI phenotype in sporadic cancers (Hawkins & Ward, 2001).

Diagnosis of Lynch syndrome specifically relies on familial genetic background and early age of cancer onset (Vasen *et al.*, 1999). After diagnosing an LS family contiguous pre-symptomatic surveillance and predictive gene counseling can be offered to its mutation carriers. Unfortunately, a drawback of the international criteria is that small LS families or LS patients without family history of cancers are not easily diagnosed. Moreover, MSI analysis and/or expression studies of MMR proteins by immunohistochemical (IHC) staining may indicate an MMR defect in a tumor but, without finding and assessing a pathogenic germline MMR gene mutation in a patient, are not enough to confirm LS (Weissman *et al.*, 2012). Many missense mutations, which cause only a single amino acid substitution in a protein, do not always segregate with a cancer phenotype in a family and the interpretation of their pathogenicity is often difficult, complicating LS diagnoses and counseling. Currently variants are classified into five groups: pathogenic, probably pathogenic, no known pathogenicity, probably no pathogenicity and effect unknown (Kansikas, Kariola & Nystrom, 2011).

Effects of biochemical defects of non-truncating *MSH2* mutations found in putative LS families were studied by analyzing protein stability, protein DNA binding and ATP mediated DNA release ability. Our previous findings showed that most studied mutations abolished MMR function in the *in vitro* MMR assay, but mechanisms for MMR malfunction remained unknown

(Ollila *et al.*, 2006). Here, we observed that mutations affecting a certain biochemical property tended to cluster together in specific domains of the MSH2 protein, suggesting that pathogenic mutations may be clustered in specific protein regions essential for MMR activity. Additionally, specific functional assays showed usefulness in cases where the variant was clinically associated with cancer, was MMR proficient but in specific functional assays expressed malfunction.

Studies on inherited cancers, e.g. LS, are important since they may clarify molecular mechanisms associated with tumorigenesis in general. Nevertheless, cancers linked to a dominantly inherited predisposition are a minority in the whole CRC burden and the etiology in most cancers seem to be strongly associated with lifestyle factors and aging (Bingham, 1996, Lipkin *et al.*, 1999, Bingham & Riboli, 2004, Fearon, 2011). Western style diet (WD), energy high, rich in fat and scarce in fiber, calcium, vitamin D and folate is one of the lifestyle factors shown to promote cancer in animals (Lipkin *et al.*, 1999, Newmark *et al.*, 2001).

Here we studied dietary effects on proteome expression in normal colonic mucosa of *Mlh1*$^{+/-}$ mice analogous to LS mutation carriers, and their wild type littermates. Remarkably, the statistically significant expression changes caused by WD occurred predominantly in mutation carriers and indicated metabolic disturbances and increased toxicity in the cells. However, our studies suggest that genetic predisposition and diet effects do not necessarily interact since we did not find any effect of Western diet on *Mlh1* expression in colonic mucosa. These two risk factors may rather drive cellular processes towards a common goal and cause changes, which will eventually push cellular homeostasis over the brink towards tumorigenesis.

REVIEW OF THE LITERATURE

CANCER

Cancer is a group of more than 100 different diseases. Nevertheless they share common features of deregulated cell growth and homeostasis. Transformation of the normal cells to cancerous ones can be described as a multi-step acquirement of several alterations in cell physiology that lead to malignant growth. Alterations that enable normal cells to turn into cancerous cells are sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, activating invasion and metastasis, reprogramming energy metabolism, evading immune destruction, tumor-promoting inflammation and genome instability and mutations. (Hanahan & Weinberg, 2011)

Further support for Hanahan's & Weinberg's list of the ten hallmarks of cancer has come from work done by the Cancer Genome Atlas consortium and an enormous sequencing project, which included various cancers (Vogelstein *et al.*, 2013). Cancer genomes harbor hundreds of mutations in different genes and seem to be highly heterogeneous. However, even phenotypically different cancers were found to resemble each other when the mutated genes found in different cancers were organized according to the molecular pathways they are involved in. For example, one cancer can have a mutation in *K-Ras*, another cancer in *Mek*, and third in *Erk*. Although, genetic changes in them are then different they all have a deregulated Ras-Mek-Erk pathway. Each separate cancer has an average of 12 different deregulated pathways (DNA damage control, TGF-β, MAPK, STAT, PI3K, RAS, cell cycle and apoptosis, NOTCH, Hedgehog, APC, chromatin modification, and transcriptional regulation) that regulate three cellular processes - cell fate, cell survival and genome maintenance (Vogelstein *et al.*, 2013). Information from cancer genome studies recognizes cancer as a *pathway disease* rather than a mere genetic disease.

Oncogenes and tumor suppressor genes

Most mutated genes in cancer can be classified as (proto-) oncogenes or tumor suppressor genes (TSG). Proto-oncogenes regulate cell differentiation and cell cycle progression through cellular signal transduction and mitogen-regulated pathways. When constitutively activated, proto-oncogenes become oncogenes and turn normal cells into rapidly dividing (pre)-cancerous cells. Mechanisms of constitutive activation are mutations, over-expression and chromosomal translocations. Mutations in a gene can affect the active site of the protein leading to permanent activation of cell division. On the other hand, mutations in the promoter region can increase gene expression, which causes higher oncoprotein abundance. Gene duplications increase gene copy number and consequently protein expression is higher. Chromosomal translocations relocate proto-oncogenes to a new locus and may activate the expression of the proto-oncogene,

especially if translocated under the promoter of a constitutively expressed gene. Translocations may also result in fusion of proto-oncogenes and other genes forming a hybrid protein, which may exhibit highly oncogenic activity. Oncogenes can be classified as growth factors, kinases, regulatory GTPases and transcription factors. Growth factors induce cell proliferation (e.g. PDGF). Kinases can be split into three sub-groups, which are tyrosine kinases, found in the cellular membrane (e.g. EGFR, PDGFR, VEGFR), cytoplasmic tyrosine kinases, which mediate signals in cell proliferation, survival and migration (e.g. Src-family, BCR-ABL, BTK) and cytoplasmic serine/threonine kinases involved in cell cycle regulation, differentiation/proliferation and apoptosis (e.g. Raf kinase and cyclin-dependent kinases). Regulatory GTPases maintain signal transduction in the cytoplasm and are involved in cellular differentiation (e.g. RAS, RHO, RAB). Transcription factors include a large group of proteins, which regulate cell proliferation (e.g. MYC). (Hanahan & Weinberg, 2011)

While proto-oncogenes promote cellular growth, TSGs keep cellular growth in control. The first tumor suppressor gene was discovered by studying families with inherited eye cancers, retinoblastomas during the ‘70s. Individuals with the disease carry a germline mutation in one allele of the retinoblastoma (*Rb1*) gene, which encodes a protein that suppresses cell proliferation and growth. Somatically gained mutation or deletion inactivates the other allele of *Rb1* during childhood and starts tumorigenesis. The event of inactivation of the other allele was named loss of heterozygosity (LOH) and the mechanism according to its founder is called Knudson’s two hit hypothesis (Knudson, 1971, Berger, Knudson & Pandolfi, 2011). Tumor suppressors can be grouped according to their functions into repressors of cell cycle progression, so called *gatekeepers* (e.g. *RB1, CD2A1*), pro-apoptotic proteins, which direct cells toward apoptosis if DNA damage is irreparable (e.g. *BAX* family), negative regulators of growth-promoting pathways (e.g. *APC*), and genes involved in DNA maintenance, so called *caretakers*. Caretakers are a large group of tumor suppressors consisting of more than 100 genes involved in different DNA repair mechanisms (e.g. *MLH1, MSH2, MSH6, BRCA1/2, ATM*). (Hanahan & Weinberg, 2011) Defects in DNA repair mechanisms directly lead to genome instability that enables oncogene activation.

Misdirected epigenetic modifications in cancer

Cancer develops as a result of complex deregulation of oncogenes and TSG. In addition to genetic factors that affect gene regulation, another layer of complexity has been added when a growing body of evidence showed that inherited epigenetic changes can affect regulation of oncogenes and TSGs. Epigenetic means of affecting gene expression in cancer have been mainly investigated in respect of DNA and chromatin methylation and different histone modifications. The classic model of CRC development by Vogelstein has been revised and abnormal DNA methylation implemented as an early event in neoplastic changes. (Allis, Jenuwein & Reinberg,

2007)

The most common way of epigenetic regulation of gene expression is by DNA methylation of cytosine (C) residues in CpG islands. CpG islands are DNA regions at least 500 base pairs in length containing more than 55% cytosine and guanine (G) consecutively following each other with an observed-to-expected CpG ratio greater than 65% (Takai & Jones, 2002). CpG islands are found in 70% of the human gene promoters (Saxonov, Berg & Brutlag, 2006). Hypermethylation of cytosines causes inactivation of the promoter and hypomethylation can cause activation. Aberrant DNA methylation is associated with CRC development. Most cancers exhibit genome wide DNA hypomethylation (Feinberg & Vogelstein, 1983, Feinberg *et al.*, 1988, Makos *et al.*, 1992), which relaxes the DNA structure and may cause over-expression of oncogenes and genome instability. Interestingly, promoters of tumor suppressors are often aberrantly hypermethylated in cancer leading to gene inactivation (Jones & Laird, 1999). The genes *MLH1* and *APC* are commonly found to be inactivated in CRC by promoter hypermethylation (Thibodeau *et al.*, 1998), especially in familial syndromes, where sometimes the second hit occurs through hypermethylation (Nakagawa *et al.*, 2001, Suter *et al.*, 2003).

Aberrant DNA methylation is related to aging and implied to be an early event in CRC development (Issa, 1999). DNA methylation is a reversible change, which makes it interesting as a target for drug treatment. In addition, diet, particularly dietary folate, and energy metabolic pathways, which involve sirtuins, isocitrate dehydrogenases, synthesis of co-enzymes (e.g. 2-oxoglutarate and FAD) seem to be profoundly involved in DNA methylation, histone modifications (methylations and acetylations) and cancer development (Slattery *et al.*, 1999, Wellen *et al.*, 2009, Figueroa *et al.*, 2010, Lu *et al.*, 2012).

Colorectal cancer (CRC)

The colon mucosa consists of crypts, which are approximately 50 cells deep. Replicating stem cells on the bottom of the crypt migrate upwards while differentiating into epithelial cells. Cells at the top of the crypt trigger apoptosis, die and shred into the colon lumen (Fig. 1). During their lifetime, the majority of people will develop some neoplastic changes in the colon epithelial tissue, but only a few of those will directly lead to CRC. (Kinzler & Vogelstein, 1996, Kwong & Dove, 2009) The most common neoplastic changes are hyperplastic and inflammatory polyps, which are not directly considered pre-cancerous, but may indicate higher risk of developing adenomatous polyps (adenomas). Adenomas may over time change into cancer and they are considered potentially malignant.

If adenomatous cells accumulate multiple mutations in oncogenes or TSGs and invade the colon or rectum wall, they can access blood and lymph vessels. Once in the lymphatic sys-

tem, cancerous cells may disseminate firstly into nearby lymph nodes and later into liver forming metastases. (Weinberg, 2007)

To date, two hypotheses of CRC morphogenesis exist, *bottom-up* and *top-down*. The most widely accepted hypothesis, *bottom-up*, proposes that stem cells gain initial mutations, which lead to polyp formation at the top of the crypt. Polyps are exposed to the genotoxic surroundings of the colon content and can more easily accumulate further mutations and progress over time to CRC (Boman & Huang, 2008). This hypothesis is challenged by a *top-down* proposal that first mutations happen in the cells at the top of the crypt causing an adenomatous polyp, which advances toward the bottom of the crypt (Shih *et al.*, 2001, Kwong & Dove, 2009).

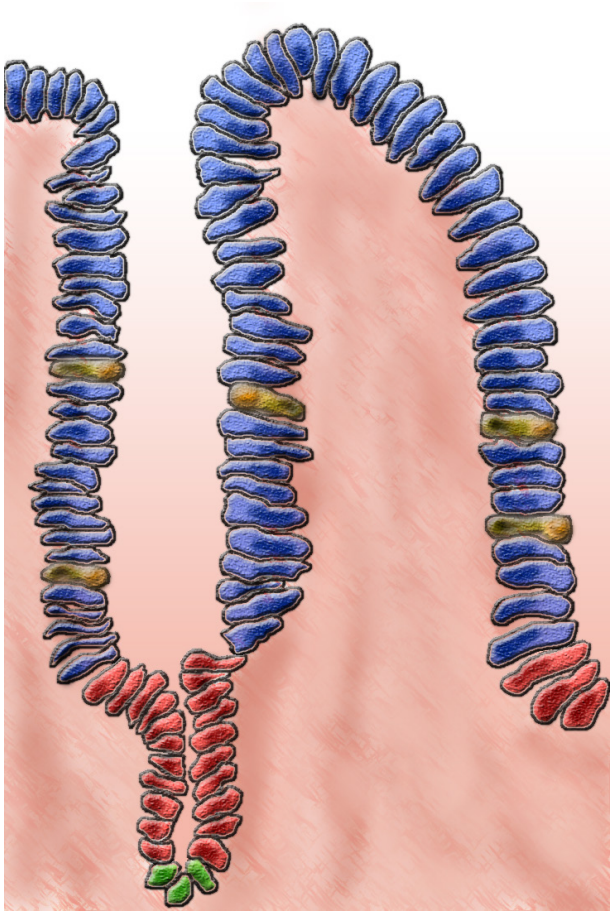

Fig. 1 Schematic structure of colon tissue shows organization of the epithelial cells (modified from Weinberg, 2007). Stem cells are at the bottom of the crypt: green, proliferative cells are found in the lower third of the crypt: red, and differentiated cells are in the remaining two thirds of the crypt: blue. Yellow cells are goblet cells. The cells at the top of the crypt undergo apoptosis and shred into the colon lumen.

At the beginning of '90s two famous CRC researchers, Fearon and Vogelstein proposed a widely accepted genetic pathway for CRC tumorigenesis (Fearon & Vogelstein, 1990), which was later revised (Kinzler & Vogelstein, 1997). Their model postulates that five to seven molecular genetic alterations are needed in a normal epithelial cell for development of carcinoma. These alterations are most commonly found in tumor suppressors such as *APC, p53* and *K-Ras* as well as genes on the chromosome 18q. In addition to genetic mutations, early epigenetic changes advance tumor progression through genome-wide loss of DNA methylation (hypomethylation) due to aging and focal gains of DNA methylation (hypermethylation). Hypomethylation of oncogene promoters can lead to biallelic expression and hypermethylation of TSG, which can lead to inactivation of genes with the ultimate goal of enabling cells to drive neoplastic processes (Fig. 2). (Allis, Jenuwein & Reinberg, 2007)

The TSG *APC* is mutated in ~80% of sporadic colon cancers (Powell *et al.*, 1992). It forms a protein complex with β-catenin preventing the translocation of β-catenin into the nucleus (Smits *et al.*, 1999). Normally, the cytoplasmic level of β-catenin is increased in the presence

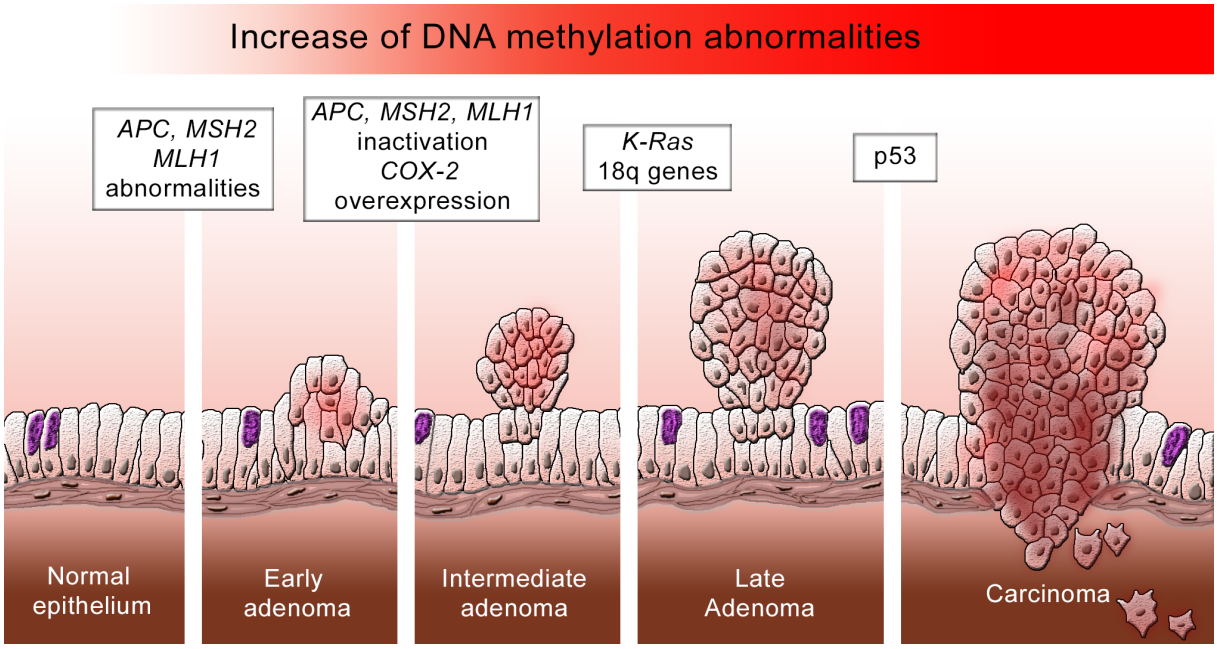

Fig. 2 Different stages of colorectal cancer (CRC) development and progression (modified from Allis, Jenuwein & Reinberg, 2007). Abnormalities in APC and mismatch repair (MMR) genes are early changes in normal epithelium leading to early adenoma. Increase of epigenetic abnormalities (increase of red in the bar) is postulated to be one of the first changes in CRC development. Complete inactivation of *APC* and/or MMR genes leads to intermediate adenoma, which later usually needs mutation in *K-Ras* or some of the genes at the 18q locus to develop into adenoma. Late adenoma develops into invasive carcinoma after obtaining mutation in *p53*.

of Wnt signal, which leads to β-catenin translocation into the nucleus and activation of several transcription factors. *APC* mutations can cause malfunctioning in APC signaling and prevent formation of the APC-β-catenin complex. Excess unbound β-catenin may trigger activation of the oncogenes *c-Myc* and cyclin D1 (He *et al.*, 1998, Tetsu & McCormick, 1999), which lead to uncontrolled cell cycle progression and proliferation if expressed constitutively.

The *p53* gene is a widely mutated tumor suppressor in several types of cancers. Most characterized mutations lead to loss of function or abnormal function of p53 (Cho & Vogelstein, 1992, Greenblatt *et al.*, 1994). It mediates the cellular response to DNA damage and prevents passing of mutations to daughter cells (Lane, 1992). Mutated p53 has also been associated with the transition of adenoma to carcinoma in CRC (Cho & Vogelstein, 1992).

Occurrence of LOH in the 18q chromosome has been commonly seen in CRC. LOH in this particular region increases in later developmental stages of CRC (it is found in 73% of carcinomas) (Cho & Vogelstein, 1992). In-depth analysis of the region that is lost in the 18q chromosome underlined tumor suppressors *SMAD4* and *SMAD2,* which are indeed frequently mutated in CRC (Thiagalingam *et al.*, 1996). SMAD4 and SMAD2 are involved in the transforming growth factor – beta (TGF-β) pathway and are responsible for cell proliferation, differentiation, apoptosis and adhesion (Xu *et al.*, 2000). Chromosomal deletion in 18q and mutations in

SMAD4 and *SMAD2* can compromise the TGF-β pathway effect on proliferation.

The *K-Ras* gene is a frequently mutated oncogene in CRC (30 - 50%) and mutations in it seem to be required quite early in adenoma-carcinoma transition. *K-Ras* codes a small 21 kDa GTPase protein, which transmits growth signals when GTP is bound to it. K-RAS is inactivated once GTP is hydrolyzed (Bos *et al.*, 1987). The majority of mutations affect the ability of K-RAS to hydrolyse GTP, which in turn causes a constitutive signal to cells for growth and proliferation (Vogelstein *et al.*, 1988).

RISK FACTORS IN CRC

Colorectal cancer and cancer in general, is considered a genetic disease. Moreover, scientific and epidemiological evidence support that diet, lifestyle and social practice substantially contribute to the development of cancer and CRC in particular. Some evidence of diet and other lifestyle effects on CRC comes from studies on the Japanese population. Incidence of CRC used to be very low in Japan before major changes in lifestyle, especially the westernization of diet, after which the incidence increased to the level in Western countries (Nagata *et al.*, 2001). Additional support to the effects of diet on CRC comes from studies on Japanese who migrated to the USA. These Japanese migrants developed similar risk to CRC as the native population (Haenszel, 1961, Buell & Dunn, 1965, Haenszel & Kurihara, 1968). Genetic factors and diet may have a pleiotropic effect on gastrointestinal tract, hormonal balance and epigenetic changes (Little, Horowitz & Feinle-Bisset, 2007). Different risk factors in CRC will be discussed in more detail later in this chapter.

Inherited predisposition to CRC

The majority of CRCs (~75%) develop sporadically in individuals without inherited susceptibility, while the rest are CRCs with a family history of CRCs. In about 10% of CRCs predisposition is caused by determined germline mutation, and these CRCs are considered hereditary/inherited. The major hereditary CRC syndromes are familial adenomatous polyposis (FAP), MYH-associated polyposis (MAP), Lynch syndrome (LS) and rare colon cancer syndromes. Although familial colon cancer syndromes are quite rare, they have greatly helped to understand cellular, biochemical and genetic mechanisms of CRC.

FAP is a rare autosomal dominant syndrome caused by germline mutations in the *APC* gene, which cause truncated APC with abnormal function (Bulow, 1987). Clinical characteristics of this disease are early age of onset and development of numerous polyps. Polyps can already occur in childhood and if left untreated some of these polyps will develop into malignant cancer. In addition to the inherited mutation a somatic mutation or LOH is needed for cancer develop-

ment. More than 1191 different variants of the *APC* gene have been reported (http://www.insight-group.org). Predominantly, variants are deletions and substitutions that lead to frameshifts and nonsense variants of the protein. The most common *APC* mutation found in 10% of FAP patients is a deletion of AAAAG in codon 1309 (Yan *et al.*, 2002). The position of the germline mutation in *APC* determines whether the second hit is a truncating mutation or loss of the second allele (Lamlum *et al.*, 1999).

In 1913 famous scientist Alderd S. Warthin described another highly penetrant syndrome as a hereditary occurrence of gastric cancer (Lynch & Krush, 1971). Later in the '70s Henry T. Lynch described the same syndrome but this time he followed colon and extracolonic cancers (uterine and stomach) of 600 descendants of an original family described by Warthin. The characteristic of this syndrome was lack of polyps and was thus originally named hereditary nonpolyposis colorectal cancer (HNPCC). Since many patients have polyps and tumors other than CRC, the syndrome was renamed as Lynch syndrome. LS is the most common inherited CRC accounting for 5 - 8% of all cases and 2 - 3% of all CRCs (Lynch & de la Chapelle, 1999, Lynch & de la Chapelle, 2003). In the early '90s germline mutations in *MSH2, MLH1, MSH6* and *PMS2* were found to segregate with cancers in LS families (Leach *et al.*, 1993, Lindblom *et al.*, 1993, Peltomaki *et al.*, 1993, Papadopoulos *et al.*, 1994, Nystrom-Lahti *et al.*, 1995).

Lynch syndrome (LS)

Once the connection between mutated MMR genes and LS was established the finding that MMR deficiency in *Saccharomyces cerevisiae* caused microsatellite instability (MSI) quickly led to the correct prediction that LS tumors also show MSI (Aaltonen *et al.*, 1993, Jiricny, 1994). Ever since, MSI has been a strong hallmark of LS cancers together with an early age of onset (~45 years) and the proximal colon being the predominant site of occurrence (Lynch & de la Chapelle, 1999, Peltomaki, Gao & Mecklin, 2001). One of the common characteristics of LS are various extra-colonic tumors, mainly found in the endometrium, ovaries, stomach and skin. Endometrial cancer has an even higher penetrance (50 - 60%) than CRC in females (Vasen *et al.*, 1996, Aarnio *et al.*, 1999).

Since the discovery of the first germline mutations in *MLH1* and *MSH2*, 2812 unique variants have been reported in five different MMR genes. *MLH1* is the most frequently mutated gene with 40% of variants (1117) followed by *MSH2* 33% (919), *MSH6* 17% (490), *PMS2* 7.5% (214) and *MLH3* 2.5% (72) (www.insight-group.org). The predominant changes in MMR genes are missense mutations and insertions/deletions.

Despite the large number of reported mutations, founder mutations, which originate from the same ancestry and thus affect several families in a specific area, are rare. The ones which have

been reported are a splice site mutation in *MLH1* exon 6 (*MLH1* c.454-1G>A) and a deletion of *MLH1* exon 16 (*MLH1* c.1732-2243_1896+404del) in Finland (Nystrom-Lahti *et al.*, 1995), a deletion of exons 1- 6 in MSH2 (*MSH2* Paracentric inversion) in the US (Wagner *et al.*, 2002), and a splice-site mutation in intron 5 of *MSH2* (c.942+3A>T) in the US and England (Froggatt *et al.*, 1999). Furthermore, about 30% of LS patients in Ashkenazy Jews carry a missense mutation in *MSH2* (*MSH2* c.1906G>C) (Guillem *et al.*, 2003, Guillem *et al.*, 2004).

The MMR genes are tumor suppressors and according to Knudson's two-hit hypothesis both of their alleles need to be inactivated in order to develop cancer. In LS patients one allele is inactivated by a germline mutation and the second allele can be inactivated by LOH or gene conversion but also by promoter hypermethylation (Yuen *et al.*, 2002, Zhang *et al.*, 2006). The second hit leads to deficient MMR activity in the cell. If DNA polymerase proofreading fails in repair there is no functional MMR mechanism, and thus DNA, especially repetitive regions (microsatellite sequences), become prone to mismatches. Changes such as insertions/deletions may also cause alterations in the reading frame of genes that contain microsatellite sequences in their coding sequence. Indeed, most affected genes in CRC are suppressors of cellular growth e.g. *TGFβRII*, *IGFIIR*, pro-apoptotic regulators e.g. *BAX* as well as MMR genes *MSH6* and *MSH3*, which contain mononucleotide repeats in their coding regions. (Markowitz *et al.*, 1995, Malkhosyan *et al.*, 1996, Souza *et al.*, 1996, Rampino *et al.*, 1997, Duval & Hamelin, 2002)

Clinicians have tried to create international criteria with an aim of unifying LS diagnostics and distinguishing LS patients from sporadic CRC. At the moment, active criteria consist of diagnostic guidelines termed Amsterdam criteria II after Amsterdam criteria I was revised in 1999 (Vasen *et al.*, 1999). Amsterdam criteria are very specific but unfortunately lack sensitivity and many LS families are not diagnosed (e.g. families with just some cancers or families without

Table 1. Amsterdam criteria II and Bethesda guidelines are existing criteria for LS dignostics (Vasen *et al.*, 1999; Umar *et al.*, 2004)

AMSTERDAM CRITERIA II*	BETHESDA GUIDELINES**
At least three relatives should have histologically verified LS cancers. One of them should be a first-degree relative of the other two.	CRC diganosed in 2 or more first- or second degree relatives with a LS cancer, one of the tumors diagnosed before age of 50
One of the relatives should be below 50 years of age when the CRC is diagnosed.	Presence of LS associated cancers, regardless of age.
At least two successive generations should be affected.	CRC diganosed in 2 or more first- or second degree relatives with a LS cancer
Familial adenomatous polyposis (FAP) should be excluded.	CRC diagnosed before age of 50
Tumors should be verified by pathological examination.	CRC with the MSI-high histology diagnosed before age of 60

* all conditions have to be fulfilled
** any of the conditions have to be fulfilled

enough information) (Rodriguez-Bigas *et al.*, 1997). Consequently, the Bethesda guidelines were established by implementation of an LS hallmark MSI (Umar *et al.*, 2004). Detailed diagnostic criteria for LS can be found in Table 1.

Although MSI is not only a specific feature of LS tumors but is also detected in ~15% of sporadic CRCs due to focal hypermethylation of the promoter of *MLH1* (Hawkins & Ward, 2001, Lawes, SenGupta & Boulos, 2003), it is widely used as a screening method for MMR defect in tumors. Generally, MSI studies employ a panel of five microsatellite markers and if two of them show instability the tumor is considered MSI-high (Boland *et al.*, 1998). Another widely used diagnostic method is to study the expression of MMR proteins in a tumor by immunohistochemistry (IHC). The tumor showing high MSI and no staining for one or several MMR proteins is MMR deficient. Although ~34% of mutations found in MMR genes are truncating, a significant amount of mutations are non-truncating (~32%) and difficult to distinguish from polymorphisms (InSiGht, 2014). The genes with non-truncating variants of uncertain significance (VUS) (Goldgar *et al.*, 2008) can still be expressed in tumors, but the expressed proteins are not necessarily functional (Mangold *et al.*, 2005, Raevaara *et al.*, 2005). Functional assays are used to assess VUS pathogenicity and thus help clinical management of families in genetic testing, counseling, and pre-symptomatic surveillance of its VUS carriers with a final aim of successful reduction in mortality (InSiGht, 2014). It is notable that failing to diagnose LS can mislead medical treatment of a CRC patient. For example standard chemotherapeutic drug treatments for cancer such as methylating agents give rise to mispairs (O^6-methylguanine (^{Me}G) – C and ^{Me}G – T) in DNA enriching mutations and induce cell death (Karran, 2001). However, MMR deficient cells are more resistant to death when they are treated with these agents. On the other hand, MMR deficient cells are more sensitive to death than MMR proficient cells when they were treated with interstrand crosslinking agents such as 1- (2-chloroethyl)-3-cyclohexyl-nitrosourea (Fiumicino *et al.*, 2000).

Mismatch repair mechanism

Mismatch repair maintains DNA integrity by detecting and repairing mismatches. A proficient MMR system distinguishes the newly synthesized DNA strand, recognizes base/base impairs or small insertion/deletion loops (IDLs), which arise during the replication and repairs damage preventing formation of mutations. First knowledge of the MMR mechanism came from the studies in *E. coli*, in which the MutS homodimer recognizes post-replicative errors and recruits the MutL homodimer forming a ternary complex. The complex activates, in an ATP-dependent manner, the latent endonuclease activity of MutH, a member of the type II family of restriction endonucleases. Activation leads to incision of a repairable DNA strand at a hemi-methylated GATC site, which is the recognition sign for the newly synthesized DNA strand in prokaryotes (Wagner & Meselson, 1976, Glickman, van den Elsen & Radman, 1978). The strand with the

impaired base is digested. DNA polymerase III corrects the mismatch and DNA ligase seals the nick. The core proteins MutS and MutL are highly conserved from prokaryotes to eukaryotes showing their importance for homeostasis of the cells. Five different MutS homologues (MSHs) have been identified in human cells, but only three of them MSH2, MSH3 and MSH6 are involved in MMR. Furthermore, human cells contain four MutL homologues (MLH) MLH1, MLH3, PMS1 (post-meiotic segregation 1) and PMS2, whereas PMS1 does not have a known role in MMR. So far, no homologue of MutH has been found in other organisms.

Contrary to prokaryotic MMR proteins, which function as homodimers, in eukaryotes they function as heterodimers. The MutS homologue proteins MSH2 and MSH6 compose the most important MMR complex MutSα, which recognizes, binds to and initiates the repair of base-base mismatches and small IDLs (Drummond *et al.*, 1995, Palombo *et al.*, 1995), while MutSβ composed of MSH2 and MSH3 repairs mainly larger IDLs (Acharya *et al.*, 1996, Palombo *et al.*, 1996). In yeast, MutSβ repairs some base-base mismatches (Harrington & Kolodner, 2007) and recently in *in vitro* studies of MMR this complex exceeds MutSα in repair of dinucleotide loops (Kantelinen *et al.*, 2010). However, in addition to the incorrectly paired bases during DNA replication e.g. G/T (Hunter *et al.*, 1987), the MMR system repairs chemically-induced DNA lesions such as alkylation-induced O^6-methylguanine paired with cytosine or thymine (Duckett *et al.*, 1996), 1,2-intrastrand (GpG) cross-links generated by cisplatin (Mello *et al.*, 1996), UV-induced photoproducts (Wang *et al.*, 1999) and 8-oxoguanine (Colussi *et al.*, 2002).

MutLα is composed of MLH1 and PMS2, MutLβ of MLH1 and PMS1 and MutLγ of MLH1 and MLH3 (Li & Modrich, 1995, Flores-Rozas & Kolodner, 1998). MutSα recruits MutLα and together they direct MMR and interact with other proteins. MutLα recruits downstream repair proteins and has an endonuclease activity substituting the MutH role in eukaryotic MMR (Yang, 2007). MutLγ seems to participate, although inefficiently, in the repair of base-base mismatches and single-nucleotide IDLs (Cannavo *et al.*, 2005, Korhonen, Vuorenmaa & Nystrom, 2008). Other factors needed for human MMR are the homo-trimeric proliferating cell nuclear antigen (PCNA) involved in DNA replication, the single-stranded binding-factor replication factor A (RPA) (Lin *et al.*, 1998) and the non-histone chromatin component high-mobility group box1 (HMGB1) (Yuan *et al.*, 2004). MSH6 and MSH3 interact with PCNA through PCNA-interacting motifs near N termini (Umar *et al.*, 1996, Kleczkowska *et al.*, 2001) and replication factor C (RFC) directs degradation reaction (Dzantiev *et al.*, 2004). Exonuclease 1 (EXO1) excises the mismatch in both directions (3'→5' and 5'→3') (Genschel, Bazemore & Modrich, 2002) interacting and forming tight heterodimeric complexes between MSH2 and MLH1 (Schmutte *et al.*, 2001, Sun, Zheng & Shen, 2002). DNA polymerase δ (Pol δ) repairs DNA (Longley, Pierce & Modrich, 1997) and DNA ligase seals the nick (Fig. 3).

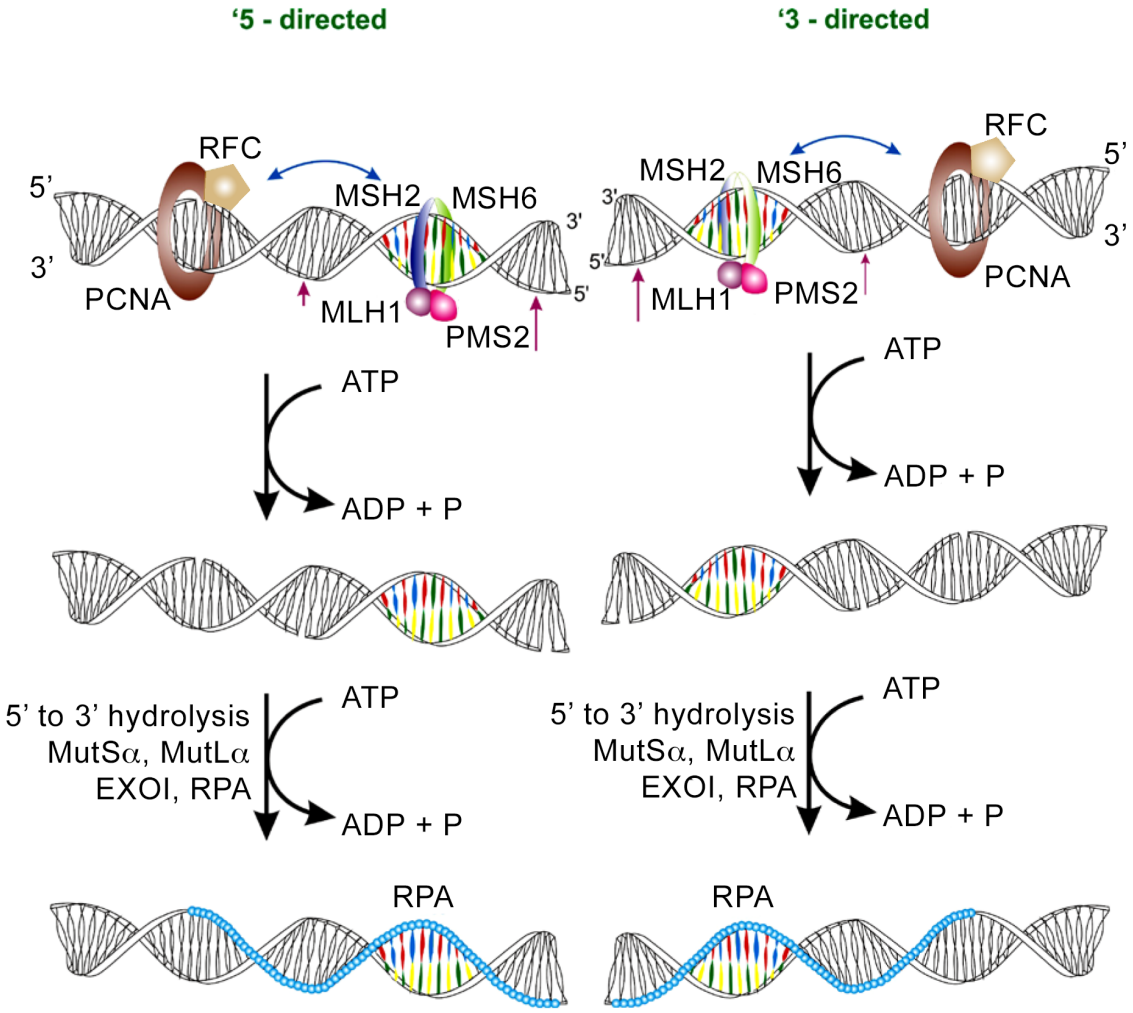

Fig. 3 Mismatch repair. MutSα (consisting of MSH2 (dark blue) and MSH6 (green) or MutSβ (consisting of MSH2 and MSH3), depending on the type, recognizes mismatch. MutSα binds to MutLα (consisting of MLH1 (violet) and PMS2 (magenta) and forms a ternary complex, which slides along the DNA in an ATP-dependent manner. EXO1 excises the mismatch, Polδ synthesizes DNA and DNA ligase ligates the nick. This process is helped by replication factor C (RFC (light brown and white pentagon), proliferative cell nuclear antigen (PCNA (brown ring) and replication protein A (RPA (light blue circles). Figure is modified from Jiricny, 2006.

Structure and function of MutSα

MutSα, a heterodimer of MSH2 and MSH6 forms an oval disc pierced by two channels like the letter θ. Both proteins of the complex contain five functional domains (Fig. 4). Domain 1 is a mismatch-binding domain with a highly conserved Phe-X-Glu pattern located at the N terminus of MSH6 that contacts the mismatch (Malkov *et al.*, 1997, Dufner *et al.*, 2000). The connector domain (domain 2) is involved in allosteric signaling and bridges the lever domain (domain 3) and the ATP binding cassette (ABC)-ATP domain (domain 5). Domain 3 spans the entire distance between domains 4 and 5, and domain 4 makes significant non-specific DNA interactions (Warren *et al.*, 2007).

The ATP domain is the most highly conserved region among MutS homologues e.g. domains of *E.coli* MutS and human MSH2 are 48% identical (Warren *et al.*, 2007). MSH proteins contain ATP binding motifs, Walker A and Walker B, which are highly conserved among prokaryotes and eukaryotes (Walker *et al.*, 1982). MutSα contains two non-equivalent ATP hydrolytic centers located at the C terminus of MSH2 and MSH6 (Martik, Baitinger & Modrich, 2004). Research in *Saccharomyces cerevisiae* has been shown that the ATPase activity of both Msh6 and Msh2 have an affect on each other (Antony *et al.*, 2006). In the presence of ATP and Mg^{2+} MutSα rapidly dissociates from complexes with either homoduplex or heteroduplex DNA. In the case of heteroduplex DNA, with ends blocked with streptavidin or the lac repressor, MutSα remains trapped on the DNA, suggesting that the complex actually slides along the DNA (Gradia, Acharya & Fishel, 1997, Blackwell *et al.*, 1998, Iaccarino *et al.*, 2000).

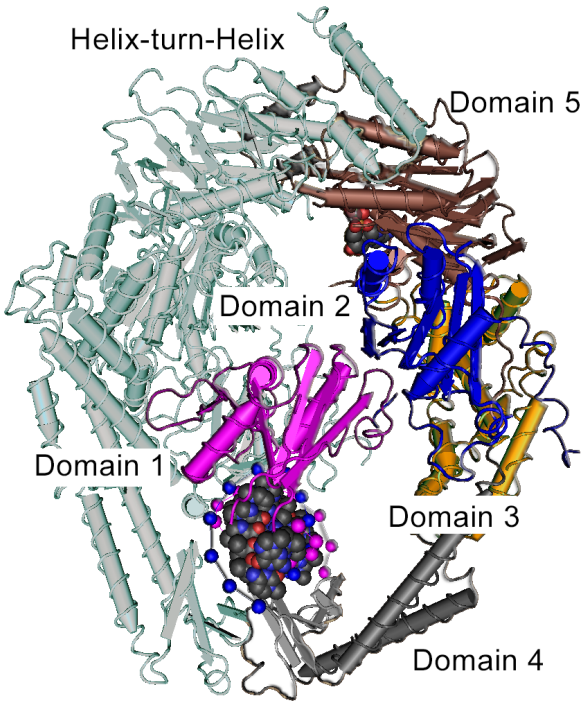

Fig. 4 3D structure of MutSα while bound to the DNA. MutSα protein complex is a heterodimer composed of MSH2 (different colors for different domains) and MSH6 (white-gray). MSH2 consists of five domains (domain 1: magenta, domain 2: blue, domain 3: yellow, domain 4: dark grey and domain 5: brown) and a structural motif helix-turn-helix (dark grey behind MSH6). Domain 1 is a DNA binding domain, domain 2 is a connector domain, which keeps MSH2 together and is responsible for allosteric regulation, domain 3 is called the lever domain and spans between domains 4 and 5. Domain 4 stabilizes DNA connection with non-specific binding sites, and domain 5 is an ATPase domain. Cn3D software (NCBI) was used for 3D visualization of the protein from NCBI:Structure.

MutS complexes recognize mismatches, IDLs or chemically modified bases in the presence of a strand break, which is ~125-150 nucleotides away (Genschel, Bazemore & Modrich, 2002) and recruits MutLα. The newly formed complex releases the mismatch in order for the mismatch to be repaired and slides along the DNA. Two models exist of the translocating and sliding MutSα, *molecular switch* and *active translocation* models. In the first model, ATPase activity and adenine nucleotide-binding domains of MutSα regulate mismatch binding as a molecular switch (Gradia, Acharya & Fishel, 1997, Gradia *et al.*, 1999). MutSα is bound to the mismatch when ADP is in the adenine nucleotide-binding domain. When ADP exchanges for ATP MutSα dissociates from the mismatch. The alternative model, active translocation, suggests that MutSα, possibly with MutLα after it is released from mismatch, translocates along the DNA hydrolyzing ATP (Blackwell *et al.*, 1998, Martik, Baitinger & Modrich, 2004). This model is based on the fact that each

subunit has two sites, which are called a latch site (L) and a D site. The L site on one subunit acts as a molecular latch, which physically blocks movement of DNA, but the D site allows DNA diffusion. Open or closed status of the L site is determined by ATP or ADP occupancy of the subunit nucleotide-binding domain. DNA diffuses through the D site for a period of time required for ATP turnover. That results in higher usage of ATP hydrolysis for DNA sliding than in the molecular switch model, which requires one ADP/ATP exchange for the conformational change of the protein complex.

Functional analysis of non-truncating MMR variants

A significant amount of reported alterations in MMR genes are non-truncating variants of uncertain significance (http://www.insight-group.org) (Goldgar *et al.*, 2008). Clinical diagnosis of Lynch syndrome consists of MMR protein expression analyses by immunochemical staining and MSI assessment. If results suggest LS, the next step is to screen for a mutation in MMR genes. When a pathogenic mutation is found, LS is confirmed, whereas in the case that no mutations are found, LS is less likely. When the found alteration is a VUS, the three steps decision tree is proposed for its assessment (Couch *et al.*, 2008). Here, after determining a VUS, the second step includes *in silico* analysis and an *in vitro* MMR assay. The variant indicating MMR deficiency with both assays suggests LS, whereas the variant showing MMR proficiency should be analyzed with other more specific biochemical assays (e.g. bandshift, protein stability, sub-localization, and protein interaction assays) (Couch *et al.*, 2008).

Functional assays determine either activity of MMR proteins in mismatch repair reactions (e.g. MMR assay) (Lahue, Au & Modrich, 1989, Nystrom-Lahti *et al.*, 2002), or a more specific function of an MMR protein (e.g. binding to each other or binding and releasing DNA). Since the *in vitro* MMR assay does not provide the biochemical reason for MMR deficiency, more specific biochemical assays in step three are needed for that knowledge. Previous studies have clearly shown that MMR proteins can be functional in the *in vitro* MMR assay, but *in vivo*, proteins are not localized in the nucleus or their expression/stability is reduced (Raevaara *et al.*, 2005, Gammie *et al.*, 2007).

Western diet as a CRC risk factor

Basic science has focused on studying genetic and molecular mechanisms causing cancer, but epidemiological studies have recognized dietary habits of Western populations as one of the high risk factors in etiology of CRC. Increased incidence of CRC correlates with the consumption of a typical Western-style diet (WD) that contains high total energy, saturated fats of animal origin and low levels of fiber (Doll & Peto, 1981). Diet scarce in vitamin D, calcium and folate

was additionally associated with CRC risk (Lipkin *et al.*, 1999). In 2007 The World Cancer Research Fund and The American Institute for Cancer Research acknowledged Western-style diet as a risk factor in CRC development estimating that 50 - 80% of CRCs can be associated with diet (http://www.dietandcancerreport.org). Epidemiological studies on the effect of diet on CRC in humans have, however, given ambiguous or even conflicting results and the mechanisms through which diet affects CRC development still needs to be studied.

One of the main characteristics of Western diet is a high intake of meat. Studies on meat intake and CRC incidence often show either a positive correlation or no significant effect. Most case studies give a strong link between meat intake and CRC, while cohort studies show no significant effect. (Norat & Riboli, 2001) Nevertheless meta-analyses utilizing different studies indicate a positive association between red meat and processed meat and CRC (Norat & Riboli, 2001, Sandhu, White & McPherson, 2001). European Prospective Investigation into Cancer showed no significant association between red meat and CRC (Linseisen *et al.*, 2002) proving inconsistency in epidemiological studies and difficulty in interpretation of their results.

Already in the ‘50s high intake of dietary fiber was associated with reduced risk of CRC (Higginson & Oettle, 1960, Burkitt, 1969). Further studies agreed on the beneficial effect of fiber, claiming that fiber can contribute up to a 40% decrease in the risk of CRC. Furthermore, the positive effect of fiber can be altered depending on the source of fiber (fiber from grains, cereals and fruit has the strongest effect) (Trock, Lanza & Greenwald, 1990, Fuchs *et al.*, 1999, Bingham *et al.*, 2003, Peters *et al.*, 2003). Studies on the effect of fiber on colonocytes implies involvement of fiber fermentation by gut microbiota and their production of short-chain fatty acids in particular butyrate, which is the principal energy substrate of the colonocyte (Roediger, 1990). *In vitro* studies also revealed that when cultured cells were treated with butyrate, growth arrest was induced. Also, caspase induced apoptosis of colon epithelial cells via altered signaling of *Fas*-ligand death receptor was noted (Heerdt, Houston & Augenlicht, 1997, Chai *et al.*, 2000, Chapkin, Fan & Lupton, 2000). Additionally, butyrate inhibits histone deacetylases and promotes pro-apoptotic effects in colonocytes *in vivo* (Boffa *et al.*, 1992, Richon *et al.*, 2000, Wilson *et al.*, 2010).

Similarly to fiber, calcium and vitamin D are considered to reduce the risk of CRC. Evidence for the effect of calcium came from several studies (Marcus & Newcomb, 1998, Mobarhan, 1999, Pietinen *et al.*, 1999) and D vitamin was later implemented based on studies, which showed that CRC incidence is higher in high geographical areas with less sunlight. Nevertheless, despite their high geographical area, the population of the Faroe Islands has lower CRC incidence, which is believed to be due to a fish- and dairy-rich diet (Dalberg *et al.*, 1999, Tangpricha *et al.*, 2001). A randomized trial with vitamin D and calcium supplements proved that these two compounds act together rather than separately (Grau *et al.*, 2003). Lower calcium and vitamin

D levels are also proposed to inhibit apoptosis and favor cell proliferation (Peters *et al.*, 2001), but the effects on early tumorigenesis remain to be elucidated. Studies on colorectal cell lines showed that extracellular calcium interacts with the parathyroid calcium-sensing receptor and promotes suppression of β-catenin/TCF signaling and activates E-cadherin expression, implying that lower calcium may affect these processes favoring CRC development (Chakrabarty *et al.*, 2003).

Folate is essential for DNA synthesis, repair, methylation and it acts as a cofactor in biological reactions (Weinstein *et al.*, 2003). Humans cannot synthesize folate and therefore, its only source is from diet. Folate becomes biologically active when converted to tetrahydrofolate (THF) in the liver. Several cohort and case studies reported that low folate intake increases the risk of CRC (Benito *et al.*, 1993, Giovannucci *et al.*, 1995, Tseng *et al.*, 1996, Giovannucci *et al.*, 1998). Controversially, recent meta-analysis of trials on folic acid supplementation showed that supplementation with folic acid does not affect occurrence of cancer (Clarke *et al.*, 2010, Vollset *et al.*, 2013). Low folate intake affects DNA maintenance and DNA or histone methylation. The folate intermediate THF accepts one carbon unit from serine or glycine and forms 5-methyl-THF (5-Me-THF) in a series of reactions. 5-Me-THF is an essential donor of methyl groups in the synthesis of methionine. S-adenosylmethionine (SAM) is synthesized from methionine, and is used by all methylation reactions in the cell as a methyl donor producing S-adenosylhomocysteine (SAH), which is further converted to homocysteine. 5-methyl-tetrahydrofolate donates a methyl group to homocysteine in a reaction catalyzed by methionine synthase. Despite the known biochemical reactions of the folate pathway, exact mechanisms, which may lead to CRC development, remain unknown.

METABOLIC STRESS AND CANCER

In 1927, Otto Warburg already discovered that cancer cells alter their glucose metabolism to produce most of their energy by glycolysis in the presence of oxygen, known as aerobic glycolysis (Warburg, Wind & Negelein, 1927, Warburg, 1956a, Warburg, 1956b). This has become the center of attention in the past two decades and has led to new insights into the nature of cancer.

Cancer cells energetically rely on aerobic glycolysis, which produces 2 molecules of ATP and pyruvate. The majority of pyruvate is further converted to lactate. In the presence of oxygen normal cells process glucose to pyruvate by glycolysis and pyruvate is further transported to mitochondria and processed in the tricarboxylic acid cycle (TCA) to carbon dioxide, resulting in higher energy production (Fain, Shestovskaia & Mamonova, 1975). Even though cancer cells seem to be inefficient in energy production, they adjust their metabolism to give them enough energy by up-regulating glucose transporters (e.g. GLUT1), which substantially increases glu-

cose import into the cytoplasm. Up-regulation of glucose transporters is associated with activated oncogenes (e.g., *RAS*, *MYC*) and mutations in tumor suppressors (e.g., *TP53*) (DeBerardinis *et al.*, 2008, Jones & Thompson, 2009).

Additionally, increased glycolysis creates molecular intermediates for biosynthetic pathways such as biosynthesis of nucleosides and amino and fatty acids, which in turn enable cancer cell proliferation (Potter, 1958, Vander Heiden, Cantley & Thompson, 2009). Recent studies revealed that cancer reliance on glycolysis depends on the hypoxia response system, which independently from RAS can signal up-regulation of glycolysis (Kroemer & Pouyssegur, 2008, Semenza, 2010). Furthermore, the hypoxia response system can be activated by reactive oxygen species (ROS) (Bonello *et al.*, 2007), the levels of which are altered in many diseases (e.g. diabetes, obesity, cancer) and may contribute to disease pathogenesis (Halliwell, 2007, Nathan, 2008, Trachootham, Alexandre & Huang, 2009).

Normal cells produce ROS at a low level in the electron transport chain, which is used in the regulation of cell signaling, proliferation and differentiation. Increase in ROS levels due to altered cellular metabolism can potentially damage cellular components and ultimately lead to cell death (Veal, Day & Morgan, 2007, Trachootham, Alexandre & Huang, 2009, Hamanaka & Chandel, 2010).

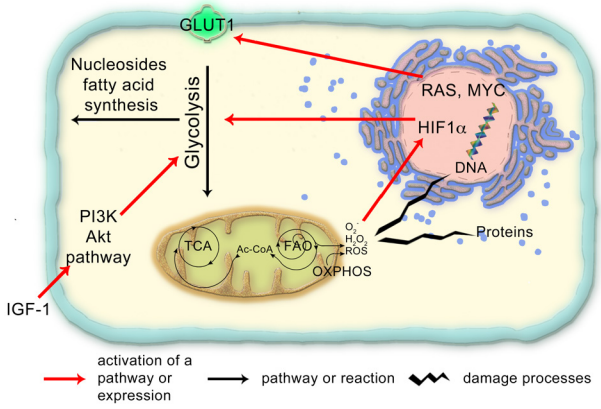

Fig. 5 Activation and regulation of metabolic pathways in cancer. Cancer cells utilize aerobic glycolysis in the presence of oxygen for production of nucleosides, fatty acids and other molecular precursors needed for cellular proliferation. Aerobic glycolysis is favored in cancer cells. RAS and MYC favor aerobic glycolysis by up-regulating the Slc2a1 gene, resulting in production of the protein glucose transporter (GLUT1). IGF-1 stimulates the PI3K-Akt pathway, which in turn activates glycolysis. Additionally high levels of ROS stabilize HIF1α, which translocates into the nucleus and activates glycolysis independently from RAS and MYC. TCA – tricarboxylic acid cycle, FAO – fatty acid oxidation, OXPHOS – oxidative phosphorylation.

Cellular levels of ROS in normal cells can be altered by nutrient excess/diet and lead to a pathological cellular state. Briefly, normal cells process eventually all nutrients by TCA, and released electrons are captured by nicotinamide adenine dinucleotide (NAD+) and flavin adenine dinucleotide (FAD). These electrons are later used in oxidative phosphorylation reactions to generate energy and the final acceptor of electrons is oxygen in the reaction that creates a molecule of water. (Wellen & Thompson, 2010) Increased concentration of electrons caused by *nutrient excess*/diet results in superoxide production, which is one of the ROS molecules (Turrens, 2003). Superoxide is converted by superoxide dismutase into hydrogen peroxide, which is used for cellular signaling (Meng, Fukada & Tonks, 2002). Hydrogen peroxide can be

converted into the highly reactive hydroxyl radical, which can damage DNA, membrane lipids, and proteins. In summary increased ROS may promote increasing DNA mutation rate, regulate signaling, transcription and promote inflammation contributing to cancer development (Fig. 5) (Wu, 2006, Lou *et al.*, 2008).

Adaptation to oxidative stress caused by ROS can in part be regulated by the unfolded protein response (UPR) (Cullinan & Diehl, 2006). UPR as a reaction to endoplasmic reticulum stress is altered in cancer and may have a role in cancer promotion (Moenner *et al.*, 2007, Lou *et al.*, 2008, Healy *et al.*, 2009).

MOUSE MODELS IN CRC STUDIES

Many different mouse models have been created for studying changes in gene expression and pathways involved in carcinogenesis in colon and other intestinal tissues. The most studied model is the Apc^{Min} strain that carries a truncating mutation at codon 850 (Moser, Pitot & Dove, 1990). The heterozygote mouse develops on average 30 polyps in the small intestine. Additionally, different mutations have been introduced in the *Apc* gene and various strains created e.g. $Apc^{\Delta 716}$ (Oshima *et al.*, 1995) and Apc^{1638N} (Fodde *et al.*, 1994). Both develop polyps in the small intestine: $Apc^{\Delta 716}$ 300 polyps on average, while Apc^{1638N} only 3. Very few polyps are formed in the colon. *Apc* mutant mice have a short life span, on average 4-5 months because of heavy tumor load in the small intestine. (Taketo & Edelmann, 2009)

MMR deficient mouse models

Due to the importance of MMR function in Lynch syndrome several mouse strains with a defective MMR gene have been created to systematically study their role in initiation and development of LS (Table 2).

Three different knock-out alleles of the *MSH2* gene have been created (de Wind *et al.*, 1995, Reitmair *et al.*, 1995, Smits *et al.*, 2000). The $Msh2^{-/-}$ mice develop normally and both sexes are fertile. The majority of the mice develop invasive T-cell lymphomas already within the first two months, and 50% of the homozygote $Msh2^{-/-}$ mice die during the first six months (Lowsky *et al.*, 1997). Mice that survive longer than six months develop adenomas and adenocarcinomas in their small intestine and gland tumors (Schwartz & Torre, 1995). Heterozygous $Msh2^{+/-}$ mice are not distinguishable from wild type littermates; some develop tumors similar to $Msh2^{-/-}$ in their second year of life (Reitmair *et al.*, 1996, de Wind *et al.*, 1998).

$Msh3^{-/-}$ mice have a mild cancer phenotype and they develop gastrointestinal tumors when the animals are older than 14 months; otherwise this mouse line is similar to the wild type mouse

Table 2. Deficient MMR mouse lines[1] with their phenotypes (adopted from Wei K *et al.,* 2002)

Genotype	50% survival [months]	Tumor		Fertility[3]	References
		Incidence	Type	male / female	
Msh2$^{-/-}$	6	High	Lymphoma, Gi, Skin	+/+	Reitmair *et al.*, 1995, de Wind *et al.*, 1995, de Wind *et al.*, 1998, Smits *et al.*, 2000
Msh3$^{-/-}$	18	Low[2]	GI tumors	+/+	de Wind *et al.*, 1999, Edelmann *et al.*, 2000
Msh6$^{-/-}$	11	High	Lymphoma, GI	+/+	Edelmann *et al.*, 1997, de Wind *et al.*, 1999, Edelmann *et al.*, 2000
Msh3$^{-/-}$ *Msh6*$^{-/-}$	6	High	Lymphoma, GI, skin	+/+	de Wind *et al.*, 1999, Edelmann *et al.*, 2000
Mlh1$^{-/-}$	6	High	Lymphoma, GI, skin	-/-	Baker *et al.*, 1996, Edelmann *et al.*, 1996, Kawate *et al.*, 1998, Prolla *et al.*, 1998,
Pms1$^{-/-}$	>18	None	None	+/+	Prolla *et al.*, 1998,
Pms2$^{-/-}$	10	High	Lymphoma & sarcoma	-/+	Prolla *et al.*, 1998,

[1] mouse lines were created on C57BL/6J and various 129 substrains genetic backgrounds
[2] Incidence of tumors was not significantly different from wild type
[3] '+' - fertile, '-' - infertile

(Edelmann *et al*., 2000).

Msh6$^{-/-}$ mice have similar cancers as *Msh2*$^{-/-}$ mice but the age of onset is slightly delayed and 50% of the mice die at the age of 8-10 months. The earliest tumors are B- and T-cell lymphomas, and at the later stage of life some mice develop gastrointestinal tumors (Edelmann *et al*., 1997, de Wind *et al*., 1999).

Double mutant mice *Msh3*$^{-/-}$/*Msh6*$^{-/-}$ show no differences to *Msh2*$^{-/-}$ mice (de Wind *et al*., 1999, Edelmann *et al*., 2000), supporting the idea that MSH3 and MSH6 are alternative binding partners of MSH2 in MMR.

Three different knock-out alleles for *MLH1* have been published, and the mice express a similar phenotype to *Msh2*$^{-/-}$ mice, except that *Mlh1*$^{-/-}$ mice are infertile (Baker *et al*., 1996, Edelmann *et al*., 1996, Kawate *et al*., 1998). *Mlh1*$^{-/-}$ mice develop lymphomas and 50% of the mice die at the age of six months. Mice that survive the first six months develop gastrointestinal and skin tumors and live no longer than twelve months. Heterozygous *Mlh1*$^{+/-}$ mice are not distinguishable from wild type mice (Prolla *et al*., 1998, Edelmann *et al*., 1999).

Pms2$^{-/-}$ mice express a less severe cancer phenotype than *Mlh1*$^{-/-}$ mice. Males are sterile but females are fertile, implying different roles of PMS2 and MLH1 in meiosis. Homozygous mice develop lymphomas and sarcomas already within the first six months, but 50% survival is until 9 – 10 months and older mice never develop gastrointestinal or skin tumors (Prolla *et al*., 1998).

Overall, inactivation of *Msh2* or *Mlh1* causes stronger cancer phenotypes than inactivation of *Msh6* or *Pms2*. However, contrary to Lynch syndrome mutation carriers, heterozygous mice are not generally distinguishable from wild type littermates and do not express early-onset tumors. This is probably because of a shorter lifespan and smaller body size. (Wei, Kucherlapati & Edelmann, 2002, Taketo & Edelmann, 2009)

Mouse as a model for diet effects in the intestine

Most correlations between the consumption of different nutrients and colon cancer in human come from epidemiological studies. However, studies of diet effects on colon epithelia over a longer period of life are quite infeasible in humans due to a long life span, a high variability in individual dietary habits and heterogeneous genetic backgrounds.

Isogenic mouse strains prove to be a useful tool in diet studies. Indeed, a Westernized diet, with high fat content, low on fiber, folate, vitamin D and calcium, has been shown to induce benign and malignant colon neoplasms in wild type mice without inherited cancer predisposition or carcinogenic treatments (Newmark *et al.*, 2001, Newmark *et al.*, 2009), suggesting that with increasing age WD itself is a considerable risk factor for CRC. Western diet has also been shown to increase the amount of intestinal neoplasms in the Apc^{1638N} mouse, which is the most commonly used rodent model for human intestinal cancer (Yang *et al.*, 1998). Many studies of different dietary components and complex diets have been performed with Apc^{Min} mice but before our studies none with MMR deficient mouse lines (Tammariello & Milner, 2010).

AIMS OF THE PRESENT STUDY

Both genetic and lifestyle factors have been shown to effect gene regulation and predisposition to cancer. The main aims in the present study were to analyze the biochemical effects of inherited MMR gene variants derived from patients with putative Lynch syndrome and effects of Western style diet on colon cancer risk.

More specific aims were:

1. To examine the biochemical defects caused by inherited non-truncating mutations in the MMR gene *MSH2*. Gene variants were expressed in *Sf9* cells. MSH2 protein variants were purified from total protien extracts and biochemically characterized. (I, II)

Furthermore, to study the effects of Western-style diet we conducted an animal diet study. CRC predisposed mouse ($Mlh1^{+/-}$), a model for Lynch syndrome, and wild type ($Mlh1^{+/+}$) as a control were selected beacuse *MLH1* gene is the most mutated gene in human LS and *MLH1* promoter can be easily hypermethylated in patients with Lynch sydrome and in general population. (III)

Here specific aims were:

2. To define the effects of Western diet on protein expression in histologically normal colon mucosa of wild type ($Mlh1^{+/+}$) and $Mlh1^{+/-}$ mice (III)

3. To use network analysis and data mining to determine which of the affected proteins may be putative players in early CRC development (III)

MATERIALS AND METHODS

MSH2 VARIANTS (I-II)

This study included eighteen non-truncating mutations in *MSH2* with the following amino acid (AA) substitutions: c.97A>C → p.T33P, c.380A>G → p.N127S, c.482T>A → p.V161D, c.484G>A → p.G162R, c.490G>A → p.G164R, c.518T>C → p.L173P, c.560T>C → p.L187P, c.815C>T → p.A272V, c.965G>A → p.G322D, c.998G>A → p.C333Y, c.1807G>A → p.D603N, c.1906G>C → p.A636P, c.2021G>C → p.G674A, c.2090G>T → p.C697F, c.2245G>A → p.E749K, c.2500G>A → p.A834T, c.2768T>A → p.V923E and deletion c.2235_2240del → p.I745-I746del. They will be further discussed as amino acid substitution in the MSH2 and called protein variants. Fifteen of them were found in putative Lynch syndrome families and thus functional assessment of pathogenicity was of the utmost importance. The variants p.N127S and p.G322D were studied because they had been shown to have frequent occurrence in patients with CRC but their classification was contradictory in the literature. The variant p.G674A , which was reported as a Lynch syndrome germline mutation (Raedle *et al.*, 2001), was included especially because it was shown to affect the biochemically interesting domain, Walker A motif, in MSH2 and was also shown to inactivate MMR in mouse cells (Lin et al. 2004).

Table 3 shows the clinical background of MSH2 protein variants. Eleven of the described variants were associated with families fulfilling the Amsterdam criteria I or II (AC). Four were found in several families, of which some fulfilled and some not the AC. Two variants were found in families, who did not fulfill the AC. IHC analysis showed loss of the MSH2 protein in tumors associated with 11 different variants, variable expression associated with one variant, and retained expression associated with 3 variants. The MSI high phenotype was present in at least one tumor connected to all 15 variants.

All the studied *MSH2* variations were expressed to a similar level as wild type *MSH2* indicating protein stability. All the variant proteins, except p.G674A (Table 3) displayed normal interaction with MSH6. (Ollila *et al.*, 2006).

Twelve variants (p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p.C333Y, p.D603N, p.A636P, p.G674A, p.C697F, p.I745-I746del and p.E749K,) showed complete loss of MMR function, p.T33P showed decreased repair efficiency, and 5 variants (p.N127S, p.A272V, p.G322D, p.A834T and p.V923E) had normal repair efficiency compared to wild type MSH2 (Ollila *et al.*, 2006) (Table 3).

Table 3. Clinical background of protein variants (modified from Ollila *et al.*, 2006)

MSH2 variant	Family code	Index Patient: Age at onset [y] / tumor site	All affected patients[1] / mean age at onset	AC I / II	Method of mutation analysis	MSI status[2]	IHC[2]			Interaction with MSH6	Expression in LoVo cells	*In vitro* MMR assay
							MSH2	MSH3	MSH6			
T33P	U01-537	45/EC	2/48	-	DS, MLPA	H	+	+	+	Normal	Normal	Decreased
V161D	FAM 10	52/CRC	3/53	+	DS, MLPA	H	-	NA	NA	Normal	Normal	Deficient
G162R	FAM C	56/EC	6/52	+	DS	H	-	+/-	+	Normal	Normal	Deficient
G164R	HNPCC 420	39/CRC	7/39	+	DS, MLPA	NA	-	NA	+	Normal	Normal	Deficient
L173P	HNPCC 183	36/CRC	9/45	+	DS, Multiplex PCR	H	-	NA	+	Normal	Normal	Deficient
L187P	HNPCC 548	41/CRC	5/42	+	DS, MLPA	H	-	NA	+	Normal	Normal	Deficient
L187P[3]	FAMILY A	42/CRC	11/48	+	DS, MLPA	H	-	-	+	Normal	Normal	Deficient
A272V	HNPCC 421	41/CRC	3/40	+	DS, MLPA	L	+	NA	+	Normal	Normal	Normal
A272V	U01-348	41/CRC	1/41	-	DS, MLPA	H	+	+	+	Normal	Normal	Normal
C333Y	HNPCC 228	41/CRC	2/41	-	DS, Multiplex PCR	NA	NA	NA	NA	Normal	Normal	Deficient
D603N	EN 13	50/EC	2/49	-	DS, MLPA	H	-	-	+	Normal	Normal	Deficient
D603N	EN 26	46/EC	5/55	+	DS	S	-	-	+	Normal	Normal	Deficient
D603N	HNPCC 122	38/CRC	1/38	-	DS	H	NA	NA	NA	Normal	Normal	Deficient
A636P[4]	HNPCC B	42/CRC, 44/EC	2/44	+	DS	H	-	+/-	+	Normal	Normal	Deficient
A636P	U01-051	43/CRC	1/43	-	DS, MLPA	H	-	-	+	Normal	Normal	Deficient
A636P	CG0074	36/EC	5/49	+	DS	H	-	NA	NA	Normal	Normal	Deficient
C697F[5]	HNPCC 62	27/CRC	5/45	+	DS, Multiplex PCR	H	-	NA	+	Normal	Normal	Deficient
C697F	HNPCC 934	33/CRC	3/38	+	DS, MLPA	H	-	NA	+	Normal	Normal	Deficient
C697F	1260	40/EC	3/49	+	DS	H	-	+	+	Normal	Normal	Deficient
DEL745-746	FAMILY 12	39/CRC	4/42	+	DS, MLPA	H	-	-	+	Normal	Normal	Deficient
E749K	HNPCC 152	29/CRC	7/29	+	DS, MLPA	H	+	NA	+/-	Normal	Normal	Deficient
A834T	HNPCC 417	28/CRC	3/39	+	DS, Multiplex PCR	H	-	-	+	Normal	Normal	Normal
V923E[6]	FAMILY 16	70/CRC	6/58	+	DS, MLPA	H	+/-	-	+	Normal	Normal	Normal

[1] Affected patients with LS tumors
[2] MSI and IHC were analyzed on the primary tumor of the index patient, except in cases labelled wit 3, 4 and 5
[3] IHC not available from index patient; MSH2 and MSH6 loss detected from sebaceous adenoma of a paternal aunt, a verified mutation carrier
[4] Tumor data are from EC of the index patient
[5] IHC not available from index patient; MSH2 and MSH6 loss detected from sebaceous carcinoma of a sister, a verified mutation carrier
[6] The index patient carries 2 mutations, MSH2 V923E and MSH6 S1188N
EC - endometrial cancer, DS - direct sequencing, NA - not available, H - high, L - low, S - stable

PROTEIN PRODUCTION AND PURIFICATION (I-II)

MutSα production

Site-directed mutagenesis, production of baculovirus vectors and production of vectors for human cell expression were completed prior to this work and will not be thoroughly discussed here (Ollila *et al.*, 2006).

Viral amplification

Sf9 cells were seeded in 20 mL of complete Grace's Insect Cell Culture Medium (Gibco). After addition of virus, cells were incubated at 27°C for 5 days. The cell suspensions were centrifuged at 1000 rpm for 10 min and supernatants with amplified viruses stored at 4°C in the dark.

Infection of the Sf9 cells for protein production

*Sf*9 cells were co-infected with equal amount of virus containing 6x His-tagged wild type (WT) *MSH6* and wild type or mutated *MSH2* to express heterodimeric MutSα protein complexes. The cells were incubated for 3 days at 27°C.

Protein extraction

Total protein extracts were extracted from infected cells in cold lysis buffer (25 mM HEPES pH 8, 2 mM 2 - mercaptoethanol, 0.5 mM spermidine, 0.15 mM spermine, 0.5 mM phenylmethylsulfonyl fluoride (PMSF) and 2x complete protease inhibitor mixture (Roche)). The cells were allowed to swell on ice for 20 min followed by addition of ice cold 100% glycerol and NaCl to a final concentration of 300mM. Suspensions were rotated at 4°C for 30 min and centrifuged at 14 000 rpm, 4°C for one hour. Supernatants were aliquoted, snap frozen in liquid nitrogen and stored at -80°C.

Ni-NTA purification

Recombinant MutSα protein complexes were purified from the total protein extract of *Sf*9 cells. Some of the variants (p.T33P, p.A636P, p.E749K, p.A834T, and p.V923E) were already purified by fast protein liquid chromatography (FPLC) by other researchers in the group. Here, the rest of the variants (p.N127S, p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p.A272V, p.G322D, p.C333Y, p.D603N, p.G674A, p.C697F p.I745-I746del) and wild type MutSα were purified from the total protein content of *Sf*9 cells with nickel-nitrilotriacetic (Ni-NTA) agarose

matrix (QIAGEN). Ni-NTA is tetradentate-chelating adsorbent, which occupies four of the six ligand-binding sites in the co-ordination sphere of the nickel ion. Two ligand sites are free to bind polyhistidine in the C-terminus end of the recombinant protein. In order to obtain active protein, native purification conditions were used. Thus, MSH2 remains bound to 6xHis tagged MSH6, enabling the co-purification of the two proteins (Gradia, Acharya & Fishel, 1997). Ni-NTA matrix was washed with PBS three times prior to incubation with proteins. The ratio of matrix and total protein extract was 1:10 and incubation was in constant rotation at +4°C for 2 hours. Suspensions of Ni-NTA and total protein extract were loaded into 1,5mL polypropylene columns (QIAGEN) and washed with 1mL of wash buffer (25 mM HEPES, 300 mM NaCl, 20 mM 1,3- diazole (imidazole), 1 μL/mL leupeptin, 1x complete EDTA free 25x, 0.5 mM PMSF) six times. Recombinant proteins were eluted with increasing concentrations of imidazole in wash buffer. Certain mutated MutSα proteins were expressed in notably weaker concentrations compared to WT. Those were purified with a slightly modified protocol. The ratio of matrix and total protein extract was 1:20 and 20mM imidazole was added already prior to incubation. Furthermore, Ni-NTA agarose was washed with PBS containing 20 mM imidazole three times. The eluted proteins were dialyzed, aliquoted, snap frozen with liquid nitrogen and stored at -80°C. The concentration of proteins was determined after purification with the Bradford assay and by comparing the band intensities in SDS-PAGE gels stained with coomassie brilliant blue.

BANDSHIFT ASSAY (I-II)

Construction of labeled homo- and heteroduplex oligomers

To produce DNA substrates for bandshift assays, single-stranded 38mer oligomers (5'-TTTCT-GACTTGGATACCA**T**CTATCTATCTATAAAATAT-3', the site of the putative mismatch in bold) were 5' labeled with 32-phosphorus (^{32}P). The kinase reaction was performed in a total volume of 10 μL containing 10 pmol of oligomer, 1x polynucleotide kinase (PNK) buffer (New England Biolabs), 100 mM DL-dithiothreitol (DTT), 10 pmol of 32- P-γ-ATP (Amersham), and 15 units of polynucleotide kinase (PNK) (New England Biolabs). The reaction was incubated at 35°C for 45 min, followed by addition of 10 units of PNK and incubated at 35°C for 30 min. Labeled oligomers were purified using ProbeQuant G-50 Micro Columns (Amersham) according to the protocol given by the manufacturer. To obtain double-stranded oligomers, the labeled single-stranded oligomers were annealed with 38 base-long complementary single-stranded oligomers. The mixture of both oligomers (labeled and non-labeled) in a 1:1 ratio was denaturated at 95°C for 5 min and annealed at 37°C for 5 min followed by incubation at room temperature for 30 min.

Mismatch binding and release assays

The bandshift assay was used to study the capability of recombinant MutSα proteins to bind to DNA heteroduplex-containing G●T mismatch, homoduplex-containing A●T, and to determine their ability to dissociate from DNA by ATP uptake.

The bandshift assay reaction mixture contained 10% glycerol, 100 mM KCl, 25 mM HEPES-KOH pH 7.5, 1 mM DTT, 0.5 mM $MgCl_2$, 0.1 mM adenosine diphosphate (ADP), 75 ng/mL bovine serum albumin (BSA), 60 ng poly-d(I-C), 25 fmol labeled oligomers and MutSα (350 ng) in a total volume of 20 μL. The salt (NaCl) and HEPES concentration was adjusted to be similar in all samples. The reaction mixtures were incubated at 37°C for 20 min and 10μL of each mixture was separated into 5% acrylamide gels.

For ATP mediated release, 1μL of ATP in the desired concentration was added into the reaction after 10 min to obtain end concentrations of 0.5 mM, 1 mM and 2 mM, respectively. All gels were dried and visualized with phosphorimager Fuji and analyzed with TINA software version 2.08e (Isotopen - meßgeräte GmbH).

MOUSE MODELS AND DIETS (III)

Heterozygote B6.129-*Mlh1*^{tm1Rak} mice (*Mlh1*$^{+/-}$) (strain 01XA2) (Edelmann *et al.*, 1996) were obtained from NCI-MMHCC; National Institutes of Health, Mouse Repository, NCI-Frederick, MD. *Mlh1*$^{+/-}$ mouse has a deletion of exon 2 in one of the two of the *Mlh1* alleles leading to a 50% decrease of MLH1 protein. Instead of the exon 2 sequence it contains a box element with a stop codon.

Diets

American Institute for Nutrition (AIN) control diet (AIN-93G) is a semi-synthetic diet designed to meet the nutritional requirements of growing rodents (Reeves, Nielsen & Fahey, 1993), while Western style diet (WD*) is a modified AIN diet, which contains high dietary fat (39% of total calories while 17.2% in AIN-93G) and reduced contents of fiber, calcium, vitamin D, and three methyl-transfer donors (i.e. folic acid, methionine, and choline). Calories from carbohydrates in WD* are reduced for 20% in comparison to AIN-93G, nevertheless WD* has increased sucrose and correspondingly less complex carbohydrates than the control diet. Previously published Western diet differs from WD* in fat source, since Newmark *et al.* used corn oil (20%) (Newmark *et al.*, 2001), whereas WD* consists of milk fat (13.3%), canola oil (5.5%) and sunflower oil (1.2%) of the total 20% fat content. The fat source in AIN-93G is exclusively soybean oil (Table 4).

Feeding experiment

Table 4. Composition of the diets

Compounds	AIN-93G [g/kg][1]	WD* [g/kg][1]
Proteins [%]	20.3	23.5
casein	200	232
L-cysteine	3	3
Carbohydrates [%]	63	52
corn starch	397.49	305.63
maltodextrin	132	95
sucrose	100	116
Fat [%]	7	20
soybean oil	70	-
anhydrous milk fat	-	**132.8**[2]
canola oil	-	**55.4**[2]
sunflower oil	-	**11.8**[2]
cellulose	50	20
calcium	5	0.5
folic acid [mg/kg]	2	0.2
vitamin D [IU/kg]	1000	100
kcal from proteins [%]	18.8	18.5
kcal from carbohydrates [%]	63.9	42.3
kcal from fat [%]	17.2	39.2

[1] if not stated differently
[2] Fat sources that differ in WD* from previously published Western diet (Newmark *et al.,* 2001, Newmark *et al.,* 2009) are in bold.

Mice were bred and treated according to the study protocol approved by the National Animal Experiment Board in Finland (ESLH-2008-06502/Ym-23). At the age of 5 - 6 weeks (time point 0, tp0), *Mlh1*$^{+/-}$ and *Mlh1*$^{+/+}$ mice were weaned from their mothers and randomly divided into two dietary groups. Mice were fed *ad libitum* with American Institute for Nutrition (AIN-93G) control diet or WD* over 12 months and kept on a 12 h light/dark cycle with controlled temperature and humidity. Mice were weighted every two weeks by personnel at the Laboratory Animal Center, University of Helsinki.

Mice were sacrificed with carbon dioxide and sampled at tp0 (*Mlh1*$^{+/-}$ N = 8, *Mlh1*$^{+/+}$ N = 7) and tp1 (12 months of age) (*Mlh1*$^{+/+}$ AIN N = 8, *Mlh1*$^{+/-}$ AIN N = 6, *Mlh1*$^{+/+}$ WD* N = 8, *Mlh1*$^{+/-}$ WD* N = 7).

Histological studies

Mice colons were examined under light microscope for possible neoplasias. All potential neoplasias were collected and analyzed at the Finnish Centre for Laboratory Animal Pathology (FCLAP), University of Helsinki, Finland.

QUANTITATIVE ANALYSIS OF PROTEOMES (III)

Sample preparation from mouse colonic mucosa

For protein and mRNA studies, the colonic mucosa (10 - 25 mg) was separated from the underlying submucosa and musculature under a dissecting microscope. Samples for protein extractions were rinsed with a solution of 10 mM Tris (Sigma-Aldrich, Finland) and 5mM magnesium acetate (Sigma-Aldrich, Germany) (pH 8.5), snap frozen and stored at -80°C. Samples for RNA extraction were stored in RNAlater (Qiagen, Valencia, CA) at -80 °C.

Protein extraction from mucosa samples

Colonic mucosa samples were mechanically homogenized in the 2-D Protein Extraction Buffer-VI (GE Healthcare) with addition of pH 8.5 Tris (30 mM) (Sigma-Aldrich) and Protease Inhibitor Mix (GE Healthcare) (1/100) (10-15 µL lysis buffer/mg tissue). After homogenization samples were vigorously shaken for 20 min at 4 °C followed by centrifugation at 15 000 g for 20 min. Supernatant was collected, snap frozen and stored at -80 °C. 2-D Quant Kit (GE Healthcare) was used to determine the protein concentrations of total extracts according to the manufacturer's instructions.

2D Difference gel electrophoresis (DIGE)

All equipment and chemicals used in the proteomic study were from GE Healthcare if not stated differently. Proteomic changes were studied using the 2D DIGE method (Fig. 6). Half of the samples from each mouse group were labeled with 1-(5-carboxypentyl)-1′-propylindocarbocyanine halide N-hydrosuccinimidyl ester (Cy3) and the other four with 1-(5-carboxypentyl)-1′-methylindodicarbocyanine halide N-hydroxysuccinimidyl ester (Cy5) (400 pmol of dye for 50 µg of sample) according to the manufacturer's instructions. An internal standard was created by pooling 25 µg of each protein sample and labeling it with 3-[(4-carboxymethyl) phenylmethyl-3′-ethyloxacarbocyanine halide N-hydroxysuccinimidyl ester (Cy2). Isoelectric focusing was performed using 24 cm long immobilized pH 3-11 non-linear gradient (IPG) strips with the Ettan IPGphor II unit. The IPG strips were rehydrated for 8 hours in an IPG box using DeStreak™ reagent and IPG Buffer, pH 3-11 NL. An equal amount (50 µg) of internal standard labeled with Cy2 combined with Cy3 and Cy5 labeled samples was introduced actively into the IPG strips using the cup-loading method according to the manufacturer's recommen-

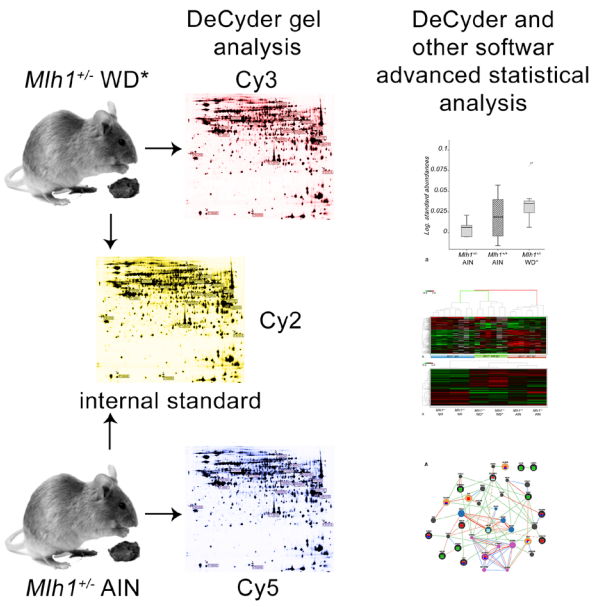

Fig. 6 Overview of the 2D-DIGE experiment. Samples are labeled with either Cy3 or Cy5. An internal standard, which contains equal amount of protein extract from all samples used in the experiment, is labeled with Cy2. Equal amounts of Cy3, Cy5 and Cy2 labeled protein extract are loaded into isoelectric focusing strips and proteins are separated according to their isoelectric point. Strips are transferred on top of large SDS-PAGE gels and proteins further separated according to their molecular weight. 2D protein spot maps from two samples (Cy3 and Cy5) and one internal standard (Cy2) are detected within the same gel after scanning with a Typhoon scanner. DeCyder software is used for analysis of images, calculation of relative protein expressions, as well as advanced analyses e.g. hierarchical clustering and principal component analysis.

dations. Conditions of isoelectric focusing were: 21°C, 75mA/strip, Step1: step and hold: 150V, 3h; step and hold: 300V, 3h; gradient: 1000V, 6h; gradient: 8000V, 1h 15min; step and hold: 8000V, 3h 45min. Isoelectric focusing was followed by equilibration of IPG strips in two steps with 1% DTT and 2.5% iodoacetamide, dissolved in equilibration buffer containing 6 M urea, (Sigma-Aldrich), 2% SDS (Sigma-Aldrich), 50 mM Tris pH 8.8 (Sigma-Aldrich), 0.02% bromophenol blue (Sigma-Aldrich) and 30% glycerol (Sigma-Aldrich) respectively. Separation of proteins in second dimension was carried out by transferring the equilibrated IPG strips to large 12.5% SDS-PAGE gels and run in homemade Laemmli buffer with constant power of 15 W per gel until the bromophenol blue front reached the lower edge of the gel. Gels were scanned immediately after the run with Typhoon 9400 scanner. Cy2, Cy3 and Cy5 images were acquired with excitation/emission values 480/530 nm, 520/590 nm and 620/680 nm respectively, and maximum values between different channels were optimized to differ less than 20 - 30%. Gel images were analyzed using the software DeCyder 2D 7.0.

Protein identification

The scanned gels were silver-stained using the PlusOne Silver staining kit (GE Healthcare). Selected protein spots were excised and proteins in-gel digested. Gel pieces were treated twice with acetonitrile (Sigma-Aldrich) (200 μL) and vacuum centrifuged until dry. Proteins were trypsinized (Trypsin gold; Promega, Sweden) and peptides were desalted using ZipTip (Millipore, Germany), eluted and mixed with α-Cyano-4-hydroxycinnamic acid (CHCA) (Sigma-Aldrich) on a MALDI plate. MALDI-MS and MALDI-MS/MS analyses were done with Autoflex III (Bruker Daltonics, Germany) using SmartBeam™ 355 nm laser. Calibration was externally performed with a peptide calibration standard (Bruker Daltonics). MALDI-MS spectra were acquired by accumulation of 5000 shots and MS/MS spectra 10000 shots.

The MS and MS/MS spectra were searched against the UniProt/SwissProt database (taxon: *Mus musculus*, release of 2011_05; www.uniprot.org) using the Mascot server (Matrix Science, www.matrixscience.com) and FlexAnalysis and BioTools software (Bruker Daltonics). The following search parameters were implemented: cleaving enzyme trypsin, maximum one missed cleavage allowed, peptide mass tolerance ± 0.1 Da for MS searches and fragment tolerance ± 0.5 – ± 1.5 Da for combined MS/MS searches. Oxidized methionine and carbamidomethylation of cysteine were set as variable and fixed modifications respectively.

WESTERN BLOT ANALYSIS (I-III)

Expression levels of purified MutSα variants (I-II)

Expression levels of 6xHis tagged MSH6 together with counterpart MSH2 were analyzed by Western blot (6% SDS-PAGE), and compared with control protein samples with known concentrations. The primary antibodies used were MSH2 (ab-2) (Oncogene) (1:250) and MSH6 (BD Bioscience Pharmingen), (1:1000) and the secondary antibody was anti α-mouse (Amersham) (1:5000) coupled with horseradish peroxidase. The chemiluminescence reaction was visualized with X-ray film.

Expression levels of MLH1, SELENB1 and IDH1 in mouse samples (III)

The expressions of the proteins found to be important in 2D DIGE analyses such as MLH1, selenium binding protein 1 (SELENB1), and isocitrate dehydrogenase 1 (IDH1) were further studied with Western blot analyses. Forty μg of sample pools containing equal amounts of total protein extracts from each mouse belonging to a respective group (MLH1) or 15 μg of extract from each mouse in the group were used for the assay. Proteins were separated on 8% SDS-PAGE for MLH1 and 14% for SELENBP1 and IDH1, blotted on Immobilon FL-PVDF (Millipore, MA, USA), and incubated with a primary antibody. The antibodies used were MLH1, clone G168-15 (1:500) (BD Pharmingen, San Diego, CA), SELENBP1, clone NBP1-54805 (1:1000) (Novusbio, UK), and IDH1, clone #8137 (1:1000) (Cell Signaling Technology, MA, USA). The α-tubulin was used as a loading control (clone DM1A) (Sigma-Aldrich).

Two different detection systems were applied. For MLH1, the secondary antibody was the peroxidase labeled anti-mouse antibody (GE Healthcare) (1:20 000) and the blot was developed using ECL Prime Western Blotting Detection Reagents (GE Healthcare), detected with X-ray film. For SELENBP1 and IDH1, secondary antibodies were IRDye 800 CW goat anti-rabbit IgG (926-32211) (1:15 000) and IRDye 680 RD goat anti-mouse IgG (926-68070) (1:15 000) (LI-COR, Germany), respectively. The blots were scanned with Odyssey (LI-COR) and quantified with the Image Studio Lite (LI-COR) software. Loading control (α-tubulin) was used to normalize protein abundance between WD* and AIN groups and additionally relative comparison of the WD* to AIN group was performed for IDH1 and SELENBP1.

RT-QPCR (III)

The RNA expressions were studied only for the genes that showed the most significant protein expression changes in 2D DIGE. The total RNA extracts were prepared using the RNeasyPlus Kit (Qiagen) with an extra DNAse treatment (Qiagen). The RNA integrity was analyzed with

Table 5. TaqMan assays for studied genes

Gene	Assay number
Acaa1b	Mm00728805_s1
Acadm	Mm01323360_g1
Aldob	Mm00523293_m1
Atp5a1	Mm00431960_m1
Car1	Mm00486717_m1
Fabp2	Mm00433188_m1
Herpud1	Mm00445600_m1
Hspa5	Mm00517690_g1
Idh1	Mm00516030_m1
Pkm	Mm00834102_gH
Slc2a1	Mm00441480_m1
Stk4[1]	Mm00451755_m1
Taldo1	Mm00807080_g1

[1]Reference gene

Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) and only high quality RNA (RNA integrity number RIN > 8) was used for cDNA synthesis reactions, which were run as duplicates and pooled for RT-qPCR reactions. Reverse transcription of pools of samples (200ng each) from 6 different mice belonging to the group *Mlh1*$^{+/-}$ AIN and 7 from *Mlh1*$^{+/-}$ WD* was achieved with Superscript III (Life technologies, Carlsbad, CA), using random primers according to the manufacturer's instructions. (More details on RNA purification (Pussila *et al.*, 2013).

The RT-qPCR analysis was done using TaqMan assays (Table 5). Both pooled samples were assayed in triplicate for the target genes using the following cycling parameters: 1 cycle of 95 °C for 10 min, 40 cycles of 95 °C for 15 s, and 60 °C for 1 min. Serine/threonine kinase 4 (*Stk4*) was used as an endogenous reference gene. Thermal cycling and fluorescence data acquisition were performed with the StepOnePlus cycler (Life Technologies) and Cq values were obtained using the Data-assist v2.0 software (Life Technologies). The mRNA expression changes in the *Mlh1*$^{+/-}$ WD* mice were analyzed using the comparative C_t ($\Delta\Delta C_t$) method, which presents the data as fold changes in gene expression normalized to the endogenous reference gene and compared to the expressions in the control group, *Mlh1*$^{+/-}$ AIN.

STATISTICAL ANALYSES (I-III)

Bandshift assays were repeated at least four times and all data were statistically analyzed in SPSS, version 12.0.1, using one-way ANOVA and tested with Tukey's test with significance level 0.05.

The weight gains between the mouse groups were compared using the Student's t-test and the Mann-Whitney U test (Lehmann & Romano, 2005).

The protein spots with significant expression changes were detected using the DeCyder 2D Software 7.0 (GE Healthcare), and the expression means between the different mouse groups compared with the Student's t-test.

The *P* value plot was created in the PASW Statistics 18 (SPSS Inc, IBM) using *p* values of all detected proteins obtained in t-test comparisons between the groups *Mlh1*$^{+/-}$ WD* and *Mlh1*$^{+/+}$ AIN, *Mlh1*$^{+/-}$ WD* and *Mlh1*$^{+/-}$ AIN, *Mlh1*$^{+/+}$ WD* and *Mlh1*$^{+/+}$ AIN, and *Mlh1*$^{+/-}$ AIN and

Mlh1$^{+/+}$ AIN. *P* value plot reveals the number of expression changes with *p* value belonging to predefined bins and are represented with columns in the graph. Here we used: bin1: $0.05 \geq p > 0$, bin2: $0.1 \geq p > 0.05$, bin3: $0.15 \geq p > 0.1$ with an increment of 0.05 until bin20: $1 \geq p > 0.95$. An increase of small *p* values ($0.05 \geq p > 0$) will result in a skewed graph towards 0 showing an effect of a treatment, whereas an even distribution of *p* values will result in a flat graph indicating no significant effect of a treatment.

The hierarchical clustering of mice was performed in the DeCyder extended data analysis using a set of proteins with significant expression changes by applying Pearson correlation and the average linkage method incorporated in the DeCyder software package.

The mean and median permutation tests (Efron *et al.*, 2001) were used to analyze the expression data from RT-qPCR and Western blot and verify data from the 2D DIGE and. Pearson's correlation was used to analyze expressions obtained from 2D-DIGE and Western blot for selected proteins.

GENE ONTOLOGY AND NETWORK ANALYSIS (III)

Gene Ontology (GO) annotations were analyzed using the Panther Protein Classification System (http://www.pantherdb.org/) to distinguish significantly enriched functional annotations. The corresponding GO terms were used to assign protein functions. Network connections between human orthologs of differentially expressed mouse genes were studied using Gene*MANIA* (www.genemania.org), which indexes 1464 association networks containing 292680904 interactions mapped to 149747 genes from 7 organisms (last update 05/2013) (Warde-Farley *et al.*, 2010). We further applied the GO biological process (GO BP) criteria for weighting of the networks and to maximize connectivity between all input genes into a given ontology class. Functional clustering of differentially expressed proteins in the *Mlh1*$^{+/-}$ WD* mice was achieved by combining protein interaction data and filtering it with medical subject heading terms (MeSH) with Human Integrated Protein-Protein Interaction rEference database, HIPPIE (http://cbdm.mdc-berlin.de/tools/hippie/information.php) (Schaefer *et al.*, 2012, Schaefer *et al.*, 2013) using medium confidence scoring (0.63). The following MesH terms in the neoplasm category were implemented for network analysis: neoplasms by site; neoplasms, hormone dependent; neoplastic processes and precancerous conditions.

RESULTS & DISCUSSION

BIOCHEMICAL ANALYSIS OF INHERITED *MSH2* VARIATIONS (I-II)

Variants in the connector or lever domain cause protein degradation

To study biochemical defects of mutations in human *MSH2* gene, we had transfected *Sf*9 cells and purified recombinant protein variants. All protein variants were expressed, although varaints had broad difference between expression levels. Protein variants could be divided into two groups according to their expression level. MSH2 p.N127S, p.A272V, p.G322D, p. G674A were expressed similarly to MSH2 wild type and thus assessed as stable (Table 6, Fig. 1/I, Fig. 3/II). The rest of variants, except p.C697F, were expressed more than 10-fold less than MSH2 wild type. Of these, the MSH2 proteins p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p.C333Y,

Table 6. The clinical and functional characteristics and interpretation of MSH2 variants

MSH2 mutation	Domain	Clinical data			Functional data[1]				Interpretation
		AC[2]	MSI	IHC[3]	Repair	Stability	Binding	Release	
T33P	DNA binding	-	+	+	+(-)	+	+	+(-)	Inconclusive[6]
N127S[4]	Connector	+/-	+	+/-	+	+	+	+	Nonpathogenic
V161D	Connector	+	+	-	-	-	NA	NA	Pathogenic
G162R	Connector	+	+	-	-	-	NA	NA	Pathogenic
G164R	Connector	+	NA	-	-	-	NA	NA	Pathogenic
L173P	Connector	+	+	-	-	-	NA	NA	Pathogenic
L187P	Connector	+	+	-	-	-	NA	NA	Pathogenic
A272V[4]	Connector	+/-	+	+	+	+	+(-)	+	Nonpathogenic
G322D[4]	Levers	+/-	+/-	+/-	+	+	+	+(-)	Nonpathogenic
C333Y	Levers	-	NA	NA	-	-	NA	NA	Pathogenic
D603N[4]	Levers	+/-	+/-	-	-	-	NA	NA	Pathogenic
A636P	ATPase	+	+	-	-	+	-	NA	Pathogenic
G674A	ATPase	NA	NA	NA	-	+	+(-)	-	Pathogenic
C697F	ATPase	+	+	-	-	+(-)	-	NA	Pathogenic
DEL745-746	ATPase	+	+	-	-	-	NA	NA	Pathogenic
E749K	ATPase	+	+	+	-	+	+	-	Pathogenic
A834T[5]	ATPase	+	+	-	+	+	+	+	Nonpathogenic
V923E	Helix-turn-helix	+	+	-	+	+	+(-)	+(-)	Inconclusive[6]

[1] '-' deficient, '+' functional, '+(-)' reduced
[2] Fulfillment of Amsterdam criteria I or II: '-' not fulfilled, '+' fulfilled, '+/-' status varies between families
[3] Immunohistochemical analysis of MSH2 protein in a primary tumor: '-' not expressed, '+' expressed, '+/-' expression status varies between reported cases
[4] Mutations reported several times with different clinical characteristics
[5] Another *MSH2* mutation was found later and confirmed as the cause of cancer predisposition in this family
[6] Variant is functional but shows reduced efficiency in one assay

p.D603N and p.I745-I746del were highly degraded and seen as a ladder on a Western blot indicating high instability (Table 6, Fig. 2/II). The same protein variants were previously characterized as pathogenic in the *in vitro* MMR assay (Ollila *et al.*, 2006). These weakly expressed proteins, p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p., contained amino acid substitutions in the connector domain, which plays a role in allosteric signaling between domains 3 and 5 and has shown to be important for the stability of the MSH2 protein (Warren *et al.*, 2007). A closer look at the amino acid (AA) substitutions revealed tremendous changes in (bio)chemical properties of AA residues. For example in p.V161D hydrophobic valine (V) is substituted into negatively charged aspartic acid (D), in p.G162R and p.G164R the smallest AA glycine (G) is substituted into positively charged arginine (R), and in p.L173P hydrophobic leucine (L) to proline (P), whose rigid structure affects the secondary structure of the protein. Any of these substitutions within the connector domain may explain decreased MSH2 stability and its lower expression level in human cells leading to MMR deficiency and cancer. Variants p.N127S and p.A272V are also within the connector domain but protein structure may not be affected since in both variants AA substitution is quite mild; in p.N127S substitution of asparagine (N) into serine (S) preserves similar polarity of the AA residue as well as in pA272V substitution of alanine (A) to valine (V). The variants p.C333Y and p.D603N were also weakly expressed. These substitutions are located in the lever domain of MSH2, which is responsible for signal transduction between the ATPase domain and the DNA binding domain (Warren *et al.*, 2007). Structural changes in the lever domain can also explain MMR deficiency and pathogenicity.

In summary, non-conservative substitutions of amino acids, affecting polarity, charge or structure in connector and lever domains disrupts stability of the protein.

Variants in the ATPase domain impair DNA binding or ATP-mediated release of protein (I-II)

MSH2 variants p.A636P and p.C697F do not bind to a mismatch

Recombinant proteins were grouped into two groups according to the purification method, since different methods could have led to different purities and thus affect bandshift assay results. The first group was purified with FPLC and contained MSH2 wild type (WT), p.T33P, p.A636P, p.E749K, p.A834T and p.V923E, whereas the second group, MSH2 - WT, p.N127S, p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p.A272V, p.G322D, p.C333Y, p.D603N, p.G674A, p.C697F and p.I745-I746del was purified with the Ni-NTA method. Ni-NTA method was used for proteins that were not sucessfully purified with FPLC. The recombinant MSH2 proteins p.V161D, p.G162R, p.G164R, p.L173P, p.L187P, p.C333Y, p.D603N and p.I745-I746del were not stable enough to obtain sufficient protein for the bandshift assay.

Results are presented as percentages of MutSα-bound oligomers out of total amount of the labeled DNA. Oligomers without mismatch containing A●T were used as the control for unspecific binding. Specific binding efficiency of purified MutSα variants for G●T was compared with binding of the wild type MutSα. FPLC purified MutSα - WT bound to the G●T oligomers with a higher efficiency (31.6 ± 12.2%) than the Ni-NTA purified WT (17.8 ± 7.6%). The recombinant proteins p.T33P, p.N127S, p.A272V, p.G322D, p.G674A, p.E749K and p.A834T bound to the labeled oligomers with G●T mismatch with similar efficiencies as the WT (Fig. 2/I, Fig. 6/II). Statistically significant reduction in mismatch binding was detected for p.V923E (15.7 ± 6.1%, $p < 0.01$), p.A636P (2.3 ± 0.5%, $p < 0.01$) and p.G697F (1.8 ± 0.4%, $p < 0.01$) (Fig. 6/II). The MSH2 variant p.A636P is frequently found in Ashkenazi Jews and reported to predispose to CRC (Guillem *et al.*, 2003). According to our study, p.A636P and p.C697F, both variants within the ATPase domain, cause MMR deficiency because they do not bind to DNA and support the already reported clinical phenotype for variant p.A636P in Ashkenazi Jews (Guillem *et al.*, 2003). The MSH2 variant p.V923E also showed reduced, but still obtainable, binding to the heteroduplex.

MSH2 variants p.G674A and p.E749K impair release from a mismatch

All the protein variants, which were able to bind to the G●T mismatch (p.T33P, p.N127S, p.A272V, p.G322D, p.G674A, p.E749K, p.A834T and p.V923E) were further studied for their ability to release mismatch in the presence of ATP. Binding efficiency is presented as the percentage of the maximum amount of protein-bound DNA measured in a reaction without any ATP.

In the presence of 0.5 mM ATP, the binding efficiencies of MutSα - WT purified with Ni-NTA and FPLC were 12.9 ± 6.8% and 12.4 ± 6.2%, respectively. With the higher ATP concentrations the binding efficiencies of both MutSα purified with Ni-NTA and FPLC were reduced (for 1 mM ATP 7.2 ± 3.2% and 6.1 ± 2.9% and for 2mM ATP 7.1 ± 4.3% and 5.0 ± 3.0%, respectively). All other recombinant proteins, except p.G674A and p.E749K, released labeled oligomers after addition of 0.5 mM ATP. No significant differences between binding efficiency in samples after addition of 1 mM or 2 mM ATP were detected. With three different ATP concentrations p.G674A showed lower DNA release than WT protein (0.5 mM: 65.5 ± 21.8%, $p<0.001$; 1 mM: 47.0 ± 19.5%, $p < 0.001$; 2 mM: 33.6 ± 23.1%, $p < 0.01$). Also p.E749K displayed reduced release from DNA in reactions with 1mM (21.2 ± 15.0%, $p < 0.05$) and 2mM ATP (21.1 ± 16.3%, $p < 0.05$). The results are presented in Fig. 7/II.

The protein variants p.G674A and p.E749K have amino acid (AA) substitutions in highly conserved motifs of the ATPase domain, which is responsible for the ATPase activity of MutSα. Substitution p.G674A is located in the Walker A motif (Fig. 7 A, B) and previous studies have

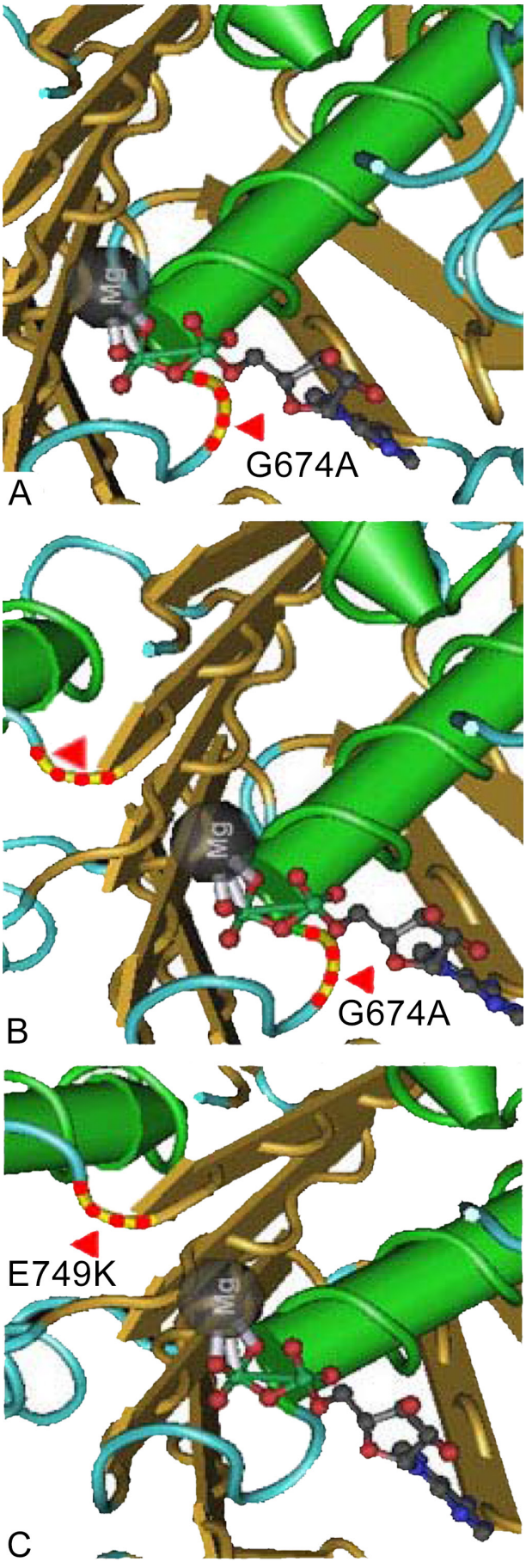

Fig. 7 Variants MSH2 p.G674A in Walker A and MSH2 p.E749K in Walker B containing a molecule of ATP. A Variant p.G674A - substitution of glycine to alanine is close to the ATP binding site, which confirms the results that protein variant p.G674A releases less DNA in the presence of ATP. B Location of both variants p.G674A and p.E749K C Variant p.E749K in the Walker B motif is in the vicinity of the Mg^{2+} ion, and ATP. Mg^{2+} is necessary for normal MMR activity (Iaccarino et al. 2000) and it is possible that this protein variant affects Mg^{2+} binding. Glutamate has a negative charge, which stabilizes Mg^{2+} unlike lysine, which is positively charged. α - helix is in green, β - sheets in gold, red arrows and yellow – red stripes show amino acid substitutions. Cn3D software (NCBI) was used for 3D visualization of the protein.

shown that variants in this motif affect MMR efficiency (Iaccarino *et al.*, 1998). Moreover, p.G674A has been shown to affect protein ability to signal DNA damage and apoptosis (Lin *et al.*, 2004). According to our results, the variant was normally expressed but MMR deficient. It was able to bind to heteroduplex DNA, but in the presence of ATP did not release normally from DNA, preventing downstream MMR proteins to proceed with repair. Our study thus shows that only one methyl group instead of hydrogen in the AA residue (G → A) can seriously impair protein function.

The MSH2 p.E749K variant, which also shows problems with mismatch release, is located in another conserved motif, Walker B (Fig. 7 B, C). The corresponding mutation in bacteria (Lamers, Winterwerp & Sixma, 2003) and *Saccharomyces cerevisiae* (Studamire *et al.*, 1999) resulted in a high expression of a mutator phenotype. Our results suggest that this variant recognizes and binds to heteroduplex DNA, remaining bound to DNA in the presence of ATP, which

disables downstream functions in MMR. Generally pathogenic *MSH2* mutations causes the loss of MSH2 in the Lynch syndrome tumor (no staining detected in IHC). In a mutation carrier of p.E749K the protein was, however, present in tumor tissue (Mangold *et al.*, 2005, Ollila *et al.*, 2006) indicating that the positive MSH2 staining in IHC is not always enough for exclusion of the MMR defect.

Inconclusive MSH2 variants

The MSH variant p.V923E was expressed similarly to the WT protein and its repair efficiency in the *in vitro* MMR assay was similar to WT (Ollila *et al.*, 2006). According to Sorting Intolerant From Tolerant (SIFT) software this variant was predicted to be tolerated (Ollila *et al.*, 2006). In our study, the variant bound normally to mismatch, but in the presence of ATP, released from it with a slightly lower efficiency; it was not significant. The colon cancer patient carrying this mutation and who also carries the mutation in *MSH6* (protein variant p.S1188N) developed cancer at a later age of onset (70 y). It is likely that the p.V923E variant has just a mild pathogenic effect and only together with another variation causes cancer.

The MSH2 variant p.T33P, which is located in the DNA binding domain of MSH2 showed decreased repair efficiency in the *in vitro* MMR assay, but was still functional (Ollila *et al.*, 2006) (Tables 3 and 6). The binding of p.T33P to heteroduplex DNA and ATP-mediated release from it was as efficient as WT´s. (Fig. 6 B & Fig. 7 B/II). This variant did not show problems in binding or release capabilities. Nevertheless, the fact that patient had an early endometrial cancer, and substitution of amino acid, from polar threonine to rigid proline, suggest deficient MMR kinetics.

Nonpathogenic MSH2 variants

The MSH2 variants p.N127S, p.A272V, p.G322D and p.A834T were stable, MMR proficient (Ollila *et al.*, 2006), and as efficient in DNA binding and ATP-mediated release as the wild type protein. MSH2 p.N127S, p.G322D and p.A834T have been reported both in healthy populations and in CRC patients, providing contradictory interpretations of their pathogenicity. Moreover, many carriers have been found to also carry another MMR gene variant (Table 1/I) (Froggatt *et al.*, 1996, Fidalgo *et al.*, 2000, Chadwick *et al.*, 2001, Samowitz *et al.*, 2001, Cravo *et al.*, 2002, Chen-Shtoyerman *et al.*, 2003, Lee *et al.*, 2005, Tanyi *et al.*, 2006). Although our results suggest that these variants do not cause MMR deficiency, we cannot exclude low penetrance effects, and connect these variants with low cancer risk.

Overall, our results classified 12 out of 18 variants as pathogenic, 4 as non-pathogenic and 2 we were not able to classify (inconclusive) (p.44 Table 6).

PROTEOMIC ANALYSIS OF THE EFFECTS OF DIET ON CRC DEVELOPMENT (III)

The lifetime risk for CRC in MMR gene mutation carriers can be as high as 80%, but disease phenotypes such as age of onset and tumor spectrum vary considerably between mutation carriers (Lindor *et al.*, 2008) suggesting that environmental factors are involved in carcinogenesis also in the CRC syndrome. Westernized dietary habits in particular have been recognized as a high risk involved in CRC tumorigenesis explaining 50 - 80% of CRCs (Nystrom & Mutanen, 2009). Diet does not necessarily generate mutations in DNA sequence as carcinogenic substances do, but most probably induces expression changes in key regulatory genes affecting normal metabolic processes in a cell.

Mice fed with Western diet show increased weight gain and trend in development of neoplasias

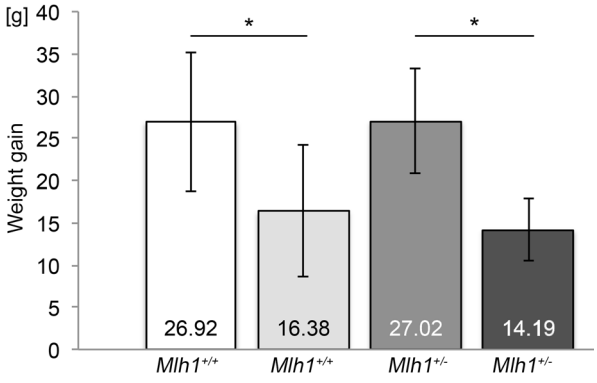

Fig. 8 Average weight gains of different mouse groups. The mice fed with WD* (both genotypes *Mlh1*$^{+/+}$ and *Mlh1*$^{+/-}$) show increased weight gain in comparison to the mice fed with AIN, whereas genetic background does not affect weight gain. * significant difference between the groups; error bars represent SD.

Table 7. Type of neoplasms in mice

Neoplasms	*Mlh1*$^{+/-}$ WD*	*Mlh1*$^{+/-}$ AIN	*Mlh1*$^{+/+}$ WD*	*Mlh1*$^{+/+}$ AIN
Hyperplastic polyp	1	1	0	0
Adenoma	2	0	1	0
Adenocarcinoma	0	0	1	0

Western diet has previously been associated with obesity (Fung *et al.*, 2001); therefore in this study we followed weight gains of the mice in different diet groups. The weight gain is defined as increase in weights between tp0 and tp1. The average end weights in groups *Mlh1*$^{+/+}$ AIN (N = 8), *Mlh1*$^{+/-}$ AIN (N = 6), *Mlh1*$^{+/+}$ WD* (N = 8) and *Mlh1*$^{+/-}$ WD* (N = 7) were 26.75 g, 29.67 g, 41.13 g, and 40.86 g, respectively. The average weight gain of the experimental groups indicate no genotype effect on weight gain (Fig. 8). Therefore, the average weight gain in mice fed with WD* and AIN was 26.97 ± 7.77 g and 15.13 ± 4.85 g, respectively, indicating a significant diet effect on weight gain (p = 0.00005).

Furthermore, macroscopic findings in colons of 29 mice at tp1 were five adenomas/hyperplastic polyps and one proximal adenocarcinoma (Table 7, Fig. 9 modified from Pussila *et al.*, 2013). Mice fed with WD* developed a substantial number of neoplasias, which

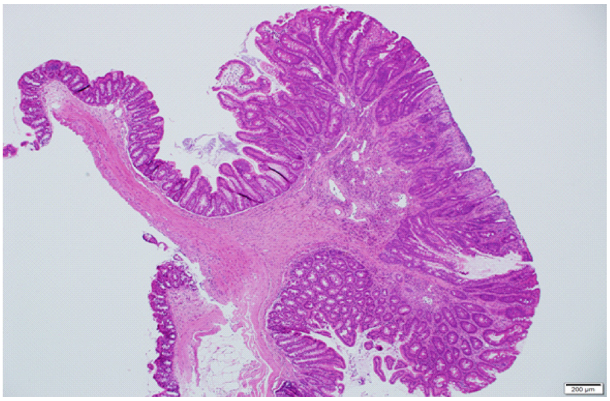

Fig. 9 Histological image of a colonic adenocarcinoma, which was found in the *Mlh1*$^{+/+}$ mouse fed with WD* (modified from Pussila *et al.* 2013).

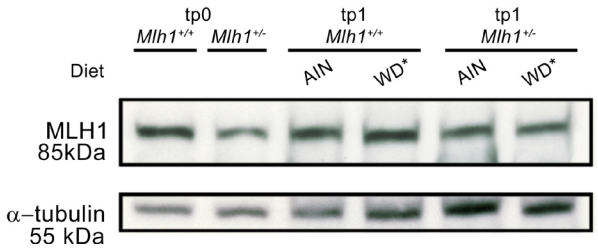

Fig. 10 The expression of the MLH1 protein analyzed with Western blot. The heterozygote *Mlh1*$^{+/-}$ mice express approximately half of the amount of MLH1 expressed in their wild type littermates, before and after the feeding experiment.

highlights the significance of WD* effect on CRC risk. The single AIN fed mouse that developed neoplasm (hyperplastic polyp) was a mutation carrier. Previous studies on Western-style diet effects on CRC risk exposed WD as an inducer of benign and malignant colonic neoplasias also in wild type mice without inherited predisposition or carcinogenic treatments (Newmark *et al.*, 2001, Newmark *et al.*, 2009). Although in that study mice were 18 months old while in our study only 12 months old, the only difference between the diets was that WD* contained a substantial amount of animal fat. Newmark & colleagues used vegetable fat, suggesting that animal fat may have accelerated tumorigenesis.

According to Knudson's two hit hypothesis (Knudson, 1971, Berger, Knudson & Pandolfi, 2011) hereditary cancers occur only after the second hit in the tumor suppressor gene. To clarify whether the second hit already occurred during the feeding experiment in the *Mlh1*$^{+/-}$ mice we analyzed the expression of the *Mlh1* gene in normal mucosa. Western blot analysis showed that *Mlh1* was still well expressed in all mice groups demonstrating that MMR deficiency was not yet detectable in colon mucosa (Fig. 10).

Differences in protein expression were mainly found between the *Mlh1*$^{+/-}$ WD* and AIN mice

In proteomic analysis we found 394 protein spots with differentially changed expression due to aging, diet and aging and diet itself. However, twelve months old *Mlh1* mutation carriers fed with WD* expressed the most significant expression differences when compared to mice fed with AIN control diet. The homozygote *Mlh1*$^{+/+}$ mice did not generally show such a strong difference between the diet groups, as can be seen in the *p*-values of all the detected expression changes, which spread uniformly between 0 and 1 when *Mlh1*$^{+/+}$ diet groups were compared to each other. On the contrary, the comparison between *Mlh1*$^{+/-}$ WD* mice and the AIN groups showed the highest number of significant *p*-values (Fig. 1/III).

Despite insignificant protein expression changes between diet groups in *Mlh1*$^{+/+}$ mice, expressions showed often the same trends in the wild type *Mlh1*$^{+/+}$ and the genetically predisposed *Mlh1*$^{+/-}$ mice, which can be seen by similar expression patterns in hierarchical clustering (Fig. 2B/III). However, the most significant protein expression differences were mainly found between the *Mlh1*$^{+/-}$ WD* and the AIN groups. Moreover, after 12 months feeding, the *Mlh1*$^{+/-}$ mice clearly segregated into two different expression clusters according to the consumed diet (Fig. 2A/III), while the *Mlh1*$^{+/+}$ mice showed a heterogeneous clustering.

Western diet affects proteins involved in energy metabolism and cellular toxicity

The statistically significant differential abundances of protein spots at tp1 were mainly found between the *Mlh1*$^{+/-}$ WD* and the two AIN groups, thus the proteins for identification were selected exclusively from those comparisons. Altogether, 18 out of the 25 selected proteins were successfully identified by peptide mass fingerprint (Table 8, Fig. 11). Eleven proteins (ACADM, ALDOB, ATP5A1, CAR1, FABP2, IDH1, KRT19, LGALS4, SELENBP1, TAL-

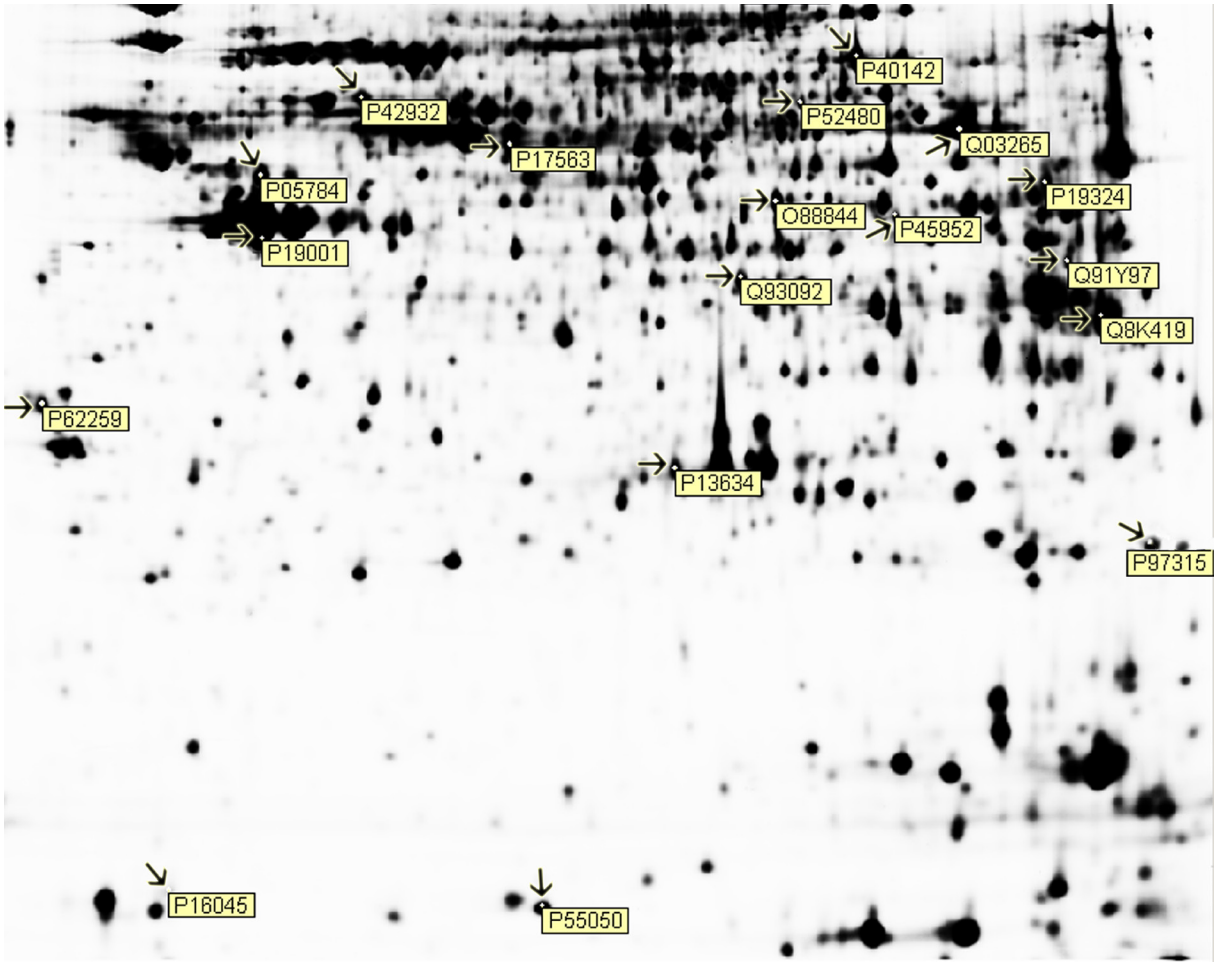

Fig. 11 The image of a representative 2D gel with positions of 18 identified proteins (indicated by arrows) and named according to UniProt ID.

Table 8. The 18 proteins showing statistically significant expression differences between the *Mlh1*$^{+/-}$ WD* and the two AIN groups

Spot No.	UniProt ID	NCBI Gene ID	NCBI Gene Symbol	UniProt Protein Name	Enzyme	*Mlh1*$^{+/-}$ WD* / *Mlh1*$^{+/-}$ AIN				*Mlh1*$^{+/-}$ WD* / *Mlh1*$^{+/+}$ AIN			
						Average ratio	*T*-test	permutation test		Average ratio	*T*-test	permutation test	
								mean	median			mean	median
1	P13634	12346	*CAR1*	Carbonic anhydrase 1	4.2.1.1	1.94	0.015	0.08	0.116	1.64	0.022	0.023	0.048
2	P17563	20341	*SELENBP1*	Selenium binding protein 1		1.55	0	0.019	0.047	1.26	0.044	0.049	0.05
3	P55050	14079	*FABP2*	Fatty acid-binding protein, intestinal		1.48	0.008	0.009	0.024	1.29	0.017	0.018	0.024
4	P45952	11364	*ACADM*	Medium-chain specific acyl-CoA dehydrogenase, mitochondrial	1.3.99.3	1.24	0.004	0.01	0	1.17	0.02	0.02	0
5	Q91Y97	230163	*ALDOB*	Fructose-bisphosphate aldolase B	4.1.2.13	1.19	0.014	0.006	0.008	1.14	0.075	0.078	0.06
6	O88844	15926	*IDH1*	Isocitrate dehydrogenase [NADP] cytoplasmic	1.1.1.42	1.18	0.001	0.003	0	1.12	0.012	0.011	0
7	Q8K419	16855	*LGALS4*	Galectin-4 (Lactose-binding lectin 4)		1.17	0.019	0.041	0.032	1.05	0.42	0.416	0.523
8	Q93092	21351	*TALDO1*	Transaldolase	2.2.1.2	1.16	0.015	0.024	0.019	1.1	0.056	0.057	0.193
9	Q03265	11946	*ATP5A1*	ATP synthase subunit alpha, mitochondrial		1.14	0.005	0.003	0.026	1.07	0.047	0.051	0.014
10	P19001	16669	*KRT19*	Keratin, type I cytoskeletal 19		1.13	0.011	0.014	0.003	1.06	0.063	0.059	0.134
11	P40142	21881	*TKT*	Transketolase	2.2.1.1	1.07	0.013	0.041	0.03	1.04	0.218	0.219	0.229
12	P42932	12469	*CCT8*	T-complex protein 1 subunit theta		-1.09	0.002	0.003	0.003	-1.02	0.726	0.735	0.617
13	P62259	22627	*YWHAE*	14-3-3 protein epsilon		-1.09	0.014	0.026	0.03	-1.05	0.055	0.056	0.039
14	P05784	16668	*KRT18*	Keratin, type I cytoskeletal 18		-1.19	0	0.002	0	-1.16	0	0	0
15	P52480	18746	*PKM2*	Pyruvate kinase isozymes M1/M2	2.7.1.40	-1.3	0.016	0.025	0.052	-1.19	0.05	0.039	0.144
16	P19324	12406	*SERPINH1*	Serpin H1		-1.33	0	0.01	0.014	-1.1	0.042	0.042	0.113
17	P16045	16852	*LGALS1*	Galectin-1		-1.33	0.001	0.028	0.011	-1.08	0.062	0.068	0.15
18	P97315	13007	*CSRP1*	Cysteine and glycine-rich protein 1		-1.59	0.005	0.008	0.023	-1.3	0.042	0.034	0.021

DO1, TKT) showed an increased and 7 proteins (CCT8, CSRP1, KRT18, LGALS1, PKM2, SERPINH1, YWHAE) a decreased expression in the *Mlh1*$^{+/-}$ WD* group when compared to the AIN groups (Table 8). All the differences were significant when the *Mlh1*$^{+/-}$ WD* group was compared to the *Mlh1*$^{+/-}$ AIN group (fold changes vary between 1.94 and -1.53). The expressions of ACADM, FABP2, IDH1, and KRT18 were also significantly changed when compared between the *Mlh1*$^{+/-}$ WD* and the *Mlh1*$^{+/+}$AIN. All expressions used are calculated as percentages and in comparisons between *Mlh1*$^{+/-}$ WD* and *Mlh1*$^{+/-}$ AIN mouse groups.

Gene ontology analysis of identified proteins expectedly associated 11 proteins with metabolic processes (GO:0008152), whereas 6 proteins were associated with cellular processes (GO:00099987) (Fig. 5A/III). Panther database analysis divided the proteins into several different classes, 4 proteins belonging to transferases (PC:002200), 3 to oxidoreductases (PC:00176), and 3 to lyases (PC:00144) (Fig. 5B/III).

Western diet disrupts cellular ROS homeostasis

Most expression changes associate with metabolic processes and neutralization of cell toxicity (Fig. 12). We detected the up-regulation of fatty acid-binding protein 2 (FABP2) in the *Mlh1*$^{+/-}$ WD* mice by 48% (p = 0.008), which most probably indicates a response to the high fat content of Western diet (20% in WD* vs. 7% in AIN). FABP2 is expressed exclusively in the intestine and its loss has been shown to cause assimilation of fatty acids in a cell (Agellon, Toth & Thomson, 2002). Free fatty acids are catabolized in fatty acid oxidation (FAO) that involves medium-chain specific acyl-CoA dehydrogenase (ACADM), which was increased by 24% (p = 0.004) in the *Mlh1*$^{+/-}$ WD* mice indicating a classical sign of starvation caused by decreased glucose levels (Klip *et al.*, 1994, Tolwani *et al.*, 2005) since the amount of carbohydrates in WD* was reduced by 20% (Table 4). Western style diet conditions cellular metabolism to utilize fatty acids for their major energy source instead of glucose, which is usually done through the ketone body pathway. Additional verification studies on glucose uptake and fatty acid metabolism showed increased mRNA expressions of solute carrier family 2 member 1 (*Scl2a1* known as *Glut1*) by 22% (p = 0) and (acetyl-Coenzyme A acyltransferase 1B) *Acaa1b*, involved in FAO of long chain fatty acids in peroxisomes, by 160% (p = 0).

Up-regulated FAO and increased ATP synthase subunit α (ATP5A1) up to 14% (p = 0.005), which is part of OXPHOS, suggests an increase of reactive oxygen species (ROS) (Turrens, 2003). Additionally, network analysis by Gene*MANIA* underlined ATP synthesis coupled proton transport (p = 7.05e-10, six proteins in the subnetwork including ATP5A1) (Fig. 6/III). Although the expression changes were detected in histologically normal colon mucosa even a moderate increase in ROS has been shown to damage proteins and DNA, regulate signaling, promote inflammation and eventually lead to cell death and cancer (Meng, Fukada & Tonks,

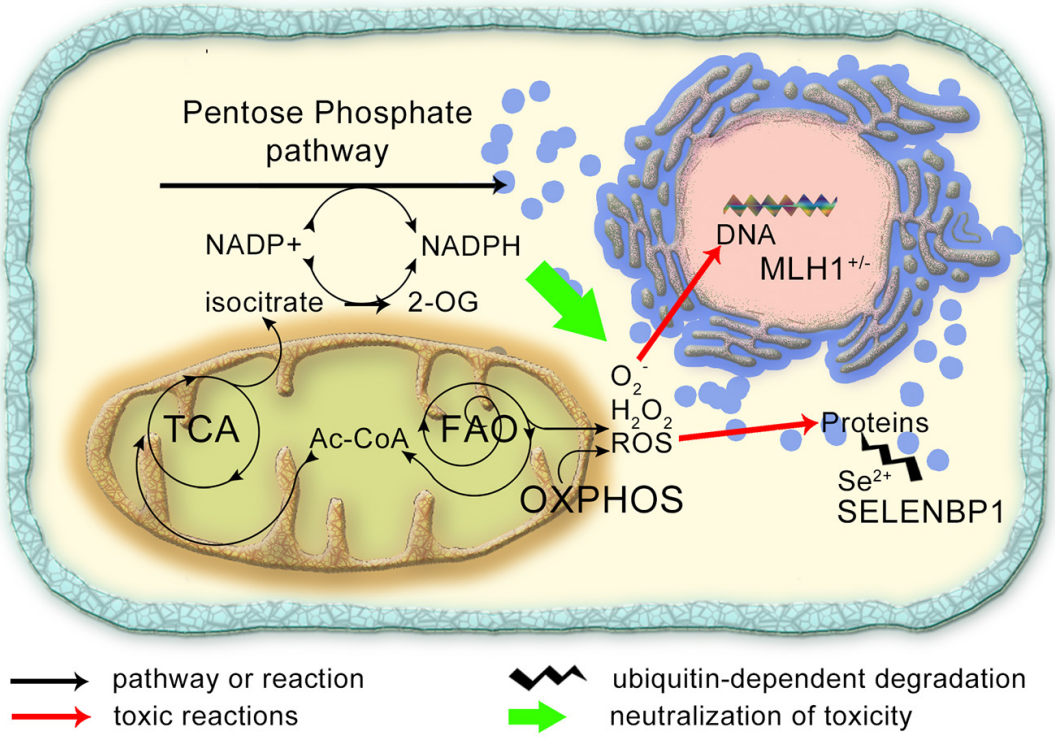

Fig. 12 Affected processes in normal colon mucosa of *Mlh1*$^{+/-}$ mice after mice were fed with WD* for twelve months. Western diet shifts energy metabolism towards fatty acid oxidation (FAO) and fatty acids become primary source of energy, which leads to upregulated oxidative phosphorylation (OXPHOS) and results in increased levels of reactive oxygen species (ROS). Increased levels of ROS damage proteins and DNA. Cellular mechanisms for reduction of ROS toxicity require nicotineamide adenine dinucleotide phosphate (NADPH). Upregulated pentose phosphate pathway and reaction catalized by isocitrate dehydrogenase 1 produce NADPH, which is used for cellular defense of ROS. Damaged proteins may be degraded in process of ubiquitination, which may be mediated in selenium dependent manner by upregulated SELENBP1.

2002, Wu, 2006, Lu, Ogasawara & Huang, 2007, Veal, Day & Morgan, 2007, Lou *et al.*, 2008, Trachootham, Alexandre & Huang, 2009, Hamanaka & Chandel, 2010, Levine & Puzio-Kuter, 2010, Sena & Chandel, 2012). Furthermore, ROS levels are altered in many diseases (e.g. diabetes, obesity, cancer) and attributed to pathology (Halliwell, 2007, Nathan, 2008, Trachootham, Alexandre & Huang, 2009).

The increased expression of *Scl2a1*, which has been reported to be increased not only because of low levels of glucose but also because of a high concentration of ROS (Bloch-Damti & Bashan, 2005), supports the elevation of ROS in the colon mucosa of the *Mlh1*$^{+/-}$ WD* mice.

Neutralization of ROS by increased expression of enzymes that generate NADPH

The cellular response to increased ROS levels is to neutralize their toxicity. In order to protect from ROS toxicity, enzymes involved in the production of NADPH are usually up-regulated (Matthews, Howarth & Butler, 2006, Kim *et al.*, 2012). We detected two enzymes, transaldolase

(TALDO1) and transketolase (TKT), involved in NADP+ reduction in the pentose phosphate pathway (PPP), which were up-regulated by 16% ($p = 0.015$) and 7% ($p = 0.013$), respectively. Additionally, isocitrate dehydrogenase 1 (IDH1), which has recently been shown to function as a tumor suppressor gene (Parsons *et al.*, 2008) and aids in total cellular production of NADPH for neutralization of ROS (Kim *et al.*, 2012), was up-regulated in the *Mlh1*$^{+/-}$ WD* mice by 18% ($p = 0.001$). We confirmed expression of IDH1 with Western blot (35% higher expression ($p = 0.035$) in *Mlh1*$^{+/-}$ WD*), which highly correlated (R = 0.8137, $p = 0.0042$) with expression changes we detected in the 2D-DIGE experiment (Fig. 3/III). Furthermore, network analysis in Gene*MANIA* revealed the importance of the oxidoreduction coenzyme metabolic processes ($p = 1.47e-4$, four proteins including ALDOB, IDH1, TALDO1 and TKT) and the nicotinamide nucleotide metabolic processes ($p = 9.77e-5$, four proteins including ALDOB, IDH1, TALDO1 and TKT) (Fig. 6 A, B/III) supporting our findings in proteomic analyses and underlining the Western diet insult on cellular ROS homeostasis.

mRNA expression of energy metabolism genes correlates with their protein expression

Energy metabolism linked genes with highest expression differences were investigated at the mRNA level. The genes, *Acadm,* fructose-bisphosphate aldolase B (*Aldob*)*, Atp5a1,* carbonic anhydrase 1 (*Car1*)*, Fabp2, Idh1* and *Taldo1* showed similar trends in expression changes at both the RNA and protein levels. The mRNA expression of *Atp5a1* was not significantly changed between the diet groups and the pyruvate kinase isozyme M1/M2 (*Pkm2*), which was shown to have a significantly lower protein expression in the *Mlh1*$^{+/-}$ WD* group than in the AIN groups, was unexpectedly up-regulated at the mRNA level, suggesting that expression may be regulated at the protein level (Fig. 4 A-H/III).

Increased expression of SELENBP1 and LGALS4 – cellular response to neoplastic processes

Since statistically significant protein expression changes were mainly found in the heterozygote *Mlh1*$^{+/-}$ mice, we took further effort to look for functional connections between the 18 proteins and MLH1 using the HIPPIE database.

In the HIPPIE network analysis, we stringently filtered protein interaction data with specific neoplastic MeSH terms. Analysis highlighted selenium binding protein 1 (SELENBP1) and galectin-4 (LGALS4) as important proteins in neoplastic/precancerous processes (Fig. 13). Both proteins directly interact with MLH1 and in the proteomic study showed an increase of 55% ($p = 0.000$) and 17% ($p = 0.019$) in *Mlh1*$^{+/-}$ WD*, respectively. SELENBP1 was additionally studied in quantitative Western blot and showed higher expression (93% increase $p = 0.02$). Moreover, Western blot expression highly correlated with expression changes we detected in 2D-DIGE experiment (R = 0.9013 $p = 0.0004$) (Fig. 3/III). The increased expression of

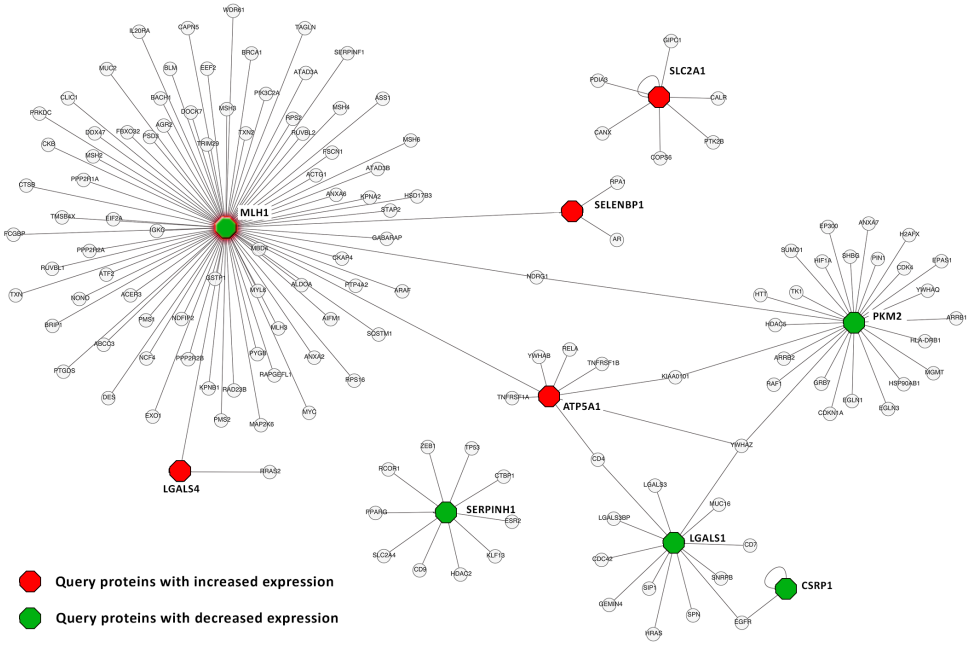

Fig. 13 The Hippie network analysis of the 18 differentially expressed proteins with the filter for neoplasms. MLH1 is directly linked to ATP5A1, LGALS4 and SELENBP1 and via bridging nodes to CSRP1, LGALS1 and PKM2 The medium confidence threshold (0.63 - second quartile in score distribution) was used for assigning the confidence of interactions.

SELENBP1 has been previously reported in precursor lesions caused by chemical carcinogens (Scortegagna *et al.*, 2009). SELENBP1 may also play a role in selenium-dependent ubiquitination/deubiquitination-mediated degradation of proteins (Jeong, Wang & Sytkowski, 2009) and have a tumor suppressor function inhibiting cell proliferation in the presence of ROS, such as H_2O_2 (Pohl *et al.*, 2009). These findings support our hypothesis that here its expression is increased to help ubiquitination of proteins in the *Mlh1*$^{+/-}$ WD* mice and possibly prevent cell proliferation. The LGALS4 is a tumor suppressor gene known to down-regulate Wnt signaling target genes (Satelli *et al.*, 2011). Thus, its increased expression also reflects a pre-tumorigenesis state where cells are trying to suppress pathogenesis.

Additionally, the HIPPIE network analysis of the protein interaction data further linked ATP5A1 directly to MLH1, while cysteine and glycine rich protein 1 (CSRP1), galectin-1 (LGALS1) and PKM2 were linked to MLH1 via bridging nodes.

Unfolded protein response may be disrupted in mice fed with Western diet

Serpin peptidase inhibitor clade H (SERPINH1, known as heat shock protein 47) associated with protein unfolding showed a 33% lower expression ($p = 0.001$) in the *Mlh1*$^{+/-}$ WD* group than in the *Mlh1*$^{+/-}$ AIN group, indicated possible changes in unfolded protein processes.

In an attempt to confirm that the found expression changes really reflect the problems in un-

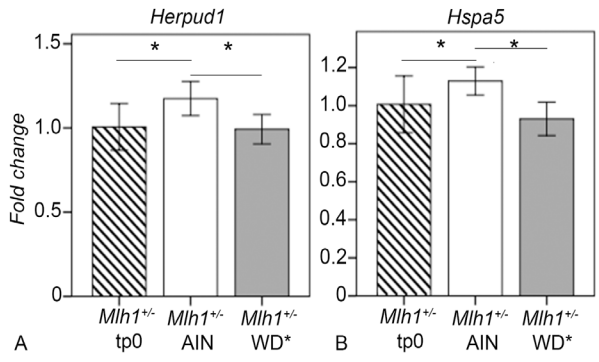

Fig. 14 The mRNA expressions of *Herpud1* (A) and *Hspa5* (B) in the different mouse groups. The fold changes of *Mlh1*$^{+/-}$ AIN and *Mlh1*$^{+/-}$ WD* were normalized and compared to *Mlh1*$^{+/-}$ tp0. Data is presented as fold change mean ± SD (n = 3). * significant difference between the groups.

folded protein processes we decided to study unfolded protein response (UPR). Heat shock protein 5 (*Hspa5*, also known as *Bip* or *Grp78*) and homocysteine-inducible and endoplasmic reticulum stress-inducible ubiquitin-like domain member 1 (*Herpud1*) were selected for quantitative mRNA expression analysis as markers for UPR. UPR markers are actively transcribed, resulting in increase of mRNA amount, upon immediate stress related to unfolded proteins. However, *Hspa5* and *Herpud1* showed 20% ($p = 0$) and 18% ($p = 0$) lower mRNA expressions in the *Mlh1*$^{+/-}$ WD* mice when compared to the *Mlh1*$^{+/-}$ AIN mice, respectively (Fig. 14). In a previous study on diabetic mice the decreased level of *Hspa5* in older mice was suggested to associate with disrupted UPR and development of insulin resistance (Yamagishi *et al.*, 2012), which may indicate that mice in our study developed insulin resistance or diabetes due to Western diet. Along with emerging studies, which imply UPR as a cellular response on ROS induced oxidative stress (Cullinan & Diehl, 2006) linking ER stress to insulin resistance (Ozcan *et al.*, 2004) and that insulin-like growth factor (IGF-I) promotes tumorigenesis by stimulation of PI3K-Akt signaling (Gallagher & LeRoith, 2010, Weissman *et al.*, 2012) our results shed light on this interesting field and underline need of further more specific mechanistic studies on this phenomenon.

Expression changes of detected proteins and CRC tumorigenesis

Many of the detected changes in protein expression (ACADM, ALDOB, ATP5A1, CAR1, LGALS1, LGALS4, and TKT) showed opposite expression trends than had previously been reported in colon cancer tissue (Barrow, Rhodes & Yu, 2011, Hu *et al.*, 2011). However, the changes we found all occur prior to the second hit in the predisposed gene (*Mlh1*) and may be the most initiative alterations in a multistep tumorigenesis. Interestingly, if we take a broad view of detected changes in normal colon mucosa, they reflect the nutritional conditioning of colon mucosa and impose a mind-provoking thought as to whether normal cells can be clonally selected by nutritional insult to eventually become cancer cells.

CONCLUSIONS

The study on biochemical defects of 18 *MSH2* variants provided the following conclusions:

- Pathogenicity of studied variants associated with disruption of different biochemical mechanisms, which correlated with the functional domains of MSH2.
- Variants in the ATPase domain did not affect expression of proteins or stability but either reduced/abolished its mismatch binding or release capability.
- The 12 variants, which were shown to be pathogenic, also associated with the Lynch syndrome phenotype. The 6 variants with low penetrance were often indistinguishable from wild type protein, suggesting that in these families, other risk factors may contribute more than the inherited variation to cancer development.

Quantitative analysis of *Mlh1*$^{+/-}$ and *Mlh1*$^{+/+}$ mice colon proteomes after feeding experiment led us to conclude the following:

- Colon proteomes of the *Mlh1*$^{+/-}$ mice showed significant expression differences after feeding with two different diets, WD* and AIN-93G.
- Network analysis underlined direct functional connections of the targeted proteins and implicated their up-regulation in early neoplastic processes.
- Western diet increases the amount of fatty acids, up-regulates fatty acid oxidation, oxidative phosphorylation, and thus also ROS, which is known to increase rate of DNA damage, supporting findings by Erdelyi *et al*. (Erdelyi *et al.*, 2009).
- Many proteins, which are involved in neutralization of toxicity, were induced, supporting the idea that increase of ROS in histologically normal tissue may eventually lead to cancer development.

FUTURE PROSPECTS

Biochemical characterization of MSH2 variants could be continued by studies on:

- Effects of mutations on MMR-mediated response to cytotoxic drugs
- Effects of mutations on MMR-mediated apoptotic signaling
- Effects of mutations on interactions between MutSα and MutLα complexes

The mouse diet study opened several interesting directions for future studies:

- UPR and ER stress linked to Western-style diet exposure should be studied in more detail and mechanically confirmed.
- Results point towards disrupted metabolic pathways in colonocytes, therefore a metabolomic study would provide a mechanistic link between WD* and changes detected on the proteome level.
- Up-regulated fatty acid oxidation seen in the study increases the cellular pool of acetyl-CoA, which affects cellular levels of acetylation modification processes. It would be interesting to access the genome wide status of histone acetylation and its possible association with oncogenic activation.

ACKNOWLEDGEMENTS

Reason and logic, base of science, are predictable and reliable, but to understand secrets of nature many times we have to confide our emotions, intuition and gut feeling.

Quite a few said to me that Acknowledgements are the most interesting part of the doctoral dissertation. Probably because you finally get to feel person behind all that dry scientific writing.

Keeping that in mind, I have to admit, I started to think about this very chapter way before I knew any line of the content of this dissertation, but left writing for a nice summer evening accompanied with few glasses of apple cider. Summer evening idea did not turn out to be that great because of all mosquitos, but that is charm of Finnish Summers, anyways.

To live as a foreigner in Finland, one of the first questions you encounter is - Why Finland? Honestly, I never gave any thoughts to Finland while living in Croatia. Only thing related to Finland was some story about Winter War, and even that came back to my mind once I was already here in Finland, after hearing heroic stories of defeated Russians. Yes I knew that Helsinki is the capital, but lived ignorantly in belief that NOKIA is a Korean high-tech company. In all my ignorance I was suggested to apply for one, just started, exchange scholarship between University of Zagreb and University of Helsinki, and I got selected. Initially thought as four months scholarship, turned out to be almost a yearlong experience of Finland and later determined past five years of my life.

It is hard not to know at least one person closely related to any of us who hasn't been affected by cancer. With two grandparents, both victims of different cancers I decided

ZAHVALE

Razum i logika, osnovica znanosti, su predvidljivi i pouzdani, no kako bismo razumjeli tajne prirode često moramo vjerovati emocijama, intuiciji i unutrašnjim osjećajima.

Mnogi su mi rekli kako su Zahvale najzanimljiviji dio doktorske dizertacije. Najvjerojatnije zato što konačno osjetiš osobu skrivenu iza suhog znanstvenog pisanja.

Moram priznati, imajući to na umu, počeo sam razmišljati o ovom poglavlju mnogo ranije nego što sam uopće znao sadržaj disertacije, no ostavio pisanje za jednu ljetnju večer uz nekoliko čaša jabukovače. Ideja o ljetnoj večeri u stvarnosti i nije bila baš sjajna zbog silnih komaraca, no to je ionako šarm finskih ljeta.

Jedno od prvih pitanja na koja naletiš kao stranac u Finskoj je - Zašto Finska? Iskreno, dok sam živio u Hrvatskoj, nikada nisam razmišljao o Finskoj. Jedino što sam znao o Finskoj bilo je nešto o Zimskom ratu, a i toga sam se sjetio tek kad sam došao u Finsku, nakon što sam čuo herojske priče o poraženim Rusima. Istina, znao sam da je Helsinki glavni grad, no živio ignoratski u uvjerenju kako je NOKIA korejska high-tech kompanija. U svom tom svojem neznanju prijavio sam se na natječaj za tek uspostavljenu bilateralnu razmjenu studenata između Sveučilišta u Helsinkiju i Sveučilišta u Zagrebu i bio odabran. Prvotna četiri mjeseca razmjene pretvorila su se u gotovo godinu dana u Finskoj i predodredili proteklih pet godina mog života.

Teško je ne znati barem jednu nam blisku osobu koja nije oboljela od raka. Činjenica da su i baka i djeda bili žrtve raka pobudila je moju želju da znanstvenu karijeru usmjer-

to focus my scientific career in the field of cancer. Only person in Viikki at that time researching cancer was Minna, very busy principal investigator. By tactically approaching her PhD student, at that time, I got my way in the group for lab practice. Person, without who I would never end up doing research with Minna; my longest Finnish friend, great mind, the best lab companion and in moments of scientific lethargy the one to be there and give a kick - Saara - thank you for teaching me that *I don't know* is not an answer for a scientist, but *let's try to figure out*, thank you for science we have shared. It is an honor to share publications with you.

Small project in autumn 2006 led to Master's thesis and already at that time Minna offered me possibility to go forward with doctoral studies. I was honored to be given that privilege and returned with enthusiasm in spring 2009. That was official beginning of Minna's and my scientific rollercoaster - an interesting ride full of unexpected curves. Supervision is not an easy task, involves human beings and there is more to it than just scientific mentoring. It is easy to troubleshoot a PCR, Western blot, but it takes a lot more to troubleshoot a person, and Minna has it. Especially, thank you for all scientific freedom.

Soon after I started with PhD, Saara left for a postdoc. Being only foreigner in the group, quite cocky for Finnish taste, rather straight-forwarded and opinionated with deep Southern Europe expressive mentality I wished for more international surrounding. Eventually, new student came to our group. It was around period when I started with proteomics and 2D DIGE optimizations in our lab. I had my hopes high when heard that she is half British half Finnish. She won me over with her straightforward personality, efficiency breaking the sky, reliability and responsibility in the lab performance, but I still don't know what was the reason why she

im u to područje. Jedina osoba u Viikki-ju koja je istraživala rak bila je Minna, vrlo zaposleni istraživač. Uspio sam ući u njezinu grupu i odraditi praksu u laboratoriju tako što sam strategijski pristupio njezinom PhD studentu. Osoba bez koje nikada ne bih istraživao s Minnom; moj najduži finski prijatelj, izvrstan um, najbolji lab drug i ona koja u momentima znanstvene letargije zna kako me pomaknuti - Saara - hvala ti što si me naučila da *ne znam* nije odgovor jednog znanstvenika, nego *hajmo pokušati pronaći odgovor*, hvala za znanost koju smo dijelili. Čast je dijeliti publikacije s tobom.

Mali projekt u jesen 2006 doveo je do diplomskog rada i već u to vrijeme Minna je ponudila da nastavim s PhD studijem. Bio sam počašćen njezinom ponudom i vratio se pun entuzijazma u proljeće 2009. To je bio početak Minninog i mog zajedničkog roller coaster-a - interesantna vožnja puna iznenadnih krivina. Mentorstvo nije jednostavan zadatak, uključuje ljudska bića i mnogo je više nego znanstveno usmjeravanje. Jednostavno je otklanjati PCR ili Western blot pogreške, no mnogo je teže usmjeravati osobu - Minna ima to u sebi. Posebno ti hvala na pruženoj znastvenoj slobodi.

Ubrzo, nakon što sam započeo s doktoratom, Saara je napustila grupu i započela s postdoc-om. Kao jedini stranac u grupi, poprilično nadmen za finski ukus, direktan i samouvjeren sa snažnim izražajnim južno-europskim mentalitetom želio sam međunarodno okruženje. Konačno smo dobili novu studenticu. Bilo je to otprilike kada sam započeo s proteomikom i 2D DIGE optimizacijom. Moja očekivanja su bila velika kada sam doznao da je pola Britanka pola Finkinja. Osvojila me svojom direktnošću, učinkovitosti bez limita, pouzdanosti i odgovornosti u laboratoriju, no još uvijek ne znam što je bio razlog zašto je ona odabrala raditi svoj dip-

chose to do her Master's project with me, occasionally mercurial personality. Long hours in the lab, so many optimizations, so many failures, but she was persistent together with me in casting numerous polyacrylamide gels. In these moments I tried to remember what I learned from Saara and give that back to Satu. I don't know if I succeeded, but definitely you were an exceptional companion in the lab, and more importantly, stayed a great friend in my life.

2D DIGE, an old proteomic method with a modernized twist, came to my life on a course of Basic Protein Chemistry and Proteomics organized by Marc. I saw 2D gels and protein spots on them (Fig.10) as a night sky with stars, and that all these spots may lead biologists to uncharted discoveries as star constellations led explorers to new continents. Eventually I got opportunity to sail a boat on my own. Marc was always supportive and ready to troubleshoot problems. He was there with good mood, cheery attitude and new technologies to broaden scientific horizons - an important segment of PhD training.

In the last year of my PhD life, I encountered rather frustrating periods, when you try to publish article and get subsequent rejections from journals - one person from Marc's lab always found some time for my issues and frustrations, despite that in his job description is not included *mentor a guy who is hanging around your MALDI*. Many times his advices reminded me on my mum's - Rabah - one big shokran! Many supervisors could take advices from you.

Once you get list of proteins - how on Earth do you interpret that? You go one-by-one. Luckily, I had less than two dozen. Reading all articles and thinking about processes in which these proteins are involved, painting larger picture took significant amount of

lomski sa mnom, koji sam ponekad nestalan poput žive. Dugi sati u laboratoriju, mnogo pokušaja optimizacije, mnogo neuspješnih eksperimenata, no ona je bila uporna zajedno sa mnom u izlijevanju nebrojenih poliakrilamidnih gelova. U tim trenucima trudio sam se sjetit što sam naučio od Saare i prenijeti isto Satu. Ne znam jesam li uspio, no definitivno si bila izuzetan kolega u laboratoriju, i još važnije, ostala veliki prijatelj u mom životu.

Na 2D DIGE, koja je modernizirana stara proteomska metoda, sam naišao na kolegiju Osnove Proteinske Kemije i Proteomike u Marc-ovoj organizaciji. 2D gelovi i razdvojeni proteini u obliku točaka (Slika 10) podsjećali su me na noćno nebo sa zvijezdama. Kako su nekoć sazvježđa navodila moreplovce do novih kontinenata tako bi sve te točke mogle voditi biologe do novih otkrića. Kad sam dobio priliku zaploviti svojim brodom, Marc je bio pun podrške, spreman pomoći u otklanjanju problema i uvijek prisutan s dobrim raspoloženjem, živahnim nastupom i novim tehnologijama kako bi proširio znastvene vidike - važan element PhDija.

U zadnjoj godini PhDja, suočavao sam se s poprilično frustrirajućim periodima kada pokušavaš objaviti članak a časopisi uzastopno šalju odbijenice. Jedna osoba iz Marc-ovog laba je uvijek našla vremena za moje probleme i frustracije, usprkos tome što mu u opisu posla nije uključeno *budi mentor osobi koja visi okolo MALDI uređaja*. Često su me njegovi savjeti podsjetili na savjete moje mame - Rabah - jedno veliko hvala! Mnogi mentori bi se mogli ugledati na tebe.

Kada konačno dobiješ listu proteina - kako protumačiti rezultate? Kreneš jedan po jedan. Srećom sam ih imao oko dvadeset. Čitanje članaka, procesuiranje činjenica i pokušaji da se uspješno sve poveže u cjelinu uzeli su poprilično vremena. Kasnije me Maciej naučio

time. Later on, Maciej taught me how to, first sketch larger picture, and then add colors by network analysis – dziękuję.

For past consecutive four years, sometime after Halloween, the period when Finland turns into Mordor I was assembling meetings with Minna and two of my follow-up advisors, Pekka and Juha. Over coffee and cookies, we would discuss project and future steps. Your steering wes always appreciated.

When we decided to establish 2D DIGE in our lab, apparent support came from head of Division of Genetics, Tapio. Thank you for lending me your equipment for 2D DIGE and supporting Minna in acquiring Typhoon scanner.

Laura, Marjaana, Minttu, Jukka, Mariann thank you for sharing office and lab time with me. Arja Ikävalko and Outi, like two angels in the lab, always happy and with smile made work in the lab many times less stressed - kiitos. Karen thank you for efficient and fast language check up.

I have to admit I traveled on many scientific conferences and most of them would not be possible without former Graduate Program in Biotechnology and Molecular Biology, Erkki and Anita, who both were always friendly and helpful.

Even though I really like working in the wet lab every now and then I need a break. In spring 2010 I came across problem based learning course CancerBio Summer School, which ended with Amazing Cancer Race event. An unusual, very interactive in its core, this course and its main organizer, Juha Klefström, have given me ever since the biggest teaching and organizing experience. An awesome and very creative course Juha, thank you for opportunity to be part of the team, grow teaching and organizing skills.

kako prvo skicirati sliku, a potom dodati boje, koristeći analizu mreža - hvala ti.

Proteklih pet godina, negdje otprilike nakon Noći vještica, u vrijeme kada Finska postane Mordor ja bih sazvao sastanak s Minnom i svojim pratećim savjetnicima, Pekkom i Juhom. Uz kavu i kekse mi bi raspravili o projektu i budućim koracima. Vaši savjeti i usmjerenje uvijek su bili cijenjeni.

Kada smo odlučili uspostaviti 2D DIGE u našem laboratoriju, velika je podrška došla od upravitelja odijela genetike, Tapia. Hvala ti što si dopustio korištenje 2D DIGE opreme i bio podrška Minni u nabavljanju Typhoon skenera.

Laura, Marjaana, Minttu, Jukka, Mariann hvala vam na vremenu koje smo dijelili u uredu i laboratoriju. Arja Ikävalko i Outi, poput dva anđela u labu, mnogo puta ste svojim veseljem učinile rad u laboratoriju manje stresnim - hvala. Karen hvala za učinkovitu i brzu provjeru engleskog jezika.

Moram priznati da sam putovao na mnogo znanstvenih simpozija i većina ih ne bi bila moguća da nije bilo Poslijediplomskog Programa u Biotehnologiji i Molekularnoj Biologiji, Erkki-ja i Anite koji su oboje bili uvijek prijateljski i od pomoći.

Iako volim raditi u laboratoriju, ponekad mi je potrebna stanka. U proljeće 2010 pronašao sam kolegij koji je baziran na problemskoj nastavi - CancerBio Ljetna Škola koja završava s Nevjerojatnom Rak Utrkom. Vrlo neobičan, interaktivan u srži, ovaj kolegij mi je zajedno sa svojim glavnim organizatorom, Juhom Klefström, pružio najveće iskustvo u nastavi i organizaciji. Izvrstan i vrlo kreativan kolegij Juha, hvala ti za pruženu mogućnost da budem dio tima i unaprijedim svoje vještine u nastavi i organizaciji.

I have had a partner in crime in organization of Amazing Cancer Race for past four years. A person who is my oldest friend in Finland, but not a Finn; a person with mentality closer to mine than any Northerner, but still Northern enough that we both like living here; a person with whom we shared many, many coffees in the early era of Helsinki coffee scene, when cafeterias were not that fine and stylish as they are now, but we both needed that table, coffee or tea and endless talks, sometimes about cancer, sometimes about philosophy; a person who dears to tell me the truth, even is aware that I'll not take it lightly but knows that I appreciate direct word more than anything; a great listener, always there for me - Maral - امش زا رکشت اب

Plato based his soulmate theory on a myth that many, many thousands of years ago, humans were androgynous creatures, which were punished by Zeus and split into two – male and female humans. According to the theory all of us are seeking our other half. If anyone proves Plato's theory, later revised by Schopenhauer and Jung, true that are Maral and Nima themselves. I am grateful of having you guys as friends and honored of having Nima as a co-author on a publication. It was pleasure to learn from you statistics, and it is always joy to share high quality discussion with you.

Process of making friendships is slow in Finland, but the most unusual beginning of a friendship was on HNPCC symposium in 2010, when we spent the whole night drinking and talking about cultural differences, high heels, Russian girls and other serious stuff. Since then these two extraordinary girls have been sincere friends, Eevi & Mervi, thank you for many inspirational talks.

Another special friendship, which started on a bus station in Viikki in late Autumn enriched my life; talks with you cheered me

Proteklih četiri godine imao sam partnera u zločinu za organizaciju Nevjerojatne Rak Utrke. Osoba koja je moj najstariji prijatelj u Finskoj, a nije Finkinja; osoba čiji je mentalitet bliži mome nego mentalitet bilo kojeg Sjevernjaka, no juš uvijek dovoljno Sjeveran da oboje volimo živjeti ovdje; osoba s kojom sam dijelio mnogo kava u ranoj eri Helsinške kava scene kada kafeterije nisu bile tako stilski uređene kao što su sada, no nas oboje smo trebali taj stol, kavu ili čaj i beskrajne razgovore, nekad o tumorima, nekad o filozofiji; osoba koja se usudi reći mi istinu iako je svjesna da to neću lagano primiti, no zna da cijenim direknu riječ više od ičega; izvrstan slušatelj, uvijek tu za mene - Maral - mnogo ti hvala.

Platon je bazirao svoju teoriju o srodnim dušama na mitu da su prije mnogo, mnogo tisuća godina ljudi bili androgena bića koja je Zeus kaznio i razdvojio u dva spola - muškarca i ženu. Po toj teoriji svatko od nas traži svoju drugu polovicu. Ako je itko dokaz Platonove teorije o srodnim dušama, koju su kasnije obradili Schopenhauer i Jung, to su onda Maral i Nima. Zahvalan sam što vas imam za prijatelje i počašćen da je Nima ko-autor na mom članku. Bilo je zadvoljstvo učiti statistiku od tebe i uvijek je veselje dijeliti visoko kvalitetnu diskusiju s tobom.

Proces stvaranja prijateljstva u Finskoj je spor, no najneobičniji početak prijateljstva dogodio se na HNPCC simpoziju 2010, kada smo proveli cijelu noć pričajući uz piće o kulturološkim razlikama, visokim petama, Ruskinjama i ostalim ozbiljnim temama. Od tada te dvije posebne curke su bile iskreni prijatelji, Eevi i Mervi hvala za mnogo inspirativnih razgovora.

Još jedno posebno prijateljstvo, koje je počelo na busnoj stanici u Viikki-ju jednog jesenjeg dana, je obogatilo moj život, razgovori

up, thank you for support when I would feel down - Ave.

Back into 2007, after I finished lab in Minna's group I stayed in Finland as a *kesämies* in Ecological Genetics Research Unit. Once I returned back to Croatia in 2008, they gave a really tempting deal to come back as a research scientist and stay later as a PhD student. I know that there is just one person to thank that for. We had a click. She was precise and demanding as a supervisor in the lab, sometimes cocky and pain in the ass :) but we had fun and more than anything for these short two months we became friends. When I came back to Finland we had our history. You experienced the best and the worst of me. I am grateful to have you Marika in my life.

Marika brought into my life hairy 27 kg of unconditional love, as Marika used to call her. I am happy for given opportunity to share runs, walks and your moments of happy, happy joy super fast after pooping runs - Iitu. When all humans failed, when world looked like falling apart, you were the one to push your head between my arms, come close and cheer me up.

Very few Croatians live in Finland, even less that studied the same faculty in Croatia, but I got lucky to share a full year with one of them. Finally I had our (Croatian) type of coffee breaks and chitchat in our language. It was an awesome year being able to speak Croatian and not care what people around would say, being able to go spontaneously for a glass of cider after work and end up in a bar until morning. Thank you for our trips, thank you for friendship, Jasna.

When Jasna, who many times was like a bigger sister to me, moved back to Croatia two persons became very close to me, and very important in my life. With one I shared

s tobom su me uvijek razveselili, hvala ti za potporu kada bih bio depresivan - Ave.

Natrag u 2007, nakon što sam završio laboratorij u Minninoj grupi ostao sam u Finskoj kao *kesämies* u Ekološko Genetičkoj Istraživačkoj Jedinici. Nakon što sam se vratio u Hrvatsku, u 2008 ponudili su mi vrlo primamljivu poziciju da se vratim kao istraživač i nastavim kao PhD student. Znam da je iza svega toga stajala samo jedna osoba. Mi smo imali klik. Kao supervizor u labu bila je ponekad trn u oku, nadmena, precizna, zahtjevna, no mi smo oboje uživali i u tih kratkih dva mjeseca postali prijatelji. Kada sam se vratio u Finsku imali smo svoju priču. Ti si iskusila najbolje i najgore u meni. Zahvalan sam što te imam u svom životu Marika

Marika je unijela u moj život dlakavih 27 kg bezuvjetne ljubavi, kako ju je Marika zvala. Sretan sam što imah priliku trčati i ići u šetnje s tobom i biti dio tvojih manično veselih trčanja nakon što bi se pokakala - Iitu. Kada bi me svi ljudi iznevjerili, kada je svijet izgledao kao da se raspada, ti bi gurnula svoju glavu između mojih ruku, došla blizu i razveselila me.

Malo Hrvata živi u Finskoj, a još manje onih koji su studirali na istom fakultetu u Hrvatskoj, no bio sam sretnik koji je proveo godinu dana s baš takvom osobom. Konačno sam imao naš (hrvatski) tip pauzi za kavu i čakule na našem jeziku. Bila je to zakon godina kada smo pričali hrvatski ne obazirući se što će ljudi okolo reći, kada bismo spontano otišli na jabukovaču poslije posla i ostali u baru do jutra. Hvala za naša putovanja, hvala za prijateljstvo, Jasna.

Nakon što se Jasna, koja mi je bila mnogo puta kao velika sestra, vratila u Hrvatsku, postao sam vrlo blizak s dvije osobe i one su postale vrlo bitne u mom životu. S jednom

neighborhood, beers in the yard, many delicious dinners, some deep talks, but more than anything true friendship. Thank you Kukka. With another one I did not share neighborhood, but direct bus line. Number 14 was main connection to many parties, gin & tonics and many more moments we've shared. Thank you for being there when I needed moments to relax, forget about work, talk about issues. Thank you Nina for taking care of my social life. Thank you for trance!

Life is a path with many crossings and each crossing represents a choice. Each choice imposes withholds. There are no words that would express my gratitude to this person who has chosen me over herself on each crossing; who has always made life situations look nicer so I wouldn't worry; who has always overstrained herself so I wouldn't feel heaviness of life; who has fought every life challenge alone and life took care that challenges are not lacking. Mum, without your support I woudn't been here where I am. Without you I wouldn't be what I am.

At first I thought that South and North couldn't be combined, but we have shared our lives for a bit more than two years, two very dynamic years consisting of a lot of traveling, many unforgettable parties, some tears, some laughs, days with rain, days with storms, mornings with rainbows… Even though you may not see, but you calmed and balanced me with your Scandinavian mentality. Thank you for going through every single sentence in this thesis with me, questioning meaning, language and style. Thank you for staying even in the moments when I couldn't stand myself, thank you for bringing the best in me - Joni.

sam dijelio susjedstvo, piva u dvorištu, mnogo ukusnih večera, poneked ozbiljne razgovore no više od ičega istinsko prijateljstvo. Hvala Kukka. S drugom pak nisam dijelio susjedstvo, već direktnu bus liniju. Bus broj 14 bio je glavna veza za mnogo partija, gin i tonike i mnogo više trenutaka koje smo dijelili. Hvala što si bila uvijek blizu kada sam trebao opuštanje, zaboraviti posao, razgovarati o problemima. Hvala Nina što si se brinula za moj socijalni život. Hvala za trans!

Život je put s mnogo raskrižja i svako raskrižje predstavlja izbor. Svaki izbor uvjetuje odricanja. Nema riječi kojima bih izrazio zahvalnost osobi koja je na svakom raskrižju odabirala mene umjesto sebe; osobi koja je uvijek uljepšavala životne situacije kako se ja ne bih brinuo; osobi koja se uvijek prenapregnula u cilju da ne osjetim težinu života; osobi koja se borila sama sa svakim životnim izazovom, a život se uvijek pobrinuo da ih ne manjka. Mama, bez tvoje podrške nikada ne bih bio ovdje gdje jesam. Bez tebe ne bih bio to što jesam.

U početku sam smatrao da jug i sjever ne mogu biti spojeni, no mi dijelimo naše živote malo više o dvije godine, dvije vrlo dinamične godine ispunjene s mnogo putovanja, mnogo nezaboravnih tuluma, ponešto suza, ponešto smijeha, dvije godine ispunjene kišnim danima, olujnim danima i jutrima s dugama... Iako ti možda ne vidiš, no ti si me smirio i uravnotežio sa svojim skandinavskim mentalitetom. Hvala ti što si sa mnom prošao kroz svaku rečenicu ove disertacije, provjeravao smisao, jezik i stil. Hvala ti što si ostajao uz mene čak i u trenucima kada sam sebe nisam mogao izdržati, hvala ti što izvučeš najbolje iz mene - Joni.

Denis

REFERENCES

Aaltonen, L. A., Peltomaki, P., Leach, F. S., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., Powell, S. M., Jen, J., Hamilton, S. R. & Et Al. 1993. Clues to the pathogenesis of familial colorectal cancer. *Science,* 260**,** 812-6.

Aarnio, M., Sankila, R., Pukkala, E., Salovaara, R., Aaltonen, L. A., De La Chapelle, A., Peltomaki, P., Mecklin, J. P. & Jarvinen, H. J. 1999. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer,* 81**,** 214-8.

Acharya, S., Wilson, T., Gradia, S., Kane, M. F., Guerrette, S., Marsischky, G. T., Kolodner, R. & Fishel, R. 1996. hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. *Proc Natl Acad Sci U S A,* 93**,** 13629-34.

Agellon, L. B., Toth, M. J. & Thomson, A. B. 2002. Intracellular lipid binding proteins of the small intestine. *Mol Cell Biochem,* 239**,** 79-82.

Allis, C. D., Jenuwein, T. & Reinberg, D. 2007. *Epigenetics,* Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory Press.

Antony, E. & Hingorani, M. M. 2003. Mismatch recognition-coupled stabilization of Msh2-Msh6 in an ATP-bound state at the initiation of DNA repair. *Biochemistry,* 42**,** 7682-93.

Antony, E., Khubchandani, S., Chen, S. & Hingorani, M. M. 2006. Contribution of Msh2 and Msh6 subunits to the asymmetric ATPase and DNA mismatch binding activities of Saccharomyces cerevisiae Msh2-Msh6 mismatch repair protein. *DNA Repair (Amst),* 5**,** 153-62.

Baker, S. M., Plug, A. W., Prolla, T. A., Bronner, C. E., Harris, A. C., Yao, X., Christie, D. M., Monell, C., Arnheim, N., Bradley, A., Ashley, T. & Liskay, R. M. 1996. Involvement of mouse Mlh1 in DNA mismatch repair and meiotic crossing over. *Nat Genet,* 13**,** 336-42.

Barrow, H., Rhodes, J. M. & Yu, L. G. 2011. The role of galectins in colorectal cancer progression. *Int J Cancer,* 129**,** 1-8.

Benito, E., Cabeza, E., Moreno, V., Obrador, A. & Bosch, F. X. 1993. Diet and colorectal adenomas: a case-control study in Majorca. *Int J Cancer,* 55**,** 213-9.

Berger, A. H., Knudson, A. G. & Pandolfi, P. P. 2011. A continuum model for tumour suppression. *Nature,* 476**,** 163-9.

Bingham, S. & Riboli, E. 2004. Diet and cancer--the European Prospective Investigation into Cancer and Nutrition. *Nat Rev Cancer,* 4**,** 206-15.

Bingham, S. A. 1996. Epidemiology and mechanisms relating diet to risk of colorectal cancer. *Nutr Res Rev,* 9**,** 197-239.

Bingham, S. A., Day, N. E., Luben, R., Ferrari, P., Slimani, N., Norat, T., Clavel-Chapelon, F., Kesse, E., Nieters, A., Boeing, H., Tjonneland, A., Overvad, K., Martinez, C., Dorronsoro, M., Gonzalez, C. A., Key, T. J., Trichopoulou, A., Naska, A., Vineis, P., Tumino, R., Krogh, V., Bueno-De-Mesquita, H. B., Peeters, P. H., Berglund, G., Hallmans, G., Lund, E., Skeie, G., Kaaks, R., Riboli, E., European Prospective Investigation Into, C. & Nutrition 2003. Dietary fibre in food and protection against colorectal cancer in the European Prospective Investigation into Cancer and Nutrition (EPIC): an observational study. *Lancet,* 361**,** 1496-501.

Blackwell, L. J., Martik, D., Bjornson, K. P., Bjornson, E. S. & Modrich, P. 1998. Nucleotide-promoted release of hMutSalpha from heteroduplex DNA is consistent with an ATP-dependent translocation mechanism. *J Biol Chem,* 273**,** 32055-62.

Bloch-Damti, A. & Bashan, N. 2005. Proposed mechanisms for the induction of insulin resistance by oxidative stress. *Antioxid Redox Signal,* 7**,** 1553-67.

Boffa, L. C., Lupton, J. R., Mariani, M. R., Ceppi, M., Newmark, H. L., Scalmati, A. & Lipkin, M. 1992. Modulation of colonic epithelial cell proliferation, histone acetylation, and luminal short chain fatty acids by variation of dietary fiber (wheat bran) in rats. *Cancer Res,* 52**,** 5906-12.

Boland, C. R., Thibodeau, S. N., Hamilton, S. R., Sidransky, D., Eshleman, J. R., Burt, R. W., Meltzer, S. J., Rodriguez-Bigas, M. A., Fodde, R., Ranzani, G. N. & Srivastava, S. 1998. 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,* 58**,** 5248-57.

Boman, B. M. & Huang, E. 2008. Human colon cancer stem cells: a new paradigm in gastrointestinal oncology. *J Clin Oncol,* 26**,** 2828-38.

Bonello, S., Zahringer, C., Belaiba, R. S., Djordjevic, T., Hess, J., Michiels, C., Kietzmann, T. & Gorlach, A. 2007. Reactive oxygen species activate the HIF-1alpha promoter via a functional NFkappaB site. *Arterioscler Thromb Vasc Biol,* 27**,** 755-61.

Bos, J. L., Fearon, E. R., Hamilton, S. R., Verlaan-De Vries, M., Van Boom, J. H., Van Der Eb, A. J. & Vogelstein, B. 1987. Prevalence of ras gene mutations in human colorectal cancers. *Nature,* 327**,** 293-7.

Buell, P. & Dunn, J. E., Jr. 1965. Cancer Mortality among Japanese Issei and Nisei of California. *Cancer,* 18**,** 656-64.

Bulow, S. 1987. Familial polyposis coli. *Dan Med Bull,* 34**,** 1-15.

Burkitt, D. P. 1969. Related disease--related cause? *Lancet,* 2**,** 1229-31.

Cannavo, E., Marra, G., Sabates-Bellver, J., Menigatti, M., Lipkin, S. M., Fischer, F., Cejka, P. & Jiricny, J. 2005. Expression of the MutL homologue hMLH3 in human cells and its role in DNA mismatch repair. *Cancer Res,* 65**,** 10759-66.

Chadwick, R. B., Pyatt, R. E., Niemann, T. H., Richards, S. K., Johnson, C. K., Stevens, M. W., Meek, J. E., Hampel, H., Prior, T. W. & De La Chapelle, A. 2001. Hereditary and somatic DNA mismatch repair gene mutations in sporadic endometrial carcinoma. *J Med Genet,* 38**,** 461-6.

Chai, F., Evdokiou, A., Young, G. P. & Zalewski, P. D. 2000. Involvement of p21(Waf1/Cip1) and its cleavage by DEVD-caspase during apoptosis of colorectal cancer cells induced by butyrate. *Carcinogenesis,* 21**,** 7-14.

Chakrabarty, S., Radjendirane, V., Appelman, H. & Varani, J. 2003. Extracellular calcium and calcium sensing receptor function in human colon carcinomas: promotion of E-cadherin expression and suppression of beta-catenin/TCF activation. *Cancer Res,* 63**,** 67-71.

Chapkin, R. S., Fan, Y. & Lupton, J. R. 2000. Effect of diet on colonic-programmed cell death: molecular mechanism of action. *Toxicol Lett,* 112-113**,** 411-4.

Chen-Shtoyerman, R., Theodor, L., Harmati, E., Friedman, E., Dacka, S., Kopelman, Y., Sternberg, A., Zarivach, R., Bar-Meir, S. & Fireman, Z. 2003. Genetic analysis of familial colorectal cancer in Israeli Arabs. *Hum Mutat,* 21**,** 446-7.

Cho, K. R. & Vogelstein, B. 1992. Suppressor gene alterations in the colorectal adenoma-carcinoma sequence. *J Cell Biochem Suppl,* 16G**,** 137-41.

Clarke, R., Halsey, J., Lewington, S., Lonn, E., Armitage, J., Manson, J. E., Bonaa, K. H., Spence, J. D., Nygard, O., Jamison, R., Gaziano, J. M., Guarino, P., Bennett, D., Mir, F., Peto, R., Collins, R. & Collaboration,

B. V. T. T. 2010. Effects of lowering homocysteine levels with B vitamins on cardiovascular disease, cancer, and cause-specific mortality: Meta-analysis of 8 randomized trials involving 37 485 individuals. *Arch Intern Med,* 170**,** 1622-31.

Colussi, C., Parlanti, E., Degan, P., Aquilina, G., Barnes, D., Macpherson, P., Karran, P., Crescenzi, M., Dogliotti, E. & Bignami, M. 2002. The mammalian mismatch repair pathway removes DNA 8-oxodGMP incorporated from the oxidized dNTP pool. *Curr Biol,* 12**,** 912-8.

Couch, F. J., Rasmussen, L. J., Hofstra, R., Monteiro, A. N., Greenblatt, M. S., De Wind, N. & Group, I. U. G. V. W. 2008. Assessment of functional effects of unclassified genetic variants. *Hum Mutat,* 29**,** 1314-26.

Cravo, M., Afonso, A. J., Lage, P., Albuquerque, C., Maia, L., Lacerda, C., Fidalgo, P., Chaves, P., Cruz, C. & Nobre-Leitao, C. 2002. Pathogenicity of missense and splice site mutations in hMSH2 and hMLH1 mismatch repair genes: implications for genetic testing. *Gut,* 50**,** 405-12.

Cullinan, S. B. & Diehl, J. A. 2006. Coordination of ER and oxidative stress signaling: the PERK/Nrf2 signaling pathway. *Int J Biochem Cell Biol,* 38**,** 317-32.

Dalberg, J., Jacobsen, O., Nielsen, N. H., Steig, B. A. & Storm, H. H. 1999. Colorectal cancer in the Faroe Islands--a setting for the study of the role of diet. *J Epidemiol Biostat,* 4**,** 31-6.

De Wind, N., Dekker, M., Berns, A., Radman, M. & Te Riele, H. 1995. Inactivation of the mouse Msh2 gene results in mismatch repair deficiency, methylation tolerance, hyperrecombination, and predisposition to cancer. *Cell,* 82**,** 321-30.

De Wind, N., Dekker, M., Claij, N., Jansen, L., Van Klink, Y., Radman, M., Riggins, G., Van Der Valk, M., Van't Wout, K. & Te Riele, H. 1999. HNPCC-like cancer predisposition in mice through simultaneous loss of Msh3 and Msh6 mismatch-repair protein functions. *Nat Genet,* 23**,** 359-62.

De Wind, N., Dekker, M., Van Rossum, A., Van Der Valk, M. & Te Riele, H. 1998. Mouse models for hereditary nonpolyposis colorectal cancer. *Cancer Res,* 58**,** 248-55.

Deberardinis, R. J., Lum, J. J., Hatzivassiliou, G. & Thompson, C. B. 2008. The biology of cancer: metabolic reprogramming fuels cell growth and proliferation. *Cell Metab,* 7**,** 11-20.

Doll, R. & Peto, R. 1981. The causes of cancer: quantitative estimates of avoidable risks of cancer in the United States today. *J Natl Cancer Inst,* 66**,** 1191-308.

Drummond, J. T., Li, G. M., Longley, M. J. & Modrich, P. 1995. Isolation of an hMSH2-p160 heterodimer that restores DNA mismatch repair to tumor cells. *Science,* 268**,** 1909-12.

Duckett, D. R., Drummond, J. T., Murchie, A. I., Reardon, J. T., Sancar, A., Lilley, D. M. & Modrich, P. 1996. Human MutSalpha recognizes damaged DNA base pairs containing O6-methylguanine, O4-methylthymine, or the cisplatin-d(GpG) adduct. *Proc Natl Acad Sci U S A,* 93**,** 6443-7.

Dufner, P., Marra, G., Raschle, M. & Jiricny, J. 2000. Mismatch recognition and DNA-dependent stimulation of the ATPase activity of hMutSalpha is abolished by a single mutation in the hMSH6 subunit. *J Biol Chem,* 275**,** 36550-5.

Duval, A. & Hamelin, R. 2002. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. *Cancer Res,* 62**,** 2447-54.

Dzantiev, L., Constantin, N., Genschel, J., Iyer, R. R., Burgers, P. M. & Modrich, P. 2004. A defined human system that supports bidirectional mismatch-provoked excision. *Mol Cell,* 15**,** 31-41.

Edelmann, W., Cohen, P. E., Kane, M., Lau, K., Morrow, B., Bennett, S., Umar, A., Kunkel, T., Cattoretti, G., Cha-

ganti, R., Pollard, J. W., Kolodner, R. D. & Kucherlapati, R. 1996. Meiotic pachytene arrest in MLH1-deficient mice. *Cell,* 85**,** 1125-34.

Edelmann, W., Umar, A., Yang, K., Heyer, J., Kucherlapati, M., Lia, M., Kneitz, B., Avdievich, E., Fan, K., Wong, E., Crouse, G., Kunkel, T., Lipkin, M., Kolodner, R. D. & Kucherlapati, R. 2000. The DNA mismatch repair genes Msh3 and Msh6 cooperate in intestinal tumor suppression. *Cancer Res,* 60**,** 803-7.

Edelmann, W., Yang, K., Kuraguchi, M., Heyer, J., Lia, M., Kneitz, B., Fan, K., Brown, A. M., Lipkin, M. & Kucherlapati, R. 1999. Tumorigenesis in Mlh1 and Mlh1/Apc1638N mutant mice. *Cancer Res,* 59**,** 1301-7.

Edelmann, W., Yang, K., Umar, A., Heyer, J., Lau, K., Fan, K., Liedtke, W., Cohen, P. E., Kane, M. F., Lipford, J. R., Yu, N., Crouse, G. F., Pollard, J. W., Kunkel, T., Lipkin, M., Kolodner, R. & Kucherlapati, R. 1997. Mutation in the mismatch repair gene Msh6 causes cancer susceptibility. *Cell,* 91**,** 467-77.

Efron, B., Tibshirani, R., Storey, J. D. & Tusher, V. 2001. Empirical Bayes analysis of a microarray experiment. *Journal of the American Statistical Association,* 96**,** 1151-1160.

Erdelyi, I., Levenkova, N., Lin, E. Y., Pinto, J. T., Lipkin, M., Quimby, F. W. & Holt, P. R. 2009. Western-style diets induce oxidative stress and dysregulate immune responses in the colon in a mouse model of sporadic colon cancer. *J Nutr,* 139**,** 2072-8.

Fain, M. A., Shestovskaia, L. G. & Mamonova, O. F. 1975. [Possibility of determining the intravital origin of lesions according to the change in protein fraction content in the skeletal muscles]. *Sud Med Ekspert,* 18**,** 17-9.

Fearon, E. R. 2011. Molecular genetics of colorectal cancer. *Annu Rev Pathol,* 6**,** 479-507.

Fearon, E. R. & Vogelstein, B. 1990. A genetic model for colorectal tumorigenesis. *Cell,* 61**,** 759-67.

Feinberg, A. P., Gehrke, C. W., Kuo, K. C. & Ehrlich, M. 1988. Reduced genomic 5-methylcytosine content in human colonic neoplasia. *Cancer Res,* 48**,** 1159-61.

Feinberg, A. P. & Vogelstein, B. 1983. Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature,* 301**,** 89-92.

Fidalgo, P., Almeida, M. R., West, S., Gaspar, C., Maia, L., Wijnen, J., Albuquerque, C., Curtis, A., Cravo, M., Fodde, R., Leitao, C. N. & Burn, J. 2000. Detection of mutations in mismatch repair genes in Portuguese families with hereditary non-polyposis colorectal cancer (HNPCC) by a multi-method approach. *Eur J Hum Genet,* 8**,** 49-53.

Figueroa, M. E., Abdel-Wahab, O., Lu, C., Ward, P. S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H. F., Tallman, M. S., Sun, Z., Wolniak, K., Peeters, J. K., Liu, W., Choe, S. E., Fantin, V. R., Paietta, E., Lowenberg, B., Licht, J. D., Godley, L. A., Delwel, R., Valk, P. J., Thompson, C. B., Levine, R. L. & Melnick, A. 2010. Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell,* 18**,** 553-67.

Fiumicino, S., Martinelli, S., Colussi, C., Aquilina, G., Leonetti, C., Crescenzi, M. & Bignami, M. 2000. Sensitivity to DNA cross-linking chemotherapeutic agents in mismatch repair-defective cells in vitro and in xenografts. *Int J Cancer,* 85**,** 590-6.

Flores-Rozas, H. & Kolodner, R. D. 1998. The Saccharomyces cerevisiae MLH3 gene functions in MSH3-dependent suppression of frameshift mutations. *Proc Natl Acad Sci U S A,* 95**,** 12404-9.

Fodde, R., Edelmann, W., Yang, K., Van Leeuwen, C., Carlson, C., Renault, B., Breukel, C., Alt, E., Lipkin, M., Khan, P. M. & Et Al. 1994. A targeted chain-termination mutation in the mouse Apc gene results in mul-

tiple intestinal tumors. *Proc Natl Acad Sci U S A,* 91**,** 8969-73.

Froggatt, N. J., Green, J., Brassett, C., Evans, D. G., Bishop, D. T., Kolodner, R. & Maher, E. R. 1999. A common MSH2 mutation in English and North American HNPCC families: origin, phenotypic expression, and sex specific differences in colorectal cancer. *J Med Genet,* 36**,** 97-102.

Froggatt, N. J., Joyce, J. A., Evans, D. G., Lunt, P. W., Koch, D. J., Ponder, B. J. & Maher, E. R. 1996. MSH2 sequence variations and inherited colorectal cancer susceptibility. *Eur J Cancer,* 32A**,** 178.

Fuchs, C. S., Giovannucci, E. L., Colditz, G. A., Hunter, D. J., Stampfer, M. J., Rosner, B., Speizer, F. E. & Willett, W. C. 1999. Dietary fiber and the risk of colorectal cancer and adenoma in women. *N Engl J Med,* 340**,** 169-76.

Fung, T. T., Rimm, E. B., Spiegelman, D., Rifai, N., Tofler, G. H., Willett, W. C. & Hu, F. B. 2001. Association between dietary patterns and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Clin Nutr,* 73**,** 61-7.

Gallagher, E. J. & Leroith, D. 2010. The proliferating role of insulin and insulin-like growth factors in cancer. *Trends Endocrinol Metab,* 21**,** 610-8.

Gammie, A. E., Erdeniz, N., Beaver, J., Devlin, B., Nanji, A. & Rose, M. D. 2007. Functional characterization of pathogenic human MSH2 missense mutations in Saccharomyces cerevisiae. *Genetics,* 177**,** 707-21.

Genschel, J., Bazemore, L. R. & Modrich, P. 2002. Human exonuclease I is required for 5' and 3' mismatch repair. *J Biol Chem,* 277**,** 13302-11.

Giovannucci, E., Rimm, E. B., Ascherio, A., Stampfer, M. J., Colditz, G. A. & Willett, W. C. 1995. Alcohol, low-methionine--low-folate diets, and risk of colon cancer in men. *J Natl Cancer Inst,* 87**,** 265-73.

Giovannucci, E., Stampfer, M. J., Colditz, G. A., Hunter, D. J., Fuchs, C., Rosner, B. A., Speizer, F. E. & Willett, W. C. 1998. Multivitamin use, folate, and colon cancer in women in the Nurses' Health Study. *Ann Intern Med,* 129**,** 517-24.

Glickman, B., Van Den Elsen, P. & Radman, M. 1978. Induced mutagenesis in dam- mutants of Escherichia coli: a role for 6-methyladenine residues in mutation avoidance. *Mol Gen Genet,* 163**,** 307-12.

Goldgar, D. E., Easton, D. F., Byrnes, G. B., Spurdle, A. B., Iversen, E. S., Greenblatt, M. S. & Group, I. U. G. V. W. 2008. Genetic evidence and integration of various data sources for classifying uncertain variants into a single model. *Hum Mutat,* 29**,** 1265-72.

Gradia, S., Acharya, S. & Fishel, R. 1997. The human mismatch recognition complex hMSH2-hMSH6 functions as a novel molecular switch. *Cell,* 91**,** 995-1005.

Gradia, S., Subramanian, D., Wilson, T., Acharya, S., Makhov, A., Griffith, J. & Fishel, R. 1999. hMSH2-hMSH6 forms a hydrolysis-independent sliding clamp on mismatched DNA. *Mol Cell,* 3**,** 255-61.

Grau, M. V., Baron, J. A., Sandler, R. S., Haile, R. W., Beach, M. L., Church, T. R. & Heber, D. 2003. Vitamin D, calcium supplementation, and colorectal adenomas: results of a randomized trial. *J Natl Cancer Inst,* 95**,** 1765-71.

Greenblatt, M. S., Bennett, W. P., Hollstein, M. & Harris, C. C. 1994. Mutations in the p53 tumor suppressor gene: clues to cancer etiology and molecular pathogenesis. *Cancer Res,* 54**,** 4855-78.

Guillem, J. G., Moore, H. G., Palmer, C., Glogowski, E., Finch, R., Nafa, K., Markowitz, A. J., Offit, K. & Ellis, N. A. 2004. A636P testing in Ashkenazi Jews. *Fam Cancer,* 3**,** 223-7.

Guillem, J. G., Rapaport, B. S., Kirchhoff, T., Kolachana, P., Nafa, K., Glogowski, E., Finch, R., Huang, H.,

Foulkes, W. D., Markowitz, A., Ellis, N. A. & Offit, K. 2003. A636P is associated with early-onset colon cancer in Ashkenazi Jews. *J Am Coll Surg,* 196**,** 222-5.

Haenszel, W. 1961. Cancer mortality among the foreign-born in the United States. *J Natl Cancer Inst,* 26**,** 37-132.

Haenszel, W. & Kurihara, M. 1968. Studies of Japanese migrants. I. Mortality from cancer and other diseases among Japanese in the United States. *J Natl Cancer Inst,* 40**,** 43-68.

Halliwell, B. 2007. Oxidative stress and cancer: have we moved forward? *Biochem J,* 401**,** 1-11.

Hamanaka, R. B. & Chandel, N. S. 2010. Mitochondrial reactive oxygen species regulate cellular signaling and dictate biological outcomes. *Trends Biochem Sci,* 35**,** 505-13.

Hanahan, D. & Weinberg, R. A. 2011. Hallmarks of cancer: the next generation. *Cell,* 144**,** 646-74.

Harrington, J. M. & Kolodner, R. D. 2007. Saccharomyces cerevisiae Msh2-Msh3 acts in repair of base-base mispairs. *Mol Cell Biol,* 27**,** 6546-54.

Hawkins, N. J. & Ward, R. L. 2001. Sporadic colorectal cancers with microsatellite instability and their possible origin in hyperplastic polyps and serrated adenomas. *J Natl Cancer Inst,* 93**,** 1307-13.

He, T. C., Sparks, A. B., Rago, C., Hermeking, H., Zawel, L., Da Costa, L. T., Morin, P. J., Vogelstein, B. & Kinzler, K. W. 1998. Identification of c-MYC as a target of the APC pathway. *Science,* 281**,** 1509-12.

Healy, S. J., Gorman, A. M., Mousavi-Shafaei, P., Gupta, S. & Samali, A. 2009. Targeting the endoplasmic reticulum-stress response as an anticancer strategy. *Eur J Pharmacol,* 625**,** 234-46.

Heerdt, B. G., Houston, M. A. & Augenlicht, L. H. 1997. Short-chain fatty acid-initiated cell cycle arrest and apoptosis of colonic epithelial cells is linked to mitochondrial function. *Cell Growth Differ,* 8**,** 523-32.

Higginson, J. & Oettle, A. G. 1960. Cancer incidence in the Bantu and "Cape Colored" races of South Africa: report of a cancer survey in the Transvaal (1953-55). *J Natl Cancer Inst,* 24**,** 589-671.

Hu, H., Deng, C., Yang, T., Dong, Q., Chen, Y., Nice, E. C., Huang, C. & Wei, Y. 2011. Proteomics revisits the cancer metabolome. *Expert Rev Proteomics,* 8**,** 505-33.

Hunter, W. N., Brown, T., Kneale, G., Anand, N. N., Rabinovich, D. & Kennard, O. 1987. The structure of guanosine-thymidine mismatches in B-DNA at 2.5-A resolution. *J Biol Chem,* 262**,** 9962-70.

Iaccarino, I., Marra, G., Dufner, P. & Jiricny, J. 2000. Mutation in the magnesium binding site of hMSH6 disables the hMutSalpha sliding clamp from translocating along DNA. *J Biol Chem,* 275**,** 2080-6.

Iaccarino, I., Marra, G., Palombo, F. & Jiricny, J. 1998. hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSalpha. *EMBO J,* 17**,** 2677-86.

Insight 2014. Application of a 5-tiered scheme for standardized classification of 2,360 unique mismatch repair gene variants in the InSiGHT locus-specific database. *Nat Genet,* 46**,** 107-15.

Issa, J. P. 1999. Aging, DNA methylation and cancer. *Crit Rev Oncol Hematol,* 32**,** 31-43.

Jeong, J. Y., Wang, Y. & Sytkowski, A. J. 2009. Human selenium binding protein-1 (hSP56) interacts with VDU1 in a selenium-dependent manner. *Biochem Biophys Res Commun,* 379**,** 583-8.

Jiricny, J. 1994. Colon cancer and DNA repair: have mismatches met their match? *Trends Genet,* 10**,** 164-8.

Jiricny, J. 2006. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol,* 7**,** 335-46.

Jones, P. A. & Laird, P. W. 1999. Cancer epigenetics comes of age. *Nat Genet,* 21**,** 163-7.

Jones, R. G. & Thompson, C. B. 2009. Tumor suppressors and cell metabolism: a recipe for cancer growth. *Genes Dev,* 23**,** 537-48.

Kansikas, M., Kariola, R. & Nystrom, M. 2011. Verification of the three-step model in assessing the pathogenicity of mismatch repair gene variants. *Hum Mutat,* 32**,** 107-15.

Kantelinen, J., Kansikas, M., Korhonen, M. K., Ollila, S., Heinimann, K., Kariola, R. & Nystrom, M. 2010. MutSbeta exceeds MutSalpha in dinucleotide loop repair. *Br J Cancer,* 102**,** 1068-73.

Karran, P. 2001. Mechanisms of tolerance to DNA damaging therapeutic drugs. *Carcinogenesis,* 22**,** 1931-7.

Kawate, H., Sakumi, K., Tsuzuki, T., Nakatsuru, Y., Ishikawa, T., Takahashi, S., Takano, H., Noda, T. & Sekiguchi, M. 1998. Separation of killing and tumorigenic effects of an alkylating agent in mice defective in two of the DNA repair genes. *Proc Natl Acad Sci U S A,* 95**,** 5116-20.

Kim, J. Y., Shin, J. Y., Kim, M., Hann, S. K. & Oh, S. H. 2012. Expression of cytosolic NADP(+)-dependent isocitrate dehydrogenase in melanocytes and its role as an antioxidant. *J Dermatol Sci,* 65**,** 118-25.

Kinzler, K. W. & Vogelstein, B. 1996. Lessons from hereditary colorectal cancer. *Cell,* 87**,** 159-70.

Kinzler, K. W. & Vogelstein, B. 1997. Cancer-susceptibility genes. Gatekeepers and caretakers. *Nature,* 386**,** 761, 763.

Kleczkowska, H. E., Marra, G., Lettieri, T. & Jiricny, J. 2001. hMSH3 and hMSH6 interact with PCNA and colocalize with it to replication foci. *Genes Dev,* 15**,** 724-36.

Klip, A., Tsakiridis, T., Marette, A. & Ortiz, P. A. 1994. Regulation of expression of glucose transporters by glucose: a review of studies in vivo and in cell cultures. *FASEB J,* 8**,** 43-53.

Knudson, A. G., Jr. 1971. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci U S A,* 68**,** 820-3.

Korhonen, M. K., Vuorenmaa, E. & Nystrom, M. 2008. The first functional study of MLH3 mutations found in cancer patients. *Genes Chromosomes Cancer,* 47**,** 803-9.

Kroemer, G. & Pouyssegur, J. 2008. Tumor cell metabolism: cancer's Achilles' heel. *Cancer Cell,* 13**,** 472-82.

Kwong, L. N. & Dove, W. F. 2009. APC and its modifiers in colon cancer. *Adv Exp Med Biol,* 656**,** 85-106.

Lahue, R. S., Au, K. G. & Modrich, P. 1989. DNA mismatch correction in a defined system. *Science,* 245**,** 160-4.

Lamers, M. H., Winterwerp, H. H. & Sixma, T. K. 2003. The alternating ATPase domains of MutS control DNA mismatch repair. *EMBO J,* 22**,** 746-56.

Lamlum, H., Ilyas, M., Rowan, A., Clark, S., Johnson, V., Bell, J., Frayling, I., Efstathiou, J., Pack, K., Payne, S., Roylance, R., Gorman, P., Sheer, D., Neale, K., Phillips, R., Talbot, I., Bodmer, W. & Tomlinson, I. 1999. The type of somatic mutation at APC in familial adenomatous polyposis is determined by the site of the germline mutation: a new facet to Knudson's 'two-hit' hypothesis. *Nat Med,* 5**,** 1071-5.

Lane, D. P. 1992. Cancer. p53, guardian of the genome. *Nature,* 358**,** 15-6.

Lawes, D. A., Sengupta, S. & Boulos, P. B. 2003. The clinical importance and prognostic implications of microsatellite instability in sporadic cancer. *Eur J Surg Oncol,* 29**,** 201-12.

Leach, F. S., Nicolaides, N. C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomaki, P., Sistonen, P., Aaltonen, L. A., Nystrom-Lahti, M. & Et Al. 1993. Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. *Cell,* 75**,** 1215-25.

Lee, S. C., Guo, J. Y., Lim, R., Soo, R., Koay, E., Salto-Tellez, M., Leong, A. & Goh, B. C. 2005. Clinical and molecular characteristics of hereditary non-polyposis colorectal cancer families in Southeast Asia. *Clin Genet,* 68**,** 137-45.

Lehmann, E. L. & Romano, J. P. 2005. *Testing statistical hypotheses,* New York, Springer.

Levine, A. J. & Puzio-Kuter, A. M. 2010. The control of the metabolic switch in cancers by oncogenes and tumor suppressor genes. *Science,* 330**,** 1340-4.

Li, G. M. & Modrich, P. 1995. Restoration of mismatch repair to nuclear extracts of H6 colorectal tumor cells by a heterodimer of human MutL homologs. *Proc Natl Acad Sci U S A,* 92**,** 1950-4.

Lin, D. P., Wang, Y., Scherer, S. J., Clark, A. B., Yang, K., Avdievich, E., Jin, B., Werling, U., Parris, T., Kurihara, N., Umar, A., Kucherlapati, R., Lipkin, M., Kunkel, T. A. & Edelmann, W. 2004. An Msh2 point mutation uncouples DNA mismatch repair and apoptosis. *Cancer Res,* 64**,** 517-22.

Lin, Y. L., Shivji, M. K., Chen, C., Kolodner, R., Wood, R. D. & Dutta, A. 1998. The evolutionarily conserved zinc finger motif in the largest subunit of human replication protein A is required for DNA replication and mismatch repair but not for nucleotide excision repair. *J Biol Chem,* 273**,** 1453-61.

Lindblom, A., Tannergard, P., Werelius, B. & Nordenskjold, M. 1993. Genetic mapping of a second locus predisposing to hereditary non-polyposis colon cancer. *Nat Genet,* 5**,** 279-82.

Lindor, N. M., Mcmaster, M. L., Lindor, C. J. & Greene, M. H. 2008. Concise handbook of familial cancer susceptibility syndromes - second edition. *J Natl Cancer Inst Monogr***,** 1-93.

Linseisen, J., Kesse, E., Slimani, N., Bueno-De-Mesquita, H. B., Ocke, M. C., Skeie, G., Kumle, M., Dorronsoro Iraeta, M., Morote Gomez, P., Janzon, L., Stattin, P., Welch, A. A., Spencer, E. A., Overvad, K., Tjonneland, A., Clavel-Chapelon, F., Miller, A. B., Klipstein-Grobusch, K., Lagiou, P., Kalapothaki, V., Masala, G., Giurdanella, M. C., Norat, T. & Riboli, E. 2002. Meat consumption in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohorts: results from 24-hour dietary recalls. *Public Health Nutr,* 5**,** 1243-58.

Lipkin, M., Reddy, B., Newmark, H. & Lamprecht, S. A. 1999. Dietary factors in human colorectal cancer. *Annu Rev Nutr,* 19**,** 545-86.

Little, T. J., Horowitz, M. & Feinle-Bisset, C. 2007. Modulation by high-fat diets of gastrointestinal function and hormones associated with the regulation of energy intake: implications for the pathophysiology of obesity. *Am J Clin Nutr,* 86**,** 531-41.

Longley, M. J., Pierce, A. J. & Modrich, P. 1997. DNA polymerase delta is required for human mismatch repair in vitro. *J Biol Chem,* 272**,** 10917-21.

Lou, Y. W., Chen, Y. Y., Hsu, S. F., Chen, R. K., Lee, C. L., Khoo, K. H., Tonks, N. K. & Meng, T. C. 2008. Redox regulation of the protein tyrosine phosphatase PTP1B in cancer cells. *FEBS J,* 275**,** 69-88.

Lowsky, R., Decoteau, J. F., Reitmair, A. H., Ichinohasama, R., Dong, W. F., Xu, Y., Mak, T. W., Kadin, M. E. & Minden, M. D. 1997. Defects of the mismatch repair gene MSH2 are implicated in the development of murine and human lymphoblastic lymphomas and are associated with the aberrant expression of rhombotin-2 (Lmo-2) and Tal-1 (SCL). *Blood,* 89**,** 2276-82.

Lu, C., Ward, P. S., Kapoor, G. S., Rohle, D., Turcan, S., Abdel-Wahab, O., Edwards, C. R., Khanin, R., Figueroa, M. E., Melnick, A., Wellen, K. E., O'rourke, D. M., Berger, S. L., Chan, T. A., Levine, R. L., Mellinghoff,

I. K. & Thompson, C. B. 2012. IDH mutation impairs histone demethylation and results in a block to cell differentiation. *Nature,* 483**,** 474-8.

Lu, W., Ogasawara, M. A. & Huang, P. 2007. Models of reactive oxygen species in cancer. *Drug Discov Today Dis Models,* 4**,** 67-73.

Lynch, H. T. & De La Chapelle, A. 1999. Genetic susceptibility to non-polyposis colorectal cancer. *J Med Genet,* 36**,** 801-18.

Lynch, H. T. & De La Chapelle, A. 2003. Hereditary colorectal cancer. *N Engl J Med,* 348**,** 919-32.

Lynch, H. T. & Krush, A. J. 1971. Cancer family "G" revisited: 1895-1970. *Cancer,* 27**,** 1505-11.

Makos, M., Nelkin, B. D., Lerman, M. I., Latif, F., Zbar, B. & Baylin, S. B. 1992. Distinct hypermethylation patterns occur at altered chromosome loci in human lung and colon cancer. *Proc Natl Acad Sci U S A,* 89**,** 1929-33.

Malkhosyan, S., Rampino, N., Yamamoto, H. & Perucho, M. 1996. Frameshift mutator mutations. *Nature,* 382**,** 499-500.

Malkov, V. A., Biswas, I., Camerini-Otero, R. D. & Hsieh, P. 1997. Photocross-linking of the NH2-terminal region of Taq MutS protein to the major groove of a heteroduplex DNA. *J Biol Chem,* 272**,** 23811-7.

Mangold, E., Pagenstecher, C., Friedl, W., Fischer, H. P., Merkelbach-Bruse, S., Ohlendorf, M., Friedrichs, N., Aretz, S., Buettner, R., Propping, P. & Mathiak, M. 2005. Tumours from MSH2 mutation carriers show loss of MSH2 expression but many tumours from MLH1 mutation carriers exhibit weak positive MLH1 staining. *J Pathol,* 207**,** 385-95.

Marcus, P. M. & Newcomb, P. A. 1998. The association of calcium and vitamin D, and colon and rectal cancer in Wisconsin women. *Int J Epidemiol,* 27**,** 788-93.

Markowitz, S., Wang, J., Myeroff, L., Parsons, R., Sun, L., Lutterbaugh, J., Fan, R. S., Zborowska, E., Kinzler, K. W., Vogelstein, B. & Et Al. 1995. Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability. *Science,* 268**,** 1336-8.

Martik, D., Baitinger, C. & Modrich, P. 2004. Differential specificities and simultaneous occupancy of human MutSalpha nucleotide binding sites. *J Biol Chem,* 279**,** 28402-10.

Matthews, G. M., Howarth, G. S. & Butler, R. N. 2006. Nutrient and antioxidant modulation of apoptosis in gastric and colon cancer cells. *Cancer Biol Ther,* 5**,** 569-72.

Mazur, D. J., Mendillo, M. L. & Kolodner, R. D. 2006. Inhibition of Msh6 ATPase activity by mispaired DNA induces a Msh2(ATP)-Msh6(ATP) state capable of hydrolysis-independent movement along DNA. *Mol Cell,* 22**,** 39-49.

Mello, J. A., Acharya, S., Fishel, R. & Essigmann, J. M. 1996. The mismatch-repair protein hMSH2 binds selectively to DNA adducts of the anticancer drug cisplatin. *Chem Biol,* 3**,** 579-89.

Meng, T. C., Fukada, T. & Tonks, N. K. 2002. Reversible oxidation and inactivation of protein tyrosine phosphatases in vivo. *Mol Cell,* 9**,** 387-99.

Mobarhan, S. 1999. Calcium and the colon: recent findings. *Nutr Rev,* 57**,** 124-6.

Moenner, M., Pluquet, O., Bouchecareilh, M. & Chevet, E. 2007. Integrated endoplasmic reticulum stress responses in cancer. *Cancer Res,* 67**,** 10631-4.

Moser, A. R., Pitot, H. C. & Dove, W. F. 1990. A dominant mutation that predisposes to multiple intestinal neopla-

sia in the mouse. *Science,* 247**,** 322-4.

Nagata, C., Shimizu, H., Kametani, M., Takeyama, N., Ohnuma, T. & Matsushita, S. 2001. Diet and colorectal adenoma in Japanese males and females. *Dis Colon Rectum,* 44**,** 105-11.

Nakagawa, H., Nuovo, G. J., Zervos, E. E., Martin, E. W., Jr., Salovaara, R., Aaltonen, L. A. & De La Chapelle, A. 2001. Age-related hypermethylation of the 5' region of MLH1 in normal colonic mucosa is associated with microsatellite-unstable colorectal cancer development. *Cancer Res,* 61**,** 6991-5.

Nathan, C. 2008. Epidemic inflammation: pondering obesity. *Mol Med,* 14**,** 485-92.

Newmark, H. L., Yang, K., Kurihara, N., Fan, K., Augenlicht, L. H. & Lipkin, M. 2009. Western-style diet-induced colonic tumors and their modulation by calcium and vitamin D in C57Bl/6 mice: a preclinical model for human sporadic colon cancer. *Carcinogenesis,* 30**,** 88-92.

Newmark, H. L., Yang, K., Lipkin, M., Kopelovich, L., Liu, Y., Fan, K. & Shinozaki, H. 2001. A Western-style diet induces benign and malignant neoplasms in the colon of normal C57Bl/6 mice. *Carcinogenesis,* 22**,** 1871-5.

Norat, T. & Riboli, E. 2001. Meat consumption and colorectal cancer: a review of epidemiologic evidence. *Nutr Rev,* 59**,** 37-47.

Nystrom, M. & Mutanen, M. 2009. Diet and epigenetics in colon cancer. *World J Gastroenterol,* 15**,** 257-63.

Nystrom-Lahti, M., Kristo, P., Nicolaides, N. C., Chang, S. Y., Aaltonen, L. A., Moisio, A. L., Jarvinen, H. J., Mecklin, J. P., Kinzler, K. W., Vogelstein, B. & Et Al. 1995. Founding mutations and Alu-mediated recombination in hereditary colon cancer. *Nat Med,* 1**,** 1203-6.

Nystrom-Lahti, M., Perrera, C., Raschle, M., Panyushkina-Seiler, E., Marra, G., Curci, A., Quaresima, B., Costanzo, F., D'urso, M., Venuta, S. & Jiricny, J. 2002. Functional analysis of MLH1 mutations linked to hereditary nonpolyposis colon cancer. *Genes Chromosomes Cancer,* 33**,** 160-7.

Ollila, S., Sarantaus, L., Kariola, R., Chan, P., Hampel, H., Holinski-Feder, E., Macrae, F., Kohonen-Corish, M., Gerdes, A. M., Peltomaki, P., Mangold, E., De La Chapelle, A., Greenblatt, M. & Nystrom, M. 2006. Pathogenicity of MSH2 missense mutations is typically associated with impaired repair capability of the mutated protein. *Gastroenterology,* 131**,** 1408-17.

Oshima, M., Oshima, H., Kitagawa, K., Kobayashi, M., Itakura, C. & Taketo, M. 1995. Loss of Apc heterozygosity and abnormal tissue building in nascent intestinal polyps in mice carrying a truncated Apc gene. *Proc Natl Acad Sci U S A,* 92**,** 4482-6.

Ozcan, U., Cao, Q., Yilmaz, E., Lee, A. H., Iwakoshi, N. N., Ozdelen, E., Tuncman, G., Gorgun, C., Glimcher, L. H. & Hotamisligil, G. S. 2004. Endoplasmic reticulum stress links obesity, insulin action, and type 2 diabetes. *Science,* 306**,** 457-61.

Palombo, F., Gallinari, P., Iaccarino, I., Lettieri, T., Hughes, M., D'arrigo, A., Truong, O., Hsuan, J. J. & Jiricny, J. 1995. GTBP, a 160-kilodalton protein essential for mismatch-binding activity in human cells. *Science,* 268**,** 1912-4.

Palombo, F., Iaccarino, I., Nakajima, E., Ikejima, M., Shimada, T. & Jiricny, J. 1996. hMutSbeta, a heterodimer of hMSH2 and hMSH3, binds to insertion/deletion loops in DNA. *Curr Biol,* 6**,** 1181-4.

Papadopoulos, N., Nicolaides, N. C., Wei, Y. F., Ruben, S. M., Carter, K. C., Rosen, C. A., Haseltine, W. A., Fleischmann, R. D., Fraser, C. M., Adams, M. D. & Et Al. 1994. Mutation of a mutL homolog in hereditary colon cancer. *Science,* 263**,** 1625-9.

Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M.,

Gallia, G. L., Olivi, A., Mclendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K., Shinjo, S. M., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E. & Kinzler, K. W. 2008. An integrated genomic analysis of human glioblastoma multiforme. *Science,* 321**,** 1807-12.

Peltomaki, P. 2005. Lynch syndrome genes. *Fam Cancer,* 4**,** 227-32.

Peltomaki, P., Aaltonen, L. A., Sistonen, P., Pylkkanen, L., Mecklin, J. P., Jarvinen, H., Green, J. S., Jass, J. R., Weber, J. L., Leach, F. S. & Et Al. 1993. Genetic mapping of a locus predisposing to human colorectal cancer. *Science,* 260**,** 810-2.

Peltomaki, P., Gao, X. & Mecklin, J. P. 2001. Genotype and phenotype in hereditary nonpolyposis colon cancer: a study of families with different vs. shared predisposing mutations. *Fam Cancer,* 1**,** 9-15.

Peters, U., Mcglynn, K. A., Chatterjee, N., Gunter, E., Garcia-Closas, M., Rothman, N. & Sinha, R. 2001. Vitamin D, calcium, and vitamin D receptor polymorphism in colorectal adenomas. *Cancer Epidemiol Biomarkers Prev,* 10**,** 1267-74.

Peters, U., Sinha, R., Chatterjee, N., Subar, A. F., Ziegler, R. G., Kulldorff, M., Bresalier, R., Weissfeld, J. L., Flood, A., Schatzkin, A., Hayes, R. B., Prostate, L. C. & Ovarian Cancer Screening Trial Project, T. 2003. Dietary fibre and colorectal adenoma in a colorectal cancer early detection programme. *Lancet,* 361**,** 1491-5.

Pietinen, P., Malila, N., Virtanen, M., Hartman, T. J., Tangrea, J. A., Albanes, D. & Virtamo, J. 1999. Diet and risk of colorectal cancer in a cohort of Finnish men. *Cancer Causes Control,* 10**,** 387-96.

Pohl, N. M., Tong, C., Fang, W., Bi, X., Li, T. & Yang, W. 2009. Transcriptional regulation and biological functions of selenium-binding protein 1 in colorectal cancer in vitro and in nude mouse xenografts. *PLoS One,* 4**,** e7774.

Potter, V. R. 1958. The biochemical approach to the cancer problem. *Fed Proc,* 17**,** 691-7.

Powell, S. M., Zilz, N., Beazer-Barclay, Y., Bryan, T. M., Hamilton, S. R., Thibodeau, S. N., Vogelstein, B. & Kinzler, K. W. 1992. APC mutations occur early during colorectal tumorigenesis. *Nature,* 359**,** 235-7.

Prolla, T. A., Baker, S. M., Harris, A. C., Tsao, J. L., Yao, X., Bronner, C. E., Zheng, B., Gordon, M., Reneker, J., Arnheim, N., Shibata, D., Bradley, A. & Liskay, R. M. 1998. Tumour susceptibility and spontaneous mutation in mice deficient in Mlh1, Pms1 and Pms2 DNA mismatch repair. *Nat Genet,* 18**,** 276-9.

Pussila, M., Sarantaus, L., Dermadi Bebek, D., Valo, S., Reyhani, N., Ollila, S., Paivarinta, E., Peltomaki, P., Mutanen, M. & Nystrom, M. 2013. Cancer-predicting gene expression changes in colonic mucosa of Western diet fed mlh1 (+/-) mice. *PLoS One,* 8**,** e76865.

Raedle, J., Trojan, J., Brieger, A., Weber, N., Schafer, D., Plotz, G., Staib-Sebler, E., Kriener, S., Lorenz, M. & Zeuzem, S. 2001. Bethesda guidelines: relation to microsatellite instability and MLH1 promoter methylation in patients with colorectal cancer. *Ann Intern Med,* 135**,** 566-76.

Raevaara, T. E., Korhonen, M. K., Lohi, H., Hampel, H., Lynch, E., Lonnqvist, K. E., Holinski-Feder, E., Sutter, C., Mckinnon, W., Duraisamy, S., Gerdes, A. M., Peltomaki, P., Kohonen-Ccorish, M., Mangold, E., Macrae, F., Greenblatt, M., De La Chapelle, A. & Nystrom, M. 2005. Functional significance and clinical phenotype of nontruncating mismatch repair variants of MLH1. *Gastroenterology,* 129**,** 537-49.

Rampino, N., Yamamoto, H., Ionov, Y., Li, Y., Sawai, H., Reed, J. C. & Perucho, M. 1997. Somatic frameshift mutations in the BAX gene in colon cancers of the microsatellite mutator phenotype. *Science,* 275**,** 967-9.

Raschle, M., Marra, G., Nystrom-Lahti, M., Schar, P. & Jiricny, J. 1999. Identification of hMutLbeta, a heterodi-

mer of hMLH1 and hPMS1. *J Biol Chem,* 274**,** 32368-75.

Reeves, P. G., Nielsen, F. H. & Fahey, G. C., Jr. 1993. AIN-93 purified diets for laboratory rodents: final report of the American Institute of Nutrition ad hoc writing committee on the reformulation of the AIN-76A rodent diet. *J Nutr,* 123**,** 1939-51.

Reitmair, A. H., Redston, M., Cai, J. C., Chuang, T. C., Bjerknes, M., Cheng, H., Hay, K., Gallinger, S., Bapat, B. & Mak, T. W. 1996. Spontaneous intestinal carcinomas and skin neoplasms in Msh2-deficient mice. *Cancer Res,* 56**,** 3842-9.

Reitmair, A. H., Schmits, R., Ewel, A., Bapat, B., Redston, M., Mitri, A., Waterhouse, P., Mittrucker, H. W., Wakeham, A., Liu, B. & Et Al. 1995. MSH2 deficient mice are viable and susceptible to lymphoid tumours. *Nat Genet,* 11**,** 64-70.

Richon, V. M., Sandhoff, T. W., Rifkind, R. A. & Marks, P. A. 2000. Histone deacetylase inhibitor selectively induces p21WAF1 expression and gene-associated histone acetylation. *Proc Natl Acad Sci U S A,* 97**,** 10014-9.

Rodriguez-Bigas, M. A., Boland, C. R., Hamilton, S. R., Henson, D. E., Jass, J. R., Khan, P. M., Lynch, H., Perucho, M., Smyrk, T., Sobin, L. & Srivastava, S. 1997. A National Cancer Institute Workshop on Hereditary Nonpolyposis Colorectal Cancer Syndrome: meeting highlights and Bethesda guidelines. *J Natl Cancer Inst,* 89**,** 1758-62.

Roediger, W. E. 1990. The starved colon--diminished mucosal nutrition, diminished absorption, and colitis. *Dis Colon Rectum,* 33**,** 858-62.

Samowitz, W. S., Curtin, K., Lin, H. H., Robertson, M. A., Schaffer, D., Nichols, M., Gruenthal, K., Leppert, M. F. & Slattery, M. L. 2001. The colon cancer burden of genetically defined hereditary nonpolyposis colon cancer. *Gastroenterology,* 121**,** 830-8.

Sandhu, M. S., White, I. R. & Mcpherson, K. 2001. Systematic review of the prospective cohort studies on meat consumption and colorectal cancer risk: a meta-analytical approach. *Cancer Epidemiol Biomarkers Prev,* 10**,** 439-46.

Satelli, A., Rao, P. S., Thirumala, S. & Rao, U. S. 2011. Galectin-4 functions as a tumor suppressor of human colorectal cancer. *Int J Cancer,* 129**,** 799-809.

Saxonov, S., Berg, P. & Brutlag, D. L. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A,* 103**,** 1412-7.

Schaefer, M. H., Fontaine, J. F., Vinayagam, A., Porras, P., Wanker, E. E. & Andrade-Navarro, M. A. 2012. HIPPIE: Integrating protein interaction networks with experiment based quality scores. *PLoS One,* 7**,** e31826.

Schaefer, M. H., Lopes, T. J., Mah, N., Shoemaker, J. E., Matsuoka, Y., Fontaine, J. F., Louis-Jeune, C., Eisfeld, A. J., Neumann, G., Perez-Iratxeta, C., Kawaoka, Y., Kitano, H. & Andrade-Navarro, M. A. 2013. Adding protein context to the human protein-protein interaction network to reveal meaningful interactions. *PLoS Comput Biol,* 9**,** e1002860.

Schmutte, C., Sadoff, M. M., Shim, K. S., Acharya, S. & Fishel, R. 2001. The interaction of DNA mismatch repair proteins with human exonuclease I. *J Biol Chem,* 276**,** 33011-8.

Schwartz, R. A. & Torre, D. P. 1995. The Muir-Torre syndrome: a 25-year retrospect. *J Am Acad Dermatol,* 33**,** 90-104.

Scortegagna, M., Martin, R. J., Kladney, R. D., Neumann, R. G. & Arbeit, J. M. 2009. Hypoxia-inducible factor-1alpha suppresses squamous carcinogenic progression and epithelial-mesenchymal transition. *Cancer Res,* 69**,** 2638-46.

Semenza, G. L. 2010. HIF-1: upstream and downstream of cancer metabolism. *Curr Opin Genet Dev,* 20**,** 51-6.

Sena, L. A. & Chandel, N. S. 2012. Physiological roles of mitochondrial reactive oxygen species. *Mol Cell,* 48**,** 158-67.

Shih, I. M., Wang, T. L., Traverso, G., Romans, K., Hamilton, S. R., Ben-Sasson, S., Kinzler, K. W. & Vogelstein, B. 2001. Top-down morphogenesis of colorectal tumors. *Proc Natl Acad Sci U S A,* 98**,** 2640-5.

Slattery, M. L., Potter, J. D., Samowitz, W., Schaffer, D. & Leppert, M. 1999. Methylenetetrahydrofolate reductase, diet, and risk of colon cancer. *Cancer Epidemiol Biomarkers Prev,* 8**,** 513-8.

Smits, R., Hofland, N., Edelmann, W., Geugien, M., Jagmohan-Changur, S., Albuquerque, C., Breukel, C., Kucherlapati, R., Kielman, M. F. & Fodde, R. 2000. Somatic Apc mutations are selected upon their capacity to inactivate the beta-catenin downregulating activity. *Genes Chromosomes Cancer,* 29**,** 229-39.

Smits, R., Kielman, M. F., Breukel, C., Zurcher, C., Neufeld, K., Jagmohan-Changur, S., Hofland, N., Van Dijk, J., White, R., Edelmann, W., Kucherlapati, R., Khan, P. M. & Fodde, R. 1999. Apc1638T: a mouse model delineating critical domains of the adenomatous polyposis coli protein involved in tumorigenesis and development. *Genes Dev,* 13**,** 1309-21.

Souza, R. F., Appel, R., Yin, J., Wang, S., Smolinski, K. N., Abraham, J. M., Zou, T. T., Shi, Y. Q., Lei, J., Cottrell, J., Cymes, K., Biden, K., Simms, L., Leggett, B., Lynch, P. M., Frazier, M., Powell, S. M., Harpaz, N., Sugimura, H., Young, J. & Meltzer, S. J. 1996. Microsatellite instability in the insulin-like growth factor II receptor gene in gastrointestinal tumours. *Nat Genet,* 14**,** 255-7.

Studamire, B., Price, G., Sugawara, N., Haber, J. E. & Alani, E. 1999. Separation-of-function mutations in Saccharomyces cerevisiae MSH2 that confer mismatch repair defects but do not affect nonhomologous-tail removal during recombination. *Mol Cell Biol,* 19**,** 7558-67.

Sun, X., Zheng, L. & Shen, B. 2002. Functional alterations of human exonuclease 1 mutants identified in atypical hereditary nonpolyposis colorectal cancer syndrome. *Cancer Res,* 62**,** 6026-30.

Suter, C. M., Norrie, M., Ku, S. L., Cheong, K. F., Tomlinson, I. & Ward, R. L. 2003. CpG island methylation is a common finding in colorectal cancer cell lines. *Br J Cancer,* 88**,** 413-9.

Takai, D. & Jones, P. A. 2002. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc Natl Acad Sci U S A,* 99**,** 3740-5.

Taketo, M. M. & Edelmann, W. 2009. Mouse models of colon cancer. *Gastroenterology,* 136**,** 780-98.

Tammariello, A. E. & Milner, J. A. 2010. Mouse models for unraveling the importance of diet in colon cancer prevention. *J Nutr Biochem,* 21**,** 77-88.

Tangpricha, V., Flanagan, J. N., Whitlatch, L. W., Tseng, C. C., Chen, T. C., Holt, P. R., Lipkin, M. S. & Holick, M. F. 2001. 25-hydroxyvitamin D-1alpha-hydroxylase in normal and malignant colon tissue. *Lancet,* 357**,** 1673-4.

Tanyi, M., Olasz, J., Lukacs, G., Csuka, O., Toth, L., Szentirmay, Z., Ress, Z., Barta, Z., Tanyi, J. L. & Damjanovich, L. 2006. Pedigree and genetic analysis of a novel mutation carrier patient suffering from hereditary nonpolyposis colorectal cancer. *World J Gastroenterol,* 12**,** 1192-7.

Tetsu, O. & Mccormick, F. 1999. Beta-catenin regulates expression of cyclin D1 in colon carcinoma cells. *Nature,* 398**,** 422-6.

Thiagalingam, S., Lengauer, C., Leach, F. S., Schutte, M., Hahn, S. A., Overhauser, J., Willson, J. K., Markowitz, S., Hamilton, S. R., Kern, S. E., Kinzler, K. W. & Vogelstein, B. 1996. Evaluation of candidate tumour

suppressor genes on chromosome 18 in colorectal cancers. *Nat Genet,* 13**,** 343-6.

Thibodeau, S. N., French, A. J., Cunningham, J. M., Tester, D., Burgart, L. J., Roche, P. C., Mcdonnell, S. K., Schaid, D. J., Vockley, C. W., Michels, V. V., Farr, G. H., Jr. & O'connell, M. J. 1998. Microsatellite instability in colorectal cancer: different mutator phenotypes and the principal involvement of hMLH1. *Cancer Res,* 58**,** 1713-8.

Tolwani, R. J., Hamm, D. A., Tian, L., Sharer, J. D., Vockley, J., Rinaldo, P., Matern, D., Schoeb, T. R. & Wood, P. A. 2005. Medium-chain acyl-CoA dehydrogenase deficiency in gene-targeted mice. *PLoS Genet,* 1**,** e23.

Trachootham, D., Alexandre, J. & Huang, P. 2009. Targeting cancer cells by ROS-mediated mechanisms: a radical therapeutic approach? *Nat Rev Drug Discov,* 8**,** 579-91.

Trock, B., Lanza, E. & Greenwald, P. 1990. Dietary fiber, vegetables, and colon cancer: critical review and meta-analyses of the epidemiologic evidence. *J Natl Cancer Inst,* 82**,** 650-61.

Tseng, M., Murray, S. C., Kupper, L. L. & Sandler, R. S. 1996. Micronutrients and the risk of colorectal adenomas. *Am J Epidemiol,* 144**,** 1005-14.

Turrens, J. F. 2003. Mitochondrial formation of reactive oxygen species. *J Physiol,* 552**,** 335-44.

Umar, A., Boland, C. R., Terdiman, J. P., Syngal, S., De La Chapelle, A., Ruschoff, J., Fishel, R., Lindor, N. M., Burgart, L. J., Hamelin, R., Hamilton, S. R., Hiatt, R. A., Jass, J., Lindblom, A., Lynch, H. T., Peltomaki, P., Ramsey, S. D., Rodriguez-Bigas, M. A., Vasen, H. F., Hawk, E. T., Barrett, J. C., Freedman, A. N. & Srivastava, S. 2004. Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. *J Natl Cancer Inst,* 96**,** 261-8.

Umar, A., Buermeyer, A. B., Simon, J. A., Thomas, D. C., Clark, A. B., Liskay, R. M. & Kunkel, T. A. 1996. Requirement for PCNA in DNA mismatch repair at a step preceding DNA resynthesis. *Cell,* 87**,** 65-73.

Vander Heiden, M. G., Cantley, L. C. & Thompson, C. B. 2009. Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science,* 324**,** 1029-33.

Vasen, H. F., Watson, P., Mecklin, J. P. & Lynch, H. T. 1999. New clinical criteria for hereditary nonpolyposis colorectal cancer (HNPCC, Lynch syndrome) proposed by the International Collaborative group on HNPCC. *Gastroenterology,* 116**,** 1453-6.

Vasen, H. F., Wijnen, J. T., Menko, F. H., Kleibeuker, J. H., Taal, B. G., Griffioen, G., Nagengast, F. M., Meijers-Heijboer, E. H., Bertario, L., Varesco, L., Bisgaard, M. L., Mohr, J., Fodde, R. & Khan, P. M. 1996. Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology,* 110**,** 1020-7.

Veal, E. A., Day, A. M. & Morgan, B. A. 2007. Hydrogen peroxide sensing and signaling. *Mol Cell,* 26**,** 1-14.

Vogelstein, B., Fearon, E. R., Hamilton, S. R., Kern, S. E., Preisinger, A. C., Leppert, M., Nakamura, Y., White, R., Smits, A. M. & Bos, J. L. 1988. Genetic alterations during colorectal-tumor development. *N Engl J Med,* 319**,** 525-32.

Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr. & Kinzler, K. W. 2013. Cancer genome landscapes. *Science,* 339**,** 1546-58.

Vollset, S. E., Clarke, R., Lewington, S., Ebbing, M., Halsey, J., Lonn, E., Armitage, J., Manson, J. E., Hankey, G. J., Spence, J. D., Galan, P., Bonaa, K. H., Jamison, R., Gaziano, J. M., Guarino, P., Baron, J. A., Logan, R. F., Giovannucci, E. L., Den Heijer, M., Ueland, P. M., Bennett, D., Collins, R., Peto, R. & Collaboration, B. V. T. T. 2013. Effects of folic acid supplementation on overall and site-specific cancer incidence during the randomised trials: meta-analyses of data on 50,000 individuals. *Lancet,* 381**,** 1029-36.

Wagner, A., Van Der Klift, H., Franken, P., Wijnen, J., Breukel, C., Bezrookove, V., Smits, R., Kinarsky, Y., Barrows, A., Franklin, B., Lynch, J., Lynch, H. & Fodde, R. 2002. A 10-Mb paracentric inversion of chromosome arm 2p inactivates MSH2 and is responsible for hereditary nonpolyposis colorectal cancer in a North-American kindred. *Genes Chromosomes Cancer,* 35**,** 49-57.

Wagner, R., Jr. & Meselson, M. 1976. Repair tracts in mismatched DNA heteroduplexes. *Proc Natl Acad Sci U S A,* 73**,** 4135-9.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. 1982. Distantly related sequences in the alpha- and beta-subunits of ATP synthase, myosin, kinases and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J,* 1**,** 945-51.

Wang, H., Lawrence, C. W., Li, G. M. & Hays, J. B. 1999. Specific binding of human MSH2.MSH6 mismatch-repair protein heterodimers to DNA incorporating thymine- or uracil-containing UV light photoproducts opposite mismatched bases. *J Biol Chem,* 274**,** 16894-900.

Warburg, O. 1956a. On respiratory impairment in cancer cells. *Science,* 124**,** 269-70.

Warburg, O. 1956b. On the origin of cancer cells. *Science,* 123**,** 309-14.

Warburg, O., Wind, F. & Negelein, E. 1927. The Metabolism of Tumors in the Body. *J Gen Physiol,* 8**,** 519-30.

Warde-Farley, D., Donaldson, S. L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C. T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G. D. & Morris, Q. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res,* 38**,** W214-20.

Warren, J. J., Pohlhaus, T. J., Changela, A., Iyer, R. R., Modrich, P. L. & Beese, L. S. 2007. Structure of the human MutSalpha DNA lesion recognition complex. *Mol Cell,* 26**,** 579-92.

Wei, K., Kucherlapati, R. & Edelmann, W. 2002. Mouse models for human DNA mismatch-repair gene defects. *Trends Mol Med,* 8**,** 346-53.

Weinberg, R. A. 2007. *The biology of cancer,* New York, Garland Science.

Weinstein, S. J., Hartman, T. J., Stolzenberg-Solomon, R., Pietinen, P., Barrett, M. J., Taylor, P. R., Virtamo, J. & Albanes, D. 2003. Null association between prostate cancer and serum folate, vitamin B(6), vitamin B(12), and homocysteine. *Cancer Epidemiol Biomarkers Prev,* 12**,** 1271-2.

Weissman, S. M., Burt, R., Church, J., Erdman, S., Hampel, H., Holter, S., Jasperson, K., Kalady, M. F., Haidle, J. L., Lynch, H. T., Palaniappan, S., Wise, P. E. & Senter, L. 2012. Identification of individuals at risk for Lynch syndrome using targeted evaluations and genetic testing: National Society of Genetic Counselors and the Collaborative Group of the Americas on Inherited Colorectal Cancer joint practice guideline. *J Genet Couns,* 21**,** 484-93.

Wellen, K. E., Hatzivassiliou, G., Sachdeva, U. M., Bui, T. V., Cross, J. R. & Thompson, C. B. 2009. ATP-citrate lyase links cellular metabolism to histone acetylation. *Science,* 324**,** 1076-80.

Wellen, K. E. & Thompson, C. B. 2010. Cellular metabolic stress: considering how cells respond to nutrient excess. *Mol Cell,* 40**,** 323-32.

Wilson, A. J., Chueh, A. C., Togel, L., Corner, G. A., Ahmed, N., Goel, S., Byun, D. S., Nasser, S., Houston, M. A., Jhawer, M., Smartt, H. J., Murray, L. B., Nicholas, C., Heerdt, B. G., Arango, D., Augenlicht, L. H. & Mariadason, J. M. 2010. Apoptotic sensitivity of colon cancer cells to histone deacetylase inhibitors is mediated by an Sp1/Sp3-activated transcriptional program involving immediate-early gene induction. *Cancer Res,* 70**,** 609-20.

Wu, W. S. 2006. The signaling mechanism of ROS in tumor progression. *Cancer Metastasis Rev,* 25**,** 695-705.

Xu, W., Angelis, K., Danielpour, D., Haddad, M. M., Bischof, O., Campisi, J., Stavnezer, E. & Medrano, E. E. 2000. Ski acts as a co-repressor with Smad2 and Smad3 to regulate the response to type beta transforming growth factor. *Proc Natl Acad Sci U S A,* 97**,** 5924-9.

Yamagishi, N., Ueda, T., Mori, A., Saito, Y. & Hatayama, T. 2012. Decreased expression of endoplasmic reticulum chaperone GRP78 in liver of diabetic mice. *Biochem Biophys Res Commun,* 417**,** 364-70.

Yan, H., Dobbie, Z., Gruber, S. B., Markowitz, S., Romans, K., Giardiello, F. M., Kinzler, K. W. & Vogelstein, B. 2002. Small changes in expression affect predisposition to tumorigenesis. *Nat Genet,* 30**,** 25-6.

Yang, K., Edelmann, W., Fan, K., Lau, K., Leung, D., Newmark, H., Kucherlapati, R. & Lipkin, M. 1998. Dietary modulation of carcinoma development in a mouse model for human familial adenomatous polyposis. *Cancer Res,* 58**,** 5713-7.

Yang, W. 2007. Human MutLalpha: the jack of all trades in MMR is also an endonuclease. *DNA Repair (Amst),* 6**,** 135-9.

Yuan, F., Gu, L., Guo, S., Wang, C. & Li, G. M. 2004. Evidence for involvement of HMGB1 protein in human DNA mismatch repair. *J Biol Chem,* 279**,** 20935-40.

Yuen, S. T., Chan, T. L., Ho, J. W., Chan, A. S., Chung, L. P., Lam, P. W., Tse, C. W., Wyllie, A. H. & Leung, S. Y. 2002. Germline, somatic and epigenetic events underlying mismatch repair deficiency in colorectal and HNPCC-related cancers. *Oncogene,* 21**,** 7585-92.

Zhang, J., Lindroos, A., Ollila, S., Russell, A., Marra, G., Mueller, H., Peltomaki, P., Plasilova, M. & Heinimann, K. 2006. Gene conversion is a frequent mechanism of inactivation of the wild-type allele in cancers from MLH1/MSH2 deletion carriers. *Cancer Res,* 66**,** 659-64.