CHARLES UNIVERSITY
Faculty of Medicine in Pilsen

Ph.D. Study Program: Anatomy, Histology and Embryology

Mgr. Michal Kroupa

Chromosomal damage and shortening of telomeres in cancer patients and healthy subjects

Studium chromozomálního poškození a zkracování délky telomer u pacientů s nádorovým onemocněním a u zdravých osob

Dissertation Thesis

Supervisor: MUDr. Pavel Vodička, CSc.
Institute of Experimental Medicine, CAS

Prohlášení:

Prohlašuji, že jsem závěrečnou práci zpracoval samostatně a že jsem řádně uvedl a citoval všechny použité prameny a literaturu. Současně prohlašuji, že práce nebyla využita k získání jiného nebo stejného titulu.

Statement of authorship:

I declare that I have prepared this Thesis by myself, referring to all cited and used literature sources. The work in the Thesis has not been submitted to obtain another academic degree or equivalent.

V Praze, 18.2.2019/ Prague, 18th February 2019

Michal Kroupa

Acknowledgements

I would like to gratefully acknowledge my supervisor MUDr. Pavel Vodička, CSc. for providing a topic of submitted Thesis. I appreciate his valuable advice, skilled leadership, support, and friendly attitude during writing the Thesis.

I would like to thank RNDr. Zdeňka Polívková for invaluable assistance and consultation.

My thanks also go to Ing. Veronika Vymetálková, Ph.D. and MUDr. Ludmila Vodičková, CSc. for consultation of the topic.

Special thanks belong to all colleagues of Department of the Molecular Biology of Cancer at the Institute of Experimental Medicine of the Czech Academy of Sciences for cooperation in the laboratory.

I would also like to acknowledge my beloved and precious wife, my family and my friends for their support.

Financial support

The entire work could not be completed without the financial support from:

the Charles University (GAUK 800216, P27/2012, UNCE 204022, PROGRES Q 28), The Czech Science Foundation (15-14789S and 18-09709S), the Internal Grant Agency of the Ministry of Health of the Czech Republic (AZV 15-27580A, AZV 17-30920A, NV18-03-00199) and the National Sustainability Program I (NPU I, Nr. LO1503) provided by the Ministry of Education Youth and Sports of the Czech Republic.

List of Contents

List of Abbreviations

53BP1	p53-Binding Protein 1
8-oxoG	8-oxo-7,8-dihydro-2′-deoxyguanosine
A-EJ	Alternative end-joining
ACs	Aberrant cells with chromosomal aberrations
APC/C	Anaphase-Promoting Complex/Cyclosome
ALT	Alternative lengthening of telomeres
APE1	Apurinic/Apyrimidinic Endonuclease 1
ATM	Ataxia-Telangiectasia Mutated
BC	Breast cancer
BER	Base excision repair
BRAF	V-Raf Murine Sarcoma Viral Oncogene Homolog B1
BUB1/3	Budding Uninhibited by Benzimidazoles 1/3
BRCA1/2	Breast Cancer Type 1/2 Susceptibility Protein
CAs	Chromosomal aberrations
CAtot	Total frequency of chromosomal aberrations
CSAs	Chromosome-type aberrations
CTAs	Chromatid-type aberrations
CDC20	Cell Division Cycle 20
CENPF	Centromere Protein-F
CHK2	Checkpoint Kinase 2
CIN	Chromosomal instability
CRC	Colorectal cancer
DNA-PK	DNA-Dependent Protein Kinase
DRC	DNA repair capacity
DSBs	DNA double-strand breaks
ERCC1	Excision Repair Cross-Complementation Group 1
ESPL1	Extra Spindle Pole Bodies Like 1 (“Separase”)
ETS	E-Twenty-Six Transformation-Specific
EXO1	Exonuclease 1
GG-NER	Global Genome Nucleotide excision repair
GIN	Genomic instability
HR	Homologous recombination
KRAS	Ki-ras2 Kirsten Rat Sarcoma Viral Oncogene Homolog
MAD2L1	Mitotic Arrest Deficient 2 Like 1
MLH1	MutL Homolog 1
MMQPCR	Multiplex Monochrome Real-Time Polymerase chain reaction
MMR	Mismatch repair
MRE11	Meiotic Recombination 11
MSH2	MutS Homolog 2
MSH6	MutS Homolog 6

MSI	Microsatellite instability
NAT	Natural Antisense Transcript
NBS1	Nijmegen Breakage Syndrome 1
NEK2	Never in Mitosis (NIMA) Related Kinase 2
NIN	Nucleotide instability
NER	Nucleotide excision repair
NHEJ	Non-homologous end joining
OGG1	8-Oxoguanine DNA Glycosylase
OS	Overall survival
PBL	Peripheral blood lymphocytes
PCNA	Proliferating Cell Nuclear Antigen
PCR	Polymerase chain reaction
PMS2	Postmeiotic Segregation Increased 1 Homolog 2
POT1	Protection of Telomeres Protein 1
PTTG1	Pituitary Tumor Transforming Gene 1 (“Securin”)
ROD	Rough Deal
CTIP	The Carboxy-Terminal Binding Protein (CtBP)-Interacting Protein
RAP1	Repressor/Activator Protein 1
RPA	Replication Protein A
RTL	Relative telomere length
SNPs	Single nucleotide polymorphisms
SSA	Single-strand annealing
SSBs	DNA Single-strand breaks
ssDNA	Single-stranded DNA
TC-NER	Transcription-Coupled Nucleotide excision repair
TERT	Telomerase Reverse Transcriptase
TERC	Telomerase RNA Component
TIN2	TRF1- and TRF2-Interacting Nuclear Protein 2
TPP1	Telomere Protection Protein 1
TP53	Tumor Protein p53
TRPA1	Transient Receptor Potential Cation Channel Subfamily A Member 1
TRF1/2	Telomeric Repeat-Binding Factor ½
XPA-XPF	Xeroderma Pigmentosum Complementation Group A-F
XRCC1-4	X-Ray Repair Cross Complementing 1-4
ZW10	Zeste White 10

Abstract

Impaired chromosome segregation during mitosis, inaccurate DNA damage response and excessive telomere shortening may all modulate the frequency of chromosomal aberrations (CAs) in peripheral blood lymphocytes (PBL). There is evidence that increased frequency of structural CAs in PBL may be considered as a marker of enhanced cancer risk. In the present Thesis, an effect of variants in genes involved in mitotic checkpoint and DNA damage response on the inter-individual differences in CAs frequency in PBL was investigated. Considering the importance of disrupted telomere structure and its function in cancer biology, a link between telomere length and clinicopathological and molecular features of cancer patients was analysed. Furthermore, the relevance of telomere length and CAs frequency as markers of patients' survival was examined.

The major outcomes of the Thesis, fully reported in detail in seven attached Manuscripts, are: I) Increased frequency of structural CAs and/or disrupted telomere length in PBL may be considered as risk factors for the different types of solid cancer; II) Telomere shortening in PBL of healthy subjects increased the frequency of structural CAs; III) Binary interactions of gene variants in mitotic checkpoint and DNA repair pathways may modulate the frequency of structural CAs in PBL of healthy subjects; IV) The application of genome-wide association study revealed novel loci associated with genes important for mitosis and linked to the frequency of CAs in PBL; V) Telomere shortening in PBL of breast and colorectal cancer (CRC) patients was associated with decreased capacity to repair mutagen-induced DNA double-strand breaks; VI) Telomere length in tumor tissue was modulated by clinicopathological features (e.g. tumor-site origin, stage of tumor development, microsatellite instability) of CRC patients and finally, it was demonstrated that VII) CRC patients with more pronounced telomere shortening in tumor tissue compared to the adjacent mucosa evinced prolonged survival.

The results may be utilized in the future, when inter-individual differences in terms of identified gene variants and disrupted telomere length maintenance may provide a prediction tool for cancer risk assessment. Furthermore, as telomerase inhibitors are currently being applied in clinical practice, it is important to understand tumor telomere length variability and its link to clinicopathological and molecular features of cancer patients.

Abstrakt

Mezi základní procesy, které mohou ovlivnit míru chromozomálních aberací (CAs) v lymfocytech periferní krve (PBL), patří chybná segregace chromozomů v průběhu mitózy, nepřesná odpověď na poškození DNA a nadměrné zkracování délky telomer. Akumulaci CAs v PBL lze považovat za ukazatel zvýšeného rizika rozvoje maligního onemocnění. Tato disertační práce je zaměřena na studium vlivu jednonukleotidových polymorfismů v genech účastnících se mitotického checkpointu a odpovědi na poškození DNA na inter-individuální rozdíly ve výskytu CAs v PBL. Dále byl zjišťován možný vliv klinicko-patologických a molekulárních charakteristik pacientů s karcinomem tlustého střeva a konečníku (CRC) na délku telomer v nádorové tkáni, jelikož narušená funkce těchto protektivních struktur chromozomů hraje klíčovou úlohu v procesu maligní transformace. Bylo také studováno, zda lze délku telomer a frekvenci CAs považovat za ukazatel celkového přežití pacientů s nádorovým onemocněním.

Hlavní poznatky této práce, založené na sedmi přiložených rukopisech, jsou: I) Zvýšený výskyt strukturních CAs a/nebo narušená regulace délky telomer v PBL mohou být považovány za rizikový faktor rozvoje nádorového onemocnění; II) Kratší délka telomer v PBL zdravých osob zvyšuje výskyt strukturních CAs; III) Binární interakce genových variant účastnících se mitotického checkpointu a opravy poškození DNA mohou ovlivňovat míru výskytu strukturních CAs v PBL zdravých osob; IV) Pomocí využití celogenomové asociační studie byly odhaleny nové genové varianty spojené s vyšší mírou CAs v PBL; V) Zkracování telomer v PBL u pacientů s karcinomem prsu a CRC bylo spojeno se sníženou kapacitou opravy dvouvláknových zlomů DNA; VI) Délka telomer v nádorové tkáni byla ovlivněna klinicko-patologickými charakteristikami pacientů s CRC (např. oblast výskytu nádoru, mikrosatelitová nestabilita a stádium rozvoje zhoubného novotvaru); VII) Kratší délka telomer v nádorové tkáni v porovnání s přilehlou mukózou u pacientů s CRC byla spojena s lepším celkovým přežitím.

Výše popsané poznatky mohou být uplatněny zejména v budoucnu, kdy informace o rizikových genových variantách a narušené regulaci délky telomer mohou být využity jako nástroj pro predikci zvýšené pravděpodobnosti rozvoje rakoviny. Vzhledem k tomu, že inhibitory telomerázy se začínají používat v klinické praxi, je také velice důležité porozumět mechanismům regulace délky telomer v nádorové tkáni.

List of Manuscripts

Manuscript I

Genetic variation in the major mitotic checkpoint genes associated with chromosomal aberrations in healthy humans. Försti A, Frank C, Smolkova B, Kazimirova A, Barancokova M, Vymetalkova V, **Kroupa M**, Naccarati A, Vodickova L, Buchancova J, Dusinska M, Musak L, Vodicka P, Hemminki K. *Cancer Lett.* 2016 Oct 1;380(2):442-6. **IF: 6.491**

Manuscript II

Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects. Vodicka P, Musak L, Frank C, Kazimirova A, Vymetalkova V, Barancokova M, Smolkova B, Dzupinkova Z, Jiraskova K, Vodenkova S, **Kroupa M**, Osina O, Naccarati A, Palitti F, Försti A, Dusinska M, Vodickova L, Hemminki K. *Carcinogenesis*. 2015 Nov;36(11):1299-306. **IF: 5.072**

Manuscript III

Genetic variation associated with chromosomal aberration frequency: A genome-wide association study. Niazi Y, Thomsen H, Smolkova B, Vodickova L, Vodenkova S, **Kroupa M**, Vymetalkova V, Kazimirova A, Barancokova M, Volkovova K, Staruchova M, Hoffmann P, Nöthen MM, Dušinská M, Musak L, Vodicka P, Hemminki K, Försti A. *Environ Mol Mutagen.* 2019 Jan;60(1):17-28. **IF: 3.254**

Manuscript IV

Chromosomal damage and telomere length in peripheral blood lymphocytes: cancer risk and patients' long-term survival. Vodenkova S, **Kroupa M**, Polivkova Z, Musak L, Ambrus M, Schneiderova M, Kozevnikovova R, Vodickova L, Rachakonda S, Hemminki K, Kumar R, Vodicka P. **Manuscript in preparation**

Manuscript V

Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients. **Kroupa M**, Polivkova Z, Rachakonda S, Schneiderova M, Vodenkova S, Buchler T, Jiraskova K, Urbanova M, Vodickova L,

Hemminki K, Kumar R, Vodicka P. *Genes Chromosomes Cancer*. 2018 Feb;57(2):61-69. **IF: 3.362**

Manuscript VI

Relationship of telomere length in colorectal cancer patients with cancer phenotype and patient prognosis. **Kroupa M**, Rachakonda S, Srinivas N, Liska V, Urbanova M, Jiraskova K, Schneiderova M, Vycital O, Vymetalkova V, Vodickova L, Kumar R, Vodicka P. **Manuscript in preparation**

Manuscript VII

Genetic variation of acquired structural chromosomal aberrations. Vodicka P, Musak L, Vodickova L, Vodenkova S, Catalano C, **Kroupa M**, Naccarati A, Polivkova Z, Vymetalkova V, Försti A, Hemminki K. *Mutat Res.* 2018 Dec;836(Pt A):13-21. **IF: 1.996**

1. Introduction

Error-free transmission of undamaged genetic information from each cell to its offspring is a critical precondition for the preservation of genome stability and perpetuation of life. At the same time, acquired mutations in DNA molecule, the raw materials of evolution, contribute to malignant transformation of the cell.

Cancer is a wide and heterogeneous group of complex disorders differing from each other in biological and molecular characteristics. One aspect of this heterogeneity, genomic instability (GIN), is a hallmark of solid tumors [1].

GIN refers to error acquisition protruding on the nucleotide or chromosomal level [2]. Despite considerable stability of DNA molecule, many chemical, biological and physical factors daily disrupt DNA structure. Robust DNA repair mechanisms remove a wide-scale of DNA damage. However, reduced DNA repair capacity (DRC) may contribute to GIN [3]. Furthermore, defective chromosome segregation during the cell division and/or telomere dysfunction may additionally cause gross chromosomal rearrangements [1, 4].

In the present thesis, we summarize several potential contributors underlying the formation of chromosomal aberrations (CAs; *Manuscript I, II, III, IV, V*), the most prevalent form of GIN [2]. Molecular mechanisms leading to the formation of CAs are still currently poorly understood. Our results may help to better understand this phenomenon.

Following text is based on 4 original research articles and a summary review paper (*Manuscript VII*), all published between 2015 to 2019. Furthermore, important parts of the present doctoral Thesis are two manuscripts (*Manuscript IV & VI)*, though under review process at the time of submitting the Thesis, focusing on telomere length in cancer patients and healthy subjects.

1.1. Genomic instability and cancer

GIN modulates tumor initiation, progression, patient prognosis and also increases mutation rate, clonal evolution, and tumor heterogeneity [5-7]. Therefore, it is a crucial prerequisite for malignant transformation and represents a driving force of cancer disease [8]. Generally, all malignancies are associated with GIN and a considerable variation in its amount and type is commonly seen among cancer cells [9].

For convenience, GIN can be divided into two major classes, nucleotide instability (NIN) and chromosomal instability (CIN). Furthermore, some authors classify microsatellite instability (MSI) as a particular case of NIN, others consider it as a distinct subcategory of GIN, though it represents insertion-deletion mutations (indels) on the nucleotide level [2, 10]

NIN comprises a spectrum of single base mutations to excessive nucleotide indels. On the other hand, CIN, a more prominent type of GIN, encompasses large-scale of numerical and structural CAs [2, 11]. Unlike NIN, the molecular basis of CIN had not yet been clarified. This may be due to the fact that a complex disruption of many mechanisms may contribute to CIN phenotype [4, 8, 12, 13] (illustrated in *Figure 1*).

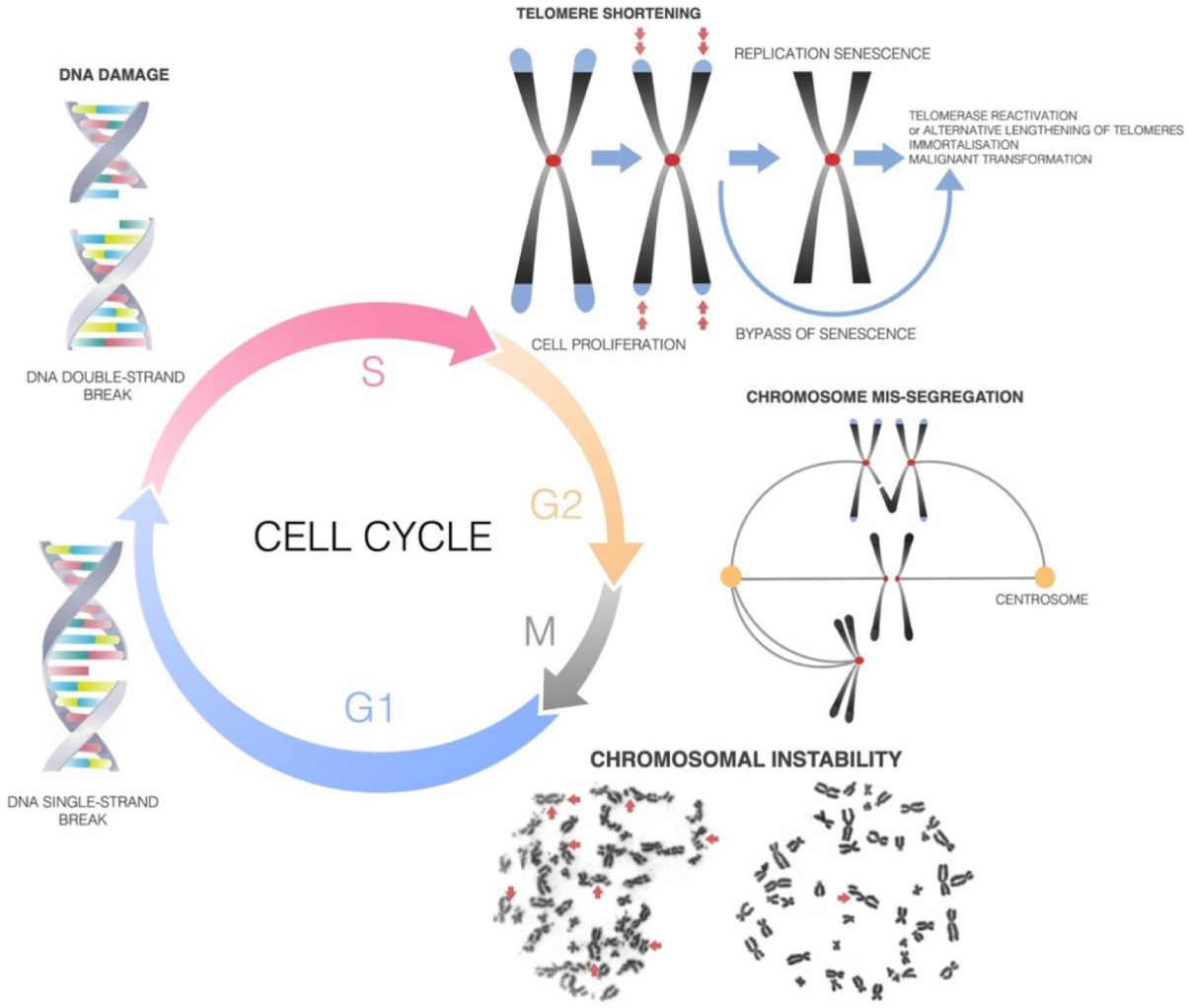


Figure 1. **Proposed potential sources of CIN** involve erroneous spindle assembly checkpoint and impaired chromosome segregation during mitosis, inaccurate DNA damage response and telomere attrition.

1.2. Chromosomal aberrations

Numerical CAs involve euploidy or aneuploidy subcategories, which refer to chromosome number that is an exact multiple of the haploid set (polyploidization; e.g. triploidy, tetraploidy) or any chromosome number that is non-euploid (e.g. monosomy, trisomy), respectively [14]. Polyploidization is a common characteristic of the several mammalian specialized cell types (e.g. liver and muscle cells [15]) and has been also documented in cancer cells [16]. Furthermore, except for some trisomies (e.g. Down or Edwards syndrome), numerical CAs of most human chromosomes are, in general, incompatible with viability of non-malignant cells. However, numerical CAs are widely present in the majority of human malignancies and represent the most common type of GIN in cancer cells. It is important to note, that even a small alteration in chromosome copy number may lead to genes dosage disequilibrium, interruption of key cell pathways (e.g. mitotic spindle checkpoint) and may ultimately result in structural CAs. Therefore, numerical CAs may contribute to the formation of structural CAs [17].

Unlike numerical CAs, structural CAs are generally compatible with cell viability and encompass specific and non-specific CAs. Specific structural CAs include inversions, translocations and deletions of DNA segments [18]. An example of specific CAs-associated disease is chronic myeloid leukemia, a neoplastic disorder resulting from the genomic reciprocal translocations [19]. Non-specific CAs, on the other hand, may be cytologically distinguished in peripheral blood lymphocytes (PBLs) and are divided into chromatid – type (CTA), affecting only one chromatid, or chromosome – type (CSA) subclass, which involves both chromatids. Both types of non-specific structural CAs have a different mechanism of origin, where a complex disruption of mitotic checkpoint, DNA repair and/or telomere length maintenance mechanisms may be critical [20-23].

In general, CTAs, including chromatid breaks and exchanges, arise as a consequence of insufficiently repaired DNA double-strand breaks (DSBs) induced during the late S or G2 phase of the cell cycle. On the other hand, CSAs, encompassing dicentric and ring chromosomes, are generated by replication of DSBs during the S phase of the cell cycle [24, 25].

CAs may be considered as an initial event in malignant transformation [26], although the relationship between non-specific CAs and cancer is not as determined as it is in case of specific CAs. However, there is compelling evidence that increased frequencies of non-

specific structural CAs, particularly CSAs, in PBL may serve as a predictive marker of cancer susceptibility [27-31], if the genetic damage in PBL reflects the DNA changes in cells undergoing carcinogenesis [27]. The first evidence of increased chromosomal damage in PBL of incident cancer patients compared with healthy controls was documented in our laboratory [31]. Currently, in case-control studies, 2 % of CAs in PBL is used as a cut-off point to distinguish between individuals with low or high level of chromosomal damage, while for CSAs and CTAs it is 1 % [29, 31], (*Manuscript I–IV*).

1.2.1. Inherited predisposition to structural CAs

There is a remarkable genetic variability among the human population. Single nucleotide polymorphisms (SNPs) are the most common type of such heterogeneity. SNPs not only affect phenotypic characteristics of each individual but they may also underlie inter-individual differences in susceptibility to a malignant transformation.

Few previous reports conducted also in our laboratory, focused on the frequency of CAs in association with SNPs in genes encoding Cyclin D1 [32], xenobiotic metabolism [21] and DNA repair enzymes [33-35] (summarized in *Manuscript VII*). Furthermore, several studies also investigated the role of SNPs in DNA repair genes for the susceptibility of cancer risk [36-40], patient's prognosis and therapy response [40]. In *Manuscript II* we aimed to bring new evidence on this research topic and focused on the influence of SNPs in genes encoding DNA repair enzymes on CAs level in PBL of healthy individuals.

Association studies investigating SNPs in genes involved in telomere-homeostasis also showed an effect on telomere length, risk of cancer disease and patient's prognosis [41-47]. However, such analyses are not frequent and the results are mostly inconclusive. Recently, meta-analysis demonstrated significant correlations between SNPs in loci coding for telomerase and cancer susceptibility, especially lung [48], pancreatic [49] and bladder cancers [50]. However, the impact of relevant gene variants on telomere length in tumor tissue of cancer patients is currently unresolved. In *Manuscript VI* we evaluated the influence of SNPs in genes important for telomere homeostasis (telomerase and shelterin complex) on telomere length in distinct tissues of colorectal cancer (CRC) patients.

Interestingly, despite the importance of mitotic checkpoint and fidelity of chromosome segregation for protection against CAs development, the studies investigating the role of SNPs in mitotic checkpoint genes and their relation to cancer are also not

common. Brendle *et al*. identified SNP in *Centromere protein-F* gene *(CENPF)* as a prognostic factor of breast cancer (BC) patients [30]. Furthermore, variants in a gene encoding Aurora kinase significantly conferred increased susceptibility to gastric cancer [51, 52]. In *Manuscript I*, we focused on SNPs in genes coding for enzymes involved in the mitotic checkpoint in relation to the frequency of CAs in PBL of healthy subjects.

Genome-wide association study (GWAS) may help to identify gene variants in relation to increased cancer risk (e.g. [53, 54]). In addition to the findings in *Manuscript I* and *II*, we performed the analysis of SNPs from across the human genome in order to identify new loci predisposing the individual to an increased frequency of CAs (*Manuscript III*). This approach may be utilized in the future, when a potential panel of SNPs may provide a prediction tool for cancer risk assessment.

1.2.2. Mitotic checkpoint and structural CAs

Following chapter of the Thesis explains the role of key enzymes involved in mitotic checkpoint. Enzymes relevant to *Manuscript I* are highlighted in the text and their function is summarized in *Figure 2*.

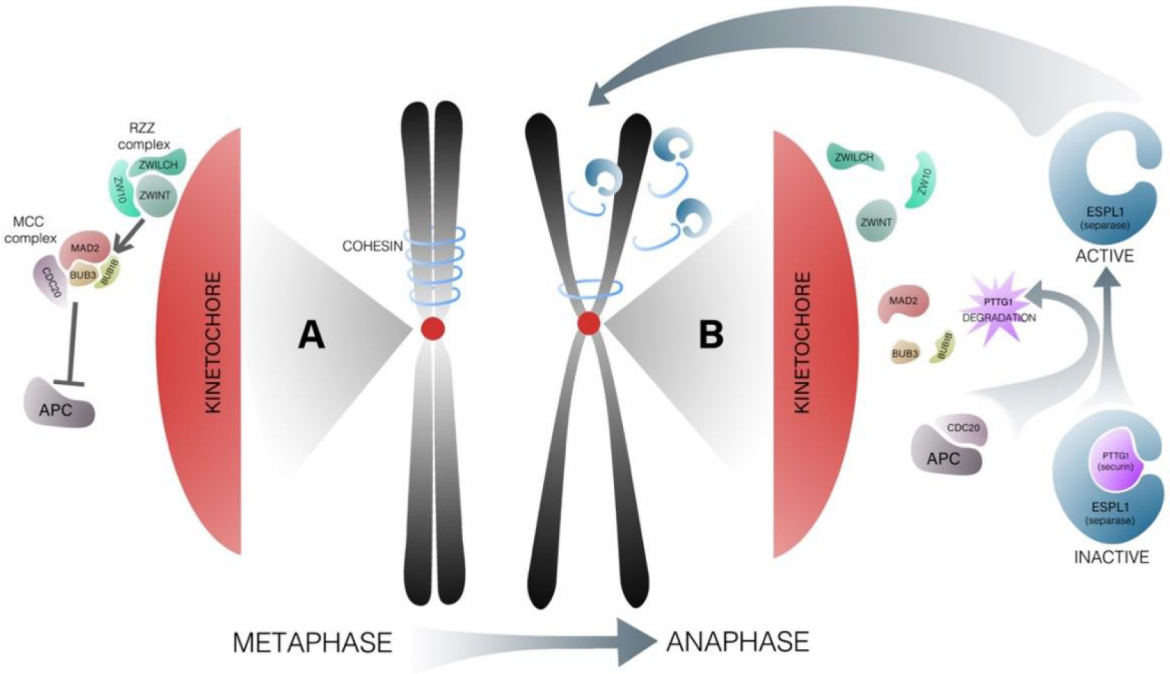


Figure 2: Metaphase – anaphase transition and major enzymes involved in mitotic checkpoint.
A) Activation of mitotic checkpoint and inhibition of chromatid segregation; **B)** Deactivation of mitotic checkpoint and consequent chromatid segregation during anaphase.

Errors at kinetochore-microtubule connection may lead to premature or impaired segregation of sister chromatids, resulting in CAs [55, 56]. To avoid this scenario, mitotic checkpoint or so-called "spindle assembly checkpoint" provides a safeguard mechanism against an erroneous distribution of genetic information during the early stages of mitosis.

After the DNA replication, sister chromatids are encircled and entrapped by Cohesin subunits. Cohesin refers to a gigantic multimeric protein with a ring-like structure, which holds together two sister chromatids until anaphase onset when all chromosomes are properly attached to the mitotic spindle. A correct connection of all chromosomes to the microtubule fibers and consequent separation of sister chromatids to the opposite spindle poles ensure accurate distribution of genetic information into daughter cells. Kinetochores, complex disc-shaped structures associated with the centromeric DNA regions of each chromatid, play an indispensable role in the whole process [57].

Mitotic checkpoint signal, emitted by affected kinetochore in case of incorrectly attached or unattached microtubules, in turn delays metaphase-anaphase transition and provides the cell an opportunity to fix the connection [58]. *Anaphase-Promoting Complex* (APC/C alias "Cyclosome"), is inhibited by mitotic checkpoint complex. APC/C represents an E3 ubiquitin ligase, that enables proteolytic degradation of several major targets, that inhibit sister chromatid segregation [59].

The question is, how does affected kinetochore generate a mitotic checkpoint signal? A key role in the initiation and also silencing of mitotic checkpoint plays RZZ complex named after the *Drosophila melanogaster* genes *Rough Deal* (*ROD*), *Zeste White 10* (*ZW10*), ***ZWINT*** and ***ZWILCH*** [60]. RZZ complex along with CENPF and *NIMA Related Kinase 2* (**NEK2**) proteins is required to recruit mitotic checkpoint complex to the kinetochore [61, 62]. Main components of mitotic checkpoint complex are *Mitotic Arrest Deficient 2* (**MAD2**), *Budding Uninhibited by Benzimidazoles 1* (**BUB1**) and *Budding Uninhibited by Benzimidazoles 3* (**BUB3**) enzymes [63]. Once mitotic checkpoint complex is assembled at kinetochores, it blocks *Cell Division Cycle 20* (CDC20), a key subunit of APC/C and sister chromatid cannot separate. However, when all kinetochores are properly connected to microtubules, CDC20 is released from inhibition by mitotic checkpoint complex, interacts with APC/C, creates active APC/C complex and enables polyubiquitination and destruction of Securin (**PTTG1**), an inhibitor of Separase (**ESPL1**). Activated Separase molecules, in turn, cleave Cohesin and sister chromatid segregation begins [64, 65].

1.2.3. DNA damage response and structural CAs

The genes encoding DNA repair enzymes are highly polymorphic and, taking into account the significance of the DNA repair defects for malignant transformation [66-69] differences in cancer risk susceptibility modulated by such SNPs may be expected. In *Manuscript II*, we studied SNPs in genes coding for DNA repair proteins in relation to the frequency of CAs in PBL of healthy subjects (*Figure 3*). The study also follows the previous research conducted at our department, for at least two decades, when substantial interindividual variability in DRC regarding relevant gene variants was observed [70-72]. Furthermore, the whole chapter is closely related to the next part of the Thesis, which focuses on the complex interplay between DNA damage response (DDR) and telomeres.

Variety of environmental and endogenous factors constantly assaults the structure of the DNA molecule and creates a diversity of DNA lesions [73]. Reactive oxygen species, alkylating chemicals, xenobiotics associated with food processing or ultraviolet light represent the most common types of DNA damaging agents, which may induce thousands of lesion per hour in each cell [74]. Those, if not removed properly, may result in DNA single-strand breaks (SSBs). If replication fork collides with SSB, DSBs may be formed [75]. DSBs represent the most harmful and potentially lethal damage. Each cell has to deal with approximately ten DSBs per day. Pathological DSBs arise for example from the effect of ionizing radiation and replication failures. Moreover, DSBs are physiologically generated by a cell itself. Those intentionally induced DSBs play a pivotal role in rearrangements of genes for antigen-receptors, which is a central feature of the adaptive immune system [76] DSBs are also abundantly created during meiotic recombination [77].

All the above examples implicate the necessity to utilize DNA repair pathways to deal with the damage, and hence to maintain genome integrity. Therefore, cells have evolved a sophisticated network of DDR. The principal DNA repair pathways in mammals comprise Nucleotide excision repair (NER), Base excision repair (BER), Mismatch repair (MMR), and DSBs repair pathways including Homologous recombination (HR) and Non-homologous end joining (NHEJ) [78]

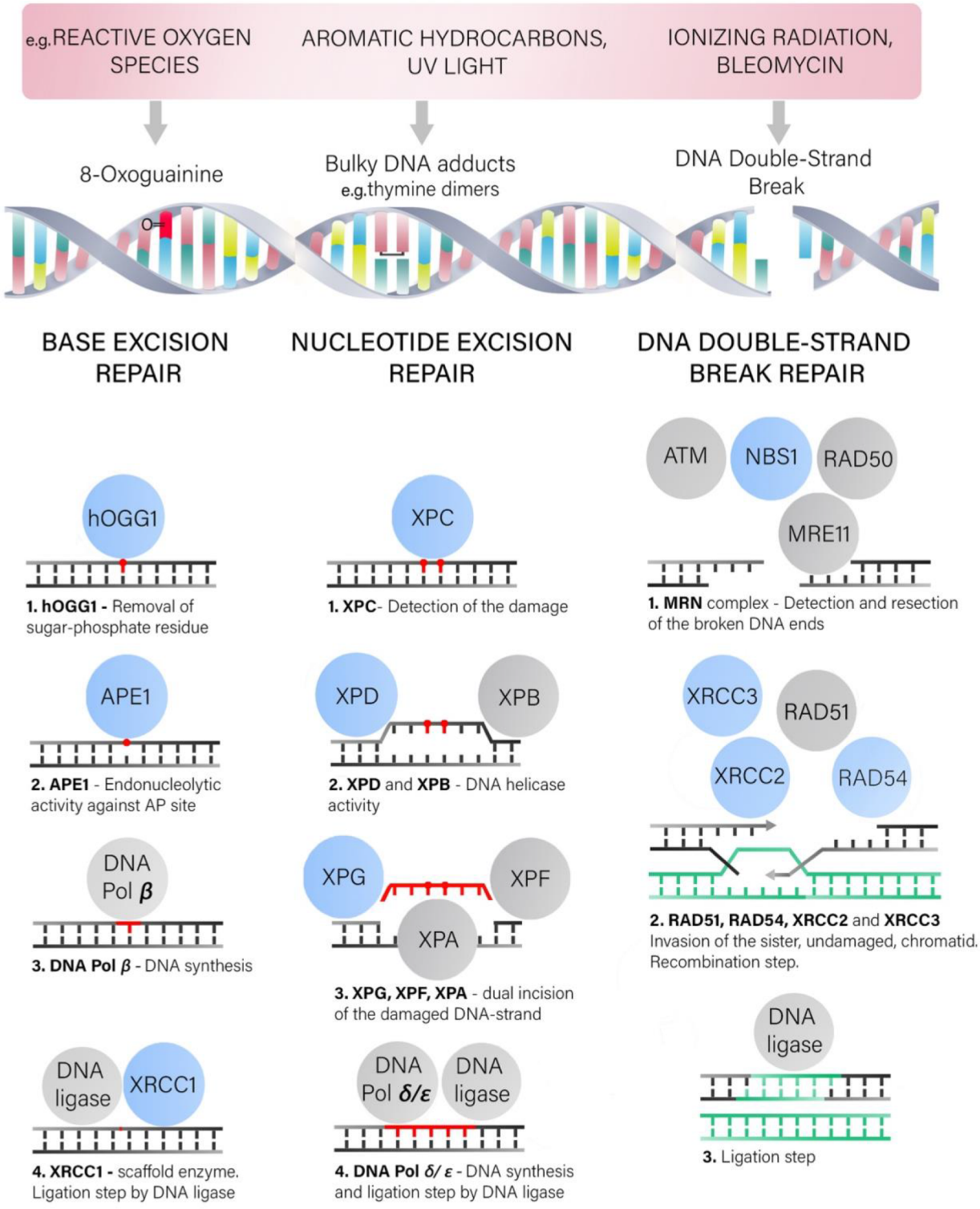


Figure 3: Simplified DNA repair pathways with the highlighted function of enzymes relevant to *Manuscript II* (brief overview below).

Except for BER, a connection between severe human syndromes, frequently associated with a high risk of cancer disease, and impaired DDR has been reported. This may be due to the critical importance of BER in maintaining cell viability [79].

Disruption of NER mechanism is associated, for illustration, with *Xeroderma pigmentosum* disorder, a condition characterized by extreme ultraviolet light sensitivity, inability to recognize and remove photoproduct from DNA and by a very large increase in the risk of skin cancer [66, 69]. *Xeroderma pigmentosum* was the first syndrome described in a link with impaired DDR [69].

Mutations in genes involved in MMR are linked to *Lynch syndrome*, a condition connected to an excessive risk of CRC [68].

Impact of defects in DSBs repair pathways on human health is also well recognized. Individuals with mutations in NHEJ-associated genes, particularly in nuclease Artemis and the catalytic subunit of *DNA-Dependent Protein Kinase* (DNA-PKcs), suffer from severe combined immunodeficiency due to the erroneous V(D)J recombination [80, 81]. Mutations in *NBS1* lead to *Nijmegen breakage syndrome*, which is associated with a high risk of malignant transformation during the individuals' life. Furthermore, *Breast Cancer Type 1/2 Susceptibility* (*BRCA1/2*) germline mutations carriers have a high probability of BC and ovarian cancers development [67].

NER pathway: NER removes large, bulky, DNA lesions. Among them, the most frequent are ultraviolet light-induced cyclobutane pyrimidine dimers, benzo[a]pyrene from tobacco smoke, aromatic amines, aflatoxin B1 or DNA adducts formed by chemotherapeutic agents [82].

NER can be divided into two distinct subpathways (Global Genome NER (GG-NER) and Transcription-Coupled NER (TC-NER)) [83, 84]. Unlike TC-NER, GG-NER works during the whole cell cycle removing strongly helix-distorting DNA lesions. TC-NER, on the other hand, focuses on DNA damage that blocks RNA polymerases during the transcription process. Although, both NER subpathways differ in recognition of the lesion, they utilize the same enzymes during the whole DNA repair process. Lesions repaired by GG-NER pathway are detected by XPC, an enzyme coded by a gene within *Xeroderma pigmentosum* (*XP*) complementation group [66, 85]. Seven remaining genes from *XP* group are *XPA* through to *XPG* as well as *XPV*. Upon the initial recognition of the damage, **XPC** enzyme recruits two DNA helicases, XPB and **XPD** [86]. Helicases unwind DNA around

the site of the lesion and allow the formation of preincision complex via accumulation of **XPA** and **XPG** [87, 88]. XPA binds *Excision Repair Cross-Complementation Group 1* (ERCC1)-XPF complex and initiates dual incision of the damaged DNA-strand [89]. Resulting 22-30 base oligonucleotide gap is subsequently filled and sealed by DNA polymerase δ/ ε and DNA ligase I, respectively [90].

BER pathway: BER removes small DNA lesions, predominantly deamination, alkylation and oxidative damage [74]. The whole process of BER shares considerable similarities with NER mechanism, including damage recognition, incision, repair, and DNA ligation.

The initial step in BER pathway represents searching for the damaged base by DNA glycosylase. Several glycosylases are present in human cells. Each of them is dedicated to specific alteration in DNA sequence. If one of those enzymes recognizes alkylation, deamination or oxidative damage, it flips and cleaves out the N-glycosyl bond of damaged base resulting in highly mutagenic apurinic/apyrimidinic (AP) site. A most common oxidative damage in a cell, 8-oxo-7,8-dihydro-2′-deoxyguanosine (8-oxoG; [74]), is specifically recognized by *8-Oxoguanine DNA Glycosylase* **(OGG1**; [91]). AP sites created by DNA glycosylase are consequently detected by *Apurinic/Apyrimidinic Endonuclease 1* **(APE1)** that cleaves phosphodiester bond to yield a 5' deoxyribose phosphate and 3' hydroxyl substrate for DNA polymerase. Specific polymerases utilized by the BER pathway are DNA polymerases β, δ and ε, respectively. Depending on the extent of the damage, one of those enzymes fills the gap with one or more new nucleotides. A final step in the BER pathway is ensured by DNA ligase. The whole enzymatic BER cascade is mediated by *X-ray Repair Cross Complementing 1* **(XRCC1)** scaffold protein [92].

MMR pathway: Proofreading activity of DNA polymerases δ and ε prevents accumulation of excessive errors during the replication of DNA molecule. Occasionally, the erroneous incorporation of the base is not recognized and removed by the exonuclease activity of DNA polymerase. Due to the increased probability of polymerase slippage events, the most prone DNA sequences to mismatch errors are nucleotide tandem repeats (e.g. microsatellites). If the polymerase undergoes a replicative mistake, it leads to the creation of base-base mismatches and indel loops, which, if not repaired properly, result in mutations in a newly synthesized DNA strand. Furthermore, some tumor suppressor genes and proto-oncogenes contain repetitive nucleotide motifs [93]. It is crucial for a cell to prevent mutations onset in these genes, that may ultimately result in the loss of their function. Thus, the MMR pathway

guards the genome against an accumulation of replication failures in regions with high homology and low complexity.

MMR pathway is initiated by the recognition of mispaired bases by *MutS Homolog 2* (MSH2) and *MutS Homolog 6* (MSH6) enzymes, followed by the recruitment of *MutL Homolog 1* (MLH1) and *Postmeiotic Segregation Increased 1 Homolog 2* (PMS2) enzymes. Consequently, *Proliferating Cell Nuclear Antigen* (PCNA) clamp activates MLH1 and PMS2 which in turn incise the daughter DNA strand. Replication error is removed by nuclease activity of *Exonuclease 1* (EXO1) or by strand-displacement synthesis. Finally, the gap is filled and the nick is sealed by the DNA ligase [94].

DSBs repair pathways: The vast majority of DSBs in eukaryotic cells is repaired by NHEJ and HR with only a minor contribution of alternative DSB repair pathway (such as Alternative end-joining (A-EJ) and Single-strand annealing (SSA)) [95]. Both A-EJ and SSA are still not well understood but they are considered as highly error-prone and mutagenic [96]. The most important dissimilarity between HR and NHEJ is a requirement of template DNA for the removal of the damage [97, 98]. HR utilizes sister, undamaged, chromatid for more precise correction of DSBs and it is used solely during the late S or G2 phase of the cell cycle following DNA replication. On the contrary, NHEJ works during the whole cell cycle without a commitment of replicated sister chromatid. Thus, NHEJ may lead to nucleotide indels and is commonly considered as an error-prone DNA repair pathway [99].

NHEJ pathway: NHEJ is initiated by sensing of broken DNA ends by Ku heterodimer, which protects DNA ends from exonucleolytic degradation. Ku subsequently forms a complex with DNA-PKcs, which in turn phosphorylates nuclease Artemis. Activated Artemis is then capable of cutting damaged nucleotides in DSBs sites [100]. In the case of blunt DSBs, commonly created from SSBs during the replication of DNA, broken ends are simply ligated without the necessity of nuclease Artemis [101]. On the contrary, up to 20 damaged nucleotides are removed while the DNA ends are jagged, resulting in very small single-stranded DNA (ssDNA) overhangs. This process is strictly controlled because more deleted nucleotides constitute a substrate for HR [101]. Inappropriate activation of the HR pathway in the absence of template DNA would lead to a great loss of genetic information. The DNA end-resection is controlled via competition between BRCA1 and *p53-Binding Protein 1* (53BP1) [102]. After the removal of damaged bases, generated small overhangs

are filled with new nucleotides by DNA Pol μ and/ or Pol λ [103, 104]. Finally, DNA nicks are sealed by DNA ligase IV along with *X-ray Repair Cross Complementing 4* (XRCC4) and Cernunnos [99].

HR pathway: Following DSB induction during the late S or G2 phase of the cell cycle, MRN heterotrimer complex, consisting of *Meiotic Recombination 11* (MRE11), RAD50 and NBS1 enzymes, recognizes the damage [105]. NBS1 then associates with *Ataxia-Telangiectasia Mutated* (ATM) kinase, which is a master regulator of the whole HR. Dozens of substrates are phosphorylated by ATM, including *Checkpoint Kinase 2* (CHK2) and P53 which consequently induce cell cycle arrest. Moreover, ATM kinase phosphorylates a histone variant 2AX [106]. Created γ-H2AX domains spread for distances up to a megabase, highlight a DSB site and allow recruitment of variety DSBs repair enzymes. Meanwhile, MRE11 nuclease in cooperation with *CTBP-interacting protein* (CTIP) and BRCA1 enzymes initiate large resection of broken DNA ends into ssDNA overhangs [105]. Generated overhangs are immediately covered by *Replication Protein A* (RPA) which stabilize the whole structure [107]. Enzyme RPA is in turn replaced by DNA recombinase RAD51 [108] via *X-Ray Repair Cross Complementing 2* (**XRCC2**). The whole RAD51 nucleofilament is stabilized by *X-Ray Repair Cross Complementing 3* (**XRCC3**) [109]. RAD51 is involved in the search for homologous DNA sequence in sister chromatid. Therefore, RAD51, along with other members of the RAD52 epistasis group (e.g. **RAD54**), facilitates strand invasion, formation and stabilisation of a so-called "D-loop structures" between a resected site of DSB and undamaged sister chromatid [110]. Thereafter, DNA polymerases fill ssDNA overhangs with new nucleotides according to the template sequence on a sister chromatid. After the DNA synthesis, junctions between chromatids are dissolved by BLM helicase and topoisomerase.

1.2.4. Telomere length and structural CAs

Enzymes involved in DDR are also linked to telomeres. Paradoxically, telomere biology studies have revealed a considerable part of DDR knowledge. Following chapter of the Thesis highlights the importance of telomere shortening in the formation of CAs (*Manuscript IV*) and it also describes current knowledge about complex interactions between DDR and telomeres (*Manuscript V*).

Telomeres are nucleoprotein structures assembled at the end of each linear eukaryotic chromosome and encompass a plethora of functions for a cell. Under physiological circumstances, telomeres help to control proliferative capacity, maintain genomic integrity and protect the ends of chromosomes against exonucleolytic deterioration [111, 112]. Proper telomere function relies on its structure formed by non-coding tandem repetitive DNA sequences (5´-TTAGGG-3´n; in vertebrates) and associated multimeric proteins of shelterin complex. The very ends of telomeres are terminated by approximately 200bp long 3´ ssDNA G-rich sequence [113]. Such overhang twists around double-stranded telomeric DNA and creates a protective lasso-like structure called telomeric "T-loop" and a smaller "D-loop" [114]. Both structures prevent recognition of DNA ends by DDR mechanisms [115] (*Figure 4*).

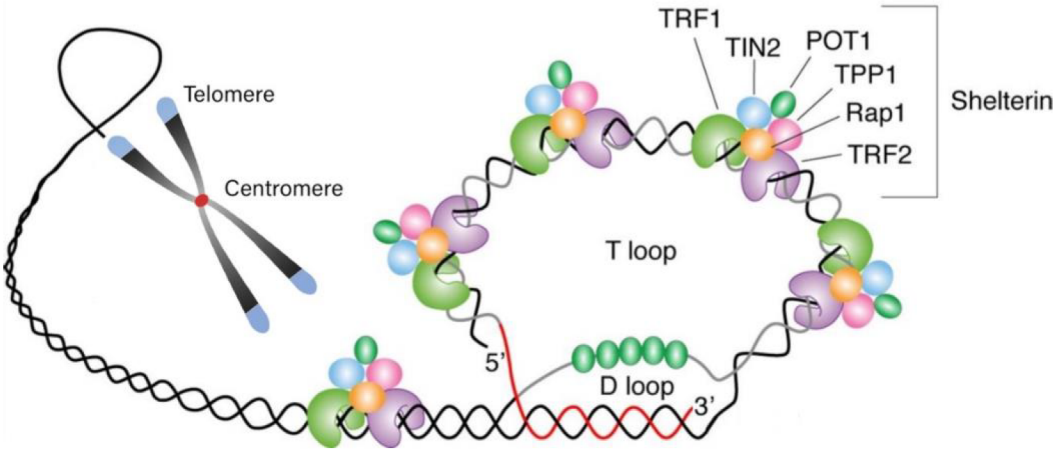


Figure 4: Simplified structure of telomeres. Edited from [116].

In humans, the length of telomeric DNA repeats is approximately 15 kilobase pairs at birth and telomere length is similar between different tissues within a newborn [117]. However, telomere length varies in cells of an adult and it also differs between homologous chromosomes [118]. Telomeres are deemed to progressively shorten through the individual´s life of about 100 nucleotides with each cell mitosis. Thereby, telomeres set the limit of cell divisions and are considered a "mitotic cell clock" mechanism [119]. Telomeric sequences undergo gradual shortening due to the unfinished DNA replication on the lagging strand [119], so-called "end-replication problem" This phenomenon is caused by the removal of the terminal RNA primer on the DNA lagging strand with each DNA replication [120].

It was shown that exactly 77 nucleotides represent a minimum length for functional telomeres. Below that certain threshold, telomeres cannot provide their end-protective

function anymore [121]. Uncapped chromosome ends resemble DSBs, which lead to the initiation of DDR and activation of cell cycle arrest and senescence state. Even though senescent cells do not proliferate any more, they may exhibit senescence-associated secretory phenotype. It leads to increased secretion of proinflammatory factors which may have an inauspicious impact on telomere length of surrounding tissues and paradoxically further contribute to telomere shortening [122]. Moreover, the senescence state may be bypassed in premalignant cells where the functional cell cycle checkpoint is absent. Those cells can continue in proliferation, which ultimately leads to the fusion of unprotected ends of chromosomes [115]. Erroneously joined chromosomes may induce the formation of anaphase bridges during cell division and initiation of DNA breakage-fusion-bridges cycles which eventually result in CIN phenotype [115].

Therefore, telomeres are a double-edged sword when it comes to cancer. On one hand, they exert a potent tumor-suppressor mechanism limiting the cell ability of uncontrolled growth. On the other hand, shortening of telomeres can lead to the telomere crisis and a formation of CAs [4, 123, 124]. In *Manuscript IV* we investigated telomere shortening in PBL of healthy subjects and cancer patients in association with frequency of CAs. Furthermore, we performed preliminary estimations of patients' survival in relation to CAs and telomere length.

1.2.4.1. Complex interactions between DDR and telomeres

Telomeric regions are highly susceptible to DNA damage either due to guanine-rich conformation or shelterin complex that inhibits DDR [125]. A relationship, if any, between SNPs in genes encoding shelterin subunits and telomere length was studied in *Manuscript VI.*

Shelterin comprises six-member protein complex (*Telomeric Repeat-Binding Factor 1* (**TRF1**), *Telomeric Repeat-Binding Factor 2* (**TRF2**), *Protection of Telomeres Protein 1* (**POT1**), *TRF1- and TRF2-Interacting Nuclear Protein 2* (**TIN2**), *Repressor/Activator Protein 1* (**RAP1**), and *Telomere Protection Protein 1* (**TPP1**) [124, 126]. TRF1 and TRF2, core subunits of a shelterin complex, interact directly with double-stranded telomeric DNA repeats, where they inhibit DDR and help to form a T-loop structure [112, 127, 128] along with RAP1 [129]. Cells with longer telomeres display a higher amount of telomere-bound TRF1 and TRF2, which may serve as an approximate indicator of

telomere length [130]. Long telomeres recruit a large number of TRF1 molecules which in turn regulate addition of other telomeric repeats by inhibition of telomerase activity [131]. Hand in hand with a negative feedback mechanism, deletion of TRF1 leads to increased telomere fragility and consequent senescence or programmed cell death [132, 133]. TIN2 also serve as a negative regulator of telomere length further underlying the entire complexity of shelterin [134]. TIN2 stabilizes the whole shelterin complex through connecting TRF1 and TRF2 [135]. Furthermore, TIN2 interacts with TPP1, which in turn mediates the binding of POT1 at telomere sites [112, 136]. TPP1, along with POT1, is important for recruitment of telomerase to ssDNA overhang [137] and regulation of telomerase activity. POT1 is also essential for the end protection against DSBs repair mechanism [138].

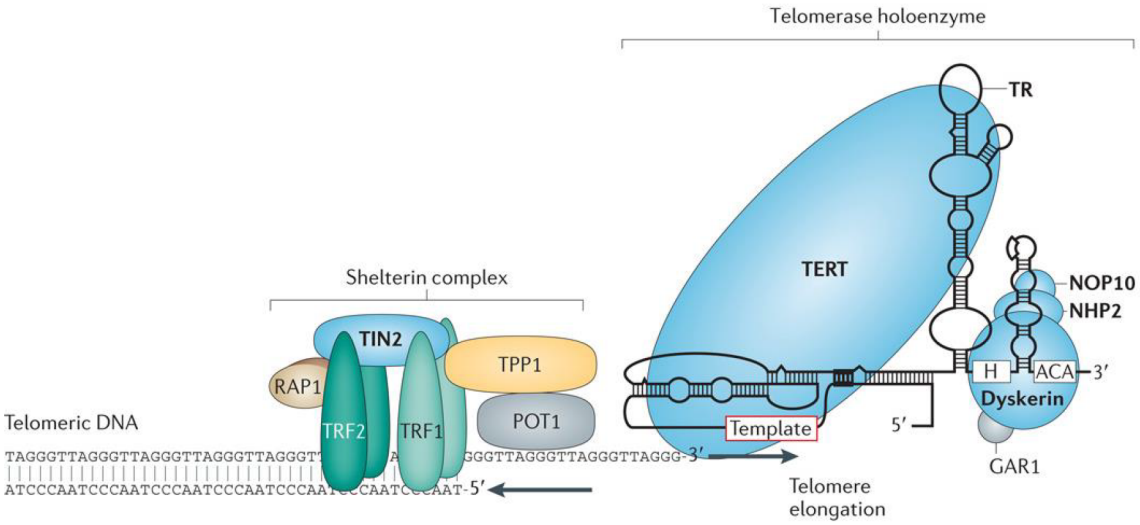


Figure 5: Interactions of shelterin subunits and telomerase holoenzyme. Adapted from [139].

Shelterin subunits recruit DDR proteins to control telomere homeostasis in terms of maintaining physiological telomere length but conversely also inhibiting improper activation of DDR at the end of chromosomes. It has been proposed that proteins important for proper DSBs DDR may control telomere length via MRN-ATM-TRF1 pathway. If TRF1 is posttranslationally modified by ATM kinase, its association with telomeres is reduced, which increases telomere fragility [140]. Shelterin complex has been also shown to inhibit NBS1/XRCC3-mediated excision of T-loop structure, which may lead to telomere deprotection and its excessive shortening [141]. Induced mutations in Ku subunits result in

massive telomere loss as well [142]. It is also known that patients with *BRCA1* mutations have longer telomeres [143].

The connection between impaired DSBs repair pathway and disrupted telomere length has been shown in patients with mutated ATM kinase. Those individuals, suffering from *Ataxia-telangiectasia* disorder, are hypersensitive to ionizing radiation, have a high amount of DSBs and furthermore, their fibroblasts display accelerated telomere shortening in comparison to fibroblasts derived from normal healthy individuals [144]. A connection between DSBs repair and telomere length maintenance pathways is also documented on Werner syndrome, which is caused by mutations in a gene encoding *WRN* helicase. Werner syndrome is marked by disrupted telomere length, high amount of CAs in fibroblasts and a high incidence of cancer [145]. WRN is important for resection of DSBs during HR [146] and it also interacts with DNA glycosylases during BER [147].

Regarding excision pathways in telomere biology [148, 149], it has been shown that accumulation of unrepaired 8-oxoG at G-rich telomere strand may lead to frequent point mutations, erroneous binding of shelterin proteins and finally a loss of function of telomeres [150-152]. Thus, erroneous BER pathway may not only influence the number of nucleotide mutations in the genome but it may also accelerate the formation of CAs via disruption of telomere length [151, 153].

It is also possible that compromised MMR pathway could diminish telomeric repeats due to the repetitive nucleotide structure of telomeres [154]. About 10-15% of sporadic CRC show MSI phenotype and, in general, MSI is caused by defects in several genes involved in MMR [155, 156]. A tumor is MSI when several loci encompassing short tandem repeats show more than two different numbers of repeat motif as would be regular in bi-allelic status [157]. CRC patients with MSI status usually evince better prognosis. It was described that those patients have up to 15% higher survival rate in comparison to microsatellite stable tumors [158, 159]. Favorable prognosis in these patients may be modulated by telomere length [160].

The above examples highlight the importance of DDR in telomere biology and genomic stability and we suggest that this area deserves more attention in the future. In *Manuscript V* we investigated a link between DSBs DRC and telomere length in PBL obtained from cancer patients and healthy individuals. The association between disrupted MMR in CRC patients and telomere length in tumor tissue was further investigated in *Manuscript VI*.

1.3. Telomere length in tumor cells

All cancer cells, without exception, have to activate telomere length maintenance mechanism to achieve replicative immortality [161]. Telomere sequence attrition can be restored by the activity of ribonucleoprotein enzyme telomerase, or by alternative telomere lengthening (ALT) mechanism. Each of these two has a different clinical impact on a cancer patient [162]. Reactivation of telomerase is a predominant telomere maintenance mechanism, observed approximately in 85-90 % of human malignancies [163]. Remaining 10-15 % of tumors utilize ALT for telomere length preservation [164-166], however, the frequency depends on the origin of tumor tissue [167].

ALT: ALT can be characterized as a de novo synthesis of telomeric sequences without the commitment of telomerase. ALT is highly prevalent in tumors of the central nervous system and in certain types of sarcomas. On the contrary, ALT mechanism has been reported in only a minority of CRC and BC patients [167]. Although little is known about ALT-mediated telomere elongation, this mechanism displays certain similarities with repair of DSBs via HR [168]. It has been proposed that ALT-positive cells can utilize the telomeric sequence from sister chromatid or also another DNA region within the same telomeric sequence as a copy template for a telomere length maintenance or elongation. The G-rich overhang of telomere basically enables "looping out", invasion of sister chromatid or the same telomere, creation of a Holliday junction and replication of the DNA template which finally leads to telomere elongation [169]. ALT-positive tumor cells display more than double the telomere length compared to the healthy somatic cells. On the contrary, telomerase-positive cells have much shorter telomeres [169]. Recently, two ALT subpathways have been described. Both of them may be a potential target for the treatment of tumors lacking telomerase activity [170].

Telomerase: RNA-dependent DNA polymerase activity of the telomerase enzyme is utilized to overcome the gradual telomere attrition [171]. The main parts of telomerase holoenzyme are *Telomerase Reverse Transcriptase* (TERT; [172]) and catalytic protein subunit *Telomerase RNA Component* (TERC; [173]). Another main part of telomerase represents Dyskerin enzyme, which stabilizes and protects TERC [174]. A core RNA sequence of

TERC is complementary to the telomeric TTAGGG repetitions and serves as a template for synthesis of telomeric repeats [171].

Even though telomerase is broadly expressed in the germline during embryogenesis, *TERT* gene is embedded in a condensed chromatin domain and tightly repressed in later development [175]. Healthy human somatic cells, with a few rare exceptions, display undetectable or very low *TERT* expression [175]. Generally, only stem and cancer cells can deal with telomere shortening in adult humans through the expression of telomerase [176].

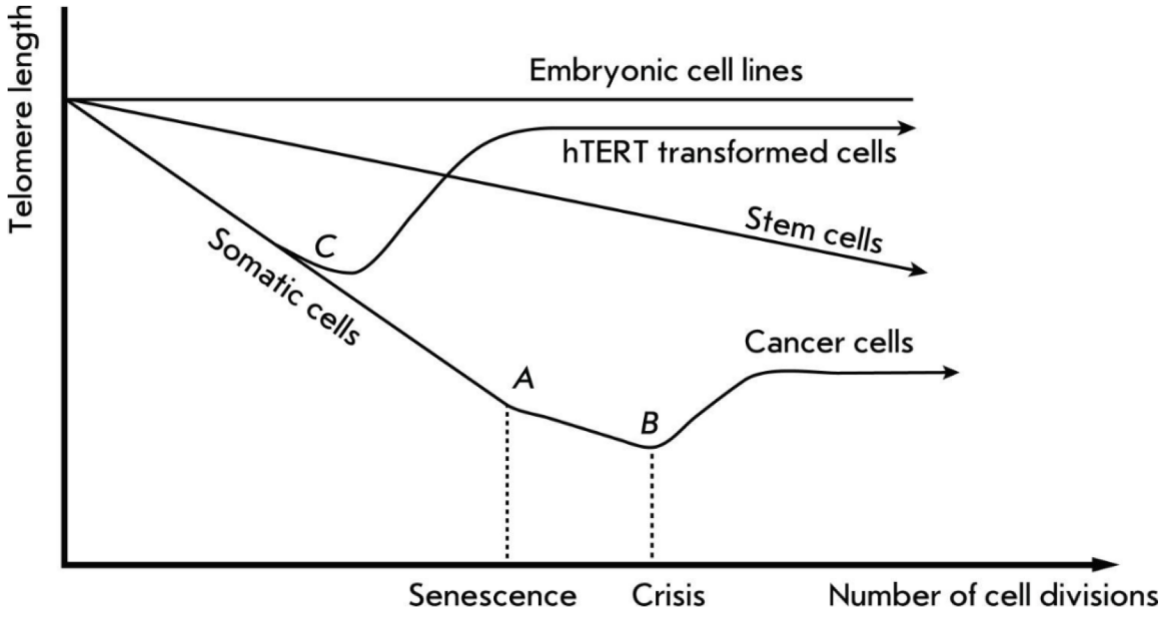


Figure 6: Relationship between telomere length and number of cell divisions. A) Replication senescence, B) Telomere crisis with subsequent apoptosis or malignant transformation, C) Cell transformation by human *TERT* gene. Adapted from [177]

Discovery of mutations within *TERT* gene promoter region in familial and sporadic melanoma cancer patients revealed a mechanism of telomerase reactivation. Horn *et al.*, described up to two-fold increase in transcription of *TERT* gene caused by binding of transcription factors from *E-twenty-six* (ETS) family in mutations-created binding motives in *TERT* promoter region. Some members of the ETS family are downstream targets of proteins from the MAPK pathway (e.g. *V-Raf Murine Sarcoma Viral Oncogene Homolog B1* (BRAF)), frequently mutated in melanoma cancer cells [178]. Those *TERT* promoter mutations, restricted to human cancer cells, have been also described in many other malignancies (summarized by [179]). Furthermore, the expression of *TERT* gene has been shown to be regulated by various hormones (e.g. estrogen) [180], which directly bind into a

promoter of *TERT* gene. Overexpression of *TERT* in solid tumors has been described in association with tumor progression, ability to metastasize and unfavorable prognosis of cancer patients [181].

As noted above, telomere length is a dynamic end-result of each eukaryotic chromosome regulated by complex telomerase-dependent or independent molecular pathways [182]. Currently, it is possible to detect telomere length of each chromosome by the Universal Single Telomere Length Analysis (U-STELA) or Telomere Shortest Length Assay (TeSLA) [183, 184]. Telomere Restriction Fragment and Multiplex Monochrome Real-Time PCR (MMQPCR) are used to investigate an average telomere length [185].

Some studies documented that, despite the acquired telomerase activity, telomere length in tumors of CRC patients is, in general, much shorter compared to adjacent non-cancerous mucosa [186-188] and only about 25 % of tumors show telomere elongation [186]. ALT mechanism is not commonly utilized by CRC cells [167], therefore the phenomenon of telomere lengthening in CRC requires further investigation. Some authors also showed more pronounced telomere shortening in tumor tissue compared to precancerous lesions [46] and the shortest telomeres in the early stage of tumor development according to TNM classification [186]. These results indicate that critical shortening of telomeres is considered as an early event of malignant transformation [189].

Molecular features of CRC patients, including *Ki-ras2 Kirsten Rat Sarcoma Viral Oncogene Homolog* (*KRAS*), *BRAF*, and *Tumor Protein p53* (*TP53*) mutations were also investigated in association with telomere length. KRAS mutations increase telomerase activity [190]. However, the results are contradictory. A negative effect of *KRAS* mutations on telomere length was observed by Balc'h *et al* [186]. In the other studies, the influence of *KRAS* and *BRAF* mutations did not have an effect on telomere length in tumor tissue of CRC patients [46, 186].

Currently, our knowledge regarding the telomere biology in tumor tissue is limited. The above examples implicate the necessity to investigate possible factors affecting telomere length in tumor tissue to improve a molecular diagnosis. An effect of clinicopathological features of CRC patients on relative telomere length (RTL) was investigated in *Manuscript VI.*

2. Aims

The aim of this Thesis was to address the following points:

- Do functionally relevant gene variants in major mitotic checkpoint and DNA repair genes modulate the frequency of structural CAs in PBL of healthy subjects?

- Does the application of GWAS disclose novel gene variants in relation to structural CAs?

- Is the frequency of structural CAs associated with telomere length in PBL of healthy subjects and cancer patients?

- Do the frequency of structural CAs and telomere length in PBL predict cancer risk and patients' long-term overall survival (OS)?

- Do cancer patients evince decreased DSBs DRC in PBL? Is telomere length in PBL connected to the DSBs DRC?

- Is telomere length in tumor tissue, adjacent mucosa, metastases and blood affected by clinicopathological features among CRC patients?

- Is telomere length in tumor tissue associated with patients' OS?

3. Material and Methods

3.1. Study populations

Manuscript I: The number of genotyped participants from the Czech and Slovak Republics, who have been included also in *Manuscript II*, ranged from 330 to 729. All subjects were sampled for peripheral blood.

Manuscript II: This study included 2196 healthy individuals from the Czech and Slovak Republics. The whole group consisted of participants with defined occupational exposure to DNA damaging agents (e.g. small organic compounds, anaesthetics, mineral fibres) and also unexposed healthy individuals. All individuals were sampled for peripheral blood.

Manuscript III: GWAS has been conducted on 576 healthy individuals followed by genotyping of two replication sample sets, which involved 482 and 1288 participants. All individuals were recruited in Slovak and Czech Republics. About 57 % and 23 % of participants, involved in GWAS and second replication set, respectively, were individuals with professional exposure to some form of genotoxic agents. About 46 % of the participants in the first replication set were newly diagnosed primary cancer patients (BC, CRC or lung cancer). For detailed characterization of the study population see *Manuscript III*, *Table I.*

Manuscript IV: The study population consisted of three groups of incident cancer patients (151 BC, 96 CRC, and 90 lung cancer) and the group of 335 healthy control individuals in order to match cases by age and gender. The patients and controls were recruited in the Czech Republic.

Manuscript V: This study involved 91 newly diagnosed cancer patients (47 BC and 44 CRC patients) and 90 matched healthy individuals. All individuals were recruited in the Czech Republic and were sampled for peripheral blood.

Manuscript VI: Two different groups of CRC patients were included in this study. Paired samples of tumor tissue and adjacent non-malignant mucosa were collected from 661 patients during the surgical resection. An additional sample of peripheral blood was available from 164 of these CRC patients and the sampling was carried out prior to the surgery. Moreover, primary tumor, adjacent mucosa, liver metastatic tissue and paired adjacent liver tissue were available from 12 patients of this group. The second group of 122

CRC patients, with the metastatic and adjacent liver tissues and without primary tumors, was included in the study.

Manuscript VII: This summary review article was based on the results observed in *Manuscript I*, *II* and *V*. Further studies, investigating the relationship between CAs and variants in genes encoding xenobiotic-metabolising enzymes, TP53 gene and Cyclin D1, were included in the review article.

3.2. Cytogenetic analysis

A standard cytogenetic procedure was utilized for evaluation of structural CAs frequency in PBL [31] (Manuscript *I–IV*). PBL cultivation was carried out in complete medium for 48 hours until the first metaphase of mitosis when the cell division was stopped using colchicine. Then, cytogenetic slides were prepared from PBL of each individual and CAs frequency per 100 metaphases was analysed by two independent scorers.

3.3. Mutagen sensitivity assay

The whole method was performed according to the previously described protocol [191]. Briefly, two whole-blood samples from each subject, phytohaemagglutinin-stimulated in complete medium, were cultivated for 72 hours. Five hours before harvesting, during the late S and G2 phase of the cell cycle, mutagen (radiomimetic drug bleomycin) was added in one of the two cultures. The second one served as a reference for detection of baseline CAs level. Two hours prior harvesting, PBL were arrested in the second metaphase of mitosis by colchicine. After the cytogenetic procedure, at least 100 cells per subject were analyzed microscopically for mutagen-induced CTAs level by two independent scorers. Chemiluminescent measurement of γ-H2AX level was used for the assessment of initial DNA damage after the bleomycin treatment.

This approach was used for measurement of mutagen sensitivity in PBL of incident CRC, BC patients and healthy subjects (*Manuscript V*) as an indirect measure of DSBs DRC.

3.4. Relative telomere length measurement

In our projects, we used MMQPCR assay for RTL assessment in tumor tissue, adjacent non-malignant mucosa, liver metastases, liver tissue (*Manuscript VI*) and blood

(*Manuscript IV, V and VI*) of incident cancer patients and healthy subjects. The whole assay was carried out according to the original methodology described previously [185], with important modifications [192]. Multiplexing of the method allowed us to use a single fluorescent dye, SYTO 9, during a real-time quantitative polymerase chain reaction (PCR) for a collection of a signal from two target sequences differing greatly in copy number (telomere and albumin amplicon in our case). All reactions were carried out in triplicates in MicroAmp Optical 384-Well Reaction Plate. MMQPCR experiments were performed on Viia 7 Real-time PCR System (Applied Biosystems) using two simultaneous programs to acquire the respective cycle threshold values for telomere sequences and also for the albumin gene. The cycle threshold values for albumin amplicon were collected above the melting temperature of the telomere product [185]. The standard curve, created by pooling DNA from 30 individuals with different age, was used to quantify the telomere and albumin product.

3.5. MSI status measurement

For assessment of MSI status of CRC patients studied in *Manuscript VI*, multiple primary microsatellite loci (*BAT-25*, *BAT-26*, *NR-21*, *NR-24*, *NR-27*) were analysed simultaneously by MMQPCR with pseudomonomorphic mononucleotide markers using primers labeled with FAM, HEX or NED. Analysis of PCR products was accomplished on 5 % denaturing gel electrophoresis as described previously by [193]. Tumor DNA samples were compared to each other, and samples showing instability at three or more investigated loci were scored as MSI.

3.6. Genotyping analysis

TaqMan allelic discrimination assay (Applied Biosystems, Foster City, California) was utilized for the detection of SNPs in target genes.

The variants were investigated in genes involved in mitotic spindle (*Manuscrip I*):
BUB1 rs1801376 (A>G SNP; a coding Arg/Glu variant), *BUB3* rs3808960 (G>T SNP; a promoter polymorphism), *MAD2L1* rs903147 (A>C SNP; a promoter polymorphism), *CENPF* rs438034 (G>A SNP: a coding Arg/Gly variant), *ESPL1* rs6580941 (C>T SNP; a promoter polymorphism), *NEK2* rs701928 (T>A SNP; a promoter polymorphism),

PTTG1 rs1862392 (T>A SNP; a promoter polymorphism), for *ZWILCH* rs3087660 (A>G SNP; a 5′UTR polymorphism) and *ZWINT* rs2241666 (G>A SNP; a coding Arg/Gly variant). All promoter SNPs were predicted to occupy transcription factor binding sites.

Following SNPs sites were investigated in genes involved in DNA repair pathways (*Manuscript II*): NER pathway: *XPD* rs13181 (T>G SNP; Lys751Gln), *XPG* rs17655 (C>G SNP; Asp1104His), *XPC* rs2228001 (A>C SNP; Lys939Gln), XPA rs1800975 (G>A SNP; a 5′UTR polymorphism), BER pathway: *XRCC1* rs1799782 (C>T SNP; Arg194Trp), rs25489 (G>A SNP; Arg280His) and rs25487 (G>A SNP; Arg399Gln), *OGG1* rs1052133 (C>G SNP; Ser326Cys), *APE1* rs1130409 (G>T SNP; Asn148Glu), DSBs pathway: *XRCC2* rs3218536 (G>A SNP; Arg188His), *RAD54L* rs1048771 (C>T SNP; Ala730Ala) and *XRCC3* rs861539 (C>T SNP; Thr241Met), *NBS1* rs1805794 (C>G SNP; Glu185Gln).

Following SNPs sites were investigated in genes coding for shelterin complex and telomerase (*Manuscript VI*): *RTEL1* rs2297440 (T>C SNP; intron variant), rs2738780 (T>C SNP; intron variant), rs7261546 (C>G/C>T SNP; intron variant), rs755017 (A>C/A>G SNP; synonymous variant), *FEN1* rs174538 (G>A SNP; a 5′UTR polymorphism), *POT1* rs7784168 (T>C SNP, intron variant), rs10250202 (A>C SNP, intron variant) *TERF1* rs2306494 (G>A SNP, intron variant), rs3863242 (C>T SNP, intron variant), *TRF2* rs251796 (A>G SNP, intron variant), rs3785074, *TERT* rs10936599 (C>T SNP, synonymous variant), rs2736118 (T>C SNP, intron variant), rs2736098 (C>T SNP, synonymous variant), rs2736100 (C>A SNP, intron variant), rs2242652 (G>A SNP, intron variant), rs2736108 (C>T SNP, intergenic variant), rs7705526 (C>A/C>T SNP, intron variant), rs10069690 (C>T SNP, intron variant), rs2736109 (C>T SNP, intergenic variant), rs2853691 (T>C SNP, synonymous variant), rs2853677 (G>A SNP, intron variant), rs2853676 (T>A/T>C SNP, intron variant), rs33954691 (G>A SNP, synonymous variant), rs3816659 (A>G SNP, intergenic variant), rs4246742 (T>A SNP, intron variant) and *TERC* rs3772190 (G>A SNP, non coding transcript variant), rs12696304 (C>G SNP, regulatory region variant).

4. Results and Discussion

The subject of the Thesis was to unveil the influence of relevant gene variants in major mitotic checkpoint and DNA repair genes on the frequency of structural CAs in PBL of healthy subjects. We also intended to identify novel gene variants in relation to structural CAs. Furthermore, we studied DSBs DRC in PBL of cancer patients, factors affecting telomere length and investigated a link between shortened telomeres and the survival of cancer patients.

Manuscript I: In our study "***Genetic variation in the major mitotic checkpoint genes associated with chromosomal aberrations in healthy humans***", we focused on 9 selected SNPs in mitotic checkpoint genes (*BUB1, BUB3, MAD2L1, ESPL1, PTTG1, CENPF, NEK2, ZWILCH,* and *ZWINT*) in association with the frequency of acquired non-specific structural CAs in PBL of 729 healthy subjects. In addition, we also tested binary gene-gene combinations because mitotic checkpoint represents a highly organized orchestra of complex interactions.

Results for individual SNPs revealed only a minor effect on a level of CAs, which is probably due to high conservation of the whole mitotic checkpoint mechanism. Only rare homozygous variant genotypes in *ZWINT* (rs2241666) and *BUB1* (rs1801376) were associated with the increased frequency of CSAs (P=0.03) and CTAs (P=0.03) in a group of controls (i.e. individuals with a frequency of CAs <2 %), respectively. However, analysis of gene pairs showed 21 significant association with an increased level of CTAs, CSAs and total CAs (CAtot). The most significant genetic models for each interaction are shown in a full version of attached *Manuscript I* (see *Table 2*). *PTTG1*, *ZWINT*, *ZWILCH*, and *MAD2L1* genes were the most frequently involved in any of two binary gene-gene interactions. Remarkably, *PTTG1-ZWILCH* & *PTTG1-ZWINT* gene interactions were significant for multiple types of CAs, including CSAs (P=0.01 & P=0.02), CTAs (P=0.006 & P=0.003) and CAtot (CAtot; P=0.01 & P=0.03), respectively. The role of PTTG1 and ZWILCH/ZWINT enzymes is very different during the mitotic checkpoint. PTTG1 is an oncogene and it functions at the level of Cohesin [194]. On the other hand, ZWILCH/ZWINT are important for activation of the mitotic checkpoint at kinetochore [63]. In a previous study, it was shown that induced single point mutation in *PTTG1* is sufficient to trigger the formation of CAs [195]. In addition, SNPs in *PTTG1* were documented to be associated with BC risk in a case-

control study conducted in 1492 healthy participants and 698 primary BC patients [196], which further highlights the relevance of PTTG1 in malignant transformation.

Presumably, low-risk alleles individually may not add up to the frequency of CAs, however, binary gene-gene interactions at different mitotic checkpoint function may multiply the effect and result in apparent malfunctions of the whole mitotic checkpoint clockwork.

Manuscript II: The manuscript entitled: "***Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects***" focused on SNPs in genes involved in DDR and their influence on the frequency of non-specific structural CAs in PBL obtained from 2196 healthy individuals. Functionally relevant SNPs in 11 genes involved in BER (*XRCC1, OGG1, APEX1*), NER (*XPA, XPD, XPG, XPC*) and DSBs (*XRCC2, XRCC3, NBN, RAD54L*) repair pathways were genotyped.

In this study, analysis of individual SNPs showed a link between *XPD* Lys751Gln homozygous variant and lower CTAs frequency (P=0.004). This result verified the previous findings of our group [34]. Moreover, we found a significant influence of *RAD54L* Ala730= heterozygous variant genotype on increased CSAs frequency (P=0.03), which is especially interesting considering the fact that synonymous substitution of the amino acid may have a potential impact on protein function [197]. RAD54 stabilizes RAD51 recombinase during DSBs repair [110] and recent evidence indicates that RAD54 deficiency is associated with CAs [198]. In general, individual gene variants modulated CAs frequency only modestly. The similar trend was described in *Manuscript I* and also in the study conducted in our collaborating laboratory (German Cancer Research Center, Heidelberg), which investigated SNPs in metabolic genes in association with CAs level [21]. However, the most important results from our study emerged from assessing pair-wise genotype interactions of studied genes. We discovered 14 gene-gene interactions significantly modulating CA frequencies, 12 modulating CSA and 9 modulating CTA frequencies. The strongest interactions, summarized in *Manuscript II* (see *Table 4*), always included gene pairs from two different DNA repair pathways. The results illustrating the most important gene-gene interactions with the influence on different types of CAs were published in our review article (see *Figure 3* in *Manuscript VII)* and the role of individual enzymes coded by those genes is shown in *Figure 7*.

Based on our results, we suggest that individual SNPs in genes coding for DNA repair proteins may have an only modest influence on the frequency of CAs in PBL. More

importantly, we have provided evidence that accumulation of CAs requires complex interplay between different DNA repair (and most likely another cellular) mechanisms.

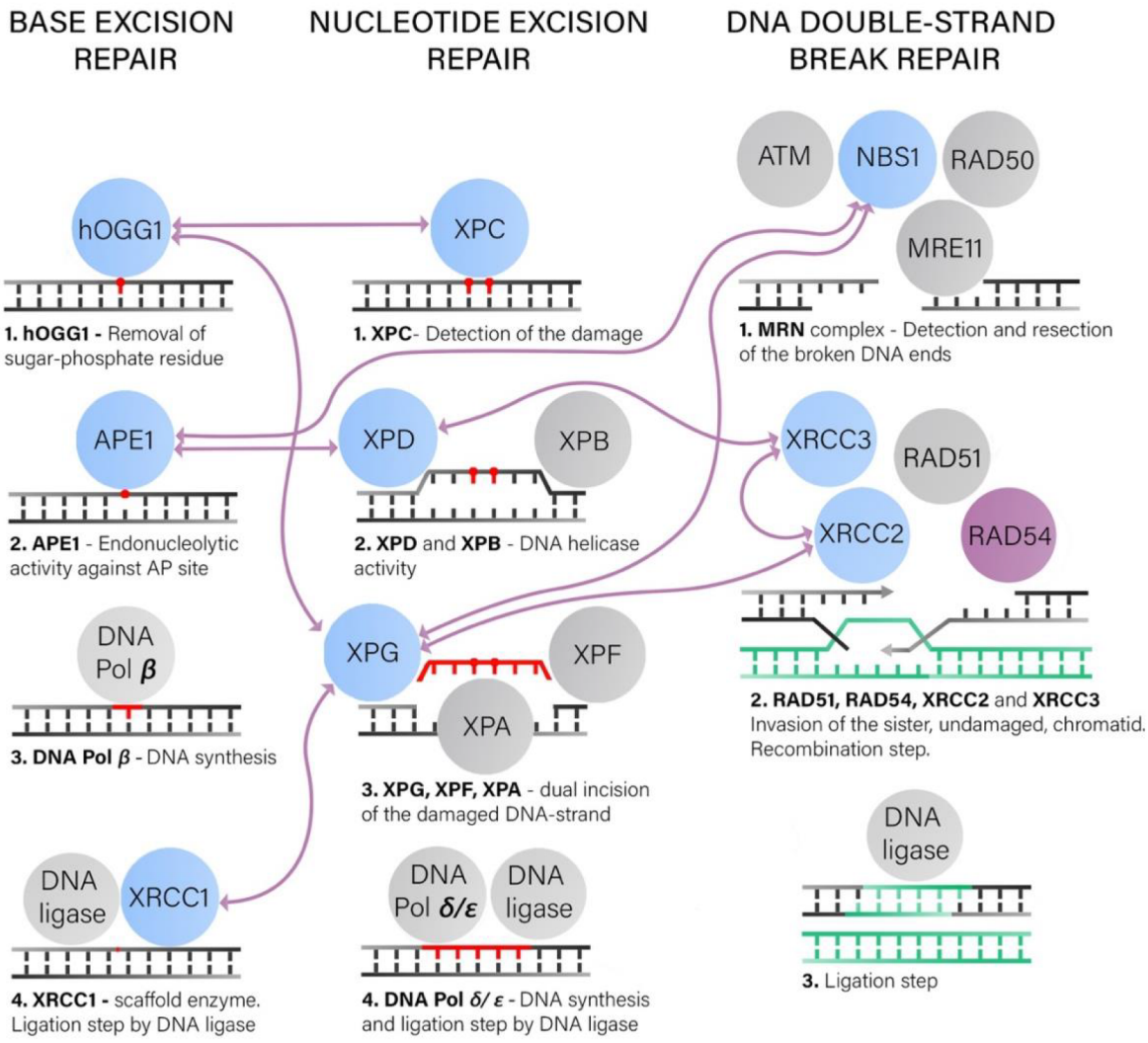


Figure 7: Complex pair-wise genotype interactions of studied genes (*Manuscript II)*

Manuscript III: The aim of the manuscript entitled: "***Genetic variation associated with chromosomal aberration frequency: A genome-wide association study***" was to identify novel inherited genetic variants predisposing individuals to higher frequency of CAs in PBL. The whole study design was based on our previous findings that showed a relationship between a frequency of CAs in PBL and SNPs in mitotic checkpoint, DNA repair (*Manuscript I, II*) and metabolic genes (summarized in *Manuscript VII*).

Regarding the characteristics of our studied populations, we confirmed a strong association between increased frequency of CAs and occupational exposure to genotoxic

agents [21, 199, 200]. We also confirmed the relationship between the increased frequency of CAs and cancer disease, which is in consistency with [29].

As a result of the GWAS analysis, 11 SNPs were selected for replication. The validation cohort supported the association of rs1383997, rs2824215, and rs983889 for increased frequency of CAs. The potential function of all those highly linked SNPs was characterized using *in silico* analysis. Rs1383997 is located in the locus that encodes *Natural Antisense Transcript* (*NAT*) with antisense complementary to *Transient Receptor Potential Cation Channel Subfamily A Member 1* (*TRPA1*) and other genes [201]. Previously, up to 28-fold up-regulation of *TRPA1* mRNA was found in kidney cancer [202]. *Transient Receptor Potential* (TRP) channels have been shown to be important for Ca^{2+} ions signaling. Dysregulation of those channels may cause changes in intracellular calcium homeostasis, consequent disruption of cellular processes, which can also promote pathophysiological cancer hallmarks [203]. Rs2824215, the second SNP from our analysis that reached a suggestive level of significance between GWAS and replication set, is located at 21q21.1 loci in long intergenic non-coding RNA. Deletion of this loci has been previously shown in association with chromosomal rearrangements [204]. Rs983889, the last discovered gene variant in our study, modulated frequency of CTAs. Based on our analysis, we found that it represents intronic SNP in a gene encoding the *F-box and Leucine-Rich Repeat Protein 7* (*FBXL7*). Interestingly, Aurora A kinase, an oncogene with indispensable function during mitosis, regulates FBXL7 [205]. Importance of FBXL7 during mitosis may explain its connection to CTAs because this type of CA arises during late S or G2 phase of the cell cycle [23].

Overall, we identified several new loci, with genes important in mitosis and intracellular signaling, in association to elevated CAs frequency.

Manuscript IV: In the study "***Chromosomal damage and telomere length in peripheral blood lymphocytes: cancer risk and patients' long-term survival***", in preparation, we investigated a prognostic value of telomere length and different types of structural chromosomal damage (CSAs, CTAs, aberrant cells with CAs (ACs) and CAtot) measured in PBL of newly diagnosed CRC, BC, and lung cancer patients. The present Manuscript represents a continuation of the previous study conducted in our laboratory where an increased level of all types of CAs in PBL of BC and lung cancer patients in comparison with healthy individuals was observed [29]. The risk of BC and lung cancer was associated with all investigated types of CAs (summarized in *Manuscript IV*, *Table II*), which further

confirmed the other results from several prospective studies (summarized by [206]). In the present Manuscript, chromosomal damage and telomere length measured in PBL of those patients were analysed to determine their prognostic utility. Furthermore, we evaluated for the first time the relationship between the frequency of non-specific structural CAs and telomere length in PBL of both, cancer patients and healthy subjects.

Regarding the prognostic value of CAs or telomere length, we did not find any association between OS or recurrence-free survival (RFS) by using univariate survival analysis. However, when we further explored the prognostic value of CAs together in association with clinicopathological data by multivariate survival model, we found that BC patients in TNM 0+I+II+III stages and CRC patients in TNM III stage with an increased level of CTAs in PBL evinced poor OS. There is currently no available literature to discuss this association and we suggest that a larger cohort of patients is necessary to clarify the prognostic value of CAs in PBL of cancer patients. With respect to telomere length, a meta-analysis published by Xu *et al*., showed a correlation between short telomeres in PBL and poor OS in patients with solid tumors [207]. Although we did not find such association in our group of patients, we further investigated the prognostic utility of telomere length measured in tumor tissue of CRC patients in *Manuscript VI.*

We also validated the negative correlation between the level of CAs and telomere length in the group of healthy individuals (ACs; P=0.0004, CAtot; P=0.0003 and CSAs; P=0.0002) [4]. This particular outcome may be explained by the fusion of excessively shortened telomeres and consequent damage of chromosomes during mitosis [115]. However, a similar result was not observed in any group of cancer patients. This phenomenon has not yet been described before and we can hypothesize that not only shortening of telomeres but also other pathways may contribute to the formation of CAs in PBL of cancer patients.

Furthermore, our group of BC patients displayed longer telomeres compared to healthy women (P<0.0001) and individuals with longer telomeres had also increased risk of the BC disease (P<0.0001). A possible explanation for longer telomeres in BC patients may be due to the presence of estrogen, which increases telomerase activity [208].

Based on our results, increased level of CTAs in PBL may be a factor affecting OS in BC and CRC patients. We discovered a negative correlation between a level of CAs and telomere length in a group of healthy individuals. Longer telomeres in PBL may be associated with an increased risk of BC.

Manuscript V: The case-control study entitled: "***Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients***" is based on other projects carried out in our laboratory, when excision DRC was measured in PBL or tumor tissue/adjacent mucosa of sporadic CRC patients [71, 209, 210]. In our present study, we focused on the measurement of mutagen-induced CTAs as an indirect phenotypic marker of DSBs DRC in PBL of CRC and BC patients and healthy controls without the disease. Moreover, we measured telomere length in PBL of these individuals to identify a potential relationship between telomere length and DSBs DRC.

We found statistically significant higher bleomycin-induced CTAs in PBL of CRC patients (P=0.03), indicating higher sensitivity towards radiomimetic drug and lower DSBs DRC in PBL. Our results are in agreement with the hypothesis of Hsu *et al.*, who suggested that patients with cancer of tissues that are directly exposed to the environment (aerodigestive tract) evince increased bleomycin sensitivity in PBL [191]. Our results are also consistent with previous studies [211, 212]. Ankathil *et al.*, investigated the bleomycin-sensitivity in familial and sporadic CRC patients and compared the frequency of induced CTAs between the unaffected family members and healthy controls without the cancer disease. Increased bleomycin-induced CTAs were observed in PBL of familial and sporadic CRC patients compared to the unaffected relatives and healthy individuals, respectively. This study demonstrated that bleomycin hypersensitivity may be related to CRC predisposition [212]. Shao *et al.*, also documented that patients with distally located tumors of colon showed higher bleomycin sensitivity in PBL compared to the right-sided malignancies, which further support the hypothesis that environmental carcinogens play a greater role in the etiology of left-sided colon tumors [211].

In our group of BC patients, the bleomycin sensitivity profile did not differ compared to healthy individuals, which is consistent with previous results [191]. However, other groups found the opposite result. In a case-control study conducted on 196 Chinese BC patients and 211 healthy controls, a greater number of bleomycin-induced CTAs was found in a group of BC cancer patients and the bleomycin sensitivity profile was modulated by exposure to the tobacco smoke [213]. Another study, including 164 BC patients and 165 healthy controls found increased bleomycin sensitivity in association with a risk of the disease [214].

Furthermore, we showed that patients with telomere length below the median, irrespective of cancer type, had a significantly higher amount of bleomycin-induced CTAs in PBL (P=0.02), which indicates a correlation between impaired DSBs and excessive

telomere shortening. This finding is novel and of particular importance. Telomeres not only inhibit erroneous induction of DDR, but they also utilize the function of many DSBs repair factors for telomere processing, telomere replication and the establishment of telomere protection [125, 215]. For illustration, Ku70 involved in NHEJ pathway stimulates fusion of dysfunctional telomeres yet protects chromosome ends from HR [216]. It is also known that 53BP1 inhibits BRCA1-mediated resection of unprotected telomeres during replication, which blocks initiation of HR and consequent disruption of telomere length [217]. Presumably, all principle DNA repair pathways are, to some extent, important for proper telomere homeostasis (reviewed in [218]). For illustration, deficiency of enzymes involved in BER pathway, APE1 and OGG1, leads to telomere length deregulation [219]. Studies on NER pathway are surprisingly scarce [218]. In *Manuscript VI* we mapped the role of impaired MMR pathway in CRC patients and its relationship to telomere length in tumor tissue.

Overall, our results support the hypothesis, that disruption of telomere length is indicative of DDR defects [220]. Still, the potentially important role of more than, currently known, four hundred proteins involved in DDR along with their complex interactions with enzymes important for telomere homeostasis remains unclear and we suggest this field deserves more attention in the future research.

Manuscript VI: In the study "***Relationship of telomere length in colorectal cancer patients with cancer phenotype and patient prognosis***", in preparation, we measured telomere length in tumor tissue, adjacent non-cancerous mucosa, peripheral blood, liver metastases and liver non-malignant tissue obtained from 661 CRC patients. We attempted to unveil the influence of clinicopathological characteristics of CRC patients on telomere length in distinct tissues. We also hypothesized that telomere length in tumor tissue may serve as a potential prognostic marker for patients OS.

Regarding the comparison of telomere length in distinct tissues of our group of patients, we observed shorter telomeres in 74 % of tumors compared to non-cancerous mucosa. This result has been previously reported [188, 221, 222] and may be explained by the proliferating activity of the rapidly growing tumors. Shorter telomeres were also identified in 83 % of liver metastases when compared to primary tumor tissues and 64 % of liver metastases had shorter telomeres compared to non-cancerous liver tissues. The vast majority of studied patients, who present with metastases, underwent chemotherapeutic treatment. This may be the explanation for shorter telomeres in metastatic tissue.

Interestingly, we found that increasing telomere length in tumor tissue correlated with TNM staging, with the shortest telomeres in TNM stage I (P=0.001). A similar result was obtained by Le Balc´h *et al*., who also documented the most pronounced telomere shortening in the early stage of tumor development, measured within 125 CRC patients [186]. The TNM system is a robust prognostic marker and tumor telomere length could serve as a possible additional factor for specification of patients' OS. Some evidence for longer telomeres as a poor prognostic factor for OS in CRC patients has been previously shown in several studies with a total of 956 CRC patients (summarized by [223]). Greater proliferation potential among tumor cells due to longer telomeres may be a reason for this result. Furthermore, we demonstrated prolonged OS in individuals with shorter telomeres in tumor tissue compared to the adjacent mucosa (P=0.02). Our results are in the agreement with a study published by [224]. We suggest that for assessment of patients OS, it is important to consider telomere length in both, tumor tissue and adjacent mucosa. Based on our results, the telomere length ratio may serve as a prognostic marker for OS in CRC patients.

Our results also showed that the proximal site of tumor origin is associated with shorter telomeres compared to the distal part of the colon (P=0.006) and rectum (P<0.0001). This observation is in accordance with previous findings [188, 222, 225]. On the contrary, other groups did not find telomere attrition in right-sided CRC tumors [186, 221]. A possible explanation for shorter telomeres in proximal colon may be the major site-specific differences of mucosal microbiota content. It has been documented that microbiota organization is a distinct feature of proximal CRC and that it enhances cellular proliferation and local inflammation [226], which may ultimately result in pronounced telomere shortening. Moreover, we found a relationship between MSI status and shorter telomeres in our group of patients and it is well recognized that proximally located colon tumors have a higher occurrence of MSI phenotype [227]. To our knowledge, there is only a limited number of studies regarding this association [188, 225] and the results of other groups are inconclusive [46, 186]. Telomeres may be considered as microsatellite sequences and individuals with MSI phenotype may evince high frequency of unfixed deletions due to disrupted MMR pathway. We also suggest that the accumulation of deletions within telomeric sequences due to MMR deficiency may result in a decrease in concentration of shelterin subunits and disrupted telomere length [130].

With respect to inherited gene variants encoding shelterin subunits and telomerase, we observed a negative association between *TERT* rs3816659 homozygous variant and telomere length in non-cancerous mucosa. This particular SNP has been previously

described in relation to overall BC risk [228]. However, not all of the samples in our study were genotyped, therefore our result should be interpreted cautiously. In another study, *TERT* rs2736100 homozygous variant was independently associated with longer telomere length in non-cancerous mucosa of CRC patients [46]. Any of the investigated SNPs in shelterin subunits did not affect telomere length.

Overall, our results demonstrate the effect of various clinicopathological features on telomere length in distinct tissues of CRC patients. We suggest that measurement of telomere length in the target tissue and adjacent non-cancerous mucosa obtained during surgical resection from CRC patients may be utilized as a biomarker for specification of patients prognosis. Still, current knowledge regarding telomere length in such tissues of CRC patients is inconclusive and as telomerase inhibitors are being applied in clinical practice, there is an urgent need to better understand telomere biology in the target tissue and to identify patients that could profit from telomerase-targeted therapy.

Manuscript VII: In the review article entitled "***Genetic variation of acquired structural chromosomal aberrations***" Vodicka *et al*., 2018, we summarized a series of recently published studies investigating the contributors behind the formation of CAs. The whole article is based on the findings published in *Manuscript I*, *II*, *III* and *V*. We concluded that instead of the effect of individual SNPs, rather their interactions (epistasis) play a significant role. An interesting finding (with still unresolved underlying molecular mechanisms) was opposite associations between *CCND1* splice site rs9344 Gly870Ala SNP and CAs in healthy individuals on one side, and lower frequency of t(11,14) translocations in myeloma cancer patients. Additional questions have to be addressed in relation to CIN and malignancies, such as particular cancer phenotype and heterogeneity and the kinetics of CIN formation in the course of malignant transformation.

5. Conclusions

The following section of the present Thesis summarizes the main outcomes and answers all questions postulated in the Aims. All of our objectives were fulfilled.

- Do functionally relevant gene variants in major mitotic checkpoint and DNA repair genes modulate the frequency of structural CAs in PBL of healthy subjects?

We have observed a modest impact of single gene variants encoding for enzymes with an important function in spindle assembly checkpoint and DDR on the frequency of structural CAs in PBL of healthy individuals. Considering the individual types of CAs, only SNPs in *XPD* Lys751Gln and *RAD54L* Ala730= had an impact on CTA and CSA frequencies, respectively. However, binary gene-gene interactions in both, mitotic checkpoint and DDR, affected the frequency of CAs. Regarding mitotic checkpoint, interactions on the level of *PTTG1* and kinetochore (including gene variants in *ZWINT*, *ZWILCH* and *MAD2L1*) modulated the frequency of all types of CAs. Furthermore, we discovered several gene-gene pairs in DDR genes, always involved in two different DNA repair pathways, in association with CAs. Based on our results and in light of existing literature, we suggest that accumulation of CAs requires complex disruption of different pathways.

- Does the application of GWAS disclose novel gene variants in relation to structural CAs?

We identified several SNPs (rs1383997, rs2824215 and rs983889) with an effect on increased CAs frequency in PBL. All of the SNPs are considered to be involved either in a cell signaling, chromosome rearrangements or regulation of mitosis.

- Is the frequency of structural CAs associated with telomere length in PBL of healthy subjects and cancer patients?

We confirmed a negative correlation between the frequency of CAs and telomere length in healthy subjects. This result did not apply for cancer patients.

- Do the frequency of structural CAs and telomere length in PBL predict cancer risk and patients' long-term OS?

We confirmed that increased frequency of CAs in PBL of healthy subjects may be considered as a marker of BC and lung cancer risk. Furthermore, we discovered that women with longer telomeres were at increased risk of BC development. When using multivariate survival analysis, we found an association between increased frequency of CTAs and poor OS in BC and CRC patients. Otherwise, patients´ OS was not affected either by telomere length or other types of CAs in PBL.

- Do cancer patients evince decreased DSBs DRC in PBL? Is telomere length in PBL connected to the DSBs DRC?

We found decreased DSBs DRC in CRC patients. This result did not apply to BC patients. Irrespective of cancer type, telomere shortening in PBL was associated with decreased DSBs DRC.

- Is telomere length in tumor tissue, adjacent mucosa, metastases and blood affected by clinicopathological features among CRC patients?

We found shorter RTL in 74 % of tumor tissues compared to non-cancerous mucosa. Significantly shorter RTL was also identified in 83 % of liver metastases when compared to primary tumor tissues. Furthermore, 64 % of liver metastases evinced shorter RTL compared to non-cancerous liver tissues. On the other hand, shorter RTL in blood did not correlate with corresponding tumor tissue. Regarding clinicopathological features of the CRC patients, we found an association between shorter RTL in tumor tissue and MSI status, proximal tumor site origin and mucinous tumor histology. Tumor RTL also gradually increased with the ascending TNM tumor stage. We also found that *TERT* intron variant rs3816659 was significantly associated with shorter RTL in adjacent non-malignant mucosa of CRC patients.

- Is telomere length in tumor tissue associated with patients' survival?

Patients with more pronounced telomere shortening in tumor tissue compared to adjacent mucosa (smaller RTL tumor/RTL non-cancerous tissue ratio) demonstrated better OS.

6. Reference List

1. Yao, Y. and W. Dai, *Genomic Instability and Cancer.* J Carcinog Mutagen, 2014. **5**.
2. Lengauer, C., K.W. Kinzler, and B. Vogelstein, *Genetic instabilities in human cancers.* Nature, 1998. **396**(6712): p. 643-9.
3. Chatterjee, N. and G.C. Walker, *Mechanisms of DNA damage, repair, and mutagenesis.* Environ Mol Mutagen, 2017. **58**(5): p. 235-263.
4. Hemminki, K., et al., *Telomere length in circulating lymphocytes: Association with chromosomal aberrations.* Genes Chromosomes Cancer, 2015. **54**(3): p. 194-6.
5. Carter, S.L., et al., *A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers.* Nat Genet, 2006. **38**(9): p. 1043-8.
6. Ferguson, L.R., et al., *Genomic instability in human cancer: Molecular insights and opportunities for therapeutic attack and prevention through diet and nutrition.* Semin Cancer Biol, 2015. **35 Suppl**: p. S5-s24.
7. Andor, N., C.C. Maley, and H.P. Ji, *Genomic Instability in Cancer: Teetering on the Limit of Tolerance.* Cancer Res, 2017. **77**(9): p. 2179-2185.
8. Pikor, L., et al., *The detection and implication of genome instability in cancer.* Cancer Metastasis Rev, 2013. **32**(3-4): p. 341-52.
9. Gisselsson, D., *Chromosome instability in cancer: how, when, and why?* Adv Cancer Res, 2003. **87**: p. 1-29.
10. Grady, W.M. and J.M. Carethers, *Genomic and epigenetic instability in colorectal cancer pathogenesis.* Gastroenterology, 2008. **135**(4): p. 1079-99.
11. Giam, M. and G. Rancati, *Aneuploidy and chromosomal instability in cancer: a jackpot to chaos.* Cell Div, 2015. **10**: p. 3.
12. Barra, V. and D. Fachinetti, *The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA.* Nat Commun, 2018. **9**(1): p. 4340.
13. Hackett, J.A. and C.W. Greider, *Balancing instability: dual roles for telomerase and telomere dysfunction in tumorigenesis.* Oncogene, 2002. **21**(4): p. 619-26.
14. Driscoll, D.A. and S. Gross, *Clinical practice. Prenatal screening for aneuploidy.* N Engl J Med, 2009. **360**(24): p. 2556-62.
15. Gentric, G. and C. Desdouets, *Polyploidization in liver tissue.* Am J Pathol, 2014. **184**(2): p. 322-31.
16. Shu, Z., S. Row, and W.M. Deng, *Endoreplication: The Good, the Bad, and the Ugly.* Trends Cell Biol, 2018. **28**(6): p. 465-474.
17. Potapova, T.A., J. Zhu, and R. Li, *Aneuploidy and chromosomal instability: a vicious cycle driving cellular evolution and cancer genome chaos.* Cancer Metastasis Rev, 2013. **32**(3-4): p. 377-89.
18. Albertini, R.J., et al., *IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. International Programme on Chemical Safety.* Mutat Res, 2000. **463**(2): p. 111-72.
19. Sawyers, C.L., *Chronic myeloid leukemia.* N Engl J Med, 1999. **340**(17): p. 1330-40.
20. Hagmar, L., et al., *The usefulness of cytogenetic biomarkers as intermediate endpoints in carcinogenesis.* Int J Hyg Environ Health, 2001. **204**(1): p. 43-7.

21. Hemminki, K., et al., *Metabolic gene variants associated with chromosomal aberrations in healthy humans.* Genes Chromosomes Cancer, 2015. **54**(4): p. 260-6.
22. Palitti, F., et al., *Formation of chromatid-type aberrations in G2 stage of the cell cycle.* Mutat Res, 1983. **110**(2): p. 343-50.
23. Durante, M., et al., *From DNA damage to chromosome aberrations: joining the break.* Mutat Res, 2013. **756**(1-2): p. 5-13.
24. Pfeiffer, P., W. Goedecke, and G. Obe, *Mechanisms of DNA double-strand break repair and their potential to induce chromosomal aberrations.* Mutagenesis, 2000. **15**(4): p. 289-302.
25. Natarajan, A.T., *Mechanisms for induction of mutations and chromosome alterations.* Environ Health Perspect, 1993. **101 Suppl 3**: p. 225-9.
26. Mitelman, F., B. Johansson, and F. Mertens, *The impact of translocations and gene fusions on cancer causation.* Nat Rev Cancer, 2007. **7**(4): p. 233-45.
27. Rossner, P., et al., *Chromosomal aberrations in lymphocytes of healthy subjects and risk of cancer.* Environ Health Perspect, 2005. **113**(5): p. 517-20.
28. Boffetta, P., et al., *Chromosomal aberrations and cancer risk: results of a cohort study from Central Europe.* Am J Epidemiol, 2007. **165**(1): p. 36-43.
29. Vodenkova, S., et al., *Structural chromosomal aberrations as potential risk markers in incident cancer patients.* Mutagenesis, 2015. **30**(4): p. 557-63.
30. Brendle, A., et al., *Single nucleotide polymorphisms in chromosomal instability genes and risk and clinical outcome of breast cancer: a Swedish prospective case-control study.* Eur J Cancer, 2009. **45**(3): p. 435-42.
31. Vodicka, P., et al., *Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls.* Carcinogenesis, 2010. **31**(7): p. 1238-41.
32. Hemminki, K., et al., *Cyclin D1 splice site variant triggers chromosomal aberrations in healthy humans.* Leukemia, 2014. **28**(3): p. 721-2.
33. Kazimirova, A., et al., *Micronuclei and chromosomal aberrations, important markers of ageing: possible association with XPC and XPD polymorphisms.* Mutat Res, 2009. **661**(1-2): p. 35-40.
34. Vodicka, P., et al., *Genetic polymorphisms in DNA repair genes and possible links with DNA repair rates, chromosomal aberrations and single-strand breaks in DNA.* Carcinogenesis, 2004. **25**(5): p. 757-63.
35. Naccarati, A., et al., *Genetic polymorphisms and possible gene-gene interactions in metabolic and DNA repair genes: effects on DNA damage.* Mutat Res, 2006. **593**(1-2): p. 22-31.
36. Naccarati, A., et al., *Sporadic colorectal cancer and individual susceptibility: a review of the association studies investigating the role of DNA repair genetic polymorphisms.* Mutat Res, 2007. **635**(2-3): p. 118-45.
37. Pardini, B., et al., *DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic.* Mutat Res, 2008. **638**(1-2): p. 146-53.
38. Jiraskova, K., et al., *Functional Polymorphisms in DNA Repair Genes Are Associated with Sporadic Colorectal Cancer Susceptibility and Clinical Outcome.* Int J Mol Sci, 2018. **20**(1).
39. Smolarz, B., et al., *Polymorphism of DNA Repair Genes via Homologous Recombination (HR) in Ovarian Cancer.* Pathol Oncol Res, 2019.
40. Pardini, B., et al., *Variation within 3'-UTRs of base excision repair genes and response to therapy in colorectal cancer patients: A potential modulation of microRNAs binding.* Clin Cancer Res, 2013. **19**(21): p. 6044-56.

41. Nan, H., et al., *Genetic variants in telomere-maintaining genes and skin cancer risk.* Hum Genet, 2011. **129**(3): p. 247-53.
42. Varadi, V., et al., *Polymorphisms in telomere-associated genes, breast cancer susceptibility and prognosis.* Eur J Cancer, 2009. **45**(17): p. 3008-16.
43. Wong, K., et al., *Association of the POT1 Germline Missense Variant p.I78T With Familial Melanoma.* JAMA Dermatol, 2018.
44. Choi, J.E., et al., *Polymorphisms in telomere maintenance genes and risk of lung cancer.* Cancer Epidemiol Biomarkers Prev, 2009. **18**(10): p. 2773-81.
45. Codd, V., et al., *Common variants near TERC are associated with mean telomere length.* Nat Genet, 2010. **42**(3): p. 197-9.
46. Suraweera, N., et al., *Relative telomere lengths in tumor and normal mucosa are related to disease progression and chromosome instability profiles in colorectal cancer.* Oncotarget, 2016. **7**(24): p. 36474-36488.
47. Llorca-Cardenosa, M.J., et al., *Long telomere length and a TERT-CLPTM1 locus polymorphism association with melanoma risk.* Eur J Cancer, 2014. **50**(18): p. 3168-77.
48. Liu, Z., et al., *Association between TERT rs2853669 polymorphism and cancer risk: A meta-analysis of 9,157 cases and 11,073 controls.* PLoS One, 2018. **13**(3): p. e0191560.
49. Campa, D., et al., *Genetic determinants of telomere length and risk of pancreatic cancer: A PANDoRA study.* Int J Cancer, 2019. **144**(6): p. 1275-1283.
50. Zhou, M., et al., *Association Between TERT rs2736098 Polymorphisms and Cancer Risk-A Meta-Analysis.* Front Physiol, 2018. **9**: p. 377.
51. Mesic, A., et al., *Single nucleotide polymorphisms rs911160 in AURKA and rs2289590 in AURKB mitotic checkpoint genes contribute to gastric cancer susceptibility.* Environ Mol Mutagen, 2017. **58**(9): p. 701-711.
52. Mesic, A., et al., *Association of the AURKA and AURKC gene polymorphisms with an increased risk of gastric cancer.* IUBMB Life, 2016. **68**(8): p. 634-44.
53. Huyghe, J.R., et al., *Discovery of common and rare genetic risk variants for colorectal cancer.* Nat Genet, 2019. **51**(1): p. 76-87.
54. Wolpin, B.M., et al., *Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer.* Nat Genet, 2014. **46**(9): p. 994-1000.
55. Thompson, S.L. and D.A. Compton, *Examining the link between chromosomal instability and aneuploidy in human cells.* J Cell Biol, 2008. **180**(4): p. 665-72.
56. Kops, G.J., B.A. Weaver, and D.W. Cleveland, *On the road to cancer: aneuploidy and the mitotic checkpoint.* Nat Rev Cancer, 2005. **5**(10): p. 773-85.
57. Foley, E.A. and T.M. Kapoor, *Microtubule attachment and spindle assembly checkpoint signalling at the kinetochore.* Nat Rev Mol Cell Biol, 2013. **14**(1): p. 25-37.
58. Lara-Gonzalez, P., F.G. Westhorpe, and S.S. Taylor, *The spindle assembly checkpoint.* Curr Biol, 2012. **22**(22): p. R966-80.
59. Pines, J., *Cubism and the cell cycle: the many faces of the APC/C.* Nat Rev Mol Cell Biol, 2011. **12**(7): p. 427-38.
60. Karess, R., *Rod-Zw10-Zwilch: a key player in the spindle checkpoint.* Trends Cell Biol, 2005. **15**(7): p. 386-92.
61. Ciossani, G., et al., *The kinetochore proteins CENP-E and CENP-F directly and specifically interact with distinct BUB mitotic checkpoint Ser/Thr kinases.* J Biol Chem, 2018. **293**(26): p. 10084-10101.
62. Liu, Q., et al., *Nek2 targets the mitotic checkpoint proteins Mad2 and Cdc20: a mechanism for aneuploidy in cancer.* Exp Mol Pathol, 2010. **88**(2): p. 225-33.

63. Liu, S.T. and H. Zhang, *The mitotic checkpoint complex (MCC): looking back and forth after 15 years.* AIMS Mol Sci, 2016. **3**(4): p. 597-634.
64. Gruber, S., C.H. Haering, and K. Nasmyth, *Chromosomal cohesin forms a ring.* Cell, 2003. **112**(6): p. 765-77.
65. Nasmyth, K., *Cohesin: a catenase with separate entry and exit gates?* Nat Cell Biol, 2011. **13**(10): p. 1170-7.
66. Lehmann, A.R., D. McGibbon, and M. Stefanini, *Xeroderma pigmentosum.* Orphanet J Rare Dis, 2011. **6**: p. 70.
67. Ha, S.M., et al., *Association of BRCA Mutation Types, Imaging Features, and Pathologic Findings in Patients With Breast Cancer With BRCA1 and BRCA2 Mutations.* AJR Am J Roentgenol, 2017. **209**(4): p. 920-928.
68. Lynch, H.T., et al., *Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications.* Clin Genet, 2009. **76**(1): p. 1-18.
69. Cleaver, J.E., *Defective repair replication of DNA in xeroderma pigmentosum.* Nature, 1968. **218**(5142): p. 652-6.
70. Slyskova, J., et al., *DNA damage and nucleotide excision repair capacity in healthy individuals.* Environ Mol Mutagen, 2011. **52**(7): p. 511-7.
71. Slyskova, J., et al., *Differences in nucleotide excision repair capacity between newly diagnosed colorectal cancer patients and healthy controls.* Mutagenesis, 2012. **27**(2): p. 225-32.
72. Slyskova, J., et al., *Functional, genetic, and epigenetic aspects of base and nucleotide excision repair in colorectal carcinomas.* Clin Cancer Res, 2012. **18**(21): p. 5878-87.
73. Dipple, A., *DNA adducts of chemical carcinogens.* Carcinogenesis, 1995. **16**(3): p. 437-41.
74. Lindahl, T. and B. Nyberg, *Rate of depurination of native deoxyribonucleic acid.* Biochemistry, 1972. **11**(19): p. 3610-8.
75. Kuzminov, A., *Single-strand interruptions in replicating chromosomes cause double-strand breaks.* Proc Natl Acad Sci U S A, 2001. **98**(15): p. 8241-6.
76. Cooper, M.D. and M.N. Alder, *The evolution of adaptive immune systems.* Cell, 2006. **124**(4): p. 815-22.
77. Lieber, M.R., *The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway.* Annu Rev Biochem, 2010. **79**: p. 181-211.
78. Roos, W.P., A.D. Thomas, and B. Kaina, *DNA damage and the balance between survival and death in cancer biology.* Nat Rev Cancer, 2016. **16**(1): p. 20-33.
79. Maynard, S., et al., *Base excision repair of oxidative DNA damage and association with cancer and aging.* Carcinogenesis, 2009. **30**(1): p. 2-10.
80. de Villartay, J.P., *Congenital defects in V(D)J recombination.* Br Med Bull, 2015. **114**(1): p. 157-67.
81. van der Burg, M., et al., *A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits Artemis activation and nonhomologous end-joining.* J Clin Invest, 2009. **119**(1): p. 91-8.
82. Spivak, G., *Nucleotide excision repair in humans.* DNA Repair (Amst), 2015. **36**: p. 13-8.
83. Gillet, L.C. and O.D. Scharer, *Molecular mechanisms of mammalian global genome nucleotide excision repair.* Chem Rev, 2006. **106**(2): p. 253-76.
84. Hanawalt, P.C. and G. Spivak, *Transcription-coupled DNA repair: two decades of progress and surprises.* Nat Rev Mol Cell Biol, 2008. **9**(12): p. 958-70.

85. Sugasawa, K., et al., *Xeroderma pigmentosum group C protein complex is the initiator of global genome nucleotide excision repair.* Mol Cell, 1998. **2**(2): p. 223-32.
86. Coin, F., V. Oksenych, and J.M. Egly, *Distinct roles for the XPB/p52 and XPD/p44 subcomplexes of TFIIH in damaged DNA opening during nucleotide excision repair.* Mol Cell, 2007. **26**(2): p. 245-56.
87. Scharer, O.D., *XPG: its products and biological roles.* Adv Exp Med Biol, 2008. **637**: p. 83-92.
88. Wakasugi, M. and A. Sancar, *Assembly, subunit composition, and footprint of human DNA repair excision nuclease.* Proc Natl Acad Sci U S A, 1998. **95**(12): p. 6669-74.
89. Fagbemi, A.F., B. Orelli, and O.D. Scharer, *Regulation of endonuclease activity in human nucleotide excision repair.* DNA Repair (Amst), 2011. **10**(7): p. 722-9.
90. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.
91. Shinmura, K. and J. Yokota, *The OGG1 gene encodes a repair enzyme for oxidatively damaged DNA and is involved in human carcinogenesis.* Antioxid Redox Signal, 2001. **3**(4): p. 597-609.
92. Whitaker, A.M., et al., *Base excision repair of oxidative DNA damage: from mechanism to disease.* Front Biosci (Landmark Ed), 2017. **22**: p. 1493-1522.
93. Kim, H.S., et al., *Mutation at intronic repeats of the ataxia-telangiectasia mutated (ATM) gene and ATM protein loss in primary gastric cancer with microsatellite instability.* PLoS One, 2013. **8**(12): p. e82769.
94. Kunkel, T.A. and D.A. Erie, *Eukaryotic Mismatch Repair in Relation to DNA Replication.* Annu Rev Genet, 2015. **49**: p. 291-313.
95. Sung, P., *Introduction to the Thematic Minireview Series: DNA double-strand break repair and pathway choice.* J Biol Chem, 2018. **293**(27): p. 10500-10501.
96. Sallmyr, A. and A.E. Tomkinson, *Repair of DNA double-strand breaks by mammalian alternative end-joining pathways.* J Biol Chem, 2018. **293**(27): p. 10536-10546.
97. Betermier, M., P. Bertrand, and B.S. Lopez, *Is non-homologous end-joining really an inherently error-prone process?* PLoS Genet, 2014. **10**(1): p. e1004086.
98. Grabarz, A., et al., *Initiation of DNA double strand break repair: signaling and single-stranded resection dictate the choice between homologous recombination, non-homologous end-joining and alternative end-joining.* Am J Cancer Res, 2012. **2**(3): p. 249-68.
99. Chiruvella, K.K., Z. Liang, and T.E. Wilson, *Repair of double-strand breaks by end joining.* Cold Spring Harb Perspect Biol, 2013. **5**(5): p. a012757.
100. Chang, H.H., G. Watanabe, and M.R. Lieber, *Unifying the DNA end-processing roles of the artemis nuclease: Ku-dependent artemis resection at blunt DNA ends.* J Biol Chem, 2015. **290**(40): p. 24036-50.
101. Chang, H.H.Y., et al., *Non-homologous DNA end joining and alternative pathways to double-strand break repair.* Nat Rev Mol Cell Biol, 2017. **18**(8): p. 495-506.
102. Bunting, S.F., et al., *53BP1 inhibits homologous recombination in Brca1-deficient cells by blocking resection of DNA breaks.* Cell, 2010. **141**(2): p. 243-54.
103. Bebenek, K., L.C. Pedersen, and T.A. Kunkel, *Structure-function studies of DNA polymerase lambda.* Biochemistry, 2014. **53**(17): p. 2781-92.
104. Moon, A.F., et al., *Sustained active site rigidity during synthesis by human DNA polymerase mu.* Nat Struct Mol Biol, 2014. **21**(3): p. 253-60.

105. Symington, L.S., *Mechanism and regulation of DNA end resection in eukaryotes.* Crit Rev Biochem Mol Biol, 2016. **51**(3): p. 195-212.
106. Rogakou, E.P., et al., *DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139.* J Biol Chem, 1998. **273**(10): p. 5858-68.
107. Wold, M.S., *Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism.* Annu Rev Biochem, 1997. **66**: p. 61-92.
108. Symington, L.S., R. Rothstein, and M. Lisby, *Mechanisms and regulation of mitotic recombination in Saccharomyces cerevisiae.* Genetics, 2014. **198**(3): p. 795-835.
109. Chun, J., E.S. Buechelmaier, and S.N. Powell, *Rad51 paralog complexes BCDX2 and CX3 act at different stages in the BRCA1-BRCA2-dependent homologous recombination pathway.* Mol Cell Biol, 2013. **33**(2): p. 387-95.
110. Symington, L.S., *Role of RAD52 epistasis group genes in homologous recombination and double-strand break repair.* Microbiol Mol Biol Rev, 2002. **66**(4): p. 630-70, table of contents.
111. Lazzerini-Denchi, E. and A. Sfeir, *Stop pulling my strings - what telomeres taught us about the DNA damage response.* Nat Rev Mol Cell Biol, 2016. **17**(6): p. 364-78.
112. Palm, W. and T. de Lange, *How shelterin protects mammalian telomeres.* Annu Rev Genet, 2008. **42**: p. 301-34.
113. Hockemeyer, D. and K. Collins, *Control of telomerase action at human telomeres.* Nat Struct Mol Biol, 2015. **22**(11): p. 848-52.
114. Griffith, J.D., et al., *Mammalian telomeres end in a large duplex loop.* Cell, 1999. **97**(4): p. 503-14.
115. de Lange, T., *How telomeres solve the end-protection problem.* Science, 2009. **326**(5955): p. 948-52.
116. Calado, R.T. and N.S. Young, *Telomere maintenance and human bone marrow failure.* Blood, 2008. **111**(9): p. 4446-55.
117. Okuda, K., et al., *Telomere length in the newborn.* Pediatr Res, 2002. **52**(3): p. 377-81.
118. Londono-Vallejo, J.A., et al., *Differences in telomere length between homologous chromosomes in humans.* Nucleic Acids Res, 2001. **29**(15): p. 3164-71.
119. Harley, C.B., A.B. Futcher, and C.W. Greider, *Telomeres shorten during ageing of human fibroblasts.* Nature, 1990. **345**(6274): p. 458-60.
120. Soudet, J., P. Jolivet, and M.T. Teixeira, *Elucidation of the DNA end-replication problem in Saccharomyces cerevisiae.* Mol Cell, 2014. **53**(6): p. 954-64.
121. Capper, R., et al., *The nature of telomere fusion and a definition of the critical telomere length in human cells.* Genes Dev, 2007. **21**(19): p. 2495-508.
122. Coppe, J.P., et al., *The senescence-associated secretory phenotype: the dark side of tumor suppression.* Annu Rev Pathol, 2010. **5**: p. 99-118.
123. Roger, L., et al., *Extensive Telomere Erosion in the Initiation of Colorectal Adenomas and Its Association With Chromosomal Instability.* Jnci-Journal of the National Cancer Institute, 2013. **105**(16): p. 1202-1211.
124. Maciejowski, J. and T. de Lange, *Telomeres in cancer: tumour suppression and genome instability.* Nat Rev Mol Cell Biol, 2017. **18**(3): p. 175-186.
125. Arnoult, N. and J. Karlseder, *Complex interactions between the DNA-damage response and mammalian telomeres.* Nat Struct Mol Biol, 2015. **22**(11): p. 859-66.
126. Wang, F., et al., *The POT1-TPP1 telomere complex is a telomerase processivity factor.* Nature, 2007. **445**(7127): p. 506-10.

127. Denchi, E.L. and T. de Lange, *Protection of telomeres through independent control of ATM and ATR by TRF2 and POT1.* Nature, 2007. **448**(7157): p. 1068-71.
128. de Lange, T., *Shelterin: the protein complex that shapes and safeguards human telomeres.* Genes Dev, 2005. **19**(18): p. 2100-10.
129. Rai, R., et al., *TRF2-RAP1 is required to protect telomeres from engaging in homologous recombination-mediated deletions and fusions.* Nat Commun, 2016. **7**: p. 10881.
130. Loayza, D. and T. De Lange, *POT1 as a terminal transducer of TRF1 telomere length control.* Nature, 2003. **423**(6943): p. 1013-8.
131. van Steensel, B. and T. de Lange, *Control of telomere length by the human telomeric protein TRF1.* Nature, 1997. **385**(6618): p. 740-3.
132. Zimmermann, M., et al., *TRF1 negotiates TTAGGG repeat-associated replication problems by recruiting the BLM helicase and the TPP1/POT1 repressor of ATR signaling.* Genes Dev, 2014. **28**(22): p. 2477-91.
133. Martinez, P., et al., *Increased telomere fragility and fusions resulting from TRF1 deficiency lead to degenerative pathologies and increased cancer in mice.* Genes Dev, 2009. **23**(17): p. 2060-75.
134. Ye, J.Z. and T. de Lange, *TIN2 is a tankyrase 1 PARP modulator in the TRF1 telomere length control complex.* Nat Genet, 2004. **36**(6): p. 618-23.
135. Xu, L., S. Li, and B.A. Stohr, *The role of telomere biology in cancer.* Annu Rev Pathol, 2013. **8**: p. 49-78.
136. Frescas, D. and T. de Lange, *Binding of TPP1 protein to TIN2 protein is required for POT1a,b protein-mediated telomere protection.* J Biol Chem, 2014. **289**(35): p. 24180-7.
137. Xin, H., et al., *TPP1 is a homologue of ciliate TEBP-beta and interacts with POT1 to recruit telomerase.* Nature, 2007. **445**(7127): p. 559-62.
138. Wu, L., et al., *Pot1 deficiency initiates DNA damage checkpoint activation and aberrant homologous recombination at telomeres.* Cell, 2006. **126**(1): p. 49-62.
139. Armanios, M. and E.H. Blackburn, *The telomere syndromes.* Nat Rev Genet, 2012. **13**(10): p. 693-704.
140. Wu, Y., S. Xiao, and X.D. Zhu, *MRE11-RAD50-NBS1 and ATM function as co-mediators of TRF1 in telomere length control.* Nat Struct Mol Biol, 2007. **14**(9): p. 832-40.
141. Wang, R.C., A. Smogorzewska, and T. de Lange, *Homologous recombination generates T-loop-sized deletions at human telomeres.* Cell, 2004. **119**(3): p. 355-68.
142. Wang, Y., G. Ghosh, and E.A. Hendrickson, *Ku86 represses lethal telomere deletion events in human somatic cells.* Proc Natl Acad Sci U S A, 2009. **106**(30): p. 12430-5.
143. Li, H., T.H. Lee, and H. Avraham, *A novel tricomplex of BRCA1, Nmi, and c-Myc inhibits c-Myc-induced human telomerase reverse transcriptase gene (hTERT) promoter activity in breast cancer.* J Biol Chem, 2002. **277**(23): p. 20965-73.
144. Metcalfe, J.A., et al., *Accelerated telomere shortening in ataxia telangiectasia.* Nat Genet, 1996. **13**(3): p. 350-3.
145. Crabbe, L., et al., *Telomere dysfunction as a cause of genomic instability in Werner syndrome.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2205-10.
146. Sturzenegger, A., et al., *DNA2 cooperates with the WRN and BLM RecQ helicases to mediate long-range DNA end resection in human cells.* J Biol Chem, 2014. **289**(39): p. 27314-26.

147. Das, A., et al., *The human Werner syndrome protein stimulates repair of oxidative DNA base damage by the DNA glycosylase NEIL1.* J Biol Chem, 2007. **282**(36): p. 26591-602.
148. Muftuoglu, M., et al., *Telomere repeat binding factor 2 interacts with base excision repair proteins and stimulates DNA synthesis by DNA polymerase beta.* Cancer Res, 2006. **66**(1): p. 113-24.
149. Vallabhaneni, H., et al., *Defective repair of oxidative base lesions by the DNA glycosylase Nth1 associates with multiple telomere defects.* PLoS Genet, 2013. **9**(7): p. e1003639.
150. Fouquerel, E., et al., *Oxidative guanine base damage regulates human telomerase activity.* Nat Struct Mol Biol, 2016. **23**(12): p. 1092-1100.
151. Ahmed, W. and J. Lingner, *Impact of oxidative stress on telomere biology.* Differentiation, 2018. **99**: p. 21-27.
152. Opresko, P.L., et al., *Oxidative damage in telomeric DNA disrupts recognition by TRF1 and TRF2.* Nucleic Acids Res, 2005. **33**(4): p. 1230-9.
153. von Zglinicki, T., *Oxidative stress shortens telomeres.* Trends Biochem Sci, 2002. **27**(7): p. 339-44.
154. Mendez-Bermudez, A. and N.J. Royle, *Deficiency in DNA mismatch repair increases the rate of telomere shortening in normal human cells.* Hum Mutat, 2011. **32**(8): p. 939-46.
155. Carethers, J.M., M. Koi, and S.S. Tseng-Rogenski, *EMAST is a Form of Microsatellite Instability That is Initiated by Inflammation and Modulates Colorectal Cancer Progression.* Genes, 2015. **6**(2): p. 185-205.
156. Aaltonen, L.A., et al., *Clues to the pathogenesis of familial colorectal cancer.* Science, 1993. **260**(5109): p. 812-6.
157. Carethers, J.M., *Microsatellite Instability Pathway and EMAST in Colorectal Cancer.* Curr Colorectal Cancer Rep, 2017. **13**(1): p. 73-80.
158. Kang, S., et al., *The significance of microsatellite instability in colorectal cancer after controlling for clinicopathological factors.* Medicine (Baltimore), 2018. **97**(9): p. e0019.
159. Popat, S., R. Hubner, and R.S. Houlston, *Systematic review of microsatellite instability and colorectal cancer prognosis.* J Clin Oncol, 2005. **23**(3): p. 609-18.
160. Bertorelle, R., et al., *Telomeres, telomerase and colorectal cancer.* World J Gastroenterol, 2014. **20**(8): p. 1940-50.
161. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.
162. Lee, M., et al., *Telomere extension by telomerase and ALT generates variant repeats by mechanistically distinct processes.* Nucleic Acids Research, 2014. **42**(3): p. 1733-1746.
163. Holt, S.E. and J.W. Shay, *Role of telomerase in cellular proliferation and cancer.* J Cell Physiol, 1999. **180**(1): p. 10-8.
164. Henson, J.D., et al., *Alternative lengthening of telomeres in mammalian cells.* Oncogene, 2002. **21**(4): p. 598-610.
165. Boardman, L.A., et al., *Correlation of Chromosomal Instability, Telomere Length and Telomere Maintenance in Microsatellite Stable Rectal Cancer: A Molecular Subclass of Rectal Cancer.* Plos One, 2013. **8**(11).
166. Cesare, A.J. and R.R. Reddel, *Alternative lengthening of telomeres: models, mechanisms and implications.* Nature Reviews Genetics, 2010. **11**(5): p. 319-330.

167. Heaphy, C.M., et al., *Prevalence of the alternative lengthening of telomeres telomere maintenance mechanism in human cancer subtypes.* Am J Pathol, 2011. **179**(4): p. 1608-15.
168. Durant, S.T., *Telomerase-independent paths to immortality in predictable cancer subtypes.* J Cancer, 2012. **3**: p. 67-82.
169. Pickett, H.A. and R.R. Reddel, *Molecular mechanisms of activity and derepression of alternative lengthening of telomeres.* Nat Struct Mol Biol, 2015. **22**(11): p. 875-80.
170. Zhang, J.M., et al., *Alternative Lengthening of Telomeres through Two Distinct Break-Induced Replication Pathways.* Cell Rep, 2019. **26**(4): p. 955-968.e3.
171. Greider, C.W. and E.H. Blackburn, *Identification of a specific telomere terminal transferase activity in Tetrahymena extracts.* Cell, 1985. **43**(2 Pt 1): p. 405-13.
172. Nakamura, T.M., et al., *Telomerase catalytic subunit homologs from fission yeast and human.* Science, 1997. **277**(5328): p. 955-9.
173. Fang, G. and T.R. Cech, *Telomerase RNA localized in the replication band and spherical subnuclear organelles in hypotrichous ciliates.* J Cell Biol, 1995. **130**(2): p. 243-53.
174. Penzo, M., et al., *Dyskerin and TERC expression may condition survival in lung cancer patients.* Oncotarget, 2015. **6**(25): p. 21755-60.
175. Wang, S. and J. Zhu, *The hTERT gene is embedded in a nuclease-resistant chromatin domain.* J Biol Chem, 2004. **279**(53): p. 55401-10.
176. Levy, M.Z., et al., *Telomere end-replication problem and cell aging.* J Mol Biol, 1992. **225**(4): p. 951-60.
177. Skvortsov, D.A., et al., *Assays for detection of telomerase activity.* Acta Naturae, 2011. **3**(1): p. 48-68.
178. Horn, S., et al., *TERT promoter mutations in familial and sporadic melanoma.* Science, 2013. **339**(6122): p. 959-61.
179. Heidenreich, B. and R. Kumar, *TERT promoter mutations in telomere biology.* Mutat Res, 2017. **771**: p. 15-31.
180. Kyo, S., et al., *Understanding and exploiting hTERT promoter regulation for diagnosis and treatment of human cancers.* Cancer Sci, 2008. **99**(8): p. 1528-38.
181. Hu, C., et al., *hTERT promotes the invasion of gastric cancer cells by enhancing FOXO3a ubiquitination and subsequent ITGB1 upregulation.* Gut, 2017. **66**(1): p. 31-42.
182. Kim, J.H., et al., *A novel telomere elongation in an adriamycin-resistant stomach cancer cell line with decreased telomerase activity.* Mol Cells, 2002. **13**(2): p. 228-36.
183. Lai, T.P., et al., *A method for measuring the distribution of the shortest telomeres in cells and tissues.* Nat Commun, 2017. **8**(1): p. 1356.
184. Aubert, G., M. Hills, and P.M. Lansdorp, *Telomere length measurement-caveats and a critical assessment of the available technologies and tools.* Mutat Res, 2012. **730**(1-2): p. 59-67.
185. Cawthon, R.M., *Telomere length measurement by a novel monochrome multiplex quantitative PCR method.* Nucleic Acids Res, 2009. **37**(3): p. e21.
186. Balc'h, E.L., et al., *Measurement of Telomere Length in Colorectal Cancers for Improved Molecular Diagnosis.* Int J Mol Sci, 2017. **18**(9).
187. Roger, L., et al., *Extensive telomere erosion in the initiation of colorectal adenomas and its association with chromosomal instability.* J Natl Cancer Inst, 2013. **105**(16): p. 1202-11.

188. Rampazzo, E., et al., *Relationship between telomere shortening, genetic instability, and site of tumour origin in colorectal cancers.* Br J Cancer, 2010. **102**(8): p. 1300-5.
189. Plentz, R.R., et al., *Telomere shortening of epithelial cells characterises the adenoma-carcinoma transition of human colorectal cancer.* Gut, 2003. **52**(9): p. 1304-7.
190. Liu, W., et al., *Kras mutations increase telomerase activity and targeting telomerase is a promising therapeutic strategy for Kras-mutant NSCLC.* Oncotarget, 2017. **8**(1): p. 179-190.
191. Hsu, T.C., et al., *Sensitivity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis.* Int J Cancer, 1989. **43**(3): p. 403-9.
192. Rachakonda, S., et al., *Telomere length and survival in primary cutaneous melanoma patients.* Sci Rep, 2018. **8**(1): p. 10947.
193. Berg, K.D., et al., *Detection of microsatellite instability by fluorescence multiplex polymerase chain reaction.* J Mol Diagn, 2000. **2**(1): p. 20-8.
194. Ren, Q. and B. Jin, *The clinical value and biological function of PTTG1 in colorectal cancer.* Biomed Pharmacother, 2017. **89**: p. 108-115.
195. Mora-Santos, M., et al., *A single mutation in Securin induces chromosomal instability and enhances cell invasion.* Eur J Cancer, 2013. **49**(2): p. 500-10.
196. Lo, Y.L., et al., *Breast cancer risk associated with genotypic polymorphism of the mitotic checkpoint genes: a multigenic study on cancer susceptibility.* Carcinogenesis, 2007. **28**(5): p. 1079-86.
197. Simonelli, V., et al., *Gene susceptibility to oxidative damage: from single nucleotide polymorphisms to function.* Mutat Res, 2012. **731**(1-2): p. 1-13.
198. Russo, A., et al., *Rad54/Rad54B deficiency is associated to increased chromosome breakage in mouse spermatocytes.* Mutagenesis, 2018. **33**(4): p. 323-332.
199. Vodicka, P., et al., *Cytogenetic markers, DNA single-strand breaks, urinary metabolites, and DNA repair rates in styrene-exposed lamination workers.* Environ Health Perspect, 2004. **112**(8): p. 867-71.
200. Musak, L., et al., *Chromosomal damage among medical staff occupationally exposed to volatile anesthetics, antineoplastic drugs, and formaldehyde.* Scand J Work Environ Health, 2013. **39**(6): p. 618-30.
201. Wight, M. and A. Werner, *The functions of natural antisense transcripts.* Essays Biochem, 2013. **54**: p. 91-101.
202. Park, Y.R., et al., *Data-driven Analysis of TRP Channels in Cancer: Linking Variation in Gene Expression to Clinical Significance.* Cancer Genomics Proteomics, 2016. **13**(1): p. 83-90.
203. Shapovalov, G., et al., *Role of TRP ion channels in cancer and tumorigenesis.* Semin Immunopathol, 2016. **38**(3): p. 357-69.
204. Haldeman-Englert, C.R., et al., *A de novo 8.8-Mb deletion of 21q21.1-q21.3 in an autistic male with a complex rearrangement involving chromosomes 6, 10, and 21.* Am J Med Genet A, 2010. **152a**(1): p. 196-202.
205. Kamran, M., et al., *Aurora kinase A regulates Survivin stability through targeting FBXL7 in gastric cancer drug resistance and prognosis.* Oncogenesis, 2017. **6**(2): p. e298.
206. Bonassi, S., et al., *Chromosomal aberration frequency in lymphocytes predicts the risk of cancer: results from a pooled cohort study of 22 358 subjects in 11 countries.* Carcinogenesis, 2008. **29**(6): p. 1178-83.
207. Xu, X., et al., *Association between telomere length and survival in cancer patients: a meta-analysis and review of literature.* Front Med, 2016. **10**(2): p. 191-203.

208. Mayer, S., et al., *Sex-specific telomere length profiles and age-dependent erosion dynamics of individual chromosome arms in humans.* Cytogenet Genome Res, 2006. **112**(3-4): p. 194-201.
209. Vodenkova, S., et al., *Base excision repair capacity as a determinant of prognosis and therapy response in colon cancer patients.* DNA Repair (Amst), 2018. **72**: p. 77-85.
210. Slyskova, J., et al., *Post-treatment recovery of suboptimal DNA repair capacity and gene expression levels in colorectal cancer patients.* Mol Carcinog, 2015. **54**(9): p. 769-78.
211. Shao, L., M. Lai, and Q. Huang, *Mutagen sensitivity and p53 expression in colorectal cancer in China.* Postgrad Med J, 2001. **77**(913): p. 713-6.
212. Ankathil, R., et al., *Deficient DNA repair capacity: a predisposing factor and high risk predictive marker in familial colorectal cancer.* J Exp Clin Cancer Res, 1999. **18**(1): p. 33-7.
213. Hu, M., et al., *Bleomycin-induced mutagen sensitivity, passive smoking, and risk of breast cancer in Chinese women: a case-control study.* Cancer Causes Control, 2013. **24**(4): p. 629-36.
214. Kosti, O., et al., *Mutagen sensitivity, tobacco smoking and breast cancer risk: a case-control study.* Carcinogenesis, 2010. **31**(4): p. 654-9.
215. Doksani, Y. and T. de Lange, *The role of double-strand break repair pathways at functional and dysfunctional telomeres.* Cold Spring Harb Perspect Biol, 2014. **6**(12): p. a016576.
216. Celli, G.B., E.L. Denchi, and T. de Lange, *Ku70 stimulates fusion of dysfunctional telomeres yet protects chromosome ends from homologous recombination.* Nat Cell Biol, 2006. **8**(8): p. 885-90.
217. Chapman, J.R., et al., *RIF1 is essential for 53BP1-dependent nonhomologous end joining and suppression of DNA double-strand break resection.* Mol Cell, 2013. **49**(5): p. 858-71.
218. Jia, P., C. Her, and W. Chai, *DNA excision repair at telomeres.* DNA Repair (Amst), 2015. **36**: p. 137-45.
219. Madlener, S., et al., *Essential role for mammalian apurinic/apyrimidinic (AP) endonuclease Ape1/Ref-1 in telomere maintenance.* Proc Natl Acad Sci U S A, 2013. **110**(44): p. 17844-9.
220. Reilly, N.M., et al., *Exploiting DNA repair defects in colorectal cancer.* Mol Oncol, 2019.
221. Gertler, R., et al., *Telomere length and human telomerase reverse transcriptase expression as markers for progression and prognosis of colorectal carcinoma.* J Clin Oncol, 2004. **22**(10): p. 1807-14.
222. Garcia-Aranda, C., et al., *Correlations of telomere length, telomerase activity, and telomeric-repeat binding factor 1 expression in colorectal carcinoma.* Cancer, 2006. **106**(3): p. 541-51.
223. Jia, H. and Z. Wang, *Telomere Length as a Prognostic Factor for Overall Survival in Colorectal Cancer Patients.* Cell Physiol Biochem, 2016. **38**(1): p. 122-8.
224. Valls, C., et al., *Telomere length is a prognostic factor for overall survival in colorectal cancer.* Colorectal Dis, 2011. **13**(11): p. 1265-72.
225. Takagi, S., et al., *Relationship between microsatellite instability and telomere shortening in colorectal cancer.* Dis Colon Rectum, 2000. **43**(10 Suppl): p. S12-7.
226. Dejea, C.M., et al., *Microbiota organization is a distinct feature of proximal colorectal cancers.* Proc Natl Acad Sci U S A, 2014. **111**(51): p. 18321-6.

227. Missiaglia, E., et al., *Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features.* Ann Oncol, 2014. **25**(10): p. 1995-2001.
228. Shen, J., et al., *Multiple genetic variants in telomere pathway genes and breast cancer risk.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(1): p. 219-28.

Manuscript I

Genetic variation in the major mitotic checkpoint genes associated with chromosomal aberrations in healthy humans.

Försti A, Frank C, Smolkova B, Kazimirova A, Barancokova M, Vymetalkova V, **Kroupa M**, Naccarati A, Vodickova L, Buchancova J, Dusinska M, Musak L, Vodicka P, Hemminki K.

Cancer Lett. 2016 Oct 1;380(2):442-6. **IF: 6.491**

Cancer Letters 380 (2016) 442–446

Contents lists available at ScienceDirect

Cancer Letters

journal homepage: www.elsevier.com/locate/canlet

Original Articles

Genetic variation in the major mitotic checkpoint genes associated with chromosomal aberrations in healthy humans

Asta Försti [a,b,*], Christoph Frank [a], Bozena Smolkova [c], Alena Kazimirova [d], Magdalena Barancokova [d], Veronika Vymetalkova [e,f], Michal Kroupa [f,g], Alessio Naccarati [e,h], Ludmila Vodickova [e,f,g], Janka Buchancova [i], Maria Dusinska [j,1], Ludovit Musak [k,1], Pavel Vodicka [e,f,1], Kari Hemminki [a,b,1]

[a] Department of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany
[b] Center for Primary Health Care Research, Lund University, Malmö, Sweden
[c] Department of Genetics, Cancer Research Institute of Slovak Academy of Sciences, Vlarska 7, 83391 Bratislava, Slovakia
[d] Faculty of Medicine, Slovak Medical University, Limbova 12, 83303 Bratislava, Slovakia
[e] Institute of Experimental Medicine, Academy of Sciences of the Czech Republic, Videnska 1083, 14200 Prague, Czech Republic
[f] Institute of Biology and Medical Genetics, 1st Medical Faculty, Charles University, Albertov 4, 12800 Prague, Czech Republic
[g] Faculty of Medicine and Biomedical Center in Pilsen, Charles University in Prague, 30605 Pilsen, Czech Republic
[h] Human Genetics Foundation, Turin, Italy
[i] Department of Public Health, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, Sklabinska 26, 03601 Martin, Slovakia
[j] Health Effects Laboratory, Department of Environmental Chemistry, NILU-Norwegian Institute for Air Research, Instituttveien 18, 2007 Kjeller, Norway
[k] Clinic of Occupational Medicine and Toxicology, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava and University Hospital Martin, Kollarova 2, 03601 Martin, Slovakia

ARTICLE INFO

Article history:
Received 19 May 2016
Received in revised form 12 July 2016
Accepted 12 July 2016

Keywords:
Chromosomal integrity
Cytogenetics
Spindle checkpoint
DNA double-stranded break

ABSTRACT

Non-specific chromosomal aberrations (CAs) are microscopically detected in about 1% of lymphocytes drawn from healthy persons. Causes of CAs in general population are not known but they may be related to risk of cancer. In view of the importance of the mitotic checkpoint machinery on maintaining chromosomal integrity we selected 9 variants in main checkpoint related genes (BUB1B, BUB3, MAD2L1, CENPF, ESPL1/separase, NEK2, PTTG1/securin, ZWILCH and ZWINT) for a genotyping study on samples from healthy individuals (N = 330 to 729) whose lymphocytes had an increased number of CAs compared to persons with a low number of CAs. Genetic variation in individual genes played a minor importance, consistent with the high conservation and selection pressure of the checkpoint system. However, gene pairs were significantly associated with CAs: PTTG1-ZWILCH and PTTG1-ZWINT. MAD2L1 and PTTG1 were the most common partners in any of the two-way interactions. The results suggest that interactions at the level of cohesin (PTTG1) and kinetochore function (ZWINT, ZWILCH and MAD2L1) contribute to the frequency of CAs, suggesting that gene variants at different checkpoint functions appeared to be required for the formation of CAs.

Introduction

Chromosomal aberrations (CAs) are markers of cancer risk and many specific clonal CAs are critical events in malignant transformation [1–4]. Non-specific CAs include missing, fragmented or fused chromosomal segments which are not clonal and may remain in lymphocytes for their life-time [5]. They are analyzed by microscopic scoring of metaphase nuclei from cultured lymphocytes and scored as chromosome-type aberrations (CSAs) and chromatid-type aberrations (CTAs). CSAs are thought to arise as a result of direct DNA damage or replication of a carcinogen-damaged DNA template; replication error may also lead to CTAs [5]. An alternative mechanism for CA formation is telomere erosion and the resulting erroneous joining of non-homologous chromosomes [6–9]. A further mechanism for CAs may be aneuploidy and chromosomal instability as a result of aberrant mitosis. Accurate chromosome segregation between two daughter cells during mitosis is supervised by a highly conserved signaling machinery termed the mitotic checkpoint which delays anaphase until all chromosomes are properly attached to the mitotic spindle [10]. Errors may interfere in many steps of the complex checkpoint control, including incorrect attachment of the kinetochore to microtubules and cohesion defects [10,11].

* Corresponding author. Fax: +49 6221 421810.
E-mail address: a.foersti@dkfz.de (A. Försti).
[1] Shared senior authorship.

http://dx.doi.org/10.1016/j.canlet.2016.07.011
0304-3835/

In view of the a priori importance of mitotic checkpoint machinery on maintaining chromosomal integrity and the lacking data on potential role of gene variants influencing non-specific CAs we selected 9 main checkpoint related genes and their putative functional SNPs for a genotyping study on samples from individuals whose lymphocytes had an increased number of CAs ('cases') and compared them to persons with a low number of CAs ('controls'). In addition to individual variants, we tested for gene–gene interactions because mitotic checkpoint is a clockwork of many genes and interactions would be expected. The genes under study were BUB1B (mitotic checkpoint serine/threonine kinase B), BUB3 (mitotic checkpoint protein), MAD2L1 (mitotic arrest deficient-like 1), CENPF (centromere protein F), ESPL1 (extra spindle poles-like 1, separase), PTTG1 (pituitary tumor-transforming gene 1, securin), NEK2 (never in mitosis gene A-related kinase 2), ZWILCH (zwilch kinetochore protein) and ZWINT (ZW10 interactor). These genes encode proteins with important functions in mitosis [10,12]. BUB1B, BUB3 and MAD2L1 delay the onset of anaphase until all chromosomes are properly aligned at the metaphase plate; they are parts of the mitotic checkpoint complex [11]. CENPF encodes a protein that associates with the centromere-kinetochore complex. ESPL1/separase cleaves cohesin at the onset of anaphase while PTTG1/securin blocks the separase function. NEK2 is essential in centrosome separation by maintaining a stable attachment of microtubules to the kinetochore. The ZWINT and ZWILCH genes encode proteins that are involved in kinetochore function and released in preparation for mitosis [11,13].

Methods

The subjects were identified from a large cohort of volunteers who have been assayed for CAs in various previous studies; for details see References [14,15]. The participants provided an informed consent. Cytogenetic analysis was performed on cultured lymphocytes, as previously described, by microscopically analyzing (two microscopists in each lab) in a double-blind fashion coded slides of 100 mitoses per person for the frequency of total CAs (CAtot), CTAs and CSAs [3,16,17]. The CAs were categorized into a high-frequency group, 'cases' (>2%) and a low-frequency group, 'controls' (≤2%); this arbitrary cut-off point was based on our previous experience [3,17]. For CTAs and CSAs, the cut-off was 1%.

Genotyping method of the present polymorphisms was based on allelic discrimination using the TaqMan technology as described by us [18,19]. The analyses included the following SNPs and numbers of tested individuals: for BUB1B rs1801376 (N = 330), for BUB3 rs3808960 (N = 330), for MAD2L1 rs903147 (N = 330), for CENPF rs438034 (N = 618), for ESPL1 rs6580941 (N = 656), for NEK2 rs701928 (N = 663), for PTTG1 rs1862392 (N = 729), for ZWILCH rs3087660 (N = 674) and for ZWINT rs2241666 (N = 662). The reason for the varying number of samples was the availability of DNA. The BUB1B rs1801376 A/G SNP is a coding Arg/Glu variant; the BUB3 rs3808960 G/T SNP is a promoter polymorphism; the MAD2L1 rs903147 A/C SNP is a promoter polymorphism; the CENPF rs438034 G/A SNP is a coding Arg/Gly variant; the ESPL1 rs6580941 C/T SNP is a promoter polymorphism; the NEK2 rs701928 T/A SNP is a promoter polymorphism; the PTTG1 rs1862392 T/A SNP is a promoter polymorphism; the ZWILCH rs3087660 A/G SNP is a 5′UTR polymorphism; the ZWINT rs2241666 G/A SNP is a coding Arg/Gly variant. All promoter polymorphisms are predicted to occupy transcription factor binding sites [18,19].

Odds ratios (ORs) from multivariable logistic regression analysis were calculated by considering simultaneous effects of putative confounders, occupational exposures, age, gender and smoking habits on the CA frequencies. The associations of these possible confounding variables have been described [14,15]. For each SNP, adjusted ORs were calculated regarding their effect on CAtot, CTA and CSA. Irrespective of whether or not a SNP appeared to be individually significant, all possible pairs of two SNPs were considered for the SNP–SNP interaction analysis. The tested genetic models were 'three genotypes' of types AA, AB and BB; for 'the dominant mode of inheritance', AB and BB were merged as one group, while AA and AB together represented the reference group for 'the recessive model'. Moreover, genotypes were converted into zero, one or two risk alleles for the additive 'allele number' model. Likelihood ratio (LR) tests were performed to assess whether considering SNP–SNP interaction yielded a significantly better fit of the data. In addition, LR test statistics were calculated for the global null hypothesis to prove the significance of the whole model. If both SNPs significantly interacted with each other for various modes of inheritance for the same pair of SNPs, the model with the lowest Akaike Information Criterion was chosen. For the best model for a variant pair, the corresponding ORs and the Wald estimates for their confidence intervals and p-values were calculated. To assess the contribution of all genetic components (both SNPs and interaction term) to the model, LR based p-values were computed and shown as 'the overall p-value'.

The studies were coordinated at the German Cancer Research Center (DKFZ) with samples and study design obtained with informed consent approved by the local Ethical Committee of the Jessenius Medical Faculty and Slovak Medical University. Sampling of peripheral blood was carried out according to the Helsinki Declaration.

Results

The number of genotyped individuals ranged from 330 to 729, as detailed in Methods. Results for individual genotypes are shown in Table 1. CAtot did not show any nominally significant associations for the variants in the 9 genes. For CSA, the rare homozygotes of ZWINT showed an OR of 0.53 and the LR test p-value was 0.03. For CTA, the trend for the ZWINT variant was similar (OR 0.56), but it was even stronger for BUB1B for which the rare homozygous variant reached an OR of 0.30 (LR test p-value 0.03).

Pair-wise interactions of the genetic models of each of the 9 genes were tested for association with CAs (Table 2). Only the most significant models with p-values <0.05, based on the interaction term analysis and the LR test, are shown in Table 2; note that only the best model for each pair is listed in Table 2. Of the 21 tests shown, 6 included CAtot with 8 variants present. ESPL1 was a partner in 3 pairs, and PTTG1 and ZWILCH in 2 pairs each. The interaction term was most significant (p = 0.002) for the BUB3-ESPL1 pair. For CSA, also 6 variant pairs interacted, involving 8 genes. MAD2L1, CENPF, PTTG1 and ZWILCH were partners in 2 associations each. The interaction term was most significant (p = 0.01) for the CENPF-ESPL1 and PTTG-ZWILCH pairs. For CTA, a total of 9 gene pairs reached a significant interaction in which all 9 genes were involved. MAD2L1 was a partner in 4 pairs, PTTG1 and ZWINT in 3 pairs each and ZWILCH was a partner in 2 pairs. The interaction term was most significant (p = 0.001) for the NEK2-ZWINT pair. For CTA an interaction of the cohesin maintaining (PTTG1/securin) and degrading (ESPL1/separase) functions was observed.

In Table 2 several of the gene combinations were found for multiple CA types, although the genetic model was not always the same. Remarkably, 2 gene pairs were significant for all 3 types of CAs: PTTG1-ZWILCH and PTTG1-ZWINT. Of the 4 remaining gene pairs for CAtot, 2 were also found in other CA types (1 in CSA and 1 in CTA). CSA and CTA shared 2 additional gene pair associations (BUB1B-MAD2L1, MAD2L1-ZWILCH). In Table 2 we show also the global null hypothesis test which proved the adequacy of the whole statistical model for almost all pairs and for most associations showed smaller p-values than the interaction term analysis.

The ORs and the significances of the models for each combination of genotypes in Table 2 are shown in Supplementary Table S1. We also show there tests for total SNP information (overall p-value). For CAtot, 2 overall p-values were nominally significant, for CSA 1 was significant and for CTA 5 were significant. The lowest p-value was always found among the gene combinations which showed the most significant interaction terms in Table 2.

Discussion

Analysis of genotype-based genetic models for 9 genes for 3 types of CAs creates obviously a mass significance problem. However, it is difficult to estimate how many completely independent tests were carried out because CAtot depends on CSA and CTA, and because genotypes and genetic models are not independent. Anyway it is clear that the smallest p-values of 0.03 in the single SNP analysis in Table 1 would not survive any correction for multiple testing. In the interaction analysis in Table 2 for each CA type at least 72 independent tests were done (9 SNPs and 2 alleles). This would translate to a Bonferroni corrected p-value of 0.0007 (0.05/72). None of the interaction term p-values would survive this correction

Table 1
Odds ratios for cases and controls by gene polymorphisms.

Variable (SNP)	CAtot					CSA					CTA				
	Frequency of cases	Frequency of controls	Odds ratio[1]	Confidence interval[2]	p-value[3]	Frequency of cases	Frequency of controls	Odds ratio	Confidence interval	p-value	Frequency of cases	Frequency of controls	Odds ratio	Confidence interval	p-value
BUB1B (AA)	93	68			1	100	61			0.65	92	69			0.03
BUB1B (AB)	85	63	1.00	(0.63–1.59)		90	58	0.95	(0.60–1.51)		72	76	0.69	(0.44–1.09)	
BUB1B (BB)	12	9	0.97	(0.39–2.52)		15	6	1.52	(0.58–4.46)		6	15	0.30	(0.10–0.79)	
BUB3 (AA)	151	117			0.49	170	98			0.51	134	134			0.67
BUB3 (AB)	37	22	1.31	(0.73–2.40)		33	26	0.73	(0.41–1.31)		34	25	1.26	(0.71–2.27)	
BUB3 (BB)	2	1	2.70	(0.25–60.28)		2	1	1.67	(0.15–37.33)		2	1	1.73	(0.15–38.91)	
MAD2L1 (AA)	71	57			0.84	72	56			0.20	66	62			0.97
MAD2L1 (AB)	93	67	1.08	(0.67–1.73)		107	53	1.56	(0.96–2.53)		82	78	0.98	(0.61–1.58)	
MAD2L1 (BB)	26	16	1.25	(0.61–2.61)		26	16	1.23	(0.60–2.56)		22	20	1.07	(0.53–2.19)	
CENPF (AA)	56	100			0.65	55	101			0.80	41	115			0.69
CENPF (AB)	113	209	0.97	(0.65–1.47)		106	216	0.91	(0.60–1.37)		94	228	1.18	(0.76–1.83)	
CENPF (BB)	54	86	1.19	(0.73–1.92)		49	91	1.03	(0.64–1.68)		36	104	1.01	(0.59–1.71)	
ESPL1 (AA)	126	217			0.97	122	221			0.87	97	246			0.64
ESPL1 (AB)	88	160	0.96	(0.68–1.36)		86	162	0.98	(0.69–1.39)		65	183	0.93	(0.64–1.35)	
ESPL1 (BB)	24	41	0.95	(0.54–1.66)		26	39	1.14	(0.65–1.98)		22	43	1.24	(0.69–2.18)	
NEK2 (AA)	140	218			0.43	130	228			0.94	112	246			0.33
NEK2 (AB)	92	172	0.87	(0.62–1.22)		94	170	1.01	(0.72–1.42)		67	197	0.78	(0.54–1.12)	
NEK2 (BB)	12	29	0.66	(0.31–1.33)		16	25	1.13	(0.57–2.21)		10	31	0.71	(0.32–1.47)	
PTTG1 (AA)	140	249			0.62	137	252			0.47	111	278			0.06
PTTG1 (AB)	105	165	1.14	(0.83–1.58)		106	164	1.20	(0.86–1.66)		86	184	1.16	(0.82–1.63)	
PTTG1 (BB)	29	41	1.22	(0.72–2.05)		29	41	1.25	(0.73–2.11)		31	39	1.91	(1.12–3.22)	
ZWILCH (AA)	82	143			0.44	78	147			0.40	65	160			0.16
ZWILCH (AB)	135	211	1.15	(0.81–1.65)		134	212	1.22	(0.85–1.75)		110	236	1.18	(0.81–1.71)	
ZWILCH (BB)	35	68	0.87	(0.52–1.42)		35	68	0.93	(0.56–1.53)		24	79	0.72	(0.41–1.24)	
ZWINT (AA)	87	133			0.21	88	132			0.03	74	146			0.07
ZWINT (AB)	117	200	0.91	(0.63–1.30)		118	199	0.90	(0.63–1.29)		88	229	0.76	(0.52–1.11)	
ZWINT (BB)	37	88	0.66	(0.41–1.06)		32	93	0.53	(0.32–0.86)		27	98	0.56	(0.33–0.92)	

[1] ORs were adjusted for age, sex, occupational environment and smoking status.
[2] Profile likelihood confidence interval.
[3] Type 3 analysis (likelihood ratio test) to assess the significance of the SNP in general.

Table 2
Pair-wise interactions of genotypes with cases and controls. Only the most significant genetic models for each interaction are shown.

	SNP 1	Mode for SNP 1	SNP 2	Mode for SNP 2	Interaction term analysis			Global null hypothesis test for significance of all covariates		
					DF	Likelihood ratio test		DF	Likelihood ratio test	
						Chi-square	p-value		Chi-square	p-value
CAtot	BUB3	Dominant	ESPL1	Dominant	1	9.26	0.002	7	19.15	0.008
	CENPF	Recessive	ESPL1	Three genotypes	2	7.48	0.02	9	37.56	<.0001
	ESPL1	Recessive	ZWILCH	Allele number	1	5.01	0.03	7	31.21	<.0001
	MAD2L1	Dominant	NEK2	Allele number	1	4.31	0.04	7	15.32	0.03
	PTTG1	Recessive	ZWILCH	Allele number	1	6.64	0.01	7	27.72	0.0002
	PTTG1	Dominant	ZWINT	Recessive	1	4.85	0.03	7	32.66	<.0001
CSA	BUB1B	Dominant	MAD2	Recessive	1	4.89	0.03	7	7.08	0.42
	CENPF	Recessive	ESPL1	Three genotypes	2	9.29	0.01	9	30.38	0.0004
	CENPF	Recessive	NEK2	Recessive	1	4.18	0.04	7	20.56	0.005
	MAD 2L1	Three genotypes	ZWILCH	Recessive	2	6.55	0.04	9	19.81	0.02
	PTTG1	Recessive	ZWILCH	Allele number	1	6.41	0.01	7	28.04	0.0002
	PTTG1	Dominant	ZWINT	Recessive	1	5.74	0.02	7	36.24	<.0001
CTA	BUB1B	Three genotypes	MAD2L1	Dominant	2	6.16	0.05	9	20.55	0.01
	BUB3	Dominant	MAD2L1	Dominant	1	6.65	0.01	7	14.65	0.04
	CENPF	Allele number	ZWINT	Allele number	1	4.32	0.04	7	25.17	0.0007
	ESPL1	Allele number	PTTG1	Dominant	1	4.63	0.03	7	25.16	0.0007
	MAD2L1	Dominant	NEK2	Recessive	1	4.18	0.04	7	14.49	0.04
	MAD2L1	Recessive	ZWILCH	Dominant	1	5.10	0.02	7	13.93	0.05
	NEK2	Three genotypes	ZWINT	Allele number	2	13.71	0.001	9	34.26	<.0001
	PTTG1	Allele number	ZWILCH	Dominant	1	7.43	0.006	7	21.45	0.003
	PTTG1	Dominant	ZWINT	Recessive	1	8.64	0.003	7	29.29	0.0001

whereas about a half of global null hypothesis p-values would remain significant.

A biological way to assess the plausibility of the results is to consider their mechanistic credibility. For example, in our previous analyses of interactions of gene variants with CAs we noted for DNA repair genes highly significant interactions when gene pairs acted on two different pathways [15]. Similarly, in analyzing genes coding for metabolic enzymes, significant associations were only noted when a GST variant was one of the pair [14]. It is probably not equally clear to understand the current interactions because the role of multiple genes in spindle checkpoint is very complex and not fully established [10]. In Fig. 1 we summarize the function of the spindle assembly checkpoint and show the most significant interactions observed in our study. Our results suggest that the PTTG1 is a key culprit in CA associations because it was a partner in the two gene pairs, PTTG1-ZWILCH and PTTG1-ZWINT, which were associated with the three types of CAs with reasonable statistical support. The PTTG1 encoded protein (alias securin) is known to play a central role in chromosomal stability because it inhibits separase function and thus prevents the proteolysis of the cohesin complex and segregation of the chromosomes [10,11]. At the onset of anaphase it is ubiquitinated and destructed and thus sister chromatids can separate. A securin mutation has been shown to lead to CAs and securin has been assumed to promote repair of double-strand DNA breaks before sister chromatid separation by interaction with the repair protein Ku [20].

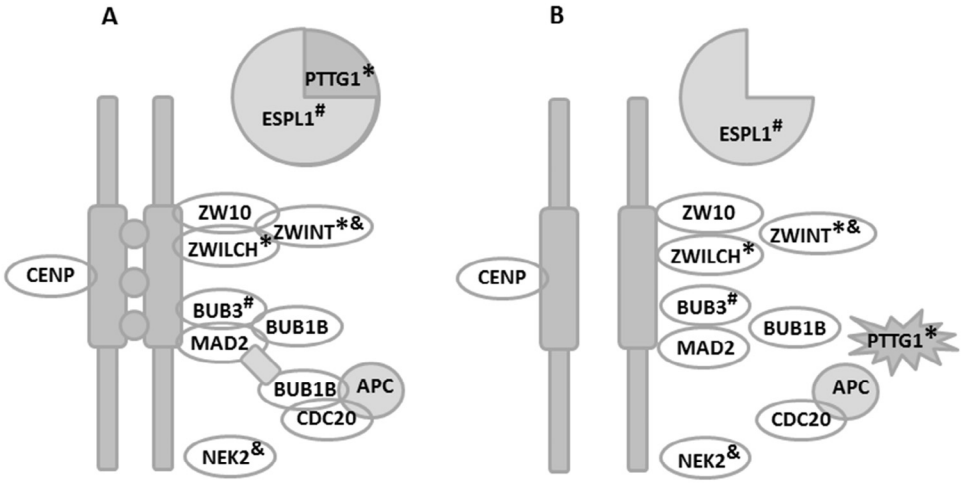


Fig. 1. Major components of the spindle assembly checkpoint. (A) In prometaphase, an inhibitory signal from unattached kinetochores induces recruitment of the mitotic checkpoint complex (MCC, including BUB3, BUB1B and MAD2) and the ZW10-ZWINT-ZWILCH complex, preventing the activation of APC/C-CDC20. (B) When all chromosomes are properly attached in the mitotic spindle, MAD2 is released from the MCC complex, resulting in the activation of APC/C-CDC20, which in turn targets PTTG1 for degradation. ESPL1 gets activated and can cleave the cohesin complexes that keep the sister chromatids together. *, # and & indicate the most significant interactions observed in our SNP–SNP interaction analysis.

Another key process in preparing for chromosomal segregation is the binding of microtubules to the kinetochore which anchors onto the sister chromatids. The ZWINT protein binds the RZZ complex, including ZWILCH, to the kinetochore. When microtubules bind to the kinetochore, dynein strips RZZ and MAD1-MAD2 complexes from the kinetochore initiating chromosomal segregation [11]. Thus, the present results suggest that interactions at the level of cohesin (PTTG1) and kinetochore function (ZWINT, ZWILCH and MAD2L1) contribute to the frequency of CAs. MAD2L1 and PTTG1 were the most common partners in any of the two-way interactions. Interestingly, for CTA an interaction of the cohesin maintaining (PTTG1/securin) and degrading (ESPL1/separase) functions was observed.

In our previous study on interactions on genes that encoded metabolic enzymes we found many gene pairs associating with CTA and only one pair associating with CSA [14]. As an explanation we suggested that metabolism processes genotoxic agents to intermediates capable of covalent binding to DNA. The binding products, DNA adducts, would rarely cause double-strand breaks suggesting a rationale for more CTAs than CSAs [5,21]. In the present study also associations with CTAs were in excess. Although the reasons are not known one can speculate that mitotic checkpoint may sense chromosomes with double-strand breaks and stop or delay entrance to anaphase.

In conclusion, the present results suggest that genetic variation in the key mitotic checkpoint genes is associated with CA formation. Importantly, variation in individual genes played a minor importance which would be consistent with the high conservation and selection pressure of the system. Thus interactions of genes at different checkpoint functions appeared to be required for the formation of CAs.

Funding

The study was supported by the Czech Science Foundation (15-14789S), Charles University of Prague, Czech Republic (P27/2012) and the Ministry of Education Youth and Sports of the Czech Republic (National Sustainability Program I (NPU I) Nr. LO1503).

Conflict of interest

None.

Appendix: Supplementary material

Supplementary data to this article can be found online at doi:10.1016/j.canlet.2016.07.011.

References

[1] F. Mitelman, B. Johansson, F. Mertens, The impact of translocations and gene fusions on cancer causation, Nat. Rev. Cancer 7 (2007) 233–245.
[2] S. Bonassi, H. Norppa, M. Ceppi, U. Stromberg, R. Vermeulen, A. Znaor, et al., Chromosomal aberration frequency in lymphocytes predicts the risk of cancer: results from a pooled cohort study of 22 358 subjects in 11 countries, Carcinogenesis 29 (2008) 1178–1183.
[3] P. Vodicka, Z. Polivkova, S. Sytarova, H. Demova, M. Kucerova, L. Vodickova, et al., Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls, Carcinogenesis 31 (2010) 1238–1241.
[4] F. Mertens, B. Johansson, T. Fioretos, F. Mitelman, The emerging complexity of gene fusions in cancer, Nat. Rev. Cancer 15 (2015) 371–381.
[5] R.J. Albertini, D. Anderson, G.R. Douglas, L. Hagmar, K. Hemminki, F. Merlo, et al., IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. International Programme on Chemical Safety, Mutat. Res. 463 (2000) 111–172.
[6] L. Xu, S. Li, B.A. Stohr, The role of telomere biology in cancer, Annu. Rev. Pathol. 8 (2013) 49–78.
[7] S.E. Artandi, R.A. DePinho, Telomeres and telomerase in cancer, Carcinogenesis 31 (2010) 9–18.
[8] H. Li, H.T. Hilmarsen, M.B. Hossain, J. Bjork, I.L. Hansteen, M. Albin, et al., Telomere length and LINE1 methylation is associated with chromosomal aberrations in peripheral blood, Genes Chromosomes Cancer 52 (2013) 1–10.
[9] K. Hemminki, S. Rachakonda, L. Musak, V. Vymetalkova, E. Halasova, A. Forsti, et al., Telomere length in circulating lymphocytes: association with chromosomal aberrations, Genes Chromosomes Cancer 54 (2015) 194–196.
[10] N. London, S. Biggins, Signalling dynamics in the spindle checkpoint response, Nat. Rev. Mol. Cell Biol. 15 (2014) 736–747.
[11] C. Sacristan, G.J. Kops, Joined at the hip: kinetochores, microtubules, and spindle assembly checkpoint signaling, Trends Cell Biol. 25 (2015) 21–28.
[12] G. Zhang, T. Lischetti, D.G. Hayward, J. Nilsson, Distinct domains in Bub1 localize RZZ and BubR1 to kinetochores to regulate the checkpoint, Nat. Commun. 6 (2015) 7162.
[13] D. Varma, X. Wan, D. Cheerambathur, R. Gassmann, A. Suzuki, J. Lawrimore, et al., Spindle assembly checkpoint proteins are positioned close to core microtubule attachment sites at kinetochores, J. Cell Biol. 202 (2013) 735–746.
[14] K. Hemminki, C. Frank, A. Forsti, L. Musak, A. Kazimirova, M. Barancokova, et al., Metabolic gene variants associated with chromosomal aberrations in healthy humans, Genes Chromosomes Cancer 54 (2015) 260–266.
[15] P. Vodicka, L. Musak, C. Frank, A. Kazimirova, V. Vymetalkova, M. Barancokova, et al., Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects, Carcinogenesis 36 (2015) 1299–1306.
[16] A. Kazimirova, M. Barancokova, Z. Dzupinkova, L. Wsolova, M. Dusinska, Micronuclei and chromosomal aberrations, important markers of ageing: possible association with XPC and XPD polymorphisms, Mutat. Res. 661 (2009) 35–40.
[17] L. Musak, Z. Smerhovsky, E. Halasova, O. Osina, L. Letkova, L. Vodickova, et al., Chromosomal damage among medical staff occupationally exposed to volatile anesthetics, antineoplastic drugs, and formaldehyde, Scand. J. Work Environ. Health 39 (2013) 618–630.
[18] A. Vaclavicek, J.L. Bermejo, B. Wappenschmidt, A. Meindl, C. Sutter, R.K. Schmutzler, et al., Genetic variation in the major mitotic checkpoint genes does not affect familial breast cancer risk, Breast Cancer Res. Treat. 106 (2007) 205–213.
[19] A. Brendle, A. Brandt, R. Johansson, K. Enquist, G. Hallmans, K. Hemminki, et al., Single nucleotide polymorphisms in chromosomal instability genes and risk and clinical outcome of breast cancer: a Swedish prospective case-control study, Eur. J. Cancer 45 (2009) 435–442.
[20] M. Mora-Santos, C. Castilla, J. Herrero-Ruiz, S. Giraldez, M.C. Limon-Mortes, C. Saez, et al., A single mutation in Securin induces chromosomal instability and enhances cell invasion, Eur. J. Cancer 49 (2013) 500–510.
[21] M. Durante, J.S. Bedford, D.J. Chen, S. Conrad, M.N. Cornforth, A.T. Natarajan, et al., From DNA damage to chromosome aberrations: joining the break, Mutat. Res. 756 (2013) 5–13.

Manuscript II

Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects.

Vodicka P, Musak L, Frank C, Kazimirova A, Vymetalkova V, Barancokova M, Smolkova B, Dzupinkova Z, Jiraskova K, Vodenkova S, **Kroupa M**, Osina O, Naccarati A, Palitti F, Försti A, Dusinska M, Vodickova L, Hemminki K.

Carcinogenesis. 2015 Nov;36(11):1299-306. **IF: 5.072**

Carcinogenesis, 2015, Vol. 36, No. 11, 1299–1306

doi:10.1093/carcin/bgv127
Advance Access publication September 8, 2015
Original Manuscript

ORIGINAL MANUSCRIPT

Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects

Pavel Vodicka[1,2,3,*], Ludovit Musak[4], Christoph Frank[5], Alena Kazimirova[6], Veronika Vymetalkova[1,2], Magdalena Barancokova[6], Bozena Smolkova[7], Zuzana Dzupinkova[8], Katerina Jiraskova[1,2], Sona Vodenkova[1,2,9], Michal Kroupa[1,10], Oto Osina[4], Alessio Naccarati[1,11], Fabrizio Palitti[12], Asta Försti[5,13], Maria Dusinska[14], Ludmila Vodickova[1,2,3] and Kari Hemminki[5,13]

[1]Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Academy of Sciences of the Czech Republic, 14220 Prague, Czech Republic, [2]Institute of Biology and Medical Genetics, 1st Faculty of Medicine, Charles University in Prague, 12800 Prague, Czech Republic, [3]Biomedical Centre, Faculty of Medicine in Pilsen, Charles University in Prague, 32300 Pilsen, Czech Republic, [4]Clinic of Occupational Medicine and Toxicology, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava and University Hospital in Martin, 03601 Martin, Slovak Republic, [5]Division of Molecular Genetic Epidemiology, German Cancer Research Centre (DKFZ), 69120 Heidelberg, Germany, [6]Faculty of Medicine, Slovak Medical University in Bratislava, 83303 Bratislava, Slovak Republic, [7]Department of Cancer Genetics, Cancer Research Institute of Slovak Academy of Sciences, 83391 Bratislava, Slovak Republic, [8]Molecular Biology Service, Research Institute for Molecular Pathology, 1030 Vienna, Austria, [9]Department of General Biology and Genetics, 3rd Faculty of Medicine, Charles University in Prague, 10000 Prague, Czech Republic, [10]Faculty of Medicine in Pilsen, Charles University in Prague, 30605 Pilsen, Czech Republic, [11]Molecular and Genetic Epidemiology Research Unit, HuGeF–Human Genetics Foundation, 10126 Torino, Italy, [12]Department of Agrobiology and Agrochemistry, University of Tuscia, 011000 Viterbo, Italy, [13]Center for Primary Health Care Research, Lund University, 20506 Malmö, Sweden and [14]Health Effects Laboratory MILK, NILU–Norwegian Institute for Air Research, 2027 Kjeller, Norway

*To whom correspondence should be addressed. Tel: +420 241 062 694; Fax: +420 241 062 782; Email: pvodicka@biomed.cas.cz

Abstract

Human cancers are often associated with numerical and structural chromosomal instability. Structural chromosomal aberrations (CAs) in peripheral blood lymphocytes (PBL) arise as consequences of direct DNA damage or due to replication on a damaged DNA template. In both cases, DNA repair is critical and inter-individual differences in its capacity are probably due to corresponding genetic variations. We investigated functional variants in DNA repair genes (base and nucleotide excision repair, double-strand break repair) in relation to CAs, chromatid-type aberrations (CTAs) and chromosome-type aberrations (CSAs) in healthy individuals. Chromosomal damage was determined by conventional cytogenetic analysis. The genotyping was performed by both restriction fragment length polymorphism and TaqMan allelic discrimination assays. Multivariate logistic regression was applied for testing individual factors on CAs, CTAs and CSAs. Pair-wise genotype interactions of 11 genes were constructed for all possible pairs of single-nucleotide polymorphisms. Analysed individually, we observed significantly lower CTA frequencies in association with *XPD* Lys751Gln homozygous variant genotype [odds ratio (OR) 0.64, 95% confidence interval (CI) 0.48–0.85, $P = 0.004$; $n = 1777$]. A significant association of heterozygous variant genotype in *RAD54L* with increased CSA frequency (OR 1.96, 95% CI 1.01–4.02, $P = 0.03$) was determined in 282 subjects with available genotype. By addressing gene–gene interactions, we discovered 14 interactions significantly modulating CAs, 9 CTAs and 12 CSAs frequencies. Highly significant interactions included always pairs from two different pathways. Although individual variants in genes encoding DNA repair proteins modulate CAs only modestly, several gene–gene interactions in DNA repair genes

Received: June 17, 2015; **Revised:** August 18, 2015; **Accepted:** August 20, 2015

evinced either enhanced or decreased CA frequencies suggesting that CAs accumulation requires complex interplay between different DNA repair pathways.

Abbreviations

BER	base excision repair
CA	chromosomal aberration
CI	confidence interval
CSA	chromosome-type aberration
CTA	chromatid-type aberration
DSB	double-strand break
NER	nucleotide excision repair
OR	odds ratio
PBL	peripheral blood lymphocytes
SNP	single-nucleotide polymorphism

Introduction

Human DNA is constantly exposed to physical (ultraviolet, ionizing radiation) and chemical (reactive oxygen species, alkylating and aralkylating) damaging agents. Efficient DNA repair machinery, comprising several distinct pathways, maintains effectively genomic integrity. Alterations in the DNA repair increase the vulnerability of the cells, resulting in an accumulation of mutations in the genome, which may ultimately lead to tumorigenesis (1). Structural chromosomal aberrations (CAs) arise as a consequence of direct DNA damage (e.g. ionizing radiation, free radicals) or due to replication on a damaged DNA template (2), in both cases DNA repair represents a key player (3). The lesions causing double-strand breaks (DSBs) are mainly responsible for chromosome-type aberrations (CSAs), whereas chromatid-type aberrations (CTAs) arise as a consequence of DNA lesions generated by genotoxic damage during G0 phase, which are insufficiently repaired prior the entering of the cell into S-phase (4). CAs in peripheral blood lymphocytes (PBL) thus reflect inter-individual sensitivity to many genotoxic substances and serve as biomarkers of an early effect of genotoxic carcinogens and carcinogenic risk (5–10).

Human cancers are often associated with chromosomal instability (both numerical and structural CAs) in the cells (11–13); these aberrations are also considered as causative events in malignant transformation (14). Frequencies of CA in PBL are predictive for cancer risk in prospective epidemiological studies (15–17), and patients with many types of cancer show elevated CAs at the time of diagnosis (10,18,19).

Individual DNA repair capacity in response to DNA damage, effectively preventing an accumulation of CAs, is often modulated by the gene variants in different DNA repair pathways (20–22). Indeed, a correlation between gene variants involved in base excision repair (BER) and the corresponding BER repair capacity has been documented (23). However, single-nucleotide polymorphisms (SNPs) in non-coding regions and even changes in wobble bases, which do not affect amino acid sequence, may be important as well (24). The investigation of DNA repair gene variants in association with DNA repair capacity or DNA damage, often performed on small study groups, does not usually bring consistent results, with the exception of the BER gene 8-oxoguanine DNA glycosylase, *OGG1* (23–26) and the nucleotide excision repair (NER) gene *XPA* (27–29). Few reports analysed effects of genetic predispositions on inter-individual variability in CAs by studying variants in genes encoding xenobiotic-metabolizing enzymes, enzymes of DNA repair or folate metabolism and DNA repair capacity (8,30–33). Most recently, by investigating SNPs in metabolic genes, modulations of DNA damage and CAs by gene–gene interactions were reported (34,35). In our recent study, we have described the significant association of rs9344 polymorphism in *Cyclin D1* at a splice site with non-specific CAs in healthy individuals (36). Interestingly, Cyclin D1 participates in DNA DSB repair pathway by binding to RAD51 that is a main recombinase involved in homologous recombination (37).

In the present study, we examined the hypothesis that functionally relevant SNPs in the BER, NER and DSB repair pathways and their gene–gene interactions may modulate frequencies of structural CAs in a large set of healthy subjects.

Materials and methods

Study population

The group of studied subjects (>2100) with measured frequencies of CAs has been very recently described in Hemminki *et al.* (35). The above healthy subjects were recruited between 2002 and 2011 in eastern Bohemia and 1997–2006 in Slovakia and consisted of unexposed controls as well as subjects with defined occupational exposures, such as small organic compounds, cytostatics, anaesthetics, metals, asbestos, mineral fibres and ionizing radiation. Peripheral blood sampling and data collection were carried out simultaneously only in the subjects apparently healthy at the time of sampling. Likewise, individuals with close relatives with any malignant diseases were excluded from the study. Otherwise, no other exclusion criteria were applied. All individuals completed a questionnaire regarding the job category, mode and duration of exposure, various exogenous factors (such as smoking, drug usage, exposure to X-ray radiation, alcohol consumption and dietary habits) prior to blood collection and provided a written consent.

The present study adheres to all principles of the Helsinki Declaration and its design was approved by the Local Ethical Committees of the Jessenius Medical Faculty (Martin, Slovakia) and of the Slovak Medical University (Bratislava, Slovakia).

Cytogenetic analysis

Cytogenetic analysis was performed in PBL stimulated to grow by phytohaemagglutinin and cultured for 48 h. Two scores conducted microscopically analysis (each evaluating half of the 100 mitoses scored per subject) in a double-blind fashion on coded slides. The frequency of CAs and the constituent CTAs and CSAs were evaluated as stated in refs. 6–9,19,32,38. The subjects were classified according to the median of CA distributions into either low-frequency (<2%) or high-frequency (≥2%) groups. Regarding CTAs and CSAs, the cut-off was set up at 1%. These arbitrary cut-off levels were introduced on the basis of the long-term experience with human biological monitoring in the Czech and Slovak Republics (9,10,39).

Genotyping

SNPs in DNA repair genes were taken into the study on the basis of predicted functional effects (SIFT and PolyPhen databases) and relevant published literature. Genotyping of DNA repair gene polymorphisms *XPD* Lys751Gln (rs13181; T > G), *XPG* Asp1104His (rs17655; C > G), *XPC* Lys939Gln (rs2228001; A > C), XPA 5′UTR (rs1800975; G > A), *XRCC1* Arg194Trp (rs1799782; C > T), Arg280His (rs25489; G > A) and Arg399Gln (rs25487G > A), *OGG1* Ser326Cys (rs1052133; C > G), *XRCC2* Arg188His (rs3218536; G > A), *RAD54L* Ala730Ala (rs1048771; C > T) and *XRCC3* Thr241Met (rs861539; C > T) was carried out using primers and conditions described previously (23,40). The amplified fragments were digested with appropriate restriction endonucleases and the digested PCR products resolved on 2% agarose gel and visualized under ultraviolet light after staining with ethidium bromide. Genetic polymorphisms in *APE1* Asn148Glu (rs1130409; G > T) and *NBS1* Glu185Gln (rs1805794; C > G) were analysed using the TaqMan allelic

discrimination assay (Applied Biosystems, Foster City, CA, Assay-on-demand, SNP Genotyping products: C 26470398 10 for *NBS1* and C 8921503 10 for *APE1*). The TaqMan genotyping reaction was amplified on a 7500 Real-Time PCR System (Applied Biosystems) using the following cycling conditions: 95°C for 10 min and 40 cycles at 92°C for 15 s and 60°C for 60 s. The results were regularly confirmed by random re-genotyping of more than 10% of the samples for each polymorphism analysed (40).

Statistical analysis

Statistical calculations have been conducted essentially as recently described in Hemminki *et al.* (35). Briefly, odds ratios (ORs) from the multivariate logistic regression analysis were used to investigate simultaneous effects of occupational exposures and main confounders on the frequencies of CAs, CTAs and CSAs. For each SNP, adjusted ORs were calculated to discern their effect on chromosomal damage. Additionally, all possible pairs of two SNPs were evaluated in binary interaction analysis. In particular, the following genetic models were tested for each pair. 'Three genotypes' assumed the SNP to be a categorical covariate comprising of genotypes AA, AB and BB. For the dominant mode of inheritance, AB and BB were merged as one group, whereas AA and AB together represented the reference group for recessive models. Moreover, genotypes were converted into zero, one or two risk alleles for the log-additive 'allele number' model. Likelihood ratio tests were performed to assess whether including the SNP–SNP interaction term yielded a significantly better fit of the data. If both SNPs significantly interacted with each other for various modes of inheritance for the same pair of SNPs, the model with the lowest Akaike information criterion was chosen. For each best model, the corresponding ORs and the Wald estimates for their confidence intervals (CIs) and *P*-values were calculated. To assess the contribution of all genetic components (both SNPs and interaction term) to the model, likelihood ratio-based *P*-values were computed. Considering multiple comparisons performed in the present study, we have also highlighted significant results after applying Dunn–Bonferroni correction. After correction, the new threshold of *P*-value significance results is 0.004.

Results

Gene polymorphisms in major DNA repair pathways included into the study with respect to their functional relevance are shown in Table 1. Distribution of subjects with high and low CA, CTA and CSA frequencies, also in relation to their main confounders (age, gender, smoking and occupational exposure), is summarized in Table 2. The frequency of individuals with particular CAs, CTAs and CSAs percentage is illustrated in Figure 1a–c. In the whole set, the mean (± standard deviation) frequencies of CAs, CTAs and CSAs were 1.54 ± 1.54%, 0.74 ± 0.98% and 0.80 ± 1.16%, respectively, with median and range being 1 (0–11), 0 (0–6) and 0 (0–10). Chromatid-type exchanges (mean ± standard deviation 0.02 ± 0.14%) were substantially less abundant than chromosome-type exchanges (mean ± standard deviation 0.09 ± 0.35%), which included mainly dicentrics and centric rings. Chromatid-type exchanges occurred in 2% of subjects with CTAs, whereas chromosome-type exchanges were detected in 7.7% of individuals with CSAs. Based on the data from 2196 investigated individuals, CAs as well as the constituent CTAs and CSAs were significantly increased in occupationally exposed subjects (ORs 2.36, 1.73 and 1.64, respectively). Age was only moderately associated with increasing CAs, whereas its association with CSAs was borderline significant (OR 1.07, 95% CI 1.00–1.13, $P = 0.04$). Smoking did not affect significantly frequencies of either CAs or the constituent CTAs and CSAs.

Results for individual genotypes, including the numbers of subjects with the particular allele in relation to the frequency of CAs, are shown in Table 3. By assessing individual DNA repair gene polymorphisms, we observed a strong association between variant GG genotype in rs13181 of *XPD* gene and decreased CTA frequency (OR 0.64, 95% CI 0.48–0.85, $P = 0.004$; $n = 1777$ subjects). This association was strong enough to withstand a correction for multiple testing (*P*-value after correction for multiple comparison is 0.004). The effect of homozygous variant G allele in *XPD* on CTAs is additionally documented in Figure 2, where this allele was associated with the lowest CTA frequencies. Further, a significant association of CT genotype in rs1048771 of *RAD54L* with increased CSAs was also observed (adjusted OR 1.96, 95% CI 1.01–4.02, $P = 0.03$; determined in 282 subjects with available genotype). None of the other studied DNA repair gene variants modulated the frequencies of CAs, CTAs and CSAs.

We performed pair-wise interactions for each of the 11 genes, which were tested for association with CA frequencies. Various genetic models were examined, and if both SNPs significantly interacted with each other for various modes of inheritance for the same pair of SNPs, the model with the lowest Akaike information criterion was chosen. The data based on the interaction term analysis and the likelihood ratio test

Table 1. DNA repair polymorphisms evaluated in this study

Gene	SNP ID	Amino acid substitution	Alleles (major/minor)	Chromosome	Location	MAF (NCBI)	MAF in controls	Function (SIFT, PolyPhen-2 algorithms)
BER								
XRCC1	rs1799782	Arg194Trp	G/A	19q13.2	Missense	0.13	0.06	Deleterious
XRCC1	rs25489	Arg280His	C/T	19q13.2	Missense	0.07	0.04	Possibly deleterious
XRCC1	rs25487	Arg399Gln	A/G	19q13.2	Missense	0.26	0.37	Benign
OGG1	rs1052133	Ser326Cys	C/G	3p26.2	Missense	0.30	0.20	Deleterious
APEX1	rs1130409	Asp148Glu	T/G	14q11.2	Missense	0.38	0.46	Ambiguous
NER								
XPA	rs1800975	—	G/A	9q22.3	5′UTR	0.35	0.34	—
XPD	rs13181	Lys751Gln	T/G	19q13.3	Missense	0.24	0.41	Deleterious
XPG	rs17655	Asp1104His	G/C	13q33	Missense	0.36	0.22	Deleterious
XPC	rs2228001	Lys939Gln	T/G	3p25	Missense	0.32	0.41	Benign
DSB								
XRCC2	rs3218536	Arg188His	C/T	7q36.1	Missense	0.05	0.06	Benign
XRCC3	rs861539	Thr241Met	G/A	14q32.3	Missense	0.22	0.36	Deleterious
NBN	rs1805794	Glu185Gln	C/G	8q21	Missense	0.36	0.33	Benign
RAD54L	rs1048771	Ala730=	C/T	1p32	Synonymous	0.19	0.10	Splicing regulation

MAF, minor allele frequency.

Table 2. Numbers of subjects with high and low frequency of CAs, CSAs and CTAs and their distribution according to basic variables

Variable	Persons	CAs			CSAs			CTAs		
		OR	95% CI	*P*	OR	95% CI	*P*	OR	95% CI	*P*
CAs (high/low)	951/1245									
CTAs (high/low)	1041/1154									
CSAs (high/low)	983/1213									
Age (minimum, maximum, mean)	18, 88, 43	1.06[a]	1.00–1.13	0.07	**1.07**	**1.00–1.13**	**0.04**	0.95	0.90–1.01	0.11
Occupational exposure (exposed/unexposed)	1207/989	**2.36**	**1.97–2.83**	**<0.01**	**1.73**	**1.45–2.06**	**<0.01**	**1.64**	**1.38–1.96**	**<0.01**
Gender (male/female)	1171/1025	1.03	0.86–1.23	0.77	1.05	0.88–1.26	0.59	**0.83**	**0.69–0.99**	**0.04**
Smoking (smokers/non-smokers)	614/1557	1.19	0.97–1.45	0.09	0.95	0.78–1.16	0.63	1.13	0.93–1.38	0.23

Significant results are highlighted in bold.
[a]ORs for age were calculated for 10 years age difference (unit = 10).
P-value was considered significant if it was <0.05.

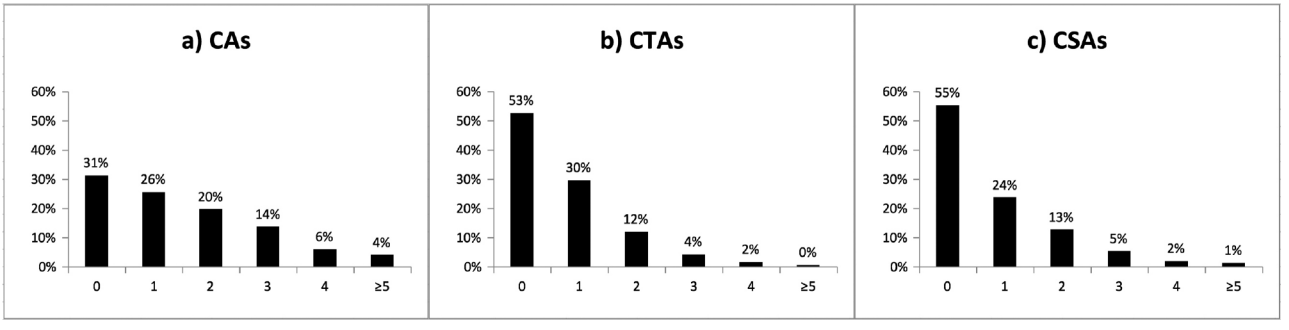


Figure 1. The frequency of individuals with particular percentage of CAs (**a**), CTAs (**b**) and CSAs (**c**). x-axis represents the percentage of chromosomal aberrations, y-axis represents the percentage of subjects with corresponding chromosomal aberrations.

are shown in Table 4. We have discovered several significant interactions, involving gene variants from BER, NER and DSB pathways: 14 interactions modulated CA, 9 CTA and 12 CSA frequencies. The overall genotype effect, considering individual SNPs, their interactions and adjustment variable combinations are shown by the global null hypothesis test. Tests for the global null hypothesis provided mostly highly significant outcomes. The ORs and the significances of the models for each combination of genotypes are showed in Supplementary Material, available at *Carcinogenesis* Online, and partially presented in Table 4.

For CAs, significant gene–gene interactions were observed mainly for genes participating in BER (*APE1*, *hOGG1*), NER (*XPC*, *XPD*) and DSB repair (*XRCC3*) together with (*NBS1*, *XRCC2* and *XPG1*). Rs1805794 in *NBS1* gene appeared most often in these interactions; although interactions with BER gene variants resulted in the higher CA frequency, opposite effect was recorded for the interactions with NER variants. Individuals both with homozygous variant GG and GG genotypes in *OGG1* and *XPG* genes, respectively, resulted in significant decrease of CAs (OR 0.59, 95% CI 0.37–0.96, *P* = 0.03).

For CTAs, the combinations of homozygous variant genotypes of *XPD* and *XPG* (OR 0.54) and/or *XRCC1* (OR 0.68) genes showed decreased frequencies of CTAs. So did the combination of variant alleles in *OGG1* and *XRCC3* (OR 0.52). On the contrary, a combination of variant alleles in DSB repair genes (*XRCC3* and *XRCC2*) resulted in the significant increase of OR to 2.56.

For CSAs, again variant alleles in *XRCC1* and *OGG1* (both BER genes) in combination with homozygous variant genotype in *XPG* resulted in significantly decreased frequencies of CSAs (OR 0.22 and 0.72, respectively).

Discussion

The onset and development of human cancer are associated with genome instability (41,42), resulting in both numerical and structural chromosomal abnormalities in cancer cells (11–13). CAs in PBL have been employed as biomarkers reflecting individual sensitivity to exogenous and endogenous genotoxic substances for many years (4,5). There are several reports pointing to the occupational exposure as a causative of enhanced CA frequencies for small molecular chemicals in plastics industry (32,43,44), for anaesthetic gases and antineoplastic drugs (9), for heavy metals (45) and for mineral fibres (6,7,38). In the present study, individuals that were occupationally exposed mainly to small molecular organic chemicals, anaesthetic gases, heavy metals and fibres (1207 subjects) indeed exhibited higher CAs, CTAs and CSAs in comparison with 989 unexposed individuals. Frequencies of chromatid- or chromosome-type exchanges contributed only marginally to the total CTAs and CSAs irrespective of exposure, as shown by us earlier (9). We recorded a borderline effect of smoking on CAs. The exposure of human DNA to above-listed agents inevitably leads to DNA damage. The capacity of DNA repair machinery to cope with these DNA alterations preserves the genomic integrity and prevents carcinogenesis (46,47). Thus, we investigated the role of functional variants in various DNA repair pathways, comprising BER (represented by *XRCC1*, *hOGG1* and *APE1*), NER (*XPA*, *XPC*, *XPD* and *XPG*) and DSB repair (*XRCC2*, *XRCC3*, *NBN* and *RAD54L*) genes. Interestingly, we observed a strong association between *XPD* Lys751Gln homozygous variant genotype and decreased CTA frequency. Our study on 1777 subjects confirmed our earlier observations on 225 healthy subjects (30) and later report on 140 subjects with higher age (8). *XPD* represents an important helicase involved in NER

Table 3. Odds ratios for high and low frequency groups by gene polymorphisms

	Cas					CTAs					CSAs				
Variables (SNP)	HF	LF	OR	95% CI	P[a]	HF	LF	OR	95% CI	P[a]	HF	LF	OR	95% CI	P[a]
XPD rs13181															
TT	271	358			0.98	301	328			**0.004**	287	342			0.87
TG	367	484	1.02	0.83–1.26		399	452	0.99	0.80–1.21		377	474	0.95	0.77–1.17	
GG	128	169	1.00	0.75–1.33		112	185	**0.64**	**0.48–0.85**		135	162	1.01	0.76–1.33	
XPG rs17655															
CC	460	601			0.67	477	584			0.43	485	576			0.57
CG	270	350	1.04	0.85–1.27		293	327	1.12	0.92–1.37		273	347	0.95	0.78–1.17	
GG	30	45	0.83	0.51–1.35		32	43	0.88	0.54–1.42		30	45	0.78	0.48–1.26	
XPC rs2228001															
AA	251	368			0.17	284	335			0.65	271	348			0.06
AC	398	467	1.20	0.97–1.49		403	462	0.99	0.81–1.23		409	456	1.12	0.91–1.38	
CC	117	171	1.00	0.75–1.33		124	164	0.88	0.66–1.17		112	176	0.81	0.61–1.08	
XRCC1 rs1799782															
CC	361	201			0.47	331	231			0.67	357	205			0.99
CT	49	25	1.09	0.66–1.85		40	34	0.80	0.49–1.31		47	27	1.02	0.62–1.71	
TT	1	2	0.25	0.01–2.66		2	1	0.97	0.09–21.1		2	1	1.04	0.10–22.7	
XRCC1 rs25489															
GG	212	120			0.70	200	132			0.18	209	123			0.13
GA	14	7	1.26	0.50–3.48		15	6	1.74	0.67–5.10		13	8	1.09	0.44–2.86	
AA	1	1	0.37	0.01–9.61		2	0	—	—		0	2	—	—	
XRCC1 rs25487															
GG	310	374			0.18	316	368			0.69	321	363			0.26
GA	352	507	0.86	0.70–1.02		385	474	0.96	0.78–1.18		384	475	0.93	0.76–1.14	
AA	98	119	1.07	0.78–1.45		102	115	1.09	0.80–1.49		86	131	0.77	0.56–1.05	
OGG1 rs1052133															
CC	495	623			0.44	526	592			0.31	511	607			0.32
CG	237	337	0.89	0.73–1.10		251	323	0.89	0.72–1.09		247	327	0.90	0.73–1.10	
GG	28	46	0.81	0.49–1.32		29	45	0.76	0.46–1.22		28	46	0.74	0.45–1.19	
XRCC3 rs861539															
CC	302	395			0.51	308	389			0.86	321	376			0.34
CT	355	454	0.98	0.79–1.21		375	434	1.06	0.86–1.30		358	451	0.91	0.74–1.11	
TT	83	126	0.83	0.60–1.14		95	114	1.04	0.76–1.42		86	123	0.80	0.58–1.09	
APE1 rs1130409															
GG	179	195			0.82	192	182			0.60	188	186			0.71
GT	316	338	1.03	0.78–1.35		313	341	0.89	0.69–1.16		321	333	0.94	0.72–1.22	
TT	144	127	1.11	0.79–1.55		145	126	1.00	0.72–1.38		146	125	1.06	0.77–1.47	
NBS1 rs1805794															
CC	245	189			0.85	242	192			0.47	249	185			0.89
CG	243	177	1.07	0.81–1.42		238	182	1.05	0.80–1.38		234	186	0.94	0.71–1.24	
GG	63	49	0.98	0.63–1.51		57	55	0.81	0.53–1.23		63	49	0.93	0.61–1.43	
XPA rs1800975															
GG	246	389			0.69	260	375			0.29	262	373			0.96
GA	236	406	0.92	0.73–1.15		274	368	1.08	0.86–1.35		262	380	0.98	0.79–1.23	
AA	57	109	0.89	0.61–1.27		59	107	0.81	0.56–1.16		68	98	1.03	0.72–1.46	
XRCC2 rs3218536															
GG	177	533			0.36	231	479			0.14	203	507			0.34
GA	23	77	0.88	0.52–1.43		39	61	1.30	0.84–2.00		31	69	1.13	0.70–1.78	
AA	0	3	—	—		0	3	—	—		0	3	—	—	
RAD54L s1048771															
CC	134	93			0.16	114	113			0.84	138	89			**0.03**
CT	30	2	0.95	0.51–1.79		26	26	1.04	0.57–1.93		39	13	**1.96**	**1.01–4.02**	
TT	3	0	—	—		2	1	2.03	0.19–44.5		3	0	—	—	

Significant results are highlighted in bold. HF, high-frequency group—cases; LF, low-frequency group—controls.
[a]Based on likelihood ratio test.
P-value was considered significant if it was <0.05.

and communicates with other DNA repair gene products in dealing with exogenous DNA damage (48), but the functional role of *XPD* Lys751Gln remains unclear. The assumption that functional SNPs in *XPD* gene modulate CTA frequency is consistent with mechanistic understanding, since CTAs arise as a consequence of DNA lesions generated by genotoxic compounds during G0 phase, which are insufficiently repaired prior to S-phase. On the other hand, scarce reports suggest increased DNA repair capacity associated with variant G allele of *XPD* gene (49,50) or are inconclusive (22). Unless the function of variant G allele in

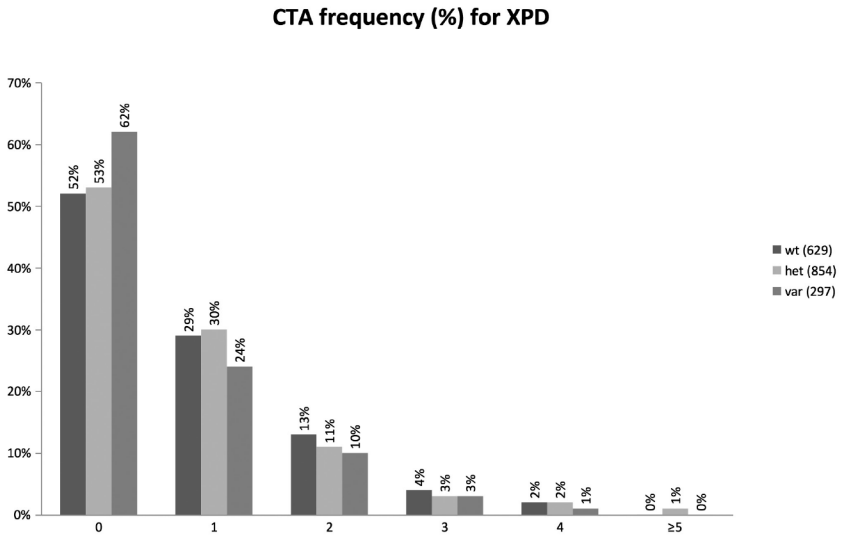


Figure 2. The frequency of individuals by CTA percent in relation to the *XPD* genotype. *x*-axis represents the frequency of individuals by CTA percent in relation to *XPD* genotype (TT, black bar; TG, light grey and GG, dark grey). *y*-axis represents the percentage of subjects with corresponding CTA frequency per particular genotype.

Table 4. Pair-wise interactions of genotypes with cases and controls

					Interaction term analysis			Global null hypothesis test for significance of all covariates		
	SNP 1	Mode for SNP 1	SNP 2	Mode for SNP 2	df	χ^2	P	df	χ^2	P
CAs	*APE1*	Three genotypes	*NBS1*	Allele number	2	11.80	0.003	9	67.45	<0.001
	hOGG1	Dominant	*XPG*	Allele number	1	7.58	0.005	7	65.78	<0.001
CSAs	*APE1*	Three genotypes	*XPD*	Recessive	2	9.78	0.008	9	94.40	<0.001
	XPD	Dominant	*XRCC3*	Allele number	1	8.33	0.004	7	31.31	<0.001
	XPG	Recessive	*NBS1*	Three genotypes	2	11.04	0.004	9	47.20	<0.001
	XRCC1	Allele number	*XPG*	Recessive	1	6.04	0.01	7	30.84	<0.001
	hOGG1	Allele number	*XPC*	Three genotypes	2	8.85	0.01	9	35.24	<0.001
	hOGG1	Dominant	*XPG*	Dominant	1	9.93	0.002	7	31.41	<0.001
CTAs	*XPG*	Allele number	*XRCC2*	Dominant	1	6.81	0.009	7	20.90	0.004
	XRCC3	Allele number	*XRCC2*	Dominant	1	6.71	0.01	7	21.06	0.003

Only the most significant genetic models for each interaction are shown ($P < 0.01$ in interaction term analysis). To assess the contribution of all genetic components (both SNPs and interaction term) to the model, likelihood ratio-based *P*-values were computed. df, degrees of freedom.
P-value was considered significant if it was <0.05.

XPD gene is clarified, the assumptions that higher level of DNA damage blocks replication fork and CTAs cannot be propagated remain speculative. An association of variant T allele in RAD54L with increased frequency of CSAs is a novel observation. *RAD54L* exhibits a DNA-dependent ATPase and supercoiling activities and plays a role in homologous recombination related repair of DSBs (51,52). However, this association is less robust due to the number of subjects with available genotype. Individually, a small risk is irrelevant, but the combination of several low-risk alleles can add up to substantial risks, even in the absence of multiplicative statistical interactions (53).

Similarly as in our recent study (35), we have addressed pair-wise interactions of the genotypes of each of the 11 DNA repair genes, which were tested for association with CA frequencies. We have discovered several highly significant interactions, involving gene variants from BER, NER and DSB pathways: 14 interactions modulated CA, 9 CTA and 12 CSA frequencies. For CAs, gene variants participating in BER (*APE1*, *hOGG1*), NER (*XPC*, *XPD*) and DSB repair (*XRCC3*) were mainly involved in significant interactions with other DNA repair gene variants (*NBS1*, *XRCC2* and *XPG1*). Interestingly, *NBS1* gene variant appears most often in these interactions. However, in interaction with their BER variants the CA frequency increases, whereas the interaction with NER gene variants shows the opposite. *NBS1* plays a relevant role in the maintenance of genomic integrity by being involved in the cellular response to DNA damage. The contradictory effect on CA frequencies in the interplay of *NBS1* variants with either BER or NER polymorphisms is certainly interesting and may reflect the specificity of these two excision repair pathways towards different kinds of DNA damage. For CTAs, the homozygous variant combinations in *XPD* with *XPG* or *XRCC1* decrease the frequencies. So does the combination of variant alleles in *OGG1* and *XRCC3*. On the contrary, a combination of variant alleles in DSB repair genes results in the significant increase. These results point again to an effect of the G allele of *XPD* Lys751Gln on CTA frequency modulation as stated above. For CSAs, variant alleles in BER genes *XRCC1* and *OGG1* in combination with homozygous variant genotype in *XPC* or *XPG* result

in significantly decreased frequencies. Interestingly, variant G allele in rs1052133 of *OGG1* in combination with variant alleles in genes involved in NER or DSB repair resulted in decreased frequencies of CAs, CTAs and CSAs, despite the fact that variant G allele is associated with the lower capacity to repair oxidative DNA damage (24). This phenomenon may be connected with the fact that 8-hydroxy-deoxyguanine adducts may block replication fork, thus preventing accumulation of CAs.

CAs arise as a consequence of the interaction between occupational exposure to various genotoxicants. In this study, we tested the impact of functional gene variants in DNA repair pathways on the frequency of CAs. Although individuals with homozygous variant genotype GG for *XPD* gene showed an association with CTAs, several gene–gene combinations in DNA repair genes evinced either enhanced or decreased frequencies of CAs, CTAs and CSAs. As suggested by Melis *et al.* (54) and now confirmed by us, the complex mechanism of CAs accumulation requires complex interplay between different DNA repair pathways. However, the mechanism may not be tracked without the knowledge of the experimentally proven functional impact of DNA repair gene variants.

Supplementary material

Supplementary Material can be found at http://carcin.oxfordjournals.org/

Funding

Projects 15-14789S (Czech Science Foundation), P27/2012 (Charles University of Prague, Czech Republic), CZ.1.05/2.1.00/03.0076, 26220220111, co-financed by EU sources and the European Regional Development Fund, 2012/25-UKMA-2 (Ministry of Health of the Slovak Republic) and APVV-0412-11.
Conflict of Interest Statement: None declared.

References

1. Hakem, R. (2008) DNA-damage repair; the good, the bad, and the ugly. EMBO J., 27, 589–605.
2. Bignold, L.P. (2009) Mechanisms of clastogen-induced chromosomal aberrations: a critical review and description of a model based on failures of tethering of DNA strand ends to strand-breaking enzymes. Mutat. Res., 681, 271–298.
3. Natarajan, A.T. et al. (2008) DNA repair and chromosomal alterations. Mutat. Res., 657, 3–7.
4. Durante, M. et al. (2013) From DNA damage to chromosome aberrations: joining the break. Mutat. Res., 756, 5–13.
5. Mateuca, R.A. et al. (2012) Cytogenetic methods in human biomonitoring: principles and uses. Methods Mol. Biol., 817, 305–334.
6. Dusinská, M. et al. (2004) Does occupational exposure to mineral fibres cause DNA or chromosome damage? Mutat. Res., 553, 103–110.
7. Dusinská, M. et al. (2004) Genotoxic effects of asbestos in humans. Mutat. Res., 553, 91–102.
8. Kazimírová, A. et al. (2009) Micronuclei and chromosomal aberrations, important markers of ageing: possible association with XPC and XPD polymorphisms. Mutat. Res., 661, 35–40.
9. Musak, L. et al. (2013) Chromosomal damage among medical staff occupationally exposed to volatile anesthetics, antineoplastic drugs, and formaldehyde. Scand. J. Work Environ. Health, 39, 618–630.
10. Vodenkova, S. et al. (2015) Structural chromosomal aberrations as potential risk markers in incident cancer patients. Mutagenesis, in 30, 557–563.
11. Futreal, P.A. et al. (2004) A census of human cancer genes. Nat. Rev. Cancer, 4, 177–183.
12. Rajagopalan, H. et al. (2004) Aneuploidy and cancer. Nature, 432, 338–341.
13. Burrell, R.A. et al. (2013) Replication stress links structural and numerical cancer chromosomal instability. Nature, 494, 492–496.
14. Mitelman, F. et al. (2007) The impact of translocations and gene fusions on cancer causation. Nat. Rev. Cancer, 7, 233–245.
15. Rossner, P. et al. (2005) Chromosomal aberrations in lymphocytes of healthy subjects and risk of cancer. Environ. Health Perspect., 113, 517–520.
16. Boffetta, P. et al. (2007) Chromosomal aberrations and cancer risk: results of a cohort study from Central Europe. Am. J. Epidemiol., 165, 36–43.
17. Rossi, A.M. et al. (2009) Association between frequency of chromosomal aberrations and cancer risk is not influenced by genetic polymorphisms in GSTM1 and GSTT1. Environ. Health Perspect., 117, 203–208.
18. Bonassi, S. et al. (2008) Chromosomal aberration frequency in lymphocytes predicts the risk of cancer: results from a pooled cohort study of 22 358 subjects in 11 countries. Carcinogenesis, 29, 1178–1183.
19. Vodicka, P. et al. (2010) Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls. Carcinogenesis, 31, 1238–1241.
20. Naccarati, A. et al. (2007) Sporadic colorectal cancer and individual susceptibility: a review of the association studies investigating the role of DNA repair genetic polymorphisms. Mutat. Res., 635, 118–145.
21. Vineis, P. et al. (2009) A field synopsis on low-penetrance variants in DNA repair genes and cancer susceptibility. J. Natl. Cancer Inst., 101, 24–36.
22. Abdel-Rahman, S.Z. et al. (2011) Evaluating the effects of genetic variants of DNA repair genes using cytogenetic mutagen sensitivity approaches. Biomarkers, 16, 393–404.
23. Vodicka, P. et al. (2007) Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects. Carcinogenesis, 28, 657–664.
24. Simonelli, V. et al. (2012) Gene susceptibility to oxidative damage: from single nucleotide polymorphisms to function. Mutat. Res., 731, 1–13.
25. Wu, L.L. et al. (2004) Urinary 8-OHdG: a marker of oxidative stress to DNA and a risk factor for cancer, atherosclerosis and diabetics. Clin. Chim. Acta, 339, 1–9.
26. Loft, S. et al. (2012) Urinary excretion of 8-oxo-7,8-dihydroguanine as biomarker of oxidative damage to DNA. Arch. Biochem. Biophys., 518, 142–150.
27. Friedberg, E.C. (2003) DNA damage and repair. Nature, 421, 436–440.
28. Hoelzl, C. et al. (2009) Use of single cell gel electrophoresis assays for the detection of DNA-protective effects of dietary factors in humans: recent results and trends. Mutat. Res., 681, 68–79.
29. Dolara, P. et al. (2012) Antioxidant vitamins and mineral supplementation, life span expansion and cancer incidence: a critical commentary. Eur. J. Nutr., 51, 769–781.
30. Vodicka, P. et al. (2004) Genetic polymorphisms in DNA repair genes and possible links with DNA repair rates, chromosomal aberrations and single-strand breaks in DNA. Carcinogenesis, 25, 757–763.
31. Naccarati, A. et al. (2006) Genetic polymorphisms and possible gene-gene interactions in metabolic and DNA repair genes: effects on DNA damage. Mutat. Res., 593, 22–31.
32. Musak, L. et al. (2008) Chromosomal aberrations in tire plant workers and interaction with polymorphisms of biotransformation and DNA repair genes. Mutat. Res., 641, 36–42.
33. Skjelbred, C.F. et al. (2011) Influence of GSTM1, GSTT1, GSTP1, NAT1, NAT2, EPHX1, MTR and MTHFR polymorphism on chromosomal aberration frequencies in human lymphocytes. Carcinogenesis, 32, 399–405.
34. Dusinska, M. et al. (2012) Are glutathione S transferases involved in DNA damage signalling? Interactions with DNA damage and repair revealed from molecular epidemiology studies. Mutat. Res., 736, 130–137.
35. Hemminki, K. et al. (2015) Metabolic gene variants associated with chromosomal aberrations in healthy humans. Genes Chromosomes Cancer, 54, 260–266.
36. Hemminki, K. et al. (2014) Cyclin D1 splice site variant triggers chromosomal aberrations in healthy humans. Leukemia, 28, 721–722.

37. Jirawatnotai, S. et al. (2011) A function for cyclin D1 in DNA repair uncovered by protein interactome analyses in human cancers. Nature, 474, 230–234.
38. Kazimirova, A. et al. (2007) Cytogenetic analysis of lymphocytes of workers occupationally exposed to rockwool and glass fibres. Molecular-epidemiologic study. Chem. Listy, 101, 192–193.
39. Srám, R.J. et al. (2004) Cytogenetic analysis and occupational health in the Czech Republic. Mutat. Res., 566, 21–48.
40. Pardini, B. et al. (2008) DNA repair genetic polymorphisms and risk of colorectal cancer in the Czech Republic. Mutat. Res., 638, 146–153.
41. Shen, Z. (2011) Genomic instability and cancer: an introduction. J. Mol. Cell Biol., 3, 1–3.
42. Abbas, T. et al. (2013) Genomic instability in cancer. Cold Spring Harb. Perspect. Biol., 5, a012914.
43. Vodicka, P. et al. (2004) Markers of individual susceptibility and DNA repair rate in workers exposed to xenobiotics in a tire plant. Environ. Mol. Mutagen., 44, 283–292.
44. Vodicka, P. et al. (2006) Styrene metabolism, genotoxicity, and potential carcinogenicity. Drug Metab. Rev., 38, 805–853.
45. Halasova, E. et al. (2012) Evaluating chromosomal damage in workers exposed to hexavalent chromium and the modulating role of polymorphisms of DNA repair genes. Int. Arch. Occup. Environ. Health, 85, 473–481.
46. Curtin, N.J. (2012) DNA repair dysregulation from cancer driver to therapeutic target. Nat. Rev. Cancer, 12, 801–817.
47. Abbotts, R. et al. (2014) DNA repair in cancer: emerging targets for personalized therapy. Cancer Manag. Res., 6, 77–92.
48. Spies, M. (2014) Two steps forward, one step back: determining XPD helicase mechanism by single-molecule fluorescence and high-resolution optical tweezers. DNA Repair (Amst.), 20, 58–70.
49. Wlodarczyk, M. et al. (2012) XPD gene rs13181 polymorphism and DNA damage in human lymphocytes. Biochem. Genet., 50, 860–870.
50. Gdowicz-Klosok, A. et al. (2013) The influence of XPD, APE1, XRCC1, and NBS1 polymorphic variants on DNA repair in cells exposed to X-rays. Mutat. Res., 755, 42–48.
51. Mjelle, R. et al. (2015) Cell cycle regulation of human DNA repair and chromatin remodeling genes. DNA Repair (Amst.), 30, 53–67.
52. Eppink, B. et al. (2011) The response of mammalian cells to UV-light reveals Rad54-dependent and independent pathways of homologous recombination. DNA Repair (Amst.), 10, 1095–1105.
53. Vineis, P. et al. (2008) Expectations and challenges stemming from genome-wide association studies. Mutagenesis, 23, 439–444.
54. Melis, J.P. et al. (2013) Oxidative DNA damage and nucleotide excision repair. Antioxid. Redox Signal., 18, 2409–2419.

Manuscript III

Genetic variation associated with chromosomal aberration frequency: A genome-wide association study.

Niazi Y, Thomsen H, Smolkova B, Vodickova L, Vodenkova S, **Kroupa M**, Vymetalkova V, Kazimirova A, Barancokova M, Volkovova K, Staruchova M, Hoffmann P, Nöthen MM, Dušinská M, Musak L, Vodicka P, Hemminki K, Försti A.

Environ Mol Mutagen. 2019 Jan;60(1):17-28. **IF: 3.254**

Environmental and Molecular Mutagenesis 60:17–28 (2019)

Research Article

Genetic Variation Associated with Chromosomal Aberration Frequency: A Genome-wide Association Study

Yasmeen Niazi,[1,14†*] Hauke Thomsen,[1†] Bozena Smolkova,[2] Ludmila Vodickova,[3,4,5] Sona Vodenkova,[3,4,6] Michal Kroupa,[3,5] Veronika Vymetalkova,[3,4] Alena Kazimirova,[7] Magdalena Barancokova,[7] Katarina Volkovova,[7] Marta Staruchova,[7] Per Hoffmann,[8,9] Markus M. Nöthen,[8,10] Maria Dušinská,[11] Ludovit Musak,[12] Pavel Vodicka,[3,4,5] Kari Hemminki[1,13] and Asta Försti[1,13]

[1]*Department of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120, Heidelberg, Germany*
[2]*Department of Molecular Oncology, Cancer Research Institute, Biomedical Research Center, Slovak Academy of Sciences, Dubravska cesta 9, 84505, Bratislava, Slovakia*
[3]*Department of Molecular Biology of Cancer, Institute of Experimental Medicine, The Czech Academy of Sciences, Videnska 1083, 142 00, Prague, Czech Republic*
[4]*Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Albertov 4, 128 00, Prague, Czech Republic*
[5]*Biomedical Centre, Faculty of Medicine in Pilsen, Charles University in Prague, Pilsen, Czech Republic*
[6]*Department of Medical Genetics, Third Faculty of Medicine, Charles University, Prague, Czech Republic*
[7]*Department of Biology, Faculty of Medicine, Slovak Medical University, Limbova 12, 833 03, Bratislava, Slovakia*
[8]*Institute of Human Genetics, University of Bonn, D-53127, Bonn, Germany*
[9]*Division of Medical Genetics, Department of Biomedicine, University of Basel, 4003, Basel, Switzerland*
[10]*Department of Genomics, Life & Brain Center, University of Bonn, D-53127, Bonn, Germany*
[11]*Health Effects Laboratory, Department of Environmental Chemistry, NILU-Norwegian Institute for Air Research, Instituttveien 18, 2007, Kjeller, Norway*
[12]*Clinic of Occupational Medicine and Toxicology, Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava and University Hospital Martin, Kollarova 2, 03601, Martin, Slovakia*
[13]*Center of Primary Health Care Research, Clinical Research Center, Lund University, 20502, Malmö, Sweden*
[14]*Medizinische Fakultät, Universität Heidelberg, Im Neuenheimer Feld 672, 69120, Heidelberg*

Chromosomal aberrations (CAs) in human peripheral blood lymphocytes (PBL) measured with the conventional cytogenetic assay have been used for human biomonitoring of genotoxic exposure for decades. CA frequency in peripheral blood is a marker of cancer susceptibility. Previous studies have shown associations between genetic variants in metabolic pathway, DNA repair and major mitotic checkpoint genes and CAs. We conducted a genome-wide association study on 576 individuals from the Czech Republic and Slovakia followed by a replication in two different sample sets of 482 (replication 1) and 1288 (replication 2) samples. To have a broad look at the genetic susceptibility associated with CA frequency, the sample sets composed of individuals either

Grant sponsor: Ministry of Health; Grant numbers: 2005/42-SZU-20; 2006/07-SZU-02; 2005/43-SZU-21.
Grant sponsor: Slovak Grant Agency; Grant numbers: APVT-21-017704; APVT-21 013202.
Grant sponsor: EC; Grant numbers: CIPA-CT94-0129; ERBICI 15-CT96-1012; QLK4-CT-1999-01629.
Grant sponsor: Charles University in Prague; Grant numbers: UNCE/MED/006; UNCE 204022; PROGRES Q 28; Nr.LO 1503.
Grant sponsor: National Science Foundation; Grant numbers: 18-09709S; 15-14789S.

*Correspondence to: Yasmeen Niazi, Department of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 580, 69120 Heidelberg, Germany. E-mail: y.niazi@dkfz.de
†These authors contributed equally.

Received 14 March 2018; provisionally accepted 18 May 2018; and in final form 3 July 2018

DOI: 10.1002/em.22236

Published online 3 October 2018 in Wiley Online Library (wileyonlinelibrary.com).

differentially exposed to smoking, occupational/environmental hazards, or they were untreated cancer patients. Phenotypes were divided into chromosome- and chromatid-type aberrations (CSAs and CTAs, respectively) and total chromosomal aberrations (CAtot). The arbitrary cutoff point between individuals with high and low CA frequency was 2% for CAtot and 1% for CSA and CTA. The data were analyzed using age, sex, occupation/cancer and smoking history as covariates. Altogether 11 loci reached the *P*-value of 10^{-5} in the GWAS. Replication 1 supported the association of rs1383997 (8q13.3) and rs2824215 (21q21.1) in CAtot and rs983889 (5p15.1) in CTA analysis. These loci were found to be associated with genes involved in mitosis, response to environmental and chemical factors and genes involved in syndromes linked to chromosomal abnormalities. Identification of new genetic variants for the frequency of CAs offers prediction tools for cancer risk in future. Environ. Mol. Mutagen. 60:17–28, 2019.

Key words: chromosome-type aberrations; chromatid-type aberrations; GWAS; single-nucleotide polymorphism

INTRODUCTION

Chromosomal aberrations (CAs) encompass structural and numerical chromosomal anomalies. Structural CAs include specific, recurrent deletions, translocations, and inversions that can only be detected by molecular cytogenetics such as fluorescent *in situ* hybridization and sequencing techniques (Albertini et al. 2000). CAs that can be cytologically distinguished at metaphase are non-specific and they can be divided into two main groups: chromosome-type aberrations (CSAs) and chromatid-type aberrations (CTAs) (Hagmar et al. 2001; Bignold 2009; Hemminki et al. 2015b; Heng et al. 2016). CSAs arise mainly as a result of direct DNA damage during G_o/G_1 phase, by clastogens such as ionizing radiation and bleomycin. The lesions acquired during G_o/G_1 phase later result in the damage of both chromatids of a chromosome which may create di-centric and ring chromosomes (Albertini et al. 2000). Apart from direct DNA damage, another important contributor to CSA frequency is telomere dysfunction. In somatic cells with critically short telomeres and low telomerase activity, telomeres shorten, become eroded and poorly end-capped. These eroded ends can be recognized by nonhomologous end joining repair and become attached to non-homologous chromosomes resulting in fused, ring or fragmented chromosomes (Gostissa et al. 2011; Jones et al. 2012). Association between relative telomere length and CA frequency, particularly that of CSAs, has been documented in our previous study (Hemminki et al. 2015a). CTAs are a result of damage by environmental or chemical clastogens during S/G_2 phase or due to replication on a damaged DNA template and involve only one chromatid of a chromosome. Examples of CTAs are chromatid breaks and exchanges (Durante et al. 2013).

Conventional cytogenetic examination for CAs in individuals exposed to mutagens and potential carcinogens has been used for decades as a surveillance mechanism for genotoxic effect (Carrano and Natarajan 1988). Many malignant and benign human tumors exhibit chromosomal abnormalities (Mitelman 2000) and an increase in the frequency of CAs has been found in the incident cancer patients thus closely linking CAs with cancer development (Vodenkova et al. 2015). Some of the CAs observed are also generated during the course of cancer development, nevertheless CA frequency in peripheral blood lymphocytes (PBLs) is considered to be an early marker of cancer susceptibility based on the hypothesis that genetic damage in PBLs reflects similar damage in other body cells undergoing carcinogenesis (Rossner et al. 2005).

Interindividual variation in the frequency of CAs, both in unexposed and in exposed individuals, has raised the question of genetic predisposition to CAs. Studies exploring the genetic causes of increased CA frequency have mainly focused on mitotic checkpoint, DNA repair, and metabolic genes and found different variants that are associated with the frequency of CAs (Hemminki et al. 2015b; Vodicka et al. 2015, 2018; Försti et al. 2016). Despite these findings, there still is a great need to explore the genetic basis of CAs. To achieve this goal, we designed a genome-wide association study (GWAS), which is the first GWAS of this nature. The study included not only individuals exposed to potential occupational and environmental carcinogens but also newly diagnosed cancer patients, who may represent a population with increased susceptibility to CAs, and individuals with no recorded exposure to carcinogens. Our aim was to find novel genetic variants predisposing to CAs and potentially to cancer and to elucidate the possible functional effects of these variants by *in silico* predictions.

MATERIALS AND METHODS

All the samples and the information in the study were obtained with written consent of the participants. The project was carried out according to the rules of the Declaration of Helsinki and ethical approval was obtained from the Ethics Committee of the Institute of Preventive and Clinical Medicine (later Slovak Medical University), the Ethics Committee of the Jessenius Faculty of Medicine in Martin, Comenius University in Bratislava, the Ethics Committee of the Institute for Clinical and Experimental Medicine and Thomayer Hospital, Czech Republic and the Ethics Committee of the VFN (General University Hospital in Prague).

Study Subjects

The GWAS sample set consisted of 639 healthy individuals; approximately 89% of them were recruited in Slovakia, while 11% came from the Czech Republic. Blood samples were taken from approximately equal

number of males and females. About 57% of these individuals were exposed to some form of genotoxins due to the nature of their professions and their exposure was assessed by personal dosimeters, the rest were office workers and local residents; 30% were self-reported smokers (Table I). The genotoxic substances included mainly small organic compounds, such as vinyl chloride, epichlorohydrine and ethylene oxide, anesthetics, heavy metals and styrene (Vodicka et al. 2004a, b, 2015, Musak et al. 2008, 2013). The replication was conducted on two different sample sets. The first replication set (replication 1) consisted of 482 individuals (Czech, n = 449, Slovak, n = 33) and the second set (replication 2) was composed of 1288 individuals (all Slovak), (Table I). About 46% of the individuals in replication 1 were newly diagnosed primary cancer patients recruited by the Department of Radiotherapy and Oncology, Faculty Hospital Kralovske Vinohrady, Prague, Czech Republic, including breast, colorectal and lung cancer patients; 29% of the individuals in this sample set were self-reported smokers (Vodicka et al. 2010; Vodenkova et al. 2015). Blood samples from the cancer patients were drawn before any treatment to avoid any treatment-related increase in the number of CAs. The healthy individuals were recruited by the Blood Center of Faculty Hospital Kralovske Vinohrady, Prague, Czech Republic.

Replication 2 included participants from six molecular-epidemiological studies. In these studies, the effects of environmental/lifestyle factors including smoking, alcohol consumption, nutrition, professional exposure (asbestos, stone wool, glass fibers, and radiation), and risk factors such as obesity and aging on health outcomes were assessed by several biomarkers, including chromosomal instability. Altogether 23% of individuals were occupationally exposed with exposure assessed by personal dosimeters and 24% were self-reported smokers. The numbers of participants from individual studies are shown in Table I; details of each study were published elsewhere (Dušinská et al. 2003, 2004a, b, 2012, Kažimírová et al. 2004, 2006, 2009; Tulinska et al. 2004; Szabová et al. 2012).

Cytogenetic Assay

Cytogenetic analysis was performed on cultured PBLs. Two short-term PBL cultures were set up for each sample. For this purpose, 0.5 ml of whole blood was added to 4.5 ml of RPMI (Roswell Park Memorial Institute) medium along with L-glutamine and $NaHCO_3$ (Gibco) supplemented with 20% fetal calf serum (Gibco), and antibiotics (penicillin and streptomycin, Gibco). Phytohaemagglutinin (0.18 mg/ml, PHA, Murex) was added as a proliferation stimulant. Incubation was done at 37 °C with 5% CO_2 for 48 hours. Colchicine (0.75 µg/ml, Sigma) was added 2 hours before harvesting the PBLs. Harvesting was followed by centrifugation of cells and hypotonic shock treatment in 0.075 M KCl for 20 min at 37 °C. PBLs were fixed twice in methanol:glacial acetic acid (3:1) and air-dried preparations were made. Staining of slides was carried out using 5% Giemsa-Romanowski solution for 5 min. Cultured PBLs were analyzed in metaphase stage under a light microscope (Dušinská et al. 2004b; Kažimírová et al. 2004; Musak et al. 2013). For each person, 100 mitoses were analyzed in a double-blind fashion and the frequency of different types of CAs (CSAs, CTAs) was recorded (Vodicka et al. 2010; Musak et al. 2013).

Phenotypes analyzed in the association analysis were divided into three categories: CAtot (total chromosomal aberrations), CSAs, and CTAs. For logistic regression analysis, the samples were divided into high CA frequency group (CA_{high}) and low CA frequency group (CA_{low}) on the basis of frequency of aberrations. The threshold for inclusion into CA_{high} in case of CAtot was CA frequency ≥ 2% while for CSA and CTA it was ≥1%. This arbitrary assignment to CA_{high} and CA_{low} groups is based on previous experience with human genotoxic monitoring in the Czech and Slovak populations [Dušinská *et al.* 2003, 2004a, b; Šrám *et al.* 2004; Vodicka *et al.* 2010; Musak *et al.* 2013].

GWAS and Quality Control

Genotyping of the 639 individuals was done using Illumina HumanOmniExpressExome8v1.3 array comprising nearly 1 million SNPs throughout the genome. General genotyping quality control assessment was done as previously described by Anderson et al. (2010). Individuals with discordant gender information, outlying heterozygosity and genotype call rates <95% were excluded. Relatedness between samples was detected by identity-by-state measures. Population stratification was assessed using principal component analysis. SNPs with one or more of the following criteria were excluded: <95% genotype call rate, minor allele frequency < 5% or Hardy–Weinberg equilibrium exact P-value $<10^{-5}$. After quality control, 576 samples and 626.004 SNPs remained. Genotypes for common variants across the genome were then imputed using data from the combined UK10K - 1000 Genomes Project (phase 3, Oct. 2014) with IMPUTE2 v2.3.2 (Howie et al. 2011) after pre-phasing with SHAPEIT software v2.12 (Delaneau et al. 2011). We set thresholds for imputation quality to retain both potential common and rare variants for validation. Specifically, poorly imputed SNPs defined by an information metric I < 0.70 were excluded. All genomic locations are given in NCBI Build 37/UCSC hg19 coordinates. All SNPs having a MAF < 5% were excluded. After imputation, the SNP set consisted of 10.258.281 genotyped and imputed SNPs. This SNP set consisting of both genotyped and imputed SNPs was used for association analysis.

Association Analysis

The consecutive association analysis of the GWAS was conducted by SNPTEST using univariate and multivariate logistic regression and linear regression models for each of the three phenotypes by including the relevant covariates, age, sex, occupational exposure and smoking status. The association data were visualized through Miami plots with generally accepted suggestive significance threshold of $P = 5.0 \times 10^{-5}$ and the genome-wide significance threshold of $P = 5.0 \times 10^{-8,}$ using the Genetic analysis package (gap) for CRAN R 2.15 and odds ratios (ORs), effect sizes and 95% confidence intervals (CIs) were obtained for the effective allele in the additive model.

In silico Analysis

In silico analysis was done using different bioinformatics tools to examine functional consequences of the highly associated SNPs. These tools included Locus zoom to plot the locus of interest, to see the orientation of genes in the region, linkage disequilibrium between the SNPs and recombination rate, (Pruim et al. 2010), UCSC genome browser (Rosenbloom et al. 2015) and Haploreg (Ward and Kellis 2012) to investigate the presence of any regulatory elements like promotors, enhancers and transcription factor binding sites, and to see the potential functions and expression effects from eQTL studies of all highly linked SNPs on candidate target genes. Regulome DB was used to predict the likely cell types of action, variant scores, regions of DNase hypersensitivity, and histone modifications (Boyle et al. 2012). A total of 11 SNPs were selected for replication as a result of *in silico* analysis.

Validation and Replication

The selected SNPs were validated in a small sample set of 149 individuals from the GWAS and replicated in 2 different replication sets. Validation and replication were carried out through TaqMan (Thermo Fisher Scientific, Darmstadt, Germany) allelic discrimination genotyping assays. Genotype detection was performed using Applied Biosystems ViiA™ 7 Real-Time PCR System, (Life technologies, Germany).

TABLE I. Characterization of the Study Population Including Distribution of High (CA_{high}) and Low Chromosomal Aberration level (CA_{low}) Among CAtot, CSA and CTA Categories and According to Major Confounders, Age, Sex, Smoking, Occupational Exposure and Cancer Status.

	GWAS	P^{12}	Replication 1	P^{12}	Replication 2	P^{12}
Mean age ±SD	44.64 ±12.70	0.01	59.63 ± 12.88	0.014	43.8 ± 15.57	2.59E-07
Female/Male %	51.7/48.3	0.57	53.6/46.4	0.13	57.7/42.3	0.02
Smoking Yes/No %	30.5/69.5	0.96	29.3/70.7	0.001	24.1/75.9	0.006
CAtot[1] no. (CA_{high}[2]/CA_{low}[3])	351/225[4]		304/178		295/993	
CTA[5] no. (CA_{high}/CA_{low})	367/209[4]		350/132		387/901	
CSA[6] no. (CA_{high}/CA_{low})	349/227[4]		268/214		367/921	
Occupational exposure %	57.6	1.21E-09			23.1	0.009
Small organic compounds%	21.7					
Anesthetics%	15.5					
Heavy metals%	12					
Styrene%	8.5					
Radiation (pilots)%					5.8	
Asbestos%					4.6	
Stone wool%					7.1	
Glass fibers%					5.8	
Cancer %			46.1	7.56E-06		
Breast cancer%			23.4			
Colorectal cancer%			14.3			
Lung cancer%			8.3			
Others	42.4[7]		53.9[8]		10.7[7]/18.8[9]/20.5[10]/26.9[11]	

[1](Total chromosomal aberrations)
[2](No. of individuals in high chromosomal aberrations group)
[3](No. of individuals in low chromosomal aberrations group)
[4](No. of individuals after the quality control)
[5](Chromatid type aberrations)
[6](Chromosome type aberrations)
[7](Office workers and local residents)
[8](Blood donors)
[9](Obesity study)
[10](Aging study)
[11](Vegetarians and nutrition studies)
[12](P values based on binary regression model exhibiting the modulation in frequency of total chromosomal aberrations by major confounders age, sex, smoking and occupational exposure).

Statistical Analysis

Post replication analysis was performed using PLINK v1.90b3.30 (Purcell et al. 2007) (http://pngu.mgh.harvard.edu/purcell/plink/). Effect sizes, 95% CIs and corresponding *P*-values were calculated by using logistic and linear regression models. All models were corrected for the same covariates as with the model above; in the analysis of replication 1, cancer status was included to the covariates. A meta-analysis for the GWAS and the two replication sets was performed using the GWAMA software (Mägi and Morris 2010). Heterogeneity was assessed by the I^2 statistics (interpreted as low <0.25, moderate 0.50 and high >0.75).

RESULTS

The number of subjects in CA_{high} and CA_{low} for CAtot, CTA and CSA, and their distribution among the covariates (age, sex, smoking, occupational exposure, and cancer status) for the three study sample sets (GWAS, replication 1 and replication 2) is summarized in Table I. These covariates were chosen because they were proven to exert a significant effect on CA frequency in previous studies (Vodicka et al. 2010; Hemminki et al. 2015b; Vodenkova et al. 2015). Association of background variables of sex, age, smoking, occupational exposure, and cancer status with CA frequency was tested with logistic regression model on CAtot. According to this analysis, occupational exposure significantly influenced CA frequency in the GWAS ($P = 1.21 \times 10^{-9}$). In replication 1, the most significant variable affecting CA frequency was cancer status ($P = 7.56 \times 10^{-6}$), while in replication 2, the effect of occupational exposure was moderate ($P = 0.009$). The effect of age was moderate in the GWAS ($P = 0.01$) and replication 1 ($P = 0.01$) sample sets but significant in the replication 2 ($P = 2.59 \times 10^{-7}$). Smoking history had a significant association with CA frequency in the replication 1 ($P = 0.001$) and replication 2 ($P = 0.006$) but not in the GWAS ($P = 0.96$). Gender was not associated with CA frequency in the GWAS and replication 1 ($P = 0.57$ and 0.13, respectively), but it was moderately associated in replication 2 ($P = 0.02$). As shown in Table II, there were some differences in the median, mean, minimum, and maximum values of CAtot, CSA, and CTA among the three data sets. In the GWAS and replication 1, the mean frequency of

TABLE II. Distribution of CAtot, CSA and CTA in the Three Data Sets

	No. of Individuals[2]		Mean±SD[1]	Median	Range
GWAS	576	CAtot	1.94 ± 1.29	2	0–7
		CSA	0.98 ± 1.04	1	0–6
		CTA	0.97 ± 0.99	1	0–6
Replication 1	482	CAtot	2.21 ± 1.57	2	0–8
		CSA	0.88 ± 1.036	1	0–5
		CTA	1.32 ± 1.22	1	0–7
Replication 2	1288	CAtot	0.95 ± 1.35	0	0–11
		CSA	0.51 ± 1.02	0	0–11
		CTA	0.44 ± 0.78	0	0–6

[1]Mean number of aberrations per hundred cells
[2]Number of individuals in the study

CAtot was about 2% and it was about 1% for CSA and CTA, while in replication 2, the frequencies were about 1% and 0.5%, respectively. Linear mixed model was also used to test the association of covariates with CA frequency and the results were very similar to those from logistic model.

Both logistic and linear regression models were applied for analysis of the phenotypes CAtot, CTA, and CSA. Altogether 11 loci, 6 from the CAtot and 5 from the CTA analysis, were chosen for replication and the most significant SNPs with $P < 1x10^{-5}$ from these loci were selected on the basis of *in silico* analyses (Table III). Selected SNPs

TABLE III. *In silico* Predictions for the Selected Variants. Functional annotations from the ENCODE based tool Haploreg v4.1 (http://www.broadinstitute.org/mammals/haploreg/haploreg.php) and eQTL analysis according to GTEx Portal (http://www.gtexportal.org/home/)

Chromosome	SNP	No. of SNPs with r2 > 0.8	Promotor histon marks	Enhancer histone marks	Proteins bound	DNAse	Motifs changed	eQTL hits	Gencode genes	Type
2	rs17215792	7	BRST[1], BRN[2]	14 tissues (BLD[3])		ESDR[4]	HMG-IY, Irf, Nkx2		KLF7	3′-UTR
3	rs340828	2		4 tissues (BLD)			16 altered motifs	1 hit	IL5RA	3′-UTR
5	rs983889	27	5 tissues	10 tissues		5 tissues	TCF12, TCF4		FBXL7	Intronic
8	rs1383997	37	21 tissues (BLD)	21 tissues (BLD)	CTCF, SMC3, EGR1, TBP	12 tissues (BLD)	BDP1, CTCF, NF-I	2 hits	RP11-383H13.1	Intronic
9	rs12002628	17					Hoxa4, Hoxb8		TRPM3	5′ -UTR
9	rs7025089	53	ESDR, BLD, LNG[5]	21 tissues (BLD)	CTCF	9 tissues	10 altered motifs	2 hits	MED27	Intronic
9	rs16931167	1		LIV[6], HRT[7], PANC			5 altered motifs		PTPRD	Intronic
9	rs7033729	9	ESDR, GI[8]	6 tissues		PLCNT[9]	Arid5b, NRSF	4 hits	FAM154A	Intronic
14	rs8003642	26	SPLN[10]	11 tissues			Irf, Sox		RP11-725G5.2	Intronic
21	rs2824215	18							AF212831.2	5′ -UTR
21	rs2837619	13					ERalpha-a, Irf, p300		DSCAM	Intronic

[1](Breast)
[2](Brain)
[3](Blood)
[4](ESC_Derived)
[5](Lung)
[6](Liver)
[7](Heart)
[8](Gastrointestinal tract)
[9](Placenta)
[10](Spleen)

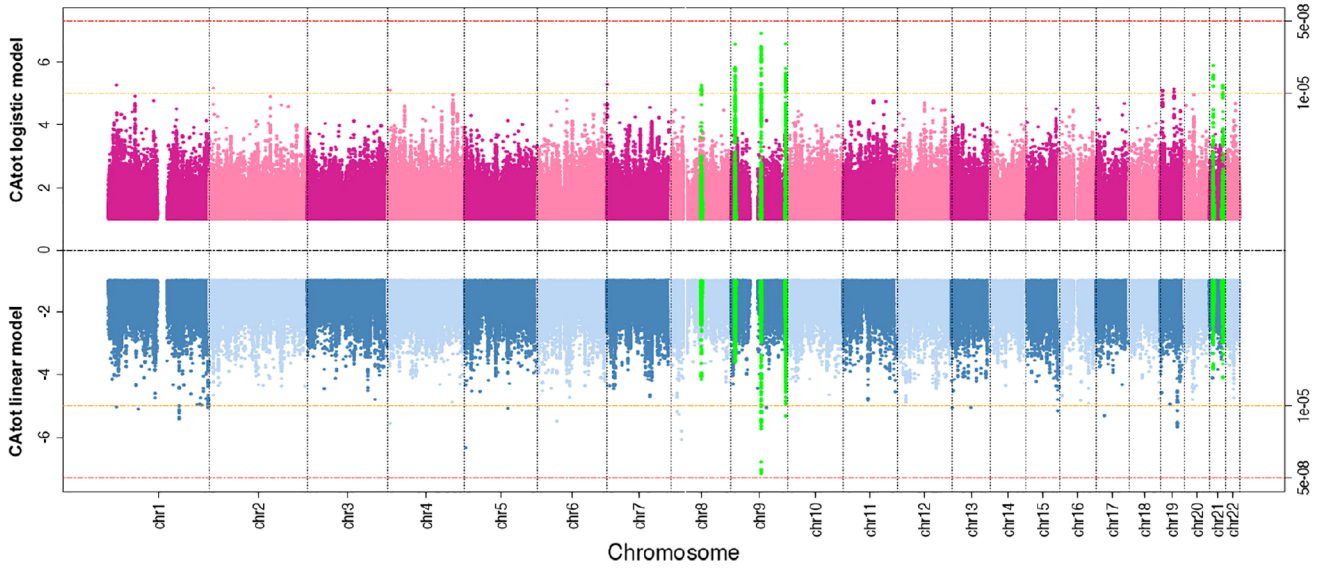


Fig. 1. Miami plot for CAtot logistic and linear models. The Y-axis shows the $-\log_{10} P$ – value of each SNP and the X-axis shows their chromosomal position. Loci selected for replication are highlighted in green. The red horizontal line represents the genome-wide significance threshold of $P = 5.0 \times 10^{-8}$ and the yellow horizontal line represents the significance threshold of $P = 5.0 \times 10^{-5}$ as suggestive level of significance.

were genotyped in the replication sets. All selected loci contained at least one directly genotyped SNP and the genotyping accuracy was confirmed in a small subset of GWAS samples. No SNP associations at the level of $P < 1 \times 10^{-5}$ were found in CSA.

In case of CAtot, the logistic regression model showed more significant associations than the linear model, however, almost all the loci showed similar trends in the linear model as well. This is evident from the Miami plot (Fig. 1). Two SNPs in chromosome 9 (rs12002628 and rs16931167) almost reached the genome-wide significance level of $5x10^{-8}$ ($P = 4.78 \times 10^{-7}$ and 2.66×10^{-7}, respectively, Table IV). For all SNPs, except for rs16931167, replication 1 showed ORs on the same direction as in the GWAS and the strongest associations in the meta-analysis were for rs1383997 at 8q13.3 (OR 0.6, 95%CI 0.49–0.73, $P = 3.44 \times 10^{-7}$) and rs2824215 at 21q21.1 (OR 1.57, 95%CI 1.29–1.91, $P = 8.7 \times 10^{-6}$). Replication 2 did not give much support for the GWAS associations, and the strongest association in the meta-analysis of all populations with $P = 4.01x10^{-5}$ was for rs12002628 at 9q21.13. The heterogeneity between the three populations was moderate to high as indicated by the I^2 values (Table IV).

For the CTA analysis, on the other hand, higher associations were found in the linear model as compared to the logistic model (Fig. 2). In CTA, five SNPs showed an association at the suggestive level of significance (Table V). Here also, the GWAS and replication 1 showed more similar associations than the GWAS and replication 2 and the heterogeneity measured by the I^2 values was high. In the meta-analysis of the GWAS and replication 1, one association, rs983889 at 5p15.1 remained statistically significant at the suggestive level ($P = 1.06 \times 10^{-5}$) and no significant associations were observed in the meta-analysis of all three populations. Although the SNPs from the GWAS were selected based on the linear model, we also calculated the ORs and 95%CIs in the logistic model (Table V). For the most significant SNP, rs983889, the OR was 0.65 (95%CI 0.52–0.80) in the meta-analysis of the GWAS and replication 1.

DISCUSSION

Inherited genetic variation may cause interindividual differences in the susceptibility to CAs. We tested this hypothesis at the whole-genome level in three populations composed of differentially exposed individuals through smoking, occupational or environmental factors, and untreated cancer patients. In general, occupational exposure and being a newly diagnosed cancer patient had a strong influence on the frequency of CAs, while the genetic factors seemed to play a minor role. The fact that the GWAS and replication 1 showed more similar associations with CAs than the GWAS and the replication 2 may be explained by the composition of these three sample sets.

In the GWAS set, more than 50% of the subjects were occupationally exposed to different compounds such as small organic compounds, anesthetics, styrene, and heavy metals (Somorovská et al. 1999; Vodicka et al. 2004a, b, 2015, Musak et al. 2008, 2013; Hemminki et al. 2015b; Försti et al. 2016). Logistic regression analysis on this sample set showed a highly significant association of increasing

TABLE IV. SNPs Selected from CAtot Model and Their Corresponding OR, 95% CI, and *P*-values in the Logistic Model in Three Sample Sets and Meta-analyses

	Logistic Model								
	SNP	Locus	Minor allele	Major allele	Model	OR	95% CI	P	I^2
GWAS	rs1383997	8q13.3	T	C	CAtot	0.56	0.44–0.71	5.44E-06	
Replication 1						0.67	0.49–0.91	0.01	
Replication 2						1.09	0.89–1.34	0.39	
Meta GWAS+ Replication 1[1]						0.6	0.49-0.73	3.44E-07	0
meta all[2]						0.8	0.70-0.93	0.002	0.89
GWAS	rs12002628	9q21.13	T	C	CAtot	0.47	0.35–0.63	4.78E-07	
Replication 1						0.93	0.64–1.33	0.67	
Replication 2						0.79	0.60–1.03	0.08	
Meta GWAS+ Replication 1						0.61	0.48–0.78	7.68E-05	0.89
meta all						0.68	0.57–0.82	4.01E-05	0.81
GWAS	rs7025089	9q34.13	C	A	CAtot	0.55	0.42–0.70	7.97E-06	
Replication 1						0.86	0.63–1.18	0.35	
Replication 2						1.19	0.95–1.47	0.13	
Meta GWAS+ Replication 1						0.66	0.53–0.81	1.02E-04	0.8
meta all						0.87	0.75–1.02	0.08	0.9
GWAS	rs16931167	9p23	T	C	CAtot	2.55	1.76–3.71	2.66E-07	
Replication 1						0.77	0.52–1.12	0.17	
Replication 2						0.96	0.73–1.28	0.8	
Meta GWAS+ Replication 1						1.41	1.07–1.86	0.02	0.95
meta all						1.17	0.96–1.42	0.13	0.92
GWAS	rs2824215	21q21.1	C	A	CAtot	1.8	1.40–2.30	3.27E-06	
Replication 1						1.27	0.94–1.7	0.12	
Replication 2						1.02	0.83–1.24	0.88	
Meta GWAS+ Replication 1						1.57	1.29–1.91	8.70E-06	0.73
meta all						1.26	1.10–1.45	0.001	0.84
GWAS	rs2837619	21q22.2	G	A	CAtot	0.61	0.47–0.77	5.51E-06	
Replication 1						0.94	0.71–1.24	0.64	
Replication 2						0.97	0.8–1.18	0.78	
Meta GWAS+ Replication 1						0.71	0.58–0.85	3.65E-04	0.86
meta all						0.83	0.72–0.95	0.01	0.84

[1](Meta-analysis between the GWAS and replication 1)
[2](Meta-analysis between all three sample sets i.e. GWAS sample set, replication 1 and replication 2)

CA frequency with occupation and a moderate association with age. In replication 1, in which 46% of individuals were incident cancer patients, a strong association between the frequency of CAs and cancer status was observed. These results are also in consistence with previous studies (Hemminki et al. 2015b; Vodicka et al. 2015). For replication 2, age was the most significantly associated factor with smoking and occupational exposure showing a moderate effect. In replication 2, the effect of occupational environment was less significant than in the GWAS sample, probably because the proportion of individuals who were occupationally exposed was only 23%. The proportion of individuals exposed to asbestos, which is equally genotoxic as the chemical compounds in the GWAS (IARC Working Group on the Evaluation of Carcinogenic Risks to Humans *et al.* 2008; IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. and International Agency for Research on Cancer. 2012) was less than 10% of the study group. The rest of the occupationally exposed individuals were exposed to stone wool and glass fibers which are comparatively less genotoxic (Dušinská et al. 2004a, b; Baan and Grosse 2004). Also, a significant part of the sample consisted of individuals from the aging, obesity, and specific food preference studies, who were nominally unexposed to genotoxic agents (Dušinská et al. 2003; Kažimírová et al. 2004, 2006, 2009; Szabová et al. 2012). These differences in the composition of the study populations were also reflected in the CA frequencies, which were about twice as high in the GWAS and replication 1 as compared to replication 2.

Due to these population and CA frequency differences and because the CAs are measured as a number of aberrations per 100 cells, we used both the logistic and the linear regression models to evaluate the associations between the genetic variants and the frequencies of CAs. We analyzed the CAs as three phenotypic categories, CAtot, CSA and CTA. However, no loci were found to be associated with CSA frequency. CSAs are also known to be affected to a lesser extent by chemical mutagens as compared to CTAs (Natarajan 1993). The SNPs in CAtot category were selected from the logistic model as it provided stronger associations. In case of CTA, linear model offered better associations than the logistic model; this difference could

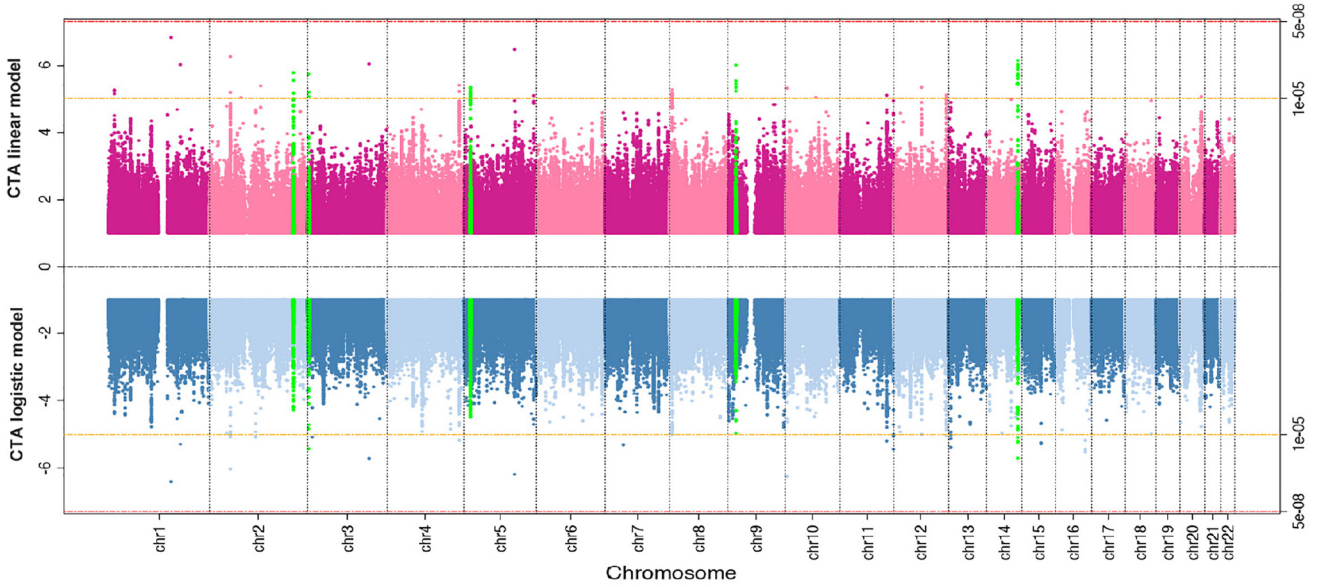


Fig. 2. Miami plot for CTA linear and logistic models. The Y-axis shows the $-\log_{10} P$ – value of each SNP and the X-axis shows their chromosomal position. Loci selected for replication are highlighted in green. The red horizontal line represents the genome-wide significance threshold of $P = 5.0 \times 10^{-8}$ and the yellow horizontal line represents the significance threshold of $P = 5.0 \times 10^{-5}$ as suggestive level of significance

be attributed to the difference in CA_{high}/CA_{low} cutoff point for CTAs (1%) and CAtot (2%). A linear regression is built on continuous variables as outcome. It has a higher precision, and it provides more statistical power with a smaller sample set (MacCallum et al. 2002).

In spite of differences in the study populations, five out of six CAtot variants showed ORs in the same direction both in the GWAS and replication 1. Two SNPs (rs1383997 and rs2824215) reached the suggestive significance *P*-value of 1×10^{-5} in the meta-analysis between the GWAS and replication 1. Addition of replication 2 to the meta-analysis resulted in only one marginal association (rs12002628, *P*-value 4.01×10^{-5}). Variants involved in CTA did not have significant support from the replications either. Only one variant (rs983889) reached the significance level of $P = 1 \times 10^{-5}$ in the meta-analysis of the GWAS and replication 1. Similar to any GWAS, all associated loci were located in the noncoding region of the genome, and we evaluated their potential functional consequences using several *in silico* tools and the existing literature data.

One of the most strongly associated SNPs in the CAtot analyses was located in the gene related to transient receptor potential (TRP) cation channels. Rs1383997 is mapped to the 8q13.3 locus that codes for an antisense transcript. This natural antisense transcript (NAT) is antisense to the musculin (*MSC*) gene and the TRP cation channel subfamily A member 1 (*TRPA1*) gene. NATs are known to regulate the expression of their corresponding sense transcript (Wight and Werner 2013). Another SNP rs12002628 on chromosome 9q21.13, which reached the suggestive level of significance in meta-analysis between GWAS and replication 1, is present at 5.8 kb 5′ to another TRP family protein *TRPM3* gene with many linked SNPs in the introns of *TRPM3*. TRP channels regulate the Ca2+ ions homeostasis in response to environmental and chemical factors. Any deregulation in Ca2+ distribution patterns can promote the signs of cancer development such as proliferation, enhanced survival and invasion (Shapovalov et al. 2016). MSC, also known as activated B-cell factor-1 (ABF-1), is a member of basic helix loop helix (bHLH) family of transcription factors which are involved in cell fate determination in several developmental processes like hematopoiesis and myogenesis (Murre et al. 1994). It is mainly expressed in activated B cells in humans and EBV-transformed lymphoblastoid cell lines (Massari et al. 1998). ABF-1 is capable of inhibiting the transactivation capability of E47 in mammalian cells. E47 is involved with several chromosomal translocations and diminished activity of E47 can lead to lymphoid malignancies (Herblot et al. 2002).

The second SNP from the CAtot analysis, rs2824215 (21q21.1) is located in a long intergenic noncoding RNA (LiNC), and deletion in this locus has been linked to autistic features with complex chromosomal rearrangements (Haldeman-Englert et al. 2010). Interestingly, two other SNPs, which we selected for replication, rs17215792 (2q33.3) and rs2837619 (21q22.2) are located in the genes associated with autism and Down syndrome, *KLF7* (Kruppel like factor 7) (Pescucci et al. 2003; Jang et al. 2015) and *DSCAM* (Down Syndrome Cell Adhesion Molecule), respectively (Yamakawa et al. 1998; Cvetkovska et al. 2013).

TABLE V. SNPs Selected from CTA Model, Their Beta and *P*-values in the Linear Model and the Corresponding OR, 95% CI, and *P*-values in the Logistic Model in Three Sample Sets and Meta-analyses

	Linear Model							Logistic Model					
	SNP	Locus	Minor allele	Major allele	Model	Beta	SE	*P*	I^2	OR	95% CI	*P*	I^2
GWAS	rs340828	3p26.2	A	G	CTA	0.26	0.06	6.45E-06		1.83	1.40–2.38	1.48E-05	
Replication 1						−0.03	0.03	0.39		0.87	0.64–1.2	0.4	
Replication 2						−0.004	0.005	0.42		1.01	0.83–1.22	0.96	
Meta GWAS+ Replication 1[1]						0.06	0.02	0.02	0.94	1.19	0.96–1.46	0.11	0.84
meta all[2]						−0.002	0.005	0.75	0.91	1.08	0.94–1.25	0.26	0.74
GWAS	rs8003642	14q32.13	C	A	CTA	−0.33	0.07	1.36E-06		0.52	0.39–0.69	1.77E-05	
Replication 1						−0.02	0.04	0.66		0.91	0.63–1.32	0.62	
Replication 2						−0.003	0.006	0.58		0.86	0.69–1.08	0.2	
Meta GWAS+ Replication 1						−0.1	0.03	3.70E-04	0.91	0.7	0.53–0.91	0.008	0.77
meta all						−0.01	0.01	0.21	0.91	0.79	0.66–0.94	0.01	0.65
GWAS	rs17215792	2q33.3	C	A	CTA	0.47	0.1	1.70E-06		2.52	1.54–4.12	1.11E-04	
Replication 1						0.01	0.06	0.82		1.08	0.59–1.99	0.81	
Replication 2						0.003	0.008	0.73		0.96	0.70–1.33	0.83	
Meta GWAS+ Replication 1						0.15	0.04	7.43E-04	0.91	1.65	1.15–2.36	0.006	0.64
meta all						0.01	0.01	0.34	0.91	1.23	0.97–1.56	0.09	0.74
GWAS	rs983889	5p15.1	G	T	CTA	−0.26	0.06	4.82E-06		0.61	0.48–0.77	3.85E-05	
Replication 1						−0.06	0.03	0.07		0.77	0.57–1.04	0.09	
Replication 2						0.008	0.005	0.09		1.05	0.88–1.26	0.61	
Meta GWAS+ Replication 1						−0.11	0.02	1.06E-05	0.82	0.65	0.52–0.80	6.61E-05	0.56
meta all						0.004	0.005	0.4	0.93	0.86	0.75–0.98	0.03	0.85
GWAS	rs7033729	9p22.1	A	G	CTA	0.53	0.11	9.99E-07		3.53	1.92–6.48	1.05E-05	
Replication 1						0.005	0.05	0.93		1.03	0.62–1.72	0.9	
Replication 2						−0.01	0.009	0.23		0.85	0.60–1.21	0.38	
Meta GWAS+ Replication 1						0.12	0.04	0.003	0.93	1.49	1.05–2.12	0.03	0.74
meta all						−0.005	0.01	0.56	0.92	1.12	0.88–1.44	0.35	0.77

[1](Meta-analysis between the GWAS and replication 1)
[2](Meta-analysis between all three sample sets i.e. GWAS sample set, replication 1 and replication 2)

Chromosomal abnormalities are an important feature of both diseases (Liao et al. 2013). Both of these SNPs, however, showed only weak, if any, association with CAs in the replication sets.

The only SNP from the CTA analysis with a suggestive level of association in the GWAS and replication 1 meta-analysis, rs983889, is an intronic SNP in the F-box and leucine-rich repeat protein 7 (*FBXL7*) gene. FBXL7 belongs to F-box proteins, which are involved in phosphorylation–dependent ubiquitination of proteins and which display proapoptotic activity (Zheng et al. 2016). Incidentally, one of the targets of FBXL7 is aurora kinase A (*AURKA*), a known oncogene, involved in regulation of mitosis (Tang et al. 2015). During late G2 phase, AURKA is recruited to centrosomes (Hanisch et al. 2006) and later on promotes centrosome maturation and bipolar spindle formation (Gruss et al. 2001). Since CTAs also arise during S/G2 phase (Durante et al. 2013), an indirect involvement of AURKA can be anticipated to affect the frequency of CTAs.

In conclusion, our GWAS identified 11 SNPs associated with CA frequency, from which three were replicated at the suggestive level of significance. *In silico* predictions of functional consequences of the identified SNPs and their loci revealed that they were directly or indirectly related to different cancers. They included genes encoding TRP cation channel proteins, which regulate the Ca2+ ions homeostasis in response to environmental and chemical factors, genes involved in autism and Down syndrome, two syndromes linked to chromosomal abnormalities, and *FBXL7*, which interacts with AURKA, an important regulator of mitosis. Although due to sample size the results of this GWAS are not definitive in terms of pointing out the exact rationale behind CAs development but they certainly point towards the probable loci that could be involved in the elevated frequency of CAs in the presence of environmental stress.

Statement of Author Contributions

PV, KH, and AF designed the study; HT and YN analyzed the data; YN and BS performed the genotyping; PV, LV, LM, SV, MK, and VV provided samples and data for GWAS and replication 1; MD, BS, AK, MB, KV, and MS provided samples and data for replication 2; PH, MMN were responsible for the GWAS; YN wrote the first draft of the manuscript; AF, KH, PV, LV, MD, and BS critically revised the manuscript; all authors read and approved the final manuscript.

ACKNOWLEDGMENTS

In the Czech Republic, the study was supported by the National Science Foundation, Grant Numbers: 15-14789S and 18-09709S; Charles University in Prague, UNCE 204022, and PROGRES Q 28; Medical Faculty in Pilsen, Charles University in Prague, National Sustainability Programme I, Nr. LO 1503 and Charles University Research Centre program UNCE/MED/006.

In the Slovak Republic, the study was supported by EC contracts QLK4-CT-1999-01629, ERBICI 15-CT96-1012, CIPA-CT94-0129, Slovak Grant Agency APVT-21 013202, APVT-21-017704 and grants from the Ministry of Health, Slovak Republic 2005/43-SZU-21 and 2006/07-SZU-02 MZ SR, 2005/42-SZU-20, SZU.

REFERENCES

Albertini RJ, Anderson D, Douglas GR, Hagmar L, Hemminki K, Merlo F, Natarajan AT, Norppa H, Shuker DE, Tice R, et al. 2000. IPCS guidelines for the monitoring of genotoxic effects of carcinogens in humans. International Programme on Chemical Safety. Mutat. Res. 463: 111–172.

Anderson CA, Pettersson FH, Clarke GM, Cardon LR, Morris AP, Zondervan KT. 2010. Data quality control in genetic case-control association studies. Nat. Protoc. 5:1564–1573.

Baan RA, Grosse Y. 2004. Man-made mineral (vitreous) fibres: evaluations of cancer hazards by the IARC Monographs Programme. Mutat. Res. Mol. Mech. Mutagen. 553:43–58.

Bignold LP. 2009. Mechanisms of clastogen-induced chromosomal aberrations: A critical review and description of a model based on failures of tethering of DNA strand ends to strand-breaking enzymes. Mutat. Res. Mutat. Res. 681:271–298.

Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, Karczewski KJ, Park J, Hitz BC, Weng S, et al. 2012. Annotation of functional variation in personal genomes using RegulomeDB. Genome Res. 22:1790–1797.

Carrano AV, Natarajan AT. 1988. International Commission for Protection Against Environmental Mutagens and Carcinogens. ICPEMC publication no. 14. Considerations for population monitoring using cytogenetic techniques. Mutat. Res. 204:379–406.

Cvetkovska V, Hibbert AD, Emran F, Chen BE. 2013. Overexpression of Down syndrome cell adhesion molecule impairs precise synaptic targeting. Nat. Neurosci. 16:677–682.

Delaneau O, Marchini J, Zagury J-F. 2011. A linear complexity phasing method for thousands of genomes. Nat. Methods 9:179–181.

Durante M, Bedford JS, Chen DJ, Conrad S, Cornforth MN, Natarajan AT, van Gent DC, Obe G. 2013. From DNA damage to chromosome aberrations: Joining the break. Mutat. Res. Toxicol. Environ. Mutagen. 756:5–13.

Dušinská M, Kažimírová A, Barančoková M, Beňo M, Smolková B, et al. 2003. Nutritional supplementation with antioxidants decreases chromosomal damage in humans. Mutagenesis 18:371–376.

Dušinská M, Barančoková M, Kažimírová A, Harrington V, Volkovová K, et al. 2004a. Does occupational exposure to mineral fibres cause DNA or chromosome damage? In: *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*,, pp. 103–110.

Dušinská M, Collins A, Kažimírová A, Barančoková M, Harrington V, et al. 2004b. Genotoxic effects of asbestos in humans. In: *Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis*, pp. 91–102.

Dušinská M, Staruchova M, Horska A, Smolkova B, Collins A, et al. 2012. Are glutathione S transferases involved in DNA damage signalling? Interactions with DNA damage and repair revealed from molecular epidemiology studies. Mutat. Res. - Fundam. Mol. Mech. Mutagen. 736:130–137.

Försti A, Frank C, Smolkova B, Kazimirova A, Barancokova M, Vymetalkova V, Kroupa M, Naccarati A, Vodickova L, Buchancova J, et al. 2016. Genetic variation in the major mitotic checkpoint genes associated with chromosomal aberrations in healthy humans. Cancer Lett. 380:442–446.

Gostissa M, Alt FW, Chiarle R. 2011. Mechanisms that Promote and Suppress Chromosomal Translocations in Lymphocytes. Annu. Rev. Immunol. 29:319–350.

Gruss OJ, Carazo-Salas RE, Schatz CA, Guarguaglini G, Kast J, Wilm M, le Bot N, Vernos I, Karsenti E, Mattaj IW. 2001. Ran induces spindle assembly by reversing the inhibitory effect of importin α on TPX2 activity. Cell 104:83–93.

Hagmar L, Strömberg U, Tinnerberg H, Mikoczy Z. 2001. The usefulness of cytogenetic biomarkers as intermediate endpoints in carcinogenesis. Int. J. Hyg. Environ. Health 204:43–47.

Haldeman-Englert CR, Chapman KA, Kruger H, Geiger EA, McDonald-McGinn DM, et al. 2010. A de novo 8.8-Mb deletion of 21q21.1-q21.3 in an autistic male with a complex rearrangement involving chromosomes 6, 10, and 21. Am. J. Med. Genet. Part A 152A:196–202.

Hanisch A, Wehner A, Nigg EA, Silljé HHW. 2006. Different Plk1 functions show distinct dependencies on Polo-Box domain-mediated targeting. Mol Biol Cell 17:448–459.

Hemminki K, Rachakonda S, Musak L, Vymetalkova L, Halasova E, et al. 2015a. Telomere length in circulating lymphocytes: Association with chromosomal aberrations. Genes Chromosom. Cancer 54: 194–196.

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

Heng HHQ, Regan SM, Liu G, Ye CJ. 2016. Why it is crucial to analyze non clonal chromosome aberrations or NCCAs? Mol. Cytogenet. 9:15.

Herblot S, Aplan PD, Hoang T. 2002. Gradient of E2A activity in B-cell development. Mol. Cell. Biol. 22:886–900.

Howie B, Marchini J, Stephens M. 2011. Genotype imputation with thousands of genomes. G3 (Bethesda) 1:457–470.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, World Health Organization, International Agency for Research on Cancer. 2008. 1,3-Butadiene, ethylene oxide, and vinyl halides (vinyl flouride, vinyl chloride, and vinyl bromide). International Agency for Research on Cancer.

IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, International Agency for Research on Cancer. 2012. A review of human carcinogens. Arsenic, metals, fibres, and dusts. International Agency for Research on Cancer.

Jang D-H, Chae H, Kim M. 2015. Autistic and Rett-like features associated with 2q33.3-q34 interstitial deletion. Am. J. Med. Genet. Part A 167:2213–2218.

Jones CH, Pepper C, Baird DM. 2012. Telomere dysfunction and its role in haematological cancer. Br. J. Haematol. 156:573–587.

Kažimírová A, Barančoková M, Volkovová K, Staruchová M, Krajcovicova-Kudlackova M, et al. 2004. Does a vegetarian diet influence genomic stability? Eur. J. Nutr. 43:32–38.

Kažimírová A, Barančoková M, Krajčovičová-Kudláčková M, Volkovová K, Staruchová M, Valachovičová M, Pauková V, Blažíček P, Wsólová L, Dušinská M. 2006. The relationship between micronuclei in human lymphocytes and selected micronutrients in vegetarians and non-vegetarians. Mutat. Res. - Genet. Toxicol. Environ. Mutagen. 611:64–70.

Kažimírová A, Barančoková M, Džupinková Z, Wsólová L, Dušinská M. 2009. Micronuclei and chromosomal aberrations, important markers of ageing: Possible association with XPC and XPD polymorphisms. Mutat. Res. Mol. Mech. Mutagen. 661:35–40.

Liao H-M, Gau SS-F, Tsai W-C, Fang J-S, Su Y-C, Chou MC, Liu SK, Chou WJ, Wu YY, Chen CH. 2013. Chromosomal Abnormalities in Patients With Autism Spectrum Disorders From Taiwan. Am. J. Med. Genet. Part B Neuropsychiatr. Genet. 162:734–741.

MacCallum RC, Zhang S, Preacher KJ, Rucker DD. 2002. On the practice of dichotomization of quantitative variables. Psychol. Methods 7: 19–40.

Mägi R, Morris AP. 2010. GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics 11:288.

Massari ME, Rivera RR, Voland JR, Quong MW, Breit TM, van Dongen JJM, de Smit O, Murre C. 1998. Characterization of ABF-1, a novel basic helix-loop-helix transcription factor expressed in activated B lymphocytes. Mol. Cell. Biol. 18:3130–3139.

Mitelman F. 2000. Recurrent chromosome aberrations in cancer. Mutat. Res. 462:247–253.

Murre C, Bain G, van Dijk MA, Engel I, Furnari BA, Massari ME, Matthews JR, Quong MW, Rivera RR, Stuiver MH. 1994. Structure and function of helix-loop-helix proteins. Biochim. Biophys. Acta 1218:129–135.

Musak L, Soucek P, Vodickova L, Naccarati A, Halasova E, Polakova V, Slyskova J, Susova S, Buchancova J, Smerhovsky Z, et al. 2008. Chromosomal aberrations in tire plant workers and interaction with polymorphisms of biotransformation and DNA repair genes. Mutat. Res. Mol. Mech. Mutagen. 641:36–42.

Musak L, Smerhovsky Z, Halasova E, Osina O, Letkova L, Vodickova L, Polakova V, Buchancova J, Hemminki K, Vodicka P. 2013. Chromosomal damage among medical staff occupationally exposed to volatile anesthetics, antineoplastic drugs, and formaldehyde. Scand. J. Work. Environ. Heal. 39:618–630.

Natarajan AT. 1993. Mechanisms for induction of mutations and chromosome alterations. Environ. Health Perspect. 101(Suppl):225–229.

Pescucci C, Meloni I, Bruttini M, Ariani F, Longo I, Mari F, Canitano R, Hayek G, Zappella M, Renieri A. 2003. Chromosome 2 deletion encompassing the MAP2 gene in a patient with autism and Rett-like features. Clin. Genet. 64:497–501.

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26:2336–2337.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, et al. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81:559–575.

Rosenbloom KR, Armstrong J, Barber GP, Casper J, Clawson H, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, et al. 2015. The UCSC Genome Browser database: 2015 update. Nucleic Acids Res. 43:D670–D681.

Rossner P, Boffetta P, Ceppi M, Bonassi S, Smerhovsky Z, Landa K, Juzova D, Šrám RJ. 2005. Chromosomal aberrations in lymphocytes of healthy subjects and risk of cancer. Environ. Health Perspect. 113:517–520.

Shapovalov G, Ritaine A, Skryma R, Prevarskaya N. 2016. Role of TRP ion channels in cancer and tumorigenesis. Semin. Immunopathol. 38:357–369.

Somorovská M, Jahnová E, Tulinská J, Zámecníková M, Sarmanová J, et al. 1999. Biomonitoring of occupational exposure to styrene in a plastics lamination plant. Mutat. Res. 428:255–269.

Šrám RJ, Rössner P, Šmerhovský Z. 2004. Cytogenetic analysis and occupational health in the Czech Republic. Mutat. Res. 566:21–48.

Szabová M, Jahnová E, Horváthová M, Ilavská S, Pružincová V, Nemessányi T, Tulinská J, Wsólová L, Volkovová K. 2012. Changes in immunologic parameters of humoral immunity and adipocytokines in obese persons are gender dependent. Hum. Immunol. 73:486–492.

Tang A, Gao K, Chu L, Zhang R, Yang J, et al. 2015. Aurora kinases: novel therapy targets in cancers. Oncotarget 8:23937–23954.

Tulinska J, Jahnova E, Dušinská M, Kuricova M, Liskova A, et al. 2004. Immunomodulatory effects of mineral fibres in occupationally exposed workers. In: Mutation Research - Fundamental and Molecular Mechanisms of Mutagenesis, pp. 111–124.

Vodenkova S, Polivkova Z, Musak L, Smerhovsky Z, Zoubkova H, Sytarova S, Kavcova E, Halasova E, Vodickova L, Jiraskova K, et al. 2015. Structural chromosomal aberrations as potential risk markers in incident cancer patients. Mutagenesis 30:557–563.

Vodicka P, Tuimala J, Stetina R, Kumar R, Manini P, et al. 2004a. Cytogenetic markers, DNA single-strand breaks, urinary metabolites, and DNA repair rates in styrene-exposed lamination workers. Environ. Health Perspect. 112:867–871.

Vodicka P, Kumar R, Stetina R, Musak L, Soucek P, et al. 2004b. Markers of individual susceptibility and DNA repair rate in workers exposed to xenobiotics in a tire plant. Environ. Mol. Mutagen. 44:283–292.

Vodicka P, Polivkova Z, Sytarova S, Demova H, Kucerova M, Vodickova L, Polakova V, Naccarati A, Smerhovsky Z, Ambrus M, et al. 2010. Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls. Carcinogenesis 31:1238–1241.

Vodicka P, Musak L, Frank C, Kazimirova A, Vymetalkova V, Barancokova M, Smolkova B, Dzupinkova Z, Jiraskova K, Vodenkova S, et al. 2015. Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects. Carcinogenesis 36:1299–1306.

Vodicka P, Musak L, Vodickova L, Vodenkova S, Catalano C, et al. 2018. Genetic variations of acquired structural chromosomal aberrations. Mutat. Res. - Genet. Toxicol. Environ. Mutagen (In Press).

Ward LD, Kellis M. 2012. HaploReg: a resource for exploring chromatin states, conservation, and regulatory motif alterations within sets of genetically linked variants. Nucleic Acids Res. 40:D930–D934.

Wight M, Werner A. 2013. The functions of natural antisense transcripts. Essays Biochem. 54:91–101.

Yamakawa K, Huot YK, Haendelt MA, Hubert R, Chen XN, Lyons GE, Korenberg JR. 1998. DSCAM: a novel member of the immunoglobulin superfamily maps in a Down syndrome region and is involved in the development of the nervous system. Hum. Mol. Genet. 7:227–237.

Zheng N, Wang Z, Wei W. 2016. Ubiquitination-mediated degradation of cell cycle-related proteins by F-box proteins. Int. J. Biochem. Cell Biol. 73:99–110.

Accepted by—
F. Marchetti

Manuscript IV

Chromosomal damage and telomere length in peripheral blood lymphocytes: cancer risk and patients' long-term survival

Vodenkova S, **Kroupa M**, Polivkova Z, Musak L, Ambrus M, Schneiderova M, Kozevnikovova R, Vodickova L, Rachakonda S, Hemminki K, Kumar R, Vodicka P.

Manuscript in preparation

International Journal of *Molecular Sciences*

MDPI

Article

Chromosomal damage and telomere length in peripheral blood lymphocytes: cancer risk and patients' long-term survival

Sona Vodenkova [1,2,3,*], Michal Kroupa [2,4], Zdenka Polivkova [1], Ludovit Musak [5], Miloslav Ambrus [6], Michaela Schneiderova [7], Renata Kozevnikovova [8], Ludmila Vodickova [2,3,4], Sivaramakrishna Rachakonda [9], Kari Hemminki [9], Rajiv Kumar [9] and Pavel Vodicka [2,3,4,*]

1 Department of Medical Genetics, Third Faculty of Medicine, Charles University, Prague, 100 00, Czech Republic; sona.vodenkova@iem.cas.cz (S.V.); zdena.polivkova@lf3.cuni.cz (Z.P.)
2 Department of Molecular Biology of Cancer, Institute of Experimental Medicine of the Czech Academy of Sciences, Prague, 142 00, Czech Republic; michal.kroupa@iem.cas.cz (M.K.)
3 Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, 128 00, Czech Republic; ludmila.vodickova@iem.cas.cz (L.V.)
4 Biomedical Centre, Faculty of Medicine in Pilsen, Charles University, Pilsen, 323 00, Czech Republic; pavel.vodicka@iem.cas.cz (P.V.)
5 Biomedical Center Martin, Jessenius Faculty of Medicine, Comenius University in Bratislava, Martin, 036 01, Slovakia; Ludovit.Musak@jfmed.uniba.sk (L.M.)
6 Department of Radiology and Oncology, Faculty Hospital Kralovske Vinohrady, Prague, 100 34, Czech Republic; miloslav.ambrus@fnkv.cz (M.A.)
7 Department of Surgery, General University Hospital in Prague, Prague, 128 00, Czech Republic; schneider77@seznam.cz (M.S.)
8 Department of Oncosurgery, MEDICON, Prague, 140 44, Czech Republic; renata.kozevnikovova@onko-centrum.cz (R.K.[1])
9 Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, 691 20, Germany; s.rachakonda@dkfz-heidelberg.de (S.R.); K.Hemminki@dkfz-heidelberg.de (K.H.); R.Kumar@dkfz-heidelberg.de (R.K.[2])
* Correspondence: sona.vodenkova@iem.cas.cz (S.V.); pavel.vodicka@iem.cas.cz (P.V.); Tel.: +420-241-062-251 (S.V.); +420-241-062-694 (P.V.)

Received: date; Accepted: date; Published: date

Abstract: Accumulation of non-specific structural chromosomal aberrations (CAs) and telomere shortening contribute to genome instability, one of the hallmarks of cancer. CAs arise as a consequence of direct DNA damage or may be driven by telomere shortening. CAs in peripheral blood lymphocytes (PBL), proven markers of exposure, may participate in pathophysiology and progression of cancer, however, the mechanisms are unresolved. The prognostic value of telomere length (TL) in cancer patients also remains unclear. We focused on the evaluation of CAs and TL in PBL of incident cancer patients (151 breast, 96 colorectal, 90 lung) and 335 controls. Obtained results were analyzed together with baseline, clinico-pathological, and follow-up data. We observed the accumulation of CAs in PBL appear as a marker of breast and lung cancer susceptibility ($P<0.0001$). Individuals with longer TL were at increased risk of breast cancer ($P<0.0001$). Accumulation of chromatid-type aberrations was associated with poor overall survival in breast and colorectal cancer patients in a multivariate survival model. In contrast to controls, patients did not exhibit any relationship between either TL and CAs or TL and age. In summary, we provided pilot data on CAs and TL in PBL in the context of cancer risk and patients' long-term survival.

Keywords: Structural chromosomal aberrations; Telomere length; Peripheral blood lymphocytes; Breast, Colorectal and Lung cancer; Cancer susceptibility; Patients' long-term survival

1. Introduction

Tumorigenesis is a multistep process driven by several genetic alterations, including chromosomal rearrangements that confer a selective advantage for a dividing cell [1]. In addition, their accumulation promotes tumor growth and progression [2]. Many cancers also exhibit epigenetic changes that can have profound effects on the expression of key genes involved in the process of malignant transformation [3].

The acquired non-specific structural chromosomal aberrations (CAs) represent multiple and diverse rearrangements within or between chromosomes. Structural CAs can be either chromatid-type (CTAs) or chromosome-type aberrations (CSAs) depending on the stage of the cell cycle and mechanism of action of clastogenic compounds. They arise during an individual's life as a consequence of direct mutagenic effect and/or due to DNA repair mechanisms dysfunction [4-6]. CAs may even be observed in peripheral blood lymphocytes (PBL) of healthy individuals. For decades, measurement of CAs in PBL has thus been conventionally used in biomonitoring of individuals' environmental and occupational exposure to clastogens [7, 8]. Several European cohort studies, both prospective and retrospective, have gathered evidence proving that the increased frequencies of CAs in PBL could serve as a predictive marker for increased cancer susceptibility [9-12].

Some CAs pose a primary event in cancer origins, other chromosomal changes are driven by the disease progression and reflect the prognosis and cancer stage [13]. Although specific chromosomal rearrangements detected in malignancies are generated during carcinogenesis, some authors assume that the non-specific structural CA frequencies in PBL may represent a marker of cancer susceptibility. Their hypothesis was based on the presumption that genetic damage in PBL reflects similar damage in different target cells undergoing malignant transformation (summarized by Boffetta *et al.* 2007 [14]).

It has recently been recognized that, except for direct DNA damage where double-strand break DNA repair is crucial, mechanisms associated with telomere shortening are important contributors to the formation of CAs as well [15-17]. Nevertheless, there is still a lack of studies addressing this issue, especially in PBL. Telomeres are special chromatin structures at the ends of linear eukaryotic chromosomes that are characterized by tandem repeats (TTAGGG)n with several associated proteins [18, 19]. Telomeres progressively shorten through cellular lifespan [20]. Their shortening can cause the proliferation arrest and cell apoptosis through the loss of protection at chromosome ends. On the other hand, the loss of telomere function can lead to telomere crisis and genome instability that can induce numerical and structural CAs [21]. Therefore, telomere length (TL) plays a critical role both in genome integrity maintenance and cancer initiation/progression.

In our present study, we focused on the evaluation of both acquired non-specific structural CAs and TL in PBL of patients with a sporadic form of the breast (BC), colorectal (CRC), and lung (LC) cancer and control individuals. Patients comprise individuals with any prior use of radio- or chemotherapy or X-rays (e.g. diagnostic) to avoid the presence of additional chromosomal damage related to the treatment. These cancer types are ranking at the top with respect to incidence and mortality worldwide [22] and therefore, there is an urgent need to search for new diagnostic, predictive and prognostic biomarkers. Since chromosomal instability (CIN) represents a marker of genotoxic exposure and a predictor of cancer risk, it is necessary to study its genetic basis as well. The novelty of this study lays in the investigation of the prognostic value of CAs and TL in PBL using survival analyses in newly diagnosed cancer patients. Moreover, correlation of CA frequencies with TL both in cancer patients and controls is addressed for the first time.

2. Results

2.1. Case-control study

The levels of all studied types of structural CAs, i.e. aberrant cells (ACs), total chromosomal aberrations (CAtot), CTAs, and CSAs were statistically significantly higher in BC and LC patients in comparison with corresponding control groups. In CRC patients, only the frequencies of ACs and CTAs were statistically significantly higher compared to controls (Table 1). Additionally, the risk of particular types of studied cancers was statistically significantly associated with CA frequencies; adjusted odds ratios (aORs) with their 95% confidence intervals (CIs) are summarized in Table 2. Individuals with higher frequencies of chromosomal damage markers were at higher risk of BC (aOR for ACs = 4.81, $P < 0.0001$; aOR for CAtot = 4.49, $P < 0.0001$; aOR for CTAs = 2.43, $P = 0.0004$; and aOR for CSAs = 2.31, $P = 0.0008$). Similarly to BC, elevated frequencies of CAs had an impact on the increased risk of LC (aOR for ACs = 3.82, $P < 0.0001$; aOR for CAtot = 3.44 $P < 0.0001$; aOR for CTAs = 3.30, $P < 0.0001$; and aOR for CSAs = 1.75, $P = 0.04$). No association between CAs and CRC risk was found in our set of patients. Both age and sex were statistically significantly associated with the risk of CRC and LC (Table 2).

Table 1. The levels of chromosomal damage (ACs, CAtot, CTAs, CSAs) and TL in each group of cancer patients and their corresponding controls.

	FEMALE CONTROLS (N=145)							BREAST CANCER PATIENTS (N=151)							
	Mean	**SD**	**95% CI**	**Range**	**Min**	**Median**	**Max**	**Mean**	**SD**	**95% CI**	**Range**	**Min**	**Median**	**Max**	**P-value**
ACs	1.63	1.20	1.43-1.82	6	0	2	6	2.62	1.57	2.36-2.87	7	0	3	7	**<0.0001**
CAtot	1.72	1.31	1.50-1.93	6	0	2	6	2.73	1.64	2.47-2.99	7	0	3	7	**<0.0001**
CTAs	1.00	0.99	0.84-1.16	4	0	1	4	1.63	1.34	1.41-1.85	6	0	1	6	**<0.0001**
CSAs	0.70	0.90	0.55-0.84	5	0	0	5	1.10	1.03	0.93-1.27	4	0	1	4	**0.0005**
TL	1.54	0.82	1.40-1.69	3.17	0.61	1.23	3.78	1.88	0.79	1.72-2.03	5.06	0.57	1.77	5.63	**<0.0001**
	CONTROLS (N=335)							**COLORECTAL CANCER PATIENTS (N=96)**							
	Mean	**SD**	**95% CI**	**Range**	**Min**	**Median**	**Max**	**Mean**	**SD**	**95% CI**	**Range**	**Min**	**Median**	**Max**	**P-value**
ACs	1.76	1.30	1.62-1.90	6	0	2	6	2.13	1.45	1.83-2.42	6	0	2	6	**0.03**
CAtot	1.89	1.45	1.73-2.04	7	0	2	7	2.26	1.67	1.92-2.60	8	0	2	8	0.06
CTAs	1.07	1.00	0.96-1.18	4	0	1	4	1.46	1.31	1.19-1.72	7	0	1	7	**0.02**
CSAs	0.81	1.10	0.69-0.93	6	0	0	6	0.80	1.00	0.60-1.00	4	0	1	4	0.78
TL	1.77	0.81	1.67-1.86	3.18	0.60	1.57	3.78	1.54	0.41	1.44-1.64	2.13	0.90	1.48	3.03	0.33
	CONTROLS (N=335)							**LUNG CANCER PATIENTS (N=90)**							
	Mean	**SD**	**95% CI**	**Range**	**Min**	**Median**	**Max**	**Mean**	**SD**	**95% CI**	**Range**	**Min**	**Median**	**Max**	**P-value**
ACs	1.76	1.30	1.62-1.90	6	0	2	6	2.84	1.49	2.53-3.16	6	0	3	6	**<0.0001**
CAtot	1.89	1.45	1.73-2.04	7	0	2	7	2.88	1.53	2.56-3.20	7	0	3	7	**<0.0001**
CTAs	1.07	1.00	0.96-1.18	4	0	1	4	1.83	1.34	1.55-2.11	6	0	2	6	**<0.0001**
CSAs	0.81	1.10	0.69-0.93	6	0	0	6	1.06	0.99	0.85-1.26	4	0	1	4	**0.0007**
TL	1.77	0.81	1.67-1.86	3.18	0.60	1.57	3.78	1.54	0.93	1.22-1.85	5.42	0.77	1.35	6.19	0.06

The table summarizes the results obtained in the case-control study calculated by Median Two-Sample Test. The levels of ACs, CAtot, CTAs, and CSAs were evaluated using the standard cytogenetic method. TL was measured as relative telomere length and its measurement was conducted using the monochrome multiplex polymerase chain reaction. Statistically significant values are in bold.

Abbreviations: ACs—aberrant cells; CAtot—total chromosomal aberrations; CI—confidence interval; CSAs—chromosome-type aberrations; CTAs—chromatid-type aberrations; N—number of samples; SD—standard deviation; TL—telomere length.

Table 2. The association of CA frequencies (ACs, CAtot, CTAs, CSAs) and TL with the risk of particular types of studied cancers.

	BREAST CANCER (N=151)			COLORECTAL CANCER (N=96)			LUNG CANCER (N=90)		
	aOR	**95% CI**	**P-value**	**aOR**	**95% CI**	**P-value**	**aOR**	**95% CI**	**P-value**
ACs	4.81	2.84-8.12	**<0.0001**	1.19	0.71-1.98	0.52	3.82	2.27-6.42	**<0.0001**
Age (years)	0.99	0.97-1.01	0.27	1.04	1.02-1.06	**<0.0001**	1.06	1.04-1.08	**<0.0001**
Sex (0 for females)	-	-	-	1.85	1.10-3.10	**0.02**	3.18	1.72-5.90	**0.0002**
CAtot	4.49	2.69-7.50	**<0.0001**	1.04	0.63-1.73	0.87	3.44	2.05-5.77	**<0.0001**
Age (years)	0.99	0.97-1.01	0.18	1.04	1.02-1.06	**<0.0001**	1.06	1.04-1.08	**<0.0001**
Sex (0 for females)	-	-	-	1.87	1.11-3.13	**0.02**	3.32	1.79-6.15	**0.0001**
CTAs	2.43	1.49-3.97	**0.0004**	1.56	0.96-2.53	0.07	3.30	1.97-5.52	**<0.0001**
Age (years)	0.99	0.98-1.01	0.49	1.04	1.02-1.06	**<0.0001**	1.06	1.04-1.08	**<0.0001**
Sex (0 for females)	-	-	-	1.84	1.10-3.09	**0.02**	3.40	1.84-6.29	**<0.0001**
CSAs	2.31	1.42-3.77	**0.0008**	0.92	0.57-1.47	0.73	1.75	1.03-2.96	**0.04**
Age (years)	0.99	0.97-1.00	0.13	1.04	1.02-1.06	**<0.0001**	1.06	1.04-1.08	**<0.0001**
Sex (0 for females)	-	-	-	1.87	1.12-3.13	**0.02**	3.70	2.02-6.76	**<0.0001**
TL	6.49	3.00-14.04	**<0.0001**	0.76	0.39-1.48	0.42	0.89	0.36-2.21	0.80
Age (years)	1.02	1.00-1.05	0.09	1.03	1.00-1.05	**0.04**	1.05	1.02-1.08	**0.003**
Sex (0 for females)	-	-	-	1.66	0.91-3.01	0.10	4.94	1.97-12.39	**0.0007**

The table summarizes the results obtained in the case-control study. aORs with 95% CIs for all studied types of CAs and TL for individual cancer types expressed by Logistic regression analysis. The levels of ACs, CAtot, CTAs, and CSAs were evaluated using the standard cytogenetic method. TL was measured as relative telomere length and its measurement was conducted using the monochrome multiplex polymerase chain reaction. Statistically significant values are in bold.

Abbreviations: ACs—aberrant cells; aOR—adjusted odds ratio; CAtot—total chromosomal aberrations; CI—confidence interval; CSAs—chromosome-type aberrations; CTAs—chromatid-type aberrations; N—number of samples; TL—telomere length.

In the current study, we have newly focused on the study of TL, measured as relative telomere length (RTL). BC patients showed statistically significantly longer TL compared to control women ($P < 0.0001$, Table 1). Longer TL was linked to higher risk of BC as well (aOR = 6.49, $P < 0.0001$, Table 2). On the contrary, CRC and LC patients had moderately shorter TL in comparison with control individuals (Table 1) but their shortening did not affect the risk of either cancer types (Table 2). We have further correlated TL with the frequencies of all types of CAs and also both TL and CAs were correlated with age. We have found negative correlation between the majority of studied CA types and TL in the whole group of controls (ACs: Spearman's rho (rs) = -0.21, P = 0.0004; CAtot: rs = -0.22, P = 0.0003; CTAs: rs = -0.10, P = 0.10; CSAs: rs = -0.22, P = 0.0002; Supplementary Figure S1a). Spearman's correlations also showed statistically significant relationship between TL and age (rs = -0.62, $P < 0.0001$) and between CSAs and age (rs = 0.17, P = 0.001) in control individuals. In contrast to controls, we have not detected the same relationship either between ACs/CAtot/CTAs/CSAs and TL (Supplementary Figure S1b for BC, S1c for CRC, S1d for LC) or between all studied parameters and age in individual groups of cancer patients (data not shown), except for LC patients who exhibited a moderately statistically significant correlation between ACs/CAtot/CSAs and age (rs = 0.23, P = 0.03; rs = 0.22, P = 0.04; rs = 0.22, P = 0.04, respectively).

2.1. Survival analyses

In our studied types of cancer, we proved well-known significant differences in both overall survival (OS) and recurrence-free survival (RFS) between BC, CRC and LC groups of patients ($P < 0.0001$). The alive/dead ratio for BC patients was 134/17 (i.e. 88.7% of patients surviving 5 years), for CRC patients 51/45 (i.e. 53.1% surviving 5 years) and for LC patients 23/63 (i.e. 26.7% surviving 5 years). Concerning RFS, patients also differed by cancer type ($P < 0.0001$) – in BC: 122 patients have survived asymptomatically and 29 patients experienced an event during the follow-up period. CRC patients were much more likely to experience an event and therefore faced decreased RFS (36 asymptomatic survival, 60 any event). In LC patients, RFS was almost identical to OS (20 asymptomatic survival, 66 any event). For this reason, we analyzed individual cancer groups separately.

2.1.1. Univariate survival analyses

All results from the univariate assessment of the impact of well-established prognostic factors (baseline and disease characteristics) as well as of ACs, CAtot, CTAs, CSAs and TL on OS and RFS, expressed by Cox regression hazard model, are presented in Table 3. Tumor-node-metastasis (TNM) stage, histopathological grade, histological classification and tumor location modulated OS and RFS in a different way. Higher age was associated with worse OS and RFS in BC and CRC patients. In BC patients, absence of both hormonal receptors and hormonal therapy decreased OS and even more RFS. Regarding CAs and TL, we did not observe any statistically significant relationship with OS and RFS in any group of cancer patients.

Table 3. Baseline and disease characteristics as well as ACs, CAtot, CTAs, CSAs and TL, and their impact on overall (OS) and recurrence-free (RFS) survival of patients with complete follow-up.

		BREAST CANCER PATIENTS (N=151)						**COLORECTAL CANCER PATIENTS (N=96)**						**LUNG CANCER PATIENTS (N=86)**					
		OS			**RFS**			**OS**			**RFS**			**OS**			**RFS**		
		HR	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**	**HR**	**95% CI**	**P-value**
Age	**Under cut-off (Ref) vs. Over cut-off[a]**	4.21	1.48-12.00	**0.007**	2.73	1.02-7.29	**0.05**	2.49	1.32-4.70	**0.005**	1.96	1.05-3.69	**0.04**	0.61	0.34-1.11	0.10	0.60	0.33-1.09	0.09
Sex	**Females (Ref) vs. Males**	-	-	-	-	-	-	1.26	0.68-2.34	0.47	1.32	0.73-2.36	0.36	1.59	0.88-2.89	0.13	1.56	0.87-2.77	0.13
Family history of cancer	**No (Ref) vs. Yes**	0.47	0.18-1.22	0.12	0.42	0.19-0.93	**0.03**	1.20	0.66-2.18	0.55	1.02	0.59-1.76	0.95	md	md	md	md	md	md
Smoking status	**Non-smokers (Ref) vs. Ex-smokers + Smokers**	1.57	0.60-4.12	0.36	1.62	0.73-3.62	0.24	1.57	0.86-2.86	0.14	1.69	0.97-2.93	0.07	1.44	0.75-2.77	0.28	1.56	0.81-3.00	0.18
TNM stage	**I (Ref) vs. II+III+IV**	3.69	1.30-10.49	**0.01**	2.41	1.08-5.37	**0.03**	*	*	0.11	1.02	0.25-4.22	0.98	3.36	0.81-13.95	0.09	1.69	0.60-4.73	0.32
	I+II (Ref) vs. III+IV	6.06	2.31-15.86	**0.0002**	3.52	1.51-8.22	**0.004**	3.50	1.72-7.15	**0.0006**	2.10	1.16-3.81	**0.01**	2.44	0.75-7.89	0.14	1.50	0.59-3.80	0.39
	I+II+III (Ref) vs. IV	27.46	9.31-80.97	**<0.0001**	10.75	3.16-36.52	**0.0001**	9.99	4.96-20.11	**<0.0001**	4.21	2.19-8.11	**<0.0001**	1.38	0.77-2.47	0.29	1.37	0.78-2.41	0.27
Histopathol. grade	**1 (Ref) vs. 2+3+4**	1.00	0.35-2.83	0.99	1.72	0.65-4.60	0.28	4.79	0.66-34.86	0.12	1.73	0.54-5.57	0.36	-	-	-	-	-	-
	1+2 (Ref) vs. 3+4	2.22	0.86-5.76	0.10	2.60	1.18-5.72	**0.02**	1.74	0.93-3.26	0.08	1.42	0.79-2.55	0.24	0.40	0.04-4.44	0.46	0.12	0.02-0.84	**0.03**
	1+2+3 (Ref) vs. 4	3.45	0.99-12.08	**0.05**	2.07	0.62-6.93	0.24	6.39	1.86-21.98	**0.003**	8.16	2.38-28.00	**0.0008**	5.34	1.05-27.12	**0.04**	2.36	0.57-9.79	0.24
Histological classification	**Ductal (Ref) vs. Lobular**	0.95	0.22-4.18	0.95	0.90	0.27-3.00	0.86	-	-	-	-	-	-	-	-	-	-	-	-
	Pulmonary (Ref) vs. Bronchogenic	-	-	-	-	-	-	-	-	-	-	-	-	1.72	1.03-2.89	**0.04**	1.70	1.02-2.82	**0.04**
	Non-small cell (Ref) vs. Small cell	-	-	-	-	-	-	-	-	-	-	-	-	0.92	0.39-2.17	0.85	0.86	0.39-1.94	0.72
Tumor location	**Right (Ref) vs. Left (BC, LC), Rectum (Ref) vs. Colon (CRC)**	1.89	0.70-5.11	0.21	1.87	0.83-4.23	0.13	0.60	0.33-1.09	0.09	1.79	1.03-3.10	**0.04**	1.79	1.06-3.01	**0.03**	1.85	1.12-3.08	**0.02**
Hormonal therapy	**Yes (Ref) vs. No**	3.74	1.37-10.17	**0.01**	3.85	1.64-9.07	**0.002**	-	-	-	-	-	-	-	-	-	-	-	-
Hormonal receptors	**ER+ (Ref) vs. ER-**	1.92	0.62-5.95	0.26	2.42	1.01-5.81	**0.05**	-	-	-	-	-	-	-	-	-	-	-	-
	PR+ (Ref) vs. PR-	2.84	1.05-7.65	**0.04**	2.59	1.16-5.78	**0.02**	-	-	-	-	-	-	-	-	-	-	-	-
	Her2+ (Ref) vs. Her2-	2.88	0.93-8.96	0.07	3.33	1.33-8.36	**0.01**	-	-	-	-	-	-	-	-	-	-	-	-

	Others (Ref) vs. ER-PR-Her2-	1.58	0.21-12.17	0.66	6.69	2.22-20.19	**0.0007**	-	-	-	-	-	-	-	-	-	-	-	-
	Others (Ref) vs. ER-PR-Her2+	4.00	1.14-14.03	**0.03**	2.40	0.72-8.05	0.16	-	-	-	-	-	-	-	-	-	-	-	-
ACs	**Under (Ref) vs. Over median value**	1.99	0.75-5.24	0.16	1.96	0.88-4.39	0.10	1.06	0.57-2.01	0.85	1.13	0.63-2.01	0.68	0.89	0.51-1.56	0.69	0.99	0.58-1.69	0.97
CAtot	**Under (Ref) vs. Over median value**	1.82	0.69-4.79	0.23	1.78	0.80-3.96	0.16	1.14	0.61-2.13	0.67	1.20	0.68-2.12	0.53	0.89	0.51-1.56	0.69	0.99	0.58-1.69	0.97
CTAs	**Under (Ref) vs. Over median value**	1.13	0.44-2.94	0.80	0.83	0.37-1.84	0.64	1.33	0.74-2.39	0.35	1.33	0.77-2.28	0.30	1.35	0.79-2.30	0.27	1.34	0.80-2.25	0.27
CSAs	**Under (Ref) vs. Over median value**	0.47	0.14-1.65	0.24	0.85	0.35-2.02	0.71	1.51	0.76-2.98	0.24	1.64	0.88-3.06	0.12	0.91	0.50-1.65	0.75	0.95	0.53-1.69	0.86
TL	**Under (Ref) vs. Over median value**	1.02	0.34-3.05	0.97	0.72	0.29-1.77	0.47	1.20	0.59-2.41	0.62	1.21	0.63-2.34	0.57	0.88	0.42-1.86	0.73	0.88	0.42-1.86	0.73

The table summarizes the results obtained by univariate assessment of the impact of well-established prognostic factors as well as ACs, CAtot, CTAs, CSAs and TL on overall (OS) and recurrence-free (RFS) survival expressed by Cox regression hazard model. Statistically significant values are in bold. aThe most suitable cut-offs for age (in years): OS and RFS (BC, CRC, LC) = 72,73,57, respectively. *Not calculated because 0 patients failed in stage I.

Abbreviations: ACs—aberrant cells; CAtot—total chromosomal aberrations; CI—confidence interval; CSAs—chromosome-type aberrations; CTAs—chromatid-type aberrations; ER+/-—α-estrogenic receptors positive/negative; Her2+/-—ErbB2 (Her2/neu) receptors positive/negative; HR—hazard ratio; md— missing data; N—number of samples; OS—overall survival; PR+/-—progesterone receptors positive/negative; RFS—recurrence-free survival; TL—telomere length TNM—tumor-node-metastasis.

2.1.1. Multivariate survival analyses

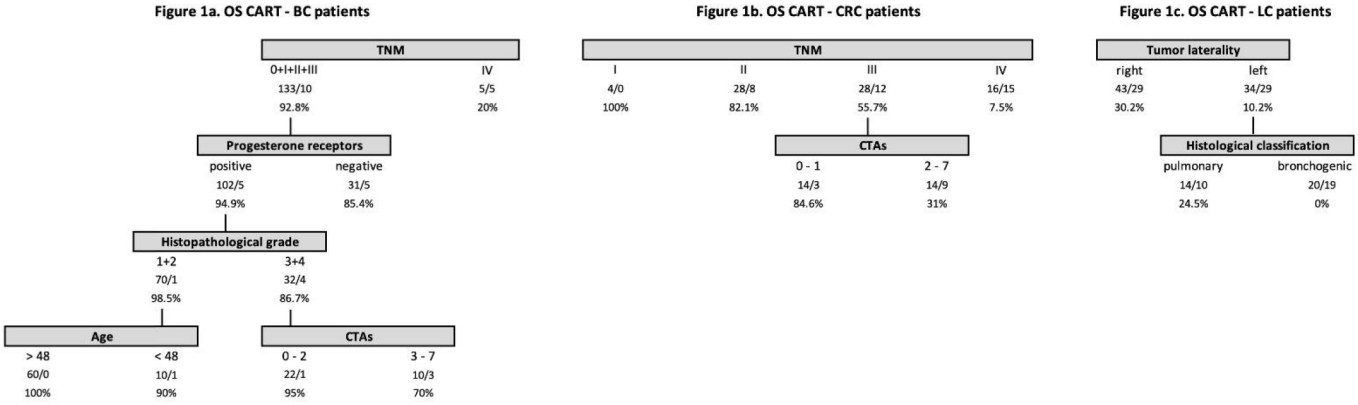
To assess the prognostic utility of all variables, we explored the interactive effects of all studied types of CAs and TL together in association with baseline and disease characteristics using a classification and regression tree (CART) analyses. The same covariates, as in the univariate analyses, were included in the multivariate model. The classification and regression trees related to both OS and RFS were generated for each group of cancer patients. The first split at the top of the tree denoted the most statistically significant prognostic factor. The percentages shown in the brackets below indicate the proportion of patients with 5-year survival. Our results indicated that TNM stage was the initial split-up factor for predicting both OS (Figure 1) and RFS (Figure 2) in BC and CRC patients, however, in LC patients the start-up point was at laterality of tumor.

In BC patients, more factors (besides TNM stage) played a key prognostic role in both OS and RFS. Variables determining the OS tree structure in BC patients were progesterone receptors, histopathological grade, age and CTAs (Figure 1a). Among TNM stage 0+I+II+III, the next split showed the interaction between progesterone positivity (positive: 94.9% vs. negative: 85.4%) and histopathological grade (grade 1+2: 98.5% vs. grade 3+4: 86.7%). Subsequently, those BC patients with grade 1+2 displayed better OS if they were over 48 years of age (>48 years: 100% vs. <48 years: 90%). On the other hand, BC patients with grade 3+4 had better OS provided their levels of CTAs were lower than 2 (0-2: 95% vs. 3-7: 70%). With regard to RFS, BC patients were divided into 2 categories after the initial split – TNM stage 0+I+II and TNM III+IV (Figure 2a). In TNM 0+I+II, the subsequent split was, again, progesterone receptors positivity/negativity (positive: 94.3% vs. negative: 65.1%) which interacted with Her2 (ErbB2 or Her2/neu) receptors positivity/negativity. The prognosis of BC patients who had negative progesterone receptors was improved by the presence of Her2 receptors (Her2+: 87.5% vs. Her2-: 50%). Similarly, this was the case in BC patients with progesterone positivity - the next most statistically significant factor in RFS was the presence of Her2 receptors (positive: 100% vs. negative: 88.8%). BC patients' RFS with Her2 negativity (after their stratification for progesterone positivity) was then influenced by age (<64 years: 90.8% vs. >64 years: 83.3%). Moreover, BC patients below 64 years of age showed CSA frequency as a terminal node of RFS CART analysis. BC patients without any CSAs displayed worse prognosis (0: 72.7% vs. 1-4: 100%).

The only statistically significant split after TNM stage was CTA frequency in CRC patients as well (Figure 1b); CRC patients in TNM stage III have experienced highly increased OS when they had CTA levels under the threshold of 1 (0-1: 84.6% vs. 2-7: 31%). The only statistically significant prognostic factor connected with RFS seemed to be TNM stage in our group of CRC patients (Figure 2b).

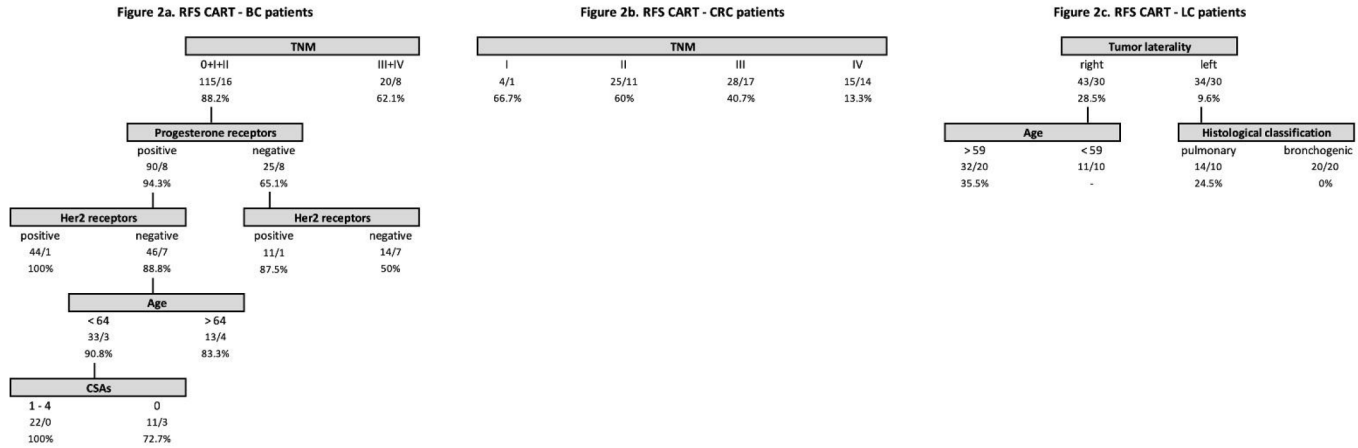
LC patients' OS was associated with tumor laterality and histological classification (Figure 1c); LC patients with left-sided tumors who also had a pulmonary type of tumor were more likely to survive 5 years than LC patients with a bronchogenic type of tumors (pulmonary: 24.5% vs. bronchogenic: 0%). Similarly to OS, LC patients with left-sided tumors and pulmonary histology had better RFS (pulmonary 24.5% vs. bronchogenic 0%, Figure 2c). Furthermore, LC patients with right-sided tumors demonstrated age-depending prognosis.

Figure 1. Overall survival (OS) classification and regression trees for BC (1a), CRC (1b), and LC (1c) patients.



Classification and regression trees represent the results of multivariate survival analysis (using Cox regression hazard model). Numbers under each node show the total number of cases in a particular subcategory/number of events and percentages of patients with 5-year OS.
Abbreviations: BC—breast cancer patients; CRC—colorectal cancer patients; CTAs—chromatid-type aberrations; LC—lung cancer patients; OS—overall survival; TNM—tumor-node-metastasis.

Figure 2. Recurrence-free survival (RFS) classification and regression trees for BC (2a), CRC (2b), and LC (2c) patients.



Classification and regression trees represent the results of multivariate survival analysis (using Cox regression hazard model). Numbers under each node show the total number of cases in a particular subcategory/number of events and percentages of patients with 5-year RFS.
Abbreviations: BC—breast cancer patients; CRC—colorectal cancer patients; CSAs—chromosome-type aberrations; LC—lung cancer patients; RFS—recurrence-free survival; TNM—tumor-node-metastasis.

3. Discussion

Both the accumulation of chromosomal damage and telomere shortening contributes to genome instability, including CIN [23], enabling the acquisition of the hallmarks of malignant diseases [24, 25]. The main purpose of our present study was to investigate the frequencies of particular types of chromosomal damage (ACs, CAtot, CTAs, and CSAs) and TL in PBL of newly diagnosed cancer patients and corresponding control individuals. Obtained results were analyzed together with collected baseline, clinico-pathological, and follow-up data. We sought to identify the impact of

increased CAs and shorter TL on the risk of the most common cancers (BC, CRC, and LC). To our best knowledge, no studies have previously addressed the relationship between frequencies of CAs with TL in PBL in both cancer patients and control individuals. For the first time, we also focused on the determination of possible prognostic utility of these biomarkers by assessment of patients' 5-year OS and RFS (with the use of univariate and multivariate models).

This report successively concludes our long-term cytogenetic studies that were launched in 2006 and initiated the collection of samples of cancer patients and control individuals. Initially, we reported the first retrospective case-control study where we have addressed the frequencies of chromosomal damage in 300 incident cancer patients matched with 300 controls [11]. In that study, the group of patients mainly consisted of individuals suffering from breast, colorectal, prostate, uterus + ovary, head + neck and bladder + kidney cancers. When studying the entire group of cancer patients and comparing it with control individuals, we found statistically significant differences in CA frequencies between these two groups. However, when we classified individual patients by tumor type, only patients with certain solid tumors differed in terms of CA frequencies from controls. For this reason, we further decided to focus on the most frequent BC, CRC and LC and extend the number of patients within the groups. In 2015, we described increased levels of chromosomal damage markers in BC, CRC and LC patients in comparison with control individuals [12]. In our present study, we conducted complex regression analyses to determine the predictive value for each type of CAs. Despite the fact that we calculated statistically significant aORs for particular types of CAs in our groups of cancer patients compared to control individuals, only the aORs over 3.0 may be considered as clinically significant. From the clinical point of view, only higher frequencies of ACs and CAtot (in the case of BC and LC) and CTAs (in the case of LC) influenced the risk of these cancers. Specifically, patients with elevated ACs and CAtot were at higher risk of BC by 48% and 45%, respectively) and LC by 38% and 34%, respectively). Higher levels of CTAs increased the risk of LC by 33%. Our results confirmed the findings from prospective studies (summarized by Bonassi *et al.* 2008 [10]). A cohort study from Central Europe suggested that rather than CTAs, CSAs are predictors of cancer risk [14]. However, a pooled analysis of thousands of control individuals revealed an equally strong cancer predictivity of both CTAs and CSAs [9].

The examination of CA frequencies in cancer patients and cancer-free individuals was newly supplemented by TL determination. Telomeres are nucleoprotein structures that protect chromosomal ends from exonucleolytic degradation and inappropriate activation of DNA repair pathways followed by end-to-end fusion of non-homologous chromosomes [26]. Li *et al.* 2013 [15] and Xu *et al.* 2013 [16] provided the evidence that CAs may arise as a consequence of telomere dysfunction and mechanisms associated with telomere shortening rather than as a result of the direct DNA damage. Moreover, sub-telomeric regions seem to be the hot spots for the formation of symmetrical exchanges between homologous chromatids. Cryptic aberrations in these regions have been shown to be associated with several human congenital abnormalities [4]. By comparison of TL between cases and controls, we found that BC patients exhibited statistically significantly longer TL than control females. Moreover, a group of BC patients had the longest TL out of all measured samples (both CRC and LC patients and all control individuals). This observation may have a possible explanation as it is widely believed that females have longer telomeres than males [27] and our set of BC patients consisted of only women. Several hypotheses have been postulated to clarify this association, one of which suggested that this is caused by the presence of estrogen [28]. An estrogen-responsive element is present in telomerase reverse transcriptase [29], which is a catalytic subunit of the enzyme telomerase. Telomerase is responsible for extending the telomeric DNA at the very tips of linear chromosomes and thus preventing the loss of genetic material in proliferating cells [18]. Estrogen may, therefore, stimulate telomerase to add telomere repeats to the chromosome ends. It is also known that overexpression of estrogen is one of the typical features of breast tumors and predispose the risk of BC. We proved this claim in our set of BC patients by discovering that individuals with longer telomeres were at increased risk of BC by 65%. However, evidence based on the results from previous studies, both prospective and retrospective, shows that associating telomere length in PBL with BC risk is still conflicting.

Certain studies found an elevated risk among women with longer telomeres [30, 31], whereas other studies showed short telomere length to be associated with increased risk of BC [32-34]. The rest of these studies did not record any significant association [35-37].

In terms of the relationship between CAs and TL, Hemminki *et al.* 2015 [17] proved that shorter TL correlates with elevated levels of CAs in control individuals. In our present study, we observed independently the same trend of statistically significant negative correlation between ACs/CAtot/CSAs and TL in the entire group of controls. It is well recognized that telomeres progressively shorten with increased age as well. Initial TL can vary between individuals, while the rate of telomere shortening reflects replicative exhaustion during ageing [38] and in the absence of a mechanism to maintain telomeres, cells eventually undergo replicative senescence [20]. In our group of controls, we also found the dependence of TL on age as telomeres gradually shortened with rising age. Age was also associated with accumulation of CSAs in control individuals. We suppose that this may be related to the lower DNA repair rate in subjects with higher age. Nevertheless, with the exception of correlation between TL and age in control individuals, all the remaining statistically significant results should be treated with caution as the Spearman's rho were small. This means that even if the relationships may be statistically significant, they may not be biologically as important. In addition, Bernadotte *et al.* 2016 [38] proposed that, instead of the measurement of average TL, it is reasonable to consider the amount of shortened telomeres as an indicator of cellular and tissue proliferative potential. This is because cells do not exhaust their proliferative potential in a simultaneous manner. Interestingly, cancer patients did not exhibit any relationship between TL and any type of CAs. Since there is no available literature to discuss this phenomenon, we can just hypothesize that in complex and multifactorial diseases such as cancer, there are more interactive pathways that contribute to the formation of DNA damage and CIN. We propose that there is a need to conduct more studies to elucidate the association between CAs and TL in PBL of cancer patients. We also did not record any clear correlation between TL and age in cancer patients. This was in contrast with the results obtained from most studies covering this topic summarized by Xu *et al.* 2016 [39], which concluded that TL negatively correlated with age in cancer patients.

CAs in PBL are proven markers of exposure, however, the question of whether and how they participate in pathophysiology and reflect progression (stage) of the disease is still unresolved today [40]. The prognostic value of TL in PBL in cancer patients remains unclear as well. The hypothesis that different frequencies of CAs and TL variations are determinants of prognosis is plausible to explain the heterogeneity in clinical outcomes of patients suffering from cancer. This problem, therefore, deserved further attention. In our present study, we focused on the determination of the prognostic utility of CAs and TL in all groups of cancer patients. We did not observe any association between any type of CAs and OS or RFS using univariate survival analysis based on median value cut-offs. However, outputs obtained from the CART analysis showed involvement of CTAs in the determination of patients' survival/mortality such that elevated CTAs occurred as a terminal node for a poor OS prognosis for BC patients in TNM 0+I+II+III stages and for CRC patients in TNM III stage. In CRC, an increase of CTAs over 1 even decreased OS by approximately 55%. Interestingly, BC patients without any CSAs exhibited worse RFS. Based on our results, it seems that different types of CAs play a different role in both cancer risk and patients' survival. Numerous authors have explored the relationships between chromosomal rearrangements/CIN and the prognosis of patients suffering from leukemia and lymphoma (summarized by Tanaka *et al.* 2016 [41]). However, there are no available publications exploring this relationship in solid tumors and therefore we propose it is necessary to carry out larger studies on this topic. In terms of TL, we did not find any association of telomere shortening with OR and RFS in our groups of cancer patients, either using univariate or multivariate survival model. There are several studies that have been studying the effect of telomere length on cancer prognosis, both in PBL and tumor tissue. Recently published meta-analysis [39] composed from 29 studies which evaluated telomere length in blood revealed the association between shortened telomere length and poor OS and RFS in solid cancers and even stronger association in hematological malignancies. Nevertheless, these TL results obtained from blood were in contrast to

the data obtained by measurement of TL in tissues [42]. By comparing the outputs obtained from both univariate and multivariate analyses between individual groups of cancer patients, LC patients' survival was not affected by alteration of any type of the studied genetic markers. Even well-established TNM stage did not influence survival, however, this was probably due to the fact that these patients were mostly diagnosed in advanced stages (TNM III+IV).

4. Materials and Methods

4.1. Study population

This study was performed by pooling data from our cytogenetic studies carried out between 2006 and 2013. The study population consisted of three groups of newly diagnosed and histologically confirmed individuals with BC (N=151), CRC (N=96) and LC (N=90), and a group of control individuals (N=335). Patients were recruited at the 3rd Faculty of Medicine, Charles University, Prague, Czech Republic (BC, CRC) and at the Jessenius Faculty of Medicine, Comenius University and University Hospital, Martin, Slovakia (LC).

Control individuals (containing 145 control females for comparison with the group of female BC patients) of similar age and sex were obtained from the Blood Centre of Faculty Hospital Kralovske Vinohrady, Prague, Czech Republic and from the Department of Surgery, General University Hospital in Prague, Czech Republic. Characteristics of the study participants were previously described in details in Vodenkova *et al.* 2015 [12]. The present study adhered to the ethical guidelines as set out in the Helsinki Declaration - all participants were sufficiently informed about all aspects of the study, agreed with the study purpose and procedures to be undertaken and provided informed consent. The design of the study was approved by the local Ethics Committees of all participating hospitals.

4.2. Samples and data collection

Two tubes of blood were drawn from control individuals (without any personal cancer history) and incident (i.e. newly diagnosed) cancer patients. During the transportation of samples, blood was kept at 4°C until the processing and/or long-term storage. Fresh heparinized blood samples were immediately after transportation used for cytogenetic analysis of chromosomes and EDTA blood samples were frozen at -20°C for subsequent DNA isolation and TL measurement.

Baseline patient characteristics such as demographics, family history of cancer, smoking habit, occupational history, body mass index, and the presence of other diseases such as hypertension, diabetes mellitus, cardiovascular disease, including their treatment, were collected at the time of diagnosis using a structured questionnaire. Disease characteristics, including tumor location, TNM stage, histopathological grade, histological classification, and the presence of hormonal receptors and administration of hormonal therapy in breast tumors, were collected after surgical resection (detailed in Vodenkova *et al.* 2015 [12]). The last update of patients' follow-up for this study was in July 2017.

The exclusion criteria for cancer patients were a personal history of any previous malignancy, any prior use of radio- or chemotherapy to avoid the presence of additional chromosomal damage related to the treatment, any prior use of X-rays (e.g. diagnostic) and any hereditary forms of cancer. For survival analyses, we further excluded patients who either had missing data about living status below 5 years, underwent a 2nd surgery for an unknown reason or had experienced duplicate tumor. As a consequence, our final set of patients was reduced by 4 LC patients. The entire data set contained 208 patients who were still alive and 125 patients who have died as a consequence of particular cancer.

4.3. Chromosomal analysis

Levels of structural CAs in PBL were evaluated using the standard cytogenetic method, as described by Vodicka *et al.* 2010 [11]. Briefly, blood cultivation was carried out in complete medium (Chromosome medium P - EKAMTB-100, EuroCloneS.p.A., Italy) for 50 hours at 37°C. After 48 hours of cultivation, cell division was stopped by colchicine (Sigma, USA) in the first metaphase of mitosis. After a cytogenetic procedure, microscopic slides were stained by conventional Giemsa staining

(Sigma, USA). Microscopical analysis of 100 metaphases with 46 ± 1 chromosomes has been blindly performed by two independent scorers. The percentages of ACs, CAtot, CTAs (i.e. chromatid breaks and exchanges), and CSAs (i.e. chromosome breaks, terminal and interstitial deletions, dicentric and ring chromosomes with their difragments, abnormal chromosomes) were detected.

With regard to CAs scoring, standardization procedure has been applied in former Czechoslovakia (and later in both separate countries) [43]. In addition, to minimize inter-laboratory and inter-scorer differences in the results, random exchanges of microscope slides for standardization between the two laboratories in Prague and Martin have been a part of the standard operating procedure.

4.4. Measurement of TL

TL was measured as RTL and its measurement was conducted using the monochrome multiplex polymerase chain reaction (PCR) assay previously described by Cawthon 2009 [44] with the implementation of slight modifications [45, 46]. All details about standard and calibration curves, DNA concentrations, negative and quality controls, master mix, conditions for telomere sequence and albumin gene amplification were recently published in Kroupa *et al.* 2018 [47]. All reactions were performed in triplicates in an optical 384-well reaction plate. Real-time PCR experiments were carried out on Viia 7 Real-time PCR System (Applied Biosystems, USA) with the use of two simultaneous programs to acquire the respective cycle threshold values for telomere sequences and the albumin gene. RTL was expressed as the ratio (T/S ratio) between telomere (T) and albumin (S; single copy gene).

4.5. Statistical analysis

Obtained results and collected baseline, clinico-pathological and follow-up data were analyzed using statistical software SAS (SAS Institute Inc., USA). The results were processed into graphs using SW Statistics (StatSoft, Inc., USA). Statistical significance of all tests was set at α = 5% threshold (P-value = 0.05). Descriptive statistical analysis was carried out for the measured parameters on the whole data set as well as on individual groups.

The differences in the levels of all investigated parameters between controls and patients were tested by nonparametric analysis of variance (Median Two-Sample Test) or Chi-square test. The relationships between the examined variables (all types of CAs, TL, and age) were investigated using Spearman's correlation, expressed by Spearman's rho and graphically plotted by linear regression. For stratification of subjects with lower or higher CAs and TL, a median value of control individuals was used. The effect of each parameter on the risk of studied cancer types was determined by logistic regression and was calculated by estimating the odds ratios (ORs) with the 95% CIs. Additionally, ORs were adjusted for age and sex (aORs) as the frequencies of some cytogenetic markers increase with age and female gender [48]. BC patients and their controls contained only women and therefore were only adjusted for age.

Clinical outcomes were evaluated by calculating patients' 5-year OS and RFS. OS was defined as the time from the date of diagnosis to the date of the patient's death – defined as a negative event or last follow-up (July 2017) – defined as censored data. RFS was calculated from the date of diagnosis to the date of occurrence of local recurrence, distant metastasis, or death, whichever came first - defined as a negative event, time to last follow-up date (patients without event) was defined as censored data. The relative risk of death and recurrence was estimated as a hazard ratio (HR) with the 95% CIs, with the use of Cox regression

Moreover, we performed a multivariate analysis, referred to as a CART [49] using the Cox regression model to identify the most prognostically significant interactions between investigated factors and patients' 5-year OS and RFS. Covariates used in the CART were baseline and clinico-pathological factors, such as age, sex, family history of cancer, smoking status, TNM stage, histopathological grade, histological classification, invasiveness, tumor location, type of subsequent therapy received and presence of hormonal receptors in breast tumors, as well as all types of CAs and

TL. The analysis comprised a set of decision rules that stratify consistent risk groups for the responsive variable. Splits for each variable were examined and the variable that provides the best split was selected. Each subgroup was further divided in the same manner.

5. Conclusions

In summary, the results of this study suggest that increased frequency of ACs, CAtot and CTAs in PBL may be considered as a marker of BC and LC risk. We discovered that individuals with longer TL in PBL were at increased risk of BC. Accumulation of CTAs in PBL was associated with poor OS in BC and CRC patients after their stratification according to disease characteristics when a multivariate survival model was used. Neither OS nor RFS was influenced by TL alterations. Interestingly, in contrast to control individuals, cancer patients did not exhibit any relationship between either TL and CA frequencies or TL and age. We are aware that our present study has certain limitations, particularly in relation to small sample sizes included in our analyses. This is probably one of the reasons that led to reduced statistical power. Moreover, outputs acquired by CART analysis should be treated with caution as only individuals with full data in all implied variables were included in the procedure which could have introduced a selection bias. On the other hand, our study is based on validated and well-established methods and well-defined clinical data. In conclusion, consideration of the increased frequency of chromosomal damage and TL in surrogate tissues (such as blood) as markers of cancer risk and progression is a topic attracting increasing interest due to its potential applications in clinical oncology and public health. This is confirmed by the dramatically increasing number of published papers. Therefore, additional studies with larger cohorts and consistent data on CAs and TL are warranted.

Supplementary Materials: Supplementary materials can be found at www.mdpi.com/xxx/s1.

Author Contributions: S.V., K.H., R.K.[2] and P.V. conceived the study; L.M., M.A., and M.S. provided hospital-based samples; R.K.[1], L.M., and L.V. performed patients' data collection; S.V. performed data curation; S.V., Z.P., and L.M. performed chromosomal analysis; M.K. and S.R. performed measurement of TL; S.V. performed data analyses and interpretation of the results with assistance from M.K., K.H., R.K.[2], and P.V.; S.V. wrote and edited the manuscript; Z.P., L.M., L.V., K.H., R.K.[2], and P.V. revised the manuscript; Z.P., R.K.[2] and P.V. supervised the project; L.V. administrated the project. All of the authors read and approved the final manuscript for publication.

Funding: This research was funded by the Grant Agency of the Charles University, GAUK 800216 (S.V.); the Grant Agency of the Czech Republic, GACR 15-14789S (P.V.); the Internal Grant Agency of the Ministry of Health of the Czech Republic, AZV 15-27580A (P.V.); the Research program of Charles University, PROGRES Q 28 Oncology (Z.P., P.V., L.V.); the project Biomedical Center Martin Slovakia co-financed from EU sources, ITMS 26220220187 (L.M.), the Charles University Research Centre program UNCE/MED/006 "University Center of Clinical and Experimental Liver Surgery" (P.V., L.V.); and the National Sustainability Program I (NPU I) provided by the Ministry of Education Youth and Sports of the Czech Republic, LO1503 (P.V., L.V.).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ACs	Aberrant cells
aOR	Adjusted odds ratio
BC	Breast cancer
CART	Classification and regression tree
CAs	Chromosomal aberrations
CAtot	Total chromosomal aberrations
CI	Confidence interval
CIN	Chromosomal instability
CRC	Colorectal cancer
CSAs	Chromosome-type aberrations
CTAs	Chromatid-type aberrations

ER	α-estrogenic receptors
Her2	ErbB2 (Her2/neu) receptors
HR	Hazard ratio
LC	Lung cancer
md	Missing data
N	Number of samples
OR	Odds ratio
OS	Overall survival
PBL	Peripheral blood lymphocytes
PCR	Polymerase chain reaction
PR	Progesterone receptors
RFS	Recurrence-free survival
rs	Spearman's rho
RTL	Relative telomere length
S	Single copy gene
SD	Standard deviation
T	Telomere
TL	Telomere length
TNM	Tumor-node-metastasis

References

1. Schvartzman, J. M.; Sotillo, R.; Benezra, R., Mitotic chromosomal instability and cancer: mouse modelling of the human disease. *Nat Rev Cancer* **2010,** 10, (2), 102-15.
2. Hasty, P.; Montagna, C., Chromosomal Rearrangements in Cancer: Detection and potential causal mechanisms. *Mol Cell Oncol* **2014,** 1, (1).
3. Chatterjee, A.; Rodger, E. J.; Eccles, M. R., Epigenetic drivers of tumourigenesis and cancer metastasis. *Semin Cancer Biol* **2018,** 51, 149-159.
4. Obe, G.; Pfeiffer, P.; Savage, J. R.; Johannes, C.; Goedecke, W.; Jeppesen, P.; Natarajan, A. T.; Martinez-Lopez, W.; Folle, G. A.; Drets, M. E., Chromosomal aberrations: formation, identification and distribution. *Mutat Res* **2002,** 504, (1-2), 17-36.
5. Natarajan, A. T.; Palitti, F., DNA repair and chromosomal alterations. *Mutat Res* **2008,** 657, (1), 3-7.
6. Bignold, L. P., Mechanisms of clastogen-induced chromosomal aberrations: a critical review and description of a model based on failures of tethering of DNA strand ends to strand-breaking enzymes. *Mutat Res* **2009,** 681, (2-3), 271-98.
7. Mateuca, R. A.; Decordier, I.; Kirsch-Volders, M., Cytogenetic methods in human biomonitoring: principles and uses. *Methods Mol Biol* **2012,** 817, 305-34.
8. Musak, L.; Smerhovsky, Z.; Halasova, E.; Osina, O.; Letkova, L.; Vodickova, L.; Polakova, V.; Buchancova, J.; Hemminki, K.; Vodicka, P., Chromosomal damage among medical staff occupationally exposed to volatile anesthetics, antineoplastic drugs, and formaldehyde. *Scand J Work Environ Health* **2013,** 39, (6), 618-30.
9. Hagmar, L.; Stromberg, U.; Bonassi, S.; Hansteen, I. L.; Knudsen, L. E.; Lindholm, C.; Norppa, H., Impact of types of lymphocyte chromosomal aberrations on human cancer risk: results from Nordic and Italian cohorts. *Cancer Res* **2004,** 64, (6), 2258-63.
10. Bonassi, S.; Norppa, H.; Ceppi, M.; Stromberg, U.; Vermeulen, R.; Znaor, A.; Cebulska-Wasilewska, A.; Fabianova, E.; Fucic, A.; Gundy, S.; Hansteen, I. L.; Knudsen, L. E.; Lazutka, J.; Rossner, P.; Sram, R. J.; Boffetta, P., Chromosomal aberration frequency in lymphocytes predicts the risk of cancer: results from a pooled cohort study of 22 358 subjects in 11 countries. *Carcinogenesis* **2008,** 29, (6), 1178-83.
11. Vodicka, P.; Polivkova, Z.; Sytarova, S.; Demova, H.; Kucerova, M.; Vodickova, L.; Polakova, V.; Naccarati, A.; Smerhovsky, Z.; Ambrus, M.; Cerna, M.; Hemminki, K., Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls. *Carcinogenesis* **2010,** 31, (7), 1238-41.
12. Vodenkova, S.; Polivkova, Z.; Musak, L.; Smerhovsky, Z.; Zoubkova, H.; Sytarova, S.; Kavcova, E.; Halasova, E.; Vodickova, L.; Jiraskova, K.; Svoboda, M.; Ambrus, M.; Hemminki, K.; Vodicka, P., Structural chromosomal aberrations as potential risk markers in incident cancer patients. *Mutagenesis* **2015,** 30, (4), 557-63.

13. Albertson, D. G.; Collins, C.; McCormick, F.; Gray, J. W., Chromosome aberrations in solid tumors. *Nat Genet* **2003,** 34, (4), 369-76.
14. Boffetta, P.; van der Hel, O.; Norppa, H.; Fabianova, E.; Fucic, A.; Gundy, S.; Lazutka, J.; Cebulska-Wasilewska, A.; Puskailerova, D.; Znaor, A.; Kelecsenyi, Z.; Kurtinaitis, J.; Rachtan, J.; Forni, A.; Vermeulen, R.; Bonassi, S., Chromosomal aberrations and cancer risk: results of a cohort study from Central Europe. *Am J Epidemiol* **2007,** 165, (1), 36-43.
15. Li, H.; Hilmarsen, H. T.; Hossain, M. B.; Bjork, J.; Hansteen, I. L.; Albin, M.; Furu Skjelbred, C.; Broberg, K., Telomere length and LINE1 methylation is associated with chromosomal aberrations in peripheral blood. *Genes Chromosomes Cancer* **2013,** 52, (1), 1-10.
16. Xu, L.; Li, S.; Stohr, B. A., The role of telomere biology in cancer. *Annu Rev Pathol* **2013,** 8, 49-78.
17. Hemminki, K.; Rachakonda, S.; Musak, L.; Vymetalkova, V.; Halasova, E.; Forsti, A.; Vodickova, L.; Buchancova, J.; Vodicka, P.; Kumar, R., Telomere length in circulating lymphocytes: Association with chromosomal aberrations. *Genes Chromosomes Cancer* **2015,** 54, (3), 194-6.
18. Heidenreich, B.; Kumar, R., TERT promoter mutations in telomere biology. *Mutat Res* **2017,** 771, 15-31.
19. Maciejowski, J.; de Lange, T., Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Biol* **2017,** 18, (3), 175-186.
20. Shay, J. W., Role of Telomeres and Telomerase in Aging and Cancer. *Cancer Discov* **2016,** 6, (6), 584-93.
21. Ernst, A.; Jones, D. T.; Maass, K. K.; Rode, A.; Deeg, K. I.; Jebaraj, B. M.; Korshunov, A.; Hovestadt, V.; Tainsky, M. A.; Pajtler, K. W.; Bender, S.; Brabetz, S.; Grobner, S.; Kool, M.; Devens, F.; Edelmann, J.; Zhang, C.; Castelo-Branco, P.; Tabori, U.; Malkin, D.; Rippe, K.; Stilgenbauer, S.; Pfister, S. M.; Zapatka, M.; Lichter, P., Telomere dysfunction and chromothripsis. *Int J Cancer* **2016,** 138, (12), 2905-14.
22. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; Jemal, A., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **2018,** 68, (6), 394-424.
23. Gisselsson, D., Chromosome instability in cancer: how, when, and why? *Adv Cancer Res* **2003,** 87, 1-29.
24. Negrini, S.; Gorgoulis, V. G.; Halazonetis, T. D., Genomic instability--an evolving hallmark of cancer. *Nat Rev Mol Cell Biol* **2010,** 11, (3), 220-8.
25. Hanahan, D.; Weinberg, R. A., Hallmarks of cancer: the next generation. *Cell* **2011,** 144, (5), 646-74.
26. Gay-Bellile, M.; Romero, P.; Cayre, A.; Veronese, L.; Privat, M.; Singh, S.; Combes, P.; Kwiatkowski, F.; Abrial, C.; Bignon, Y. J.; Vago, P.; Penault-Llorca, F.; Tchirkov, A., ERCC1 and telomere status in breast tumours treated with neoadjuvant chemotherapy and their association with patient prognosis. *J Pathol Clin Res* **2016,** 2, (4), 234-246.
27. Gardner, M.; Bann, D.; Wiley, L.; Cooper, R.; Hardy, R.; Nitsch, D.; Martin-Ruiz, C.; Shiels, P.; Sayer, A. A.; Barbieri, M.; Bekaert, S.; Bischoff, C.; Brooks-Wilson, A.; Chen, W.; Cooper, C.; Christensen, K.; De Meyer, T.; Deary, I.; Der, G.; Diez Roux, A.; Fitzpatrick, A.; Hajat, A.; Halaschek-Wiener, J.; Harris, S.; Hunt, S. C.; Jagger, C.; Jeon, H. S.; Kaplan, R.; Kimura, M.; Lansdorp, P.; Li, C.; Maeda, T.; Mangino, M.; Nawrot, T. S.; Nilsson, P.; Nordfjall, K.; Paolisso, G.; Ren, F.; Riabowol, K.; Robertson, T.; Roos, G.; Staessen, J. A.; Spector, T.; Tang, N.; Unryn, B.; van der Harst, P.; Woo, J.; Xing, C.; Yadegarfar, M. E.; Park, J. Y.; Young, N.; Kuh, D.; von Zglinicki, T.; Ben-Shlomo, Y.; Halcyon study, t., Gender and telomere length: systematic review and meta-analysis. *Exp Gerontol* **2014,** 51, 15-27.
28. Mayer, S.; Bruderlein, S.; Perner, S.; Waibel, I.; Holdenried, A.; Ciloglu, N.; Hasel, C.; Mattfeldt, T.; Nielsen, K. V.; Moller, P., Sex-specific telomere length profiles and age-dependent erosion dynamics of individual chromosome arms in humans. *Cytogenet Genome Res* **2006,** 112, (3-4), 194-201.
29. Nawrot, T. S.; Staessen, J. A.; Gardner, J. P.; Aviv, A., Telomere length and possible link to X chromosome. *Lancet* **2004,** 363, (9408), 507-10.
30. Svenson, U.; Nordfjall, K.; Stegmayr, B.; Manjer, J.; Nilsson, P.; Tavelin, B.; Henriksson, R.; Lenner, P.; Roos, G., Breast cancer survival is associated with telomere length in peripheral blood cells. *Cancer Res* **2008,** 68, (10), 3618-23.
31. Gramatges, M. M.; Telli, M. L.; Balise, R.; Ford, J. M., Longer relative telomere length in blood from women with sporadic and familial breast cancer compared with healthy controls. *Cancer Epidemiol Biomarkers Prev* **2010,** 19, (2), 605-13.
32. Levy, T.; Agoulnik, I.; Atkinson, E. N.; Tong, X. W.; Gause, H. M.; Hasenburg, A.; Runnebaum, I. B.; Stickeler, E.; Mobus, V. J.; Kaplan, A. L.; Kieback, D. G., Telomere length in human white blood cells remains constant with age and is shorter in breast cancer patients. *Anticancer Res* **1998,** 18, (3A), 1345-9.
33. Pooley, K. A.; Sandhu, M. S.; Tyrer, J.; Shah, M.; Driver, K. E.; Luben, R. N.; Bingham, S. A.; Ponder, B. A.; Pharoah, P. D.; Khaw, K. T.; Easton, D. F.; Dunning, A. M., Telomere length in prospective and retrospective cancer case-control studies. *Cancer Res* **2010,** 70, (8), 3170-6.

34. Qu, S.; Wen, W.; Shu, X. O.; Chow, W. H.; Xiang, Y. B.; Wu, J.; Ji, B. T.; Rothman, N.; Yang, G.; Cai, Q.; Gao, Y. T.; Zheng, W., Association of leukocyte telomere length with breast cancer risk: nested case-control findings from the Shanghai Women's Health Study. *Am J Epidemiol* **2013,** 177, (7), 617-24.
35. Shen, J.; Terry, M. B.; Gurvich, I.; Liao, Y.; Senie, R. T.; Santella, R. M., Short telomere length and breast cancer risk: a study in sister sets. *Cancer Res* **2007,** 67, (11), 5538-44.
36. Shen, J.; Gammon, M. D.; Terry, M. B.; Wang, Q.; Bradshaw, P.; Teitelbaum, S. L.; Neugut, A. I.; Santella, R. M., Telomere length, oxidative damage, antioxidants and breast cancer risk. *Int J Cancer* **2009,** 124, (7), 1637-43.
37. Zheng, Y. L.; Ambrosone, C.; Byrne, C.; Davis, W.; Nesline, M.; McCann, S. E., Telomere length in blood cells and breast cancer risk: investigations in two case-control studies. *Breast Cancer Res Treat* **2010,** 120, (3), 769-75.
38. Bernadotte, A.; Mikhelson, V. M.; Spivak, I. M., Markers of cellular senescence. Telomere shortening as a marker of cellular senescence. *Aging (Albany NY)* **2016,** 8, (1), 3-11.
39. Xu, X.; Qu, K.; Pang, Q.; Wang, Z.; Zhou, Y.; Liu, C., Association between telomere length and survival in cancer patients: a meta-analysis and review of literature. *Front Med* **2016,** 10, (2), 191-203.
40. Venkatesan, S.; Natarajan, A. T.; Hande, M. P., Chromosomal instability--mechanisms and consequences. *Mutat Res Genet Toxicol Environ Mutagen* **2015,** 793, 176-84.
41. Tanaka, K.; Hirota, T., Chromosomal instability: A common feature and a therapeutic target of cancer. *Biochim Biophys Acta* **2016,** 1866, (1), 64-75.
42. Ennour-Idrissi, K.; Maunsell, E.; Diorio, C., Telomere Length and Breast Cancer Prognosis: A Systematic Review. *Cancer Epidemiol Biomarkers Prev* **2017,** 26, (1), 3-10.
43. Rossner, P.; Cerna, M.; Bavorova, H.; Pastorkova, A.; Ocadlikova, D., Monitoring of human exposure to occupational genotoxicants. *Cent Eur J Public Health* **1995,** 3, (4), 219-23.
44. Cawthon, R. M., Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res* **2009,** 37, (3), e21.
45. Heidenreich, B.; Rachakonda, P. S.; Hosen, I.; Volz, F.; Hemminki, K.; Weyerbrock, A.; Kumar, R., TERT promoter mutations and telomere length in adult malignant gliomas and recurrences. *Oncotarget* **2015,** 6, (12), 10617-33.
46. Hosen, I.; Rachakonda, P. S.; Heidenreich, B.; de Verdier, P. J.; Ryk, C.; Steineck, G.; Hemminki, K.; Kumar, R., Mutations in TERT promoter and FGFR3 and telomere length in bladder cancer. *Int J Cancer* **2015,** 137, (7), 1621-9.
47. Kroupa, M.; Polivkova, Z.; Rachakonda, S.; Schneiderova, M.; Vodenkova, S.; Buchler, T.; Jiraskova, K.; Urbanova, M.; Vodickova, L.; Hemminki, K.; Kumar, R.; Vodicka, P., Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients. *Genes Chromosomes Cancer* **2018,** 57, (2), 61-69.
48. Fenech, M., Important variables that influence base-line micronucleus frequency in cytokinesis-blocked lymphocytes-a biomarker for DNA damage in human populations. *Mutat Res* **1998,** 404, (1-2), 155-65.
49. Lemon, S. C.; Roy, J.; Clark, M. A.; Friedmann, P. D.; Rakowski, W., Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. *Ann Behav Med* **2003,** 26, (3), 172-81.

Manuscript V

Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients.

Kroupa M, Polivkova Z, Rachakonda S, Schneiderova M, Vodenkova S, Buchler T, Jiraskova K, Urbanova M, Vodickova L, Hemminki K, Kumar R, Vodicka P.

Genes Chromosomes Cancer. 2018 Feb;57(2):61-69. **IF: 3.362**

Received: 30 June 2017 | Revised: 2 October 2017 | Accepted: 11 October 2017

DOI: 10.1002/gcc.22508

RESEARCH ARTICLE

WILEY

Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients

Michal Kroupa[1,2] | Zdenka Polivkova[3] | Sivaramakrishna Rachakonda[4] | Michaela Schneiderova[5] | Sona Vodenkova[2,3,7] | Tomas Buchler[6] | Katerina Jiraskova[2,7] | Marketa Urbanova[2,7] | Ludmila Vodickova[1,2,7] | Kari Hemminki[4] | Rajiv Kumar[4] | Pavel Vodicka[1,2,7]

[1]Faculty of Medicine and Biomedical Center in Pilsen, Charles University, Pilsen, 30605, Czech Republic

[2]Department of Molecular Biology of Cancer, Institute of Experimental Medicine, The Czech Academy of Sciences, Prague, 14220, Czech Republic

[3]Department of Medical Genetics, Third Faculty of Medicine, Charles University, Prague, 10000, Czech Republic

[4]Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, 69120, Germany

[5]Department of Surgery, General University Hospital in Prague, Prague, 12800, Czech Republic

[6]Department of Oncology, First Faculty of Medicine, Charles University and Thomayer Hospital, Prague, 14059, Czech Republic

[7]Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, 12800, Czech Republic

Correspondence
Pavel Vodicka, Department of Molecular Biology of Cancer, Institute of Experimental Medicine, The Czech Academy of Sciences, 14220 Prague, Czech Republic.
Email: pvodicka@biomed.cas.cz

Funding information
National Science Foundation, Grant Number: 15-14789S; Czech Ministry of Health, Grant Number: 15-27580A; Univerzita Karlova v Praze, Grant Numbers: COST LD-14050, SVV 260390, UNCE 204022, and PROGRES Q 28; Lékařská Fakulta v Plzni, Univerzita Karlova, Grant Number: SVV 260 390; European Cooperation in Science and Technology, Grant Number: LD-14050; TRANSCAN through German Ministry of Education and Research (BMBF), Grant number: 01KT15511

Abstract

Disruption of genomic integrity due to deficient DNA repair capacity and telomere shortening constitute hallmarks of malignant diseases. Incomplete or deficient repair of DNA double-strand breaks (DSB) is manifested by chromosomal aberrations and their frequency reflects inter-individual differences of response to exposure to mutagenic compounds. In this study, we investigated chromosomal integrity in peripheral blood lymphocytes (PBL) from newly diagnosed cancer patients, including 47 breast (BC) and 44 colorectal cancer (CRC) patients and 90 matched healthy controls. Mutagen sensitivity was evaluated by measuring chromatid breaks (CTAs) induced by bleomycin and supplemented by the chemiluminescent measurement of γ-H2AX phosphorylation in 19 cancer patients (11 BC, 8 CRC). Relative telomere length (RTL) was determined in 22 BC, 32 CRC, and 64 controls. We observed statistically significant increased level of CTAs ($P = .03$) and increased percentage of aberrant cells (ACs) with CTAs ($P = .05$) in CRC patients compared with controls after bleomycin treatment. No differences were observed between BC cases and corresponding controls. CRC and BC patients with shorter RTL (below median) exhibited significantly higher amount of ACs ($P = .02$), CTAs ($P = .02$), and cells with high frequency of CTAs (≥12 CTAs/PBL; $P = .03$) after bleomycin treatment. No such associations were observed in healthy controls. γ-H2AX phosphorylation after bleomycin treatment in PBL did not differ between CRC and BC patients. Our results suggest that altered DSB repair measured by sensitivity towards mutagen in PBL occurs particularly in CRC carcinogenesis. Irrespective of cancer type, telomere shortening may be associated with a decreased capacity to repair DSB.

1 | INTRODUCTION

Tumor development, a multistep process, is driven through the accumulation of genetic and epigenetic changes. Chromosomal aberrations (CAs) play a crucial role in the carcinogenic process.[1] CAs may directly trigger tumorigenesis through rearrangements leading to increased proto-oncogene expression or alternation of its structure as in Burkitt's lymphoma and chronic myeloid leukemia. Other CAs, like chromosomal losses or gains, can be either as causative events in malignant disease or as secondary consequence. Elevated frequency of CAs has been consistently described in various prospective studies as a marker and a predictor of subsequent malignancies. Increased CAs have also been described in newly diagnosed cancer patients in retrospective studies.[1–3]

All CAs arise as a consequence of unrepaired or insufficiently repaired DNA double-strand breaks (DSB). Dependent on the cell cycle stage and type of DNA damage, chromatid-type (CTAs) or chromosome-type aberrations (CSAs) occur in somatic cells: CSAs originate as a consequence of DSB generated by genotoxic damage during G0 or G1 phase, whereas CTAs arise during late S or G2 phase of the cell cycle mainly due to the insufficiently repaired DNA breaks.[4] CAs can ultimately result in chromosome deletions, translocations, and gene amplifications observed frequently in human cancers.

Mechanisms associated with telomere shortening are important contributors to formation of CAs. Telomeres are nucleoprotein structures composed of TTAGGG tandem repeats and associated proteins that cap ends of eukaryotic chromosomes. Telomere complex regulates a "cellular mitotic clock" and protects chromosomes against exonucleolytic degradation, DNA damage and chromosomal instability.[5–7] Recent studies report shorter telomeres associated with higher frequency of CAs in peripheral blood.[8,9] Telomere shortening has been for long hypothesized to be associated with cancer development,[10,11] primarily through the induction of genomic and chromosomal instability.[12,13]

Telomere shortens at each round of replication and critically shortened telomeres may be poorly end-capped and recognized as DSB by repair machinery.[14] These processes lead to end-to-end fusions of chromosomes, initiation of breakage-fusion-breakage cycles and development of numerical and structural CAs.[15–17] Reactivation of telomerase leads to stabilization of telomeres and immortalization of premalignant cells, thus enabling cancer progression.[18,19]

Mutagen sensitivity can serve as functional biomarker of DNA repair capacity. Mutagen sensitivity is measured in cells (usually peripheral blood lymphocytes; PBL) from subjects with disease and from healthy subjects after induction of DNA damage by chemical or physical agents.[20,21] Such challenge assay is based on the determination of CTAs.[22–27] Increased amount of CTAs usually indicates an impaired ability to repair DSB.[28]

The main aim of this exploratory study was to employ the mutagen sensitivity for assessment of the potential interactions between induced DSB in PBL and cancer predisposition, particularly to breast (BC) and colorectal (CRC) cancer. Inter-individual variation in DNA repair capacity was evaluated in vitro utilizing the G2 chromosomal mutagen sensitivity assay to quantify bleomycin-induced CTAs.[22] Furthermore, we also measured relative telomere length (RTL) and compared mutagen sensitivity with γ-H2AX phosphorylation which is a hallmark of DSB.[29]

This study is the first to link telomere shortening and the radiomimetic effect of bleomycin in PBL of cancer patients and raises the possibility that telomere shortening may be a proxy for underlying mutagen inter-individual sensitivity.

Above parameters followed in cancer patients may also be informative about the disease prognosis and prediction of the therapy.

2 | MATERIALS AND METHODS

2.1 | Study population

Blood samples from newly diagnosed, untreated, and histologically confirmed BC and CRC patients were collected at Thomayer Hospital and General University Hospital, both in Prague, Czech Republic. All patients signed informed consent. Sampling of peripheral blood was carried out according to the Helsinki Declaration and the study design was approved by the appropriate Ethics committees. In all participating cancer patients basic clinico-pathological characteristics were collected.

The control study population was made up of hospital healthy volunteers and healthy individuals without any previous cancer disease recruited between 2012 and 2015 in central Bohemia (Faculty Hospital Kralovske Vinohrady and Department of Surgery, General University Hospital, both in Prague). The control subjects enrolled in the study filled in a questionnaire, in which they declared an occurrence of serious diseases (malignancies, cardiovascular, and diabetes). Any individual apparently suffering from above diseases were excluded from the study. Peripheral blood sampling was preferentially carried out only in the subjects apparently healthy at the time of sampling.

Formal matching for age and sex was performed between group of patients and control population.

The description of the studied population including fundamental characteristics of particular cancer and background variables are presented in **Table** 1.

2.2 | Blood collection

Five ml of heparinized peripheral blood was collected from all individuals and used for the bleomycin test and γ-H2AX assay, additional 5 ml of blood collected into ethylenediaminetetraacetic acid was used for DNA extraction and further RTL measurement. All the blood samples in our study were drawn before any chemotherapy and radiotherapy treatments to exclude reverse causality for our results. Genomic DNA was isolated from PBL using standard procedures and the aliquots were used for RTL measurements. The time between blood collection and the chromosome analysis did not exceed 2 h, this interval was standardized. Blood was stored on ice (during transport) or in a refrigerator prior to the processing. DNA damage was induced by the treatment of the samples using bleomycin. All blood samples were cultured also without the bleomycin treatment and served as a reference for determination of background chromosome damage.

TABLE 1 The characteristics of the studied population

		CRC cases	Controls to CRC cases	BC cases	Controls to BC cases
All subjects		$N = 44$	$N = 44$	$N = 47$	$N = 46$
Sex	Females	23	23	47	46
	Males	21	21	0	0
Age (y)		65.3 ± 13.8	59.3 ± 16.6	63.0 ± 11.2	62.7 ± 18.5
Diagnosis	Colon cancer (C18)	24	...	...	...
	Rectosigmoid cancer (C19)	6	...	...	...
	Rectal cancer (C20)	9	...	...	...
	Anus cancer (C21)	1	...	...	...
	Unspecified colon cancer	4	...	...	...
	Central portion of breast (C501)	-	...	1	...
	Outer upper quadrant of breast (C504)	-	...	29	...
	Unspecified breast cancer (C509)	-	...	17	...
Smoking	Non-smokers	17	17	23	19
	Ex-smokers	11	11	8	1
	Current smokers	5	5	11	7

Abbreviations: BC, breast cancer; CRC, colorectal cancer

2.3 | Chromosome analysis (mutagen sensitivity assay)

Mutagen sensitivity assay was performed according to the previously described protocol with minor modifications (Hsu et al., 1989). Briefly, two whole-blood samples from each subject were cultured for 72 h in EKAMTB-100 complete medium with phytohaemagglutinin (EuroClone S.p.A., Italy). Five hours before harvesting (late S and G2 phase of the cell cycle), bleomycin (Sigma) was added (final concentration 30 μg/ml) in one of the two cultures, the other culture served as a reference. Two hours prior to harvest, Colcemid (Calbiochem, 0.1 μg/ml) was added to arrest cells in metaphase. The PBL were harvested and after cytogenetic procedure dropped onto slides and stained with Giemsa. At least 100 cells per subject were analyzed microscopically for CTAs. Two scorers conducted microscopic analysis (each evaluating 50 cells per subject) in a double-blind fashion on coded slides. CTAs (breaks and exchanges) were cytogenetically assessed in samples of all groups after the bleomycin treatment, while not affected samples served for detection of baseline CAs level. Extremely low values of spontaneous CTAs (common in peripheral blood lymphocytes) in all control slides were considered as negligible and data were not included for the whole analysis. Only metaphases with 46 ± 1 chromosomes were included. Gaps were scored but not included to the number of CTAs. Mutagen sensitivity was expressed as the average number of breaks per cell (CTAs/cell, according Hsu et al.[22]). By recording CTAs, we followed the recommendation of the Chatham Bars Inn Workshop Conference (1971), that is, an achromatic lesion whose length is smaller than the diameter of the chromatid was classified as a chromatid gap and a lesion whose length was equal to or longer than the diameter of the chromatid was regarded as a CTA.[22] Cells with more than (≥12) CTAs were assessed as those with high chromosome damage, as previously reported.[22]

2.4 | RTL measurement

RTL was measured in DNA aliquots of subjects prior to bleomycin treatment using the monochrome multiplex polymerase chain reaction (PCR) assay as described previously[30] with minor modifications.[31,32] We carried out two-fold serial dilution standard curve which produced seven known concentrations (genomic DNA pooled from 10 healthy individuals). Range of DNA concentration for calibration curve descended from 28.5 to 0.5 ng/μl. Based on the known concentration of DNA, we could prepare standard curve for telomere (T) and also for albumin (S) using multiplex assay. Samples lacking any genomic DNA were included as negative controls. All reactions were carried out in triplicates in an optical 384-well reaction plate with final volume of master mix in each well 1.75 μl. Master mix was prepared using 740 μl of ×5 HOT FIREPol Probe qPCR Mix Plus with ROX (Solis BioDyne, Tartu, Estonia), 5.5 μl, 1.5 μM of Syto 9 (Invitrogen, Carlsbad, California, USA), 7.4 μl Telg primer (at 200 nM), 14.8 μl Telc primer (at 400 nM), 7.4 μl primer albugcr2 (at 200 nM), albdgcr2 primer (at 400 nM) and finally 2262 μl of double-distilled water. Sequences of primers are given in Supporting Information, Table S1. Pipetted volume of DNA samples in each well was 1.75 μl with approximate concentration 2 ng/μl. Real-time PCR experiments were performed on Viia 7 Real-time PCR System (Applied Biosystems, Foster City, California, USA) using two simultaneous programs to acquire the respective CT values for

telomere sequences and the albumin gene. The conditions for telomere sequence amplification were 95°C/15 min, 2 cycles of 95°C/20 sec and 49°C/1 min, followed by 25 cycles of 85°C/20 sec with signal acquisition at 59°C/30 sec. The conditions for albumin gene were 35 cycles of 95°C/15 sec, 85°C/30 sec with signal acquisition at 84°C/30 sec. The specificity of all amplifications was determined by melting curve analysis done at default settings (95°C/15 sec, 60°C/1 min with continuous signal acquisition at 0.05°C/sec ramping, 95°C/15 sec). The quality control was performed using the Applied Biosystems software. The standard curve was used to quantify the telomere and albumin genes based on the respective CT values and the obtained triplicate values were averaged. RTL was expressed as the ratio between T/S. Inter-assay and intra-assay variations were determined by duplicating the reference DNA for all the dilutions in all the assays performed.[31] If T/S ratio exceeded 1, the DNA sample had longer telomeres. On the contrary if T/S ratio was lower than 1, the DNA sample had shorter telomeres. PCR efficiency varied between 95% and 102%. Inter-plate variation for T and S was 3.5% and 3.37%, respectively. Intra-plate variation for T was 0.42% and 0.28% for S.

2.5 | γ-H2AX concentration measurement

Similarly as in "Chromosome analysis" blood samples were processed for γ-H2AX concentration measurement. Briefly, whole-blood samples from each subject were cultured for 72 h in EKAMTB-100 complete medium with phytohaemagglutinin (EuroClone S.p.A., Italy). Five hours before harvesting (late S and G2 phase of the cell cycle), bleomycin (Sigma) was added (final concentration 30 μg/ml) in one of the two cultures, the other culture served as a reference.

Histone 2AX is a 14 kDa ubiquitous member of the H2A histone family that contains an evolutionarily conserved SQ motif at the C-terminus in eukaryotes. Serine 139 within this motif becomes rapidly phosphorylated by ATM and ATR kinases to yield a form known as γ-H2AX in response to DSB and apoptosis.[29]

This assay documents differences of γ-H2AX levels in human PBL. Immobilized γ-H2AX antibody in the wells of a 96-well plate captures γ-H2AX from sample lysate. Incubation with a H2AX detecting antibody (Trevigen, Gaithersburg, MD), followed by addition of a Goat anti-mouse horseradish peroxidase (HRP) conjugate and a chemiluminescent HRP substrate yields relative light units that directly correlates with the amount of γ-H2AX in the sample.

2.6 | Statistical evaluation

The differences in RTL and mutagen sensitivity parameters between the particular groups were assessed by non-parametric tests (Mann-Whitney *U* test, Kruskal-Wallis test). For stratification of subjects with lower or higher RTL a median value of control individuals was used. Associations between the RTL and mutagen sensitivity parameters were determined by Logistic regression adjusted for age and sex, χ^2-test and by Spearman correlation. A difference was considered statistically significant if the *P* value will be .05 or smaller.

3 | RESULTS

The results from mutagen sensitivity assay in individual groups of cancer patients and corresponding controls, which serve as an indirect functional biomarker of DNA repair capacity, are summarized in **Table 2**. We observed significantly higher amount of bleomycin-induced CTAs ($P = .03$) and aberrant cells (ACs; $P = .05$) in PBL of 44 CRC patients in comparison with 44 corresponding healthy controls, whereas there were no significant differences between BC patients and female controls ($P = .39$ for CTAs and $P = 0.41$ for ACs, respectively, **Figure 1**). Neither CTAs nor ACs were significantly affected by age and gender in any group of patients or controls. Smoking status was the only variable associated with lower amount of ACs after bleomycin treatment in smoking patients in comparison with non-smoking cases ($P = .03$).

Results from mutagen sensitivity assay were compared with γ-H2AX concentration measurement in 11 BC and 8 CRC cases. No correlation between γ-H2AX concentration and amount of ACs and CTAs was observed in either group of patients ($r = -0.14$; $P = .64$). We did not find any difference between BC and CRC cases in γ-H2AX concentration after bleomycin treatment (**Figure 2**; $P = .72$).

RTL was measured in 118 patients and healthy controls. We observed an inverse association between RTL and age ($P = .005$). Additionally, the RTL was longer in women than men ($P = .0008$). Therefore, the results for associations between RTL and mutagen sensitivity assay outcomes were adjusted for age and sex.

All individuals with shorter RTL (below median = 1.5) showed significantly increased CTAs in PBL after bleomycin treatment ($P = .05$) and also increased number of cells with ≥12 CTAs ($P = .03$). After combining patients (both CRC and BC; $N = 54$) with shorter RTL (below median RTL) we observed significantly increased number of ACs ($P = .02$), CTAs ($P = .02$) and cells with high amount of CTAs (≥12 CTAs/PBL; $P = .03$) after bleomycin treatment. However, the same parameters were not affected in the control group. **Figure 3** documents Spearman correlation of RTL with CTAs in a pooled group of cases (upper box, $R = -0.36$) and in controls (lower box, $R = -0.09$). A separate analysis showed that lower amount of CTAs ($P = .03$) is associated with higher than median RTL in BC ($N = 22$), whereas shorter than median RTL resulted in a tendency to higher amounts of ACs ($P = .09$), CTAs ($P = 0.09$) and cells with ≥12 CTAs ($P = .05$) in CRC ($N = 32$) patients.

4 | DISCUSSION

In the present study, we investigated chromosomal integrity in PBL from newly diagnosed BC and CRC patients and matched healthy controls. We observed significantly increased levels of CTAs and ACs only in CRC patients compared with controls after bleomycin treatment. Interestingly, CRC and BC patients with shorter RTL exhibited significantly higher amount of ACs, CTAs and cells with high frequency of CTAs after bleomycin treatment. We observed that altered DSB repair measured by sensitivity toward mutagen in PBL occurs particularly in

TABLE 2 Mutagen sensitivity and RTL in individual groups of cancer patients and corresponding controls

Mutagen sensitivity	Subjects	CRC cases $N = 44$	Controls to CRC cases $N = 44$	BC cases $N = 47$	Controls to BC cases $N = 46$
ACs (%)	Untreated culture	0.89 ± 1.4	1.14 ± 1.2	0.87 ± 1.2	1.00 ± 1.2
	Bleomycin treatment	50.8 ± 11.4	46.2 ± 9.5	47.6 ± 10.0	49.8 ± 10.7
CTAs	Untreated culture	0.93 ± 1.6	1.21 ± 1.3	1.02 ± 1.4	1.11 ± 1.5
	Bleomycin treatment	158 ± 0.6	130 ± 0.4	136 ± 0.5	146 ± 0.5
Cells with ≥12 CTAs (%)	Untreated culture	0	0	0	0
	Bleomycin treatment	3.70 ± 2.6	2.57 ± 1.7	2.83 ± 2.1	3.30 ± 2.5
RTL	Subjects	$N = 32$	$N = 31$	$N = 22$	$N = 33$
	Mean	1.35 ± 0.5	1.59 ± 0.5	1.77 ± 0.4	1.52 ± 0.5

Abbreviations: AC, aberrant cells; BC, breast cancer; CRC, colorectal cancer; CTA, chromatid-type aberrations; RTL, relative telomere length

CRC. Irrespective of cancer type, telomere shortening may be associated with a decreased capacity to repair DSB.

Unlike comet assay or host cell reactivation,[33,34] the mutagen sensitivity has been employed as one of the well-established marker for cancer susceptibility.[35] Mutagen sensitivity, expressed as a mean number of CTAs per cell at metaphase PBLs, is measured by following bleomycin exposure in the G2 phase of the cell cycle. Several studies have considered that this functional test may reflect the inter-individual differences in DNA damage response.[22,23,25–27,35–38] The reported increased amount of CTAs correlates with suboptimal ability to repair DSB, suggesting that the outcome from mutagen sensitivity assay may reflect DSB repair capacity.[28] Our present results suggest that altered DSB repair measured by sensitivity toward mutagen in PBL occurs particularly in CRC patients. It was postulated that mutagen sensitivity may particularly be associated with carcinogenesis of organs and tissues that are in direct contact with the external environment (eg, respiratory, digestive, and bowel cancers) but not in those which are lacking direct environmental exposure (eg, BC and brain).[22,37] This is supported by the role of bleomycin sensitivity phenotype as an important predisposition factor in relation to lung, CRC, and head and neck cancers.[39–42] However, other authors attributed the results from mutagen sensitivity rather to the effect of life-style factors.[43,44] Moreover, most recent reports suggest important roles of somatic mutations not only induced by environmental factors, but those connected with replication errors.[45] These data may further pinpoint the importance of mutagen sensitivity together with mapping telomere biology.

Here, we provide further evidence that bleomycin sensitivity is significantly associated with CRC. However, the bleomycin sensitivity

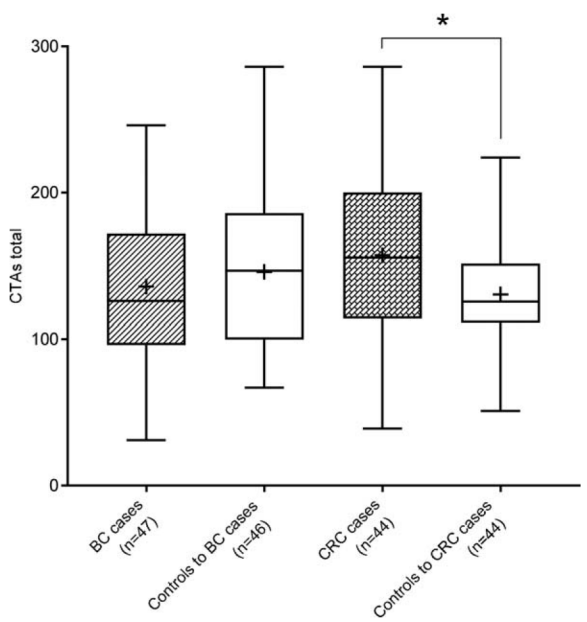


FIGURE 1 CTAs in all investigated groups; *$P = .03$. Axis "y" depicts the number of chromatid-type aberrations (we scored 100 cells/individual; mean ± SD) in all investigated groups (axis "x"). Significant difference is marked by an appropriate P value

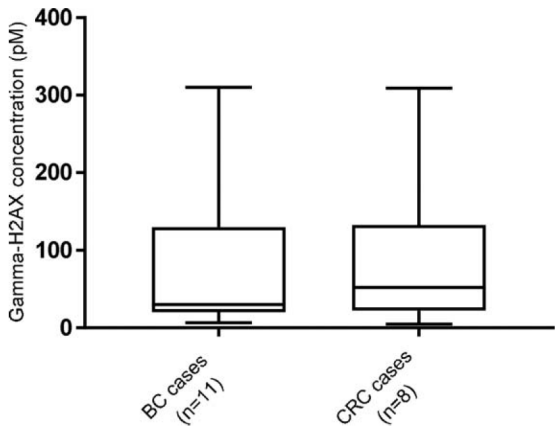
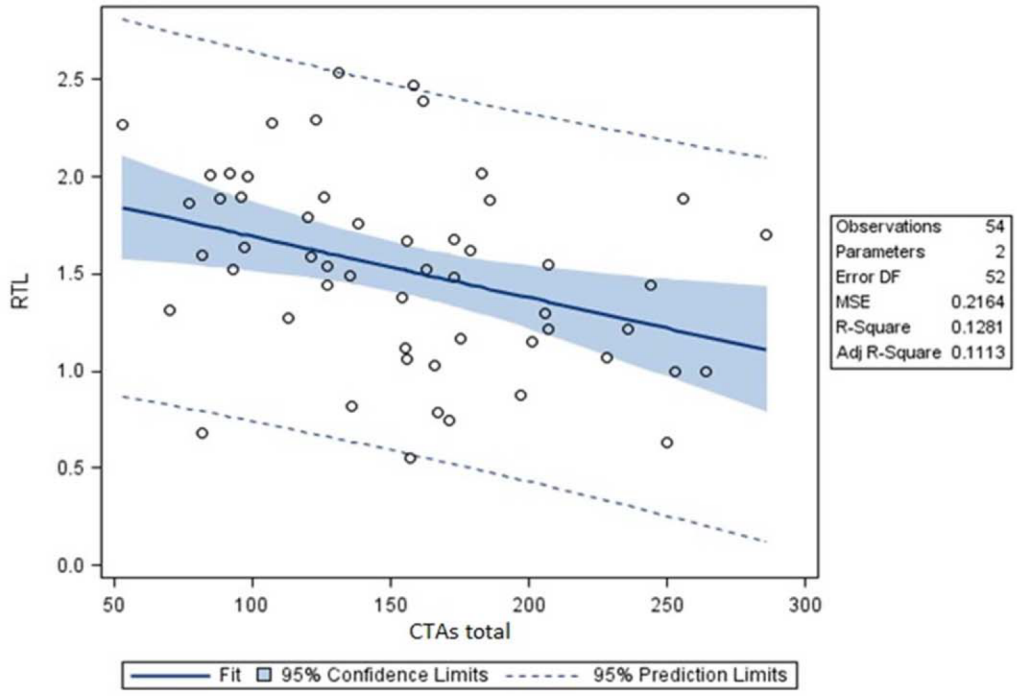


FIGURE 2 γ-H2AX concentration (pM) in breast cancer (BC) a colorectal cancer (CRC) patients; $R = -0.14$; $P = .64$. γ-H2AX concentration (pM) in BC a CRC patients. Axis "y" shows the concentration of γ-H2AX (mean ± SD) in individual groups of cancer patients (axis "x")



Spearman correlation of RTL with CTAs in a pooled group of controls (R=-0.09; p=0.87).

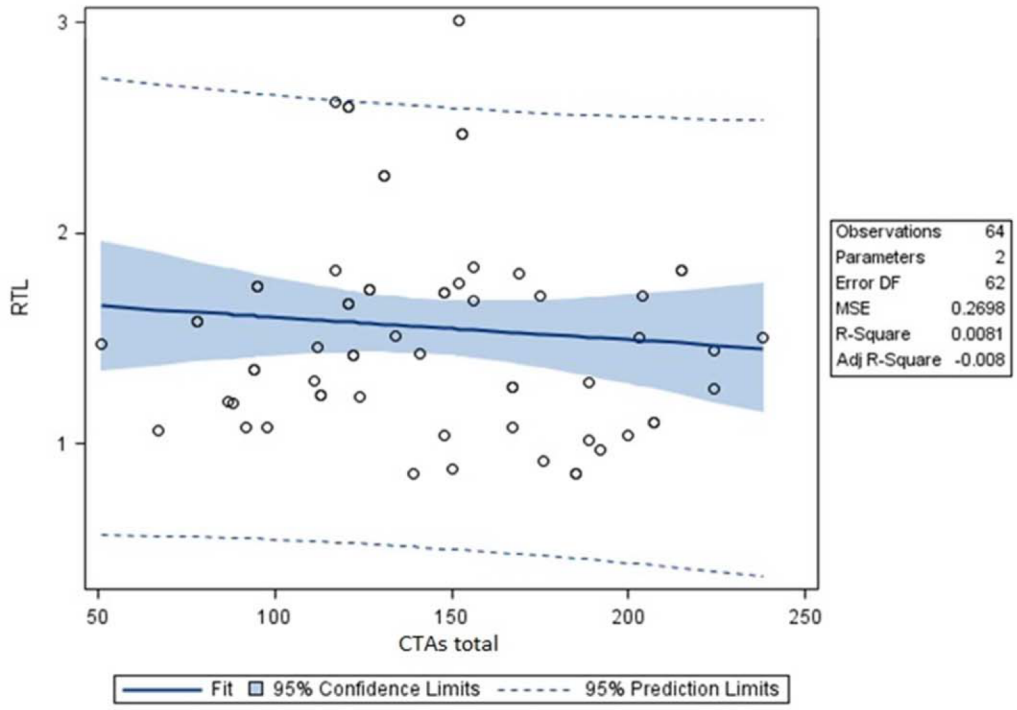


FIGURE 3 Spearman correlation of relative telomere length (RTL) with chromatid-type aberrations (CTAs) in a pooled group of cases (upper graph; $R = -0.36$; $^*P = .02$) and controls (lower graph; $R = -0.09$; $P = .87$) [Color figure can be viewed at wileyonlinelibrary.com]

profile of BC patients was similar to that of the control population. Our results are in agreement with Hsu et al.,[22] who confirmed that bleomycin response profile in PBL differs in CRC, lung, and head/neck cancer patients (but not in BC) compared with healthy controls. Some studies[27,46,47] have shown strong bleomycin sensitivity not only for familial BC, but also for sporadic patients. Possible explanation for these discordant results could be different aetiology of familial BC as compared with the sporadic form and the heterogeneity of BC. Mutated BRCA1 and BRCA2 genes in familial BC, which are involved in DSB repair machinery,[48] could cause higher DNA damage in bleomycin exposed PBL. Natarajan et al.[46] postulated that mutagen sensitivity phenotype is a risk factor for BC.

Mutagen sensitivity was not affected either by age or gender in all groups of patients and controls, which is in agreement with Hsu et al.[22] However, we observed lower ACs in smoking patients compared with non-smoking relatives. Some studies[27,49] showed that the effect of tobacco smoking may modulate mutagen sensitivity status. Others postulated that mutagen sensitivity may be a tool for characterization of genotoxic exposure to heavy tobacco and alcohol use rather than for individual susceptibility toward cancer.[44,50] Our results indicate that smoking status may exert an influence on mutagen sensitivity probably due to induced DSB repair. Similar phenomenon has been described by us on induced base excision repair capacities among smokers and workers exposed to industrial xenobiotics.[51,52]

Mutagen sensitivity profile from 19 BC and CRC patients were compared with the measurement of γ-H2AX as a hypothesized alternative for mutagen sensitivity assay. γ-H2AX is a hallmark of DSB[29] and we postulated that quantification of chromosomal damage after bleomycin treatment in PBL may be comparable with the concentration of γ-H2AX. Our results did not show any correlation between mutagen sensitivity and γ-H2AX. In contrast, the concentration of γ-H2AH was higher in CRC, which exhibit higher amount of CTAs, suggesting a lower DSB repair capacity in CRC. We assume that γ-H2AX is the first acute response of the cell to cope with DSB, however CTAs are the final results of unrepaired DSB with the absence of γ-H2AX. Therefore, γ-H2AX measurement could not be used instead of mutagen sensitivity assay.

The investigation of mutagen sensitivity in cases and controls was supplemented by RTL determination in the group of 118 CRC and BC patients and corresponding healthy controls. According to Valls-Bautista et al.,[53] RTL in PBL reflects changes connected with tumor and differs from that in healthy control subjects. The authors suggested that changes in RTL take place in the early stages of the tumor development. RTL was measured in PBL due to its non-invasive availability and could therefore be used as an initial marker for CRC.[53]

Telomere shortening has also been linked to reduced DSB repair capacity.[54] Our results showed statistically significant correlation between shortening of telomeres and mutagen sensitivity profile in a group of BC and CRC patients. To our knowledge, this is the first study describing DSB repair capacity in BC and CRC patients in correlation with a shortening of telomeres. It has recently been reviewed that PBL from Fanconi anemia patients bearing mutations in homologous recombination repair gene Fanconi anemia complementation group D2 (FANCD2) show short telomeres and telomere dysfunction induced foci.[7] Recently, Pooley et al. and Thorvaldsdottir et al.[55,56] suggested that BC BRCA2 mutation carriers have significantly shorter or longer telomere length than unaffected women. Above observation add to the chain of evidence on the interplay between telomere complex and DSB repair.

ACKNOWLEDGMENTS

The study was supported by grants 15–14789S (National Science Foundation), 15–27580A (Czech Ministry of Health), COST LD-14050, SVV 260390, UNCE 204022 and PROGRES Q 28. The study was also supported by a grant from TRANSCAN through German Ministry of Education and Research (BMBF) under grant number 01KT15511.

ORCID

Michal Kroupa http://orcid.org/0000-0001-9112-7047
Rajiv Kumar http://orcid.org/0000-0002-6093-0395

REFERENCES

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

[2] Vodicka P, Polivkova Z, Sytarova S, et al.Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls. *Carcinogenesis*. 2010;31:1238–1241.

[3] Vodenkova S, Polivkova Z, Musak L, et al.Structural chromosomal aberrations as potential risk markers in incident cancer patients. *Mutagenesis*. 2015;30:557–563.

[4] Durante M, Bedford JS, Chen DJ, et al.From DNA damage to chromosome aberrations: joining the break. *Mutat Res*. 2013;756:5–13.

[5] Blackburn EH. Switching and signaling at the telomere. *Cell*. 2001; 106:661–673.

[6] Maciejowski J, de Lange T. Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Biol*. 2017;18:175–186.

[7] Heidenreich B, Kumar R. TERT promoter mutations in telomere biology. *Mutat Res*. 2017;771:15–31.

[8] Li H, Hilmarsen HT, Hossain MB, et al. Telomere length and LINE1 methylation is associated with chromosomal aberrations in peripheral blood. *Genes Chromosomes Cancer*. 2013;52:1–10.

[9] Hemminki K, Rachakonda S, Musak L, et al. Telomere length in circulating lymphocytes: association with chromosomal aberrations. *Genes Chromosomes Cancer*. 2015;54:194–196.

[10] Broberg K, Bjork J, Paulsson K, Hoglund M, Albin M. Constitutional short telomeres are strong genetic susceptibility markers for bladder cancer. *Carcinogenesis*. 2005;26:1263–1271.

[11] Jang JS, Choi YY, Lee WK, et al. Telomere length and the risk of lung cancer. *Cancer Sci*. 2008;99:1385–1389.

[12] DePinho RA. The age of cancer. *Nature*. 2000;408:248–254.

[13] Shay JW. Role of telomeres and telomerase in aging and cancer. *Cancer Discov*. 2016;6:584–593.

[14] Gostissa M, Alt FW, Chiarle R. Mechanisms that promote and suppress chromosomal translocations in lymphocytes. *Ann Rev Immunol*. 2011;29:319–350.

[15] Maser RS, DePinho RA. Connecting chromosomes, crisis, and cancer. *Science*. 2002;297:565–569.

[16] Meeker AK, Hicks JL, Iacobuzio-Donahue CA. Telomere length abnormalities occur early in the initiation of epithelial carcinogenesis. *Clin Cancer Res*. 2004;10:3317–3326.

[17] Maciejowski J, Li Y, Bosco N, Campbell PJ, de Lange T. Chromothripsis and kataegis induced by telomere crisis. *Cell*. 2015;163: 1641–1654.

[18] Shay JW. Molecular pathogenesis of aging and cancer: are telomeres and telomerase the connection? *J Clin Pathol*. 1997;50:799–800.

[19] Vallarelli AF, Rachakonda PS, André J, et al. TERT promoter mutations in melanoma render TERT expression dependent on MAPK pathway activation. *Oncotarget*. 2016;7:53127–53136.

[20] Valdiglesias V, Pasaro E, Mendez J, Laffon B. Assays to determine DNA repair ability. *J Toxicol Environ Health A*. 2011;74:1094–1109.

[21] Slyskova J, Naccarati A, Polakova V, et al. DNA damage and nucleotide excision repair capacity in healthy individuals. *Environ Mol Mutagen*. 2011;52:511–517.

[22] Hsu TC, Johnston DA, Cherry LM, et al. Sensitivity to genotoxic effects of bleomycin in humans: possible relationship to environmental carcinogenesis. *Int J Cancer*. 1989;43:403–409.

[23] Cloos J, Steen I, Joenje H, et al. Association between bleomycin genotoxicity and non-constitutional risk factors for head and neck cancer. *Cancer Lett*. 1993;74:161–165.

[24] Cloos J, Spitz MR, Schantz SP, et al. Genetic susceptibility to head and neck squamous cell carcinoma. *J Natl Cancer Inst*. 1996;88:530–535.

[25] Spitz MR, Hsu TC, Wu X, Fueger JJ, Amos CI, Roth JA. Mutagen sensitivity as a biological marker of lung cancer risk in African Americans. *Cancer Epidemiol Biomarkers Prev*. 1995;4:99–103.

[26] Zheng YL, Loffredo CA, Yu Z, et al. Bleomycin-induced chromosome breaks as a risk marker for lung cancer: a case-control study with population and hospital controls. *Carcinogenesis*. 2003;24:269–274.

[27] Hu M, Han D, Sun S, Yan Y, Zhang J, Zhou Y. Bleomycin-induced mutagen sensitivity, passive smoking, and risk of breast cancer in Chinese women: a case-control study. *Cancer Causes Control*. 2013;24:629–636.

[28] Laczmanska I, Gil J, Karpinski P, et al. Polymorphism in nucleotide excision repair gene XPC correlates with bleomycin-induced chromosomal aberrations. *Environ Mol Mutagen*. 2007;48:666–671.

[29] Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem*. 1998;273:5858–5868.

[30] Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res*. 2009;37:e21.

[31] Heidenreich B, Rachakonda PS, Hosen I, et al. TERT promoter mutations and telomere length in adult malignant gliomas and recurrences. *Oncotarget*. 2015;6:10617–10633.

[32] Hosen I, Rachakonda PS, Heidenreich B, et al. Mutations in TERT promoter and FGFR3 and telomere length in bladder cancer. *Int J Cancer*. 2015;137:1621–1629.

[33] Sigurdson AJ, Jones IM, Wei Q, et al. Prospective analysis of DNA damage and repair markers of lung cancer risk from the prostate, lung, colorectal and ovarian (PLCO) cancer screening trial. *Carcinogenesis*. 2011;32:69–73.

[34] Koppen G, Azqueta A, Pourrut B, Brunborg G, Collins AR, Langie SAS. The next three decades of the comet assay: a report of the 11th International Comet Assay Workshop. *Mutagenesis*. 2017;32:397–408.

[35] Au WW. Mutagen sensitivity assays in population studies. *Mutat Res*. 2003;544:273–277.

[36] Cloos J, de BWP, Snel MH, et al. Microarray analysis of bleomycin-exposed lymphoblastoid cells for identifying cancer susceptibility genes. *Mol Cancer Res*. 2006;4:71–77.

[37] Spitz MR, Fueger JJ, Beddingfield NA, et al. Chromosome sensitivity to bleomycin-induced mutagenesis, an independent risk factor for upper aerodigestive tract cancers. *Cancer Res*. 1989;49:4626–4628.

[38] Angelini S, Kumar R, Carbone F, et al. Inherited susceptibility to bleomycin-induced micronuclei: correlating polymorphisms in GSTT1, GSTM1 and DNA repair genes with mutagen sensitivity. *Mutat Res*. 2008;638:90–97.

[39] Wu X, Lin J, Etzel CJ, et al. Interplay between mutagen sensitivity and epidemiological factors in modulatinglung cancer risk. *Int J Cancer*. 2007;120:2687–2695.

[40] Shao L, Lai M, Huang Q. Mutagen sensitivity and p53 expression in colorectal cancer in China. *Postgrad Med J*. 2001;77:713–716.

[41] Fireman Z, Shabtai F, Lurie B. Chromosome sensitivity to bleomycin-induced mutagenesis in lymphocytes from colorectal cancer patients under 40 years of age. *Dis Colon Rectum*. 1994;37:1317–1320.

[42] Ankathil R, Jyothish B, Madhavan J, Nair MK. Deficient DNA repair capacity: a predisposing factor and high risk predictive marker in familial colorectal cancer. *J Exp Clin Cancer Res*. 1999;18:33–37.

[43] Szekely G, Remenar E, Kasler M, Gundy S. Mutagen sensitivity of patients with cancer at different sites of the head and neck. *Mutagenesis*. 2005;20:381–385.

[44] Szekely G, Remenar E, Kasler M, Gundy S. Does the bleomycin sensitivity assay express cancer phenotype? *Mutagenesis*. 2003;18:59–63.

[45] Tomasetti C, Li L, Vogelstein B. Stem cell divisions, somatic mutations, cancer etiology, and cancer prevention. *Science*. 2017;355:1330–1334.

[46] Natarajan TG, Ganesan N, Carter NP, et al. γ-Radiation-induced chromosomal mutagen sensitivity is associated with breast cancer risk in African-American women: caffeine modulates the outcome of mutagen sensitivity assay. *Cancer Epidemiol Biomarkers Prev*. 2006;15:437–442.

[47] Jyothish B, Ankathil R, Chandini R, et al. DNA repair proficiency: a potential marker for identification of high risk members in breast cancer families. *Cancer Lett*. 1998;24:9–13.

[48] Turnbull C, Rahman N. Genetic predisposition to breast cancer: past, present, and future. *Ann Rev Genomics Hum Genet*. 2008;9:321–345.

[49] Kosti O, Byrne C, Meeker KL, et al. Mutagen sensitivity, tobacco smoking and breast cancer risk: a case-control study. *Carcinogenesis*. 2010;31:654–659.

[50] Szekely G, Remenar E, Kasler M, Bodrog A, Gundy S. Validity of chromosome analysis and bleomycin sensitivity assay in the prevention of head and neck cancer in Hungary. *Orv Hetil*. 2001;142:611–616.

[51] Vodicka P, Tuimala J, Stetina R, et al. Cytogenetic markers, DNA single-strand breaks, urinary metabolites, and DNA repair rates in styrene-exposed lamination workers. *Env Health Perspect*. 2004;112:867–871.

[52] Vodicka P, Kumar R, Stetina R, et al. Markers of individual susceptibility and DNA repair rate in workers exposed to xenobiotics in a tire plant. *Env Mol Mutagen*. 2004;44:283–292.

[53] Valls-Bautista C, Piñol-Felis C, Reñé-Espinet JM, Buenestado-García J, Viñas-Salas J. In colon cancer, normal colon tissue and blood cells have altered telomere lengths. *J Surg Oncol*. 2015;111:899–904.

[54] Uziel O, Yerushalmi R, Zuriano L, et al. BRCA1/2 mutations perturb telomere biology: characterization of structural and functional abnormalities in vitro and in vivo. *Oncotarget*. 2016;7:2433–2454.

[55] Thorvaldsdottir B, Aradottir M, Stefansson OA, Bodvarsdottir SK, Eyfjord JE. Telomere length is predictive of breast cancer risk in BRCA2 mutation carriers. *Cancer Epidemiol Biomarkers Prev*. 2017;26(8):1248–1254. https://doi.org/10.1158/1055-9965.EPI-16-0946.

[56] Pooley KA, McGuffog L, Barrowdale D, et al. Lymphocyte telomere length is long in BRCA1 and BRCA2 mutation carriers regardless of

cancer-affected status. *Cancer Epidemiol Biomarkers Prev*. 2014;23:1018–1024.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Kroupa M, Polivkova Z, Rachakonda S, et al. Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients. *Genes Chromosomes Cancer*. 2018;57:61–69. https://doi.org/10.1002/gcc.22508

Manuscript VI

Relationship of telomere length in colorectal cancer patients with cancer phenotype and patient prognosis.

Kroupa M, Rachakonda S, Liska V, Srinivas N, Urbanova M, Jiraskova K, Schneiderova M, Vycital O, Vymetalkova V, Vodickova L, Kumar R, Vodicka P.

Manuscript in preparation

Relationship of telomere length in colorectal cancer patients with cancer phenotype and patient prognosis

Kroupa Michal[1,2], Rachakonda Sivarama Krishna[3], Liska Vaclav[4], Srinivas Nalini[3], Urbanova Marketa[1,5], Jiraskova Katerina[1,5], Schneiderova Michaela[6], Vycital Ondrej[4], Vymetalkova Veronika[1,4,5], Vodickova Ludmila[1,4,5], Kumar Rajiv[3], Vodicka Pavel[1,4,5*]

1 Department of Molecular Biology of Cancer, Institute of Experimental Medicine, The Czech Academy of Sciences, Prague, Czech Republic

2 Department of Histology and Embryology, Faculty of Medicine in Pilsen, Charles University, Czech Republic

3 Division of Molecular Genetic Epidemiology, German Cancer Research Center, Heidelberg, Germany

4 Faculty of Medicine and Biomedical Center in Pilsen, Charles University, Pilsen, Czech Republic

5 Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, Czech Republic

6 Department of Surgery, General University Hospital in Prague, Prague, Czech Republic.

⋆ Corresponding author: Department of Molecular Biology of Cancer, Institute of Experimental Medicine of the Czech Academy of Sciences, Videnska 1083, 142 20, Prague, Czech Republic.

E-mail address: pavel.vodicka@iem.cas.cz (P. Vodicka).

Keywords: Telomere length, Colorectal cancer, Microsatellite instability,

Abstract

Telomeres, repetitive DNA capping ends of eukaryotic chromosomes, are important in maintenance of genomic integrity. Perturbed telomeres are common features of many human malignancies. The aim of the present study was to investigate correlation of telomere length (TL) in tumor tissues, adjacent mucosa and blood from patients with colorectal cancer with different clinicopathological features and impact on patient survival. In addition, we also measured TL in a limited number of liver metastases, non-cancerous liver tissues or corresponding tissues from the same patients. We observed that TL in tumor tissues was shorter than the adjacent mucosa (Wilcoxon test, $P<0.0001$). Shorter TL was observed in tumors with lower stage than in those exhibiting advanced stages (Wilcoxon test $P=0.001$). Telomeres were shorter in tumors at the proximal than at the distal sites of the colon (Wilcoxon test, $P<0.0001$). Shorter telomeres were also associated with microsatellite instability (Wilcoxon test, $P=0.001$), and mucinous tumor histology (Wilcoxon test, $P<0.0001$). Patients with smaller TL ratio between tumor tissues and adjacent mucosa were associated with an increased overall survival (Log-rank test, $P=0.022$). Metastasized tumors had shorter telomeres than the adjacent non-cancerous liver tissues (Wilcoxon test, $P=0.0005$). Overall, our results demonstrate differences in TL between tumors and adjacent mucosa, between tumors located at different sites and association with patient survival.

Introduction

Telomeres, tandem G-rich hexanucleotide repeats that are involved in maintenance of genome integrity, undergo a progressive shortening though successive cell division. Gradual telomeric attrition is caused by incomplete DNA replication of a lagging strand. Telomere length (TL) is also affected by the genotoxic effect of environmental and intra-cellular DNA-damaging agents [1-3]. Telomere shortening correlates with age [4]. Tumor cells undergo faster telomeric attrition than non-cancerous somatic cells due to increased proliferation. Telomere shortening can act as potent tumor-suppressing mechanism limiting cells from an uncontrolled growth. However, most cancers evolve mechanism to overcome proliferative barrier due to telomere attrition through telomerase rejuvenation. The rejuvenated telomerase preferentially stabilizes shortest telomeres and critically short telomeres can lead to the formation of anaphase bridges through breakage-fusion-bridge cycles that contributes to chromosome instability (CIN) [5].

CIN along with microsatellite instability (MSI) represent two major pathways prevalent in the genesis of CRC [6]. Impaired mismatch repair (MMR) leading to MSI is characterized by excessive in-del mutations in microsatellite sequences [7]. It is possible that a compromised MMR pathway could as well affect telomeric repeats. MSI accounts for 10-15% of all CRC cases [8-10], and are usually associated with better prognosis in patients [11, 12].

Due to importance in cancer initiation, progression and patient prognosis, it is imperative to understand the status of TL and the factors affecting it in tumor tissue of CRC patients. In this study, we determined TL in tumors and adjacent mucosa in order to find a correlation, if any. We also investigated the impact of tumor location on TL. The impact of TL in tumor/mucosa on overall patient survival was also determined.

Materials and Methods

Population characteristic

Sporadic CRC patients (N=721) with histologically confirmed tumors were included in the study. Personal data that included date of birth, sex, and diabetes mellitus were obtained using a structured questionnaire. For all patients, clinical data including tumor-related parameters, such as the tumor location, international union against cancer (UICC) TNM stage

system status, degree of tumor differentiation, were collected along with information about distant metastases, relapse, and date of death.

Patients were recruited from 2004 to 2014 in different oncological and gastroenterological departments of various hospitals within the Czech Republic. The last update of patient follow-up for this study was December 2015. The description of the studied population including fundamental characteristics of CRC and background variables is presented in **Table I.**

The study included paired tumor tissues and adjacent non-malignant mucosa from 721 patients collected during surgical resection. In addition, peripheral blood was available from 164 of those patients, sampled prior to surgery. Further, we also had access to primary tumors, adjacent mucosa, liver metastatic tissues, and paired adjacent liver tissue from 12 patients from the original group of patients. We also included metastatic and adjacent liver tissues from 122 patients with CRC without primary tumors in the study.

DNA was extracted from tumor tissues, non-affected adjacent mucosa, blood, metastatic tissues, and liver samples using the DNeasy Blood and Tissue Kit (Qiagen, Courtaboeuf, France). The study was approved by the local ethics committee of each participating hospital. Written informed consent to participate in the study and to approve the use of biological samples for genetic analyses was obtained from all patients, according to the Helsinki declaration.

Relative telomere length

TL was measured as relative telomere length (RTL) as described previously with some modifications [8-13]. Syto 9, a single fluorescent dye, was used in real-time qPCR for the collection of a signal from telomere (T; repeat copy number) and albumin (S; human single-copy gene) amplicon, two target sequences greatly differing in copy number. The values of cycle threshold (Ct) for albumin amplicon were collected above the melting temperature of the telomere product. Measurement of TL from all DNA samples was carried out in triplicate using a MicroAmp Optical 384-Well Reaction Plate (Applied Biosystems, Foster City, California, USA). A standard two-fold serial dilution was prepared based on the known initial concentration of reference DNA (genomic DNA pooled from 30 individuals with 40-55 years age range and without any specific sex ratio [9]). Real-time qPCR experiments were performed on a Viia 7 Real-time PCR System (Applied Biosystems) using two simultaneous programs to acquire the respective Ct values for telomere sequences and the albumin gene (primer sequences in **Supplementary Table I**). The standard curve was used to quantify telomere and albumin

genes based on respective Ct values, and the obtained triplicate values were averaged [14]. TL was expressed as the ratio between T/ S. Inter-assay and intra-assay variations were determined by duplicating the reference DNA for all dilutions in each assay performed. If T/ S ratio exceeded 1, the DNA sample had longer telomeres. On the contrary, if T/ S ratio was lower than 1, the DNA sample had shorter telomeres. PCR efficiency for TL measurement varied between 95% and 102%. Inter plate variation for T & S was 3.5% & 3.4%, respectively. Intra-plate variation for T was 0.42% & 0.28% for S.

Microsatellite instability

Microsatellite instability (MSI) status was determined by the molecular testing of 5 mononucleotide repeat markers (Bethesda consensus panel, BAT 25, BAT 26, NR 21, NR 24, NR 27) that were run as a pentaplex using fluorescently labeled primers and standard PCR. Fragment analysis was performed on ABI 3130 (Applied Biosystems). Final comparison between tumor and non-tumor DNA short tandem repetition profiles was analyzed with GeneMapper v4.1 software (Applied Biosystems). Tumor specimen was classified as MSI when 2 or more loci were unstable.

Statistical analysis

TL measured in the tissues and blood cells of CRC patients were expressed as median and range to characterize individual groups (by means of SAS descriptive statistics, see below). Statistical analyses were conducted on natural data by using non-parametrical tests. Differences between TL in two distinct tissues were analyzed using non-parametrical ANOVA (Wilcoxon signed-rank test). An interquartile range was defined as the distribution of TL values between 25th and 75th percentiles. The relationship between the patient age at diagnosis and TL values was calculated by Pearson correlation coefficient. TL ratio was expressed as TL in tumor/TL in adjacent mucosa. If TL ratio exceeded 1, telomeres were longer in tumor tissue than the adjacent mucosa. The curves for overall survival (OS) were derived by the Kaplan–Meier log-rank test. OS was defined as the time from the surgery to the date of death, or the date of last follow up. Based on the median TL cut-off, all CRC patients were stratified into two groups. Statistical analyses were conducted using SAS Institute Inc. software (Cary, NC, USA). The final results were graphically illustrated using Prism8 GraphPad software (San Diego, California, USA). Statistical significance for all tests was set at P-value=0.05.

Results

The median age for the group of CRC patients, for whom tissue pairs were available, at the time of diagnosis was 68 years (range 33-96 years); of those 62.9% were men and 37.1% were women. The patients for whom metastasized tissues were available without primary tumors, had median age of 63 (range 39-79 years); 66.1% men and 33.9% women. **Table I** summarizes TL data in tumor and adjacent mucosa tissues of CRC patients along with various personal and clinicopathological characteristics.

Telomere length in tumor tissue, adjacent mucosa, peripheral blood lymphocytes, liver metastases, and adjacent liver tissues

TL in tumor tissues (n=696) was statistically significantly shorter (median [inter quartile range]: 0.99 [0.65-1.50]) than the adjacent mucosa (n=677, 1.29 [0.83-2.26], Wilcoxon test, P<0.0001; **Figure I**). Shorter TL in tumor tissue (i.e. TL ratio <1) was observed in tumors from 74% patients, while for the remaining 26% of CRC patients, tumors had longer TL in tumors than the adjacent mucosa (based on 661 comparisons, where the data for TL in both tissues were available). We did not find any difference between TL in peripheral blood lymphocytes (PBL, n=164, 0.76 [0.56-1.04]) and corresponding tumor tissues (0.78 [0.58-1.04], P=0.2).

TL was statistically significantly higher in 10 (83%) primary tumor tissue samples (1.62 [1.34-1.94]) than in respective liver metastases (0.71 [0.52-1.01], P<0.0001), which was measured in tumors from only 12 CRC patients (**Figure II**).

TL was statistically significantly (P=0.0005) shorter in a group of liver metastasized tumors (0.76 [0.56-0.98]) than in the adjacent non-cancerous liver tissues (0.86 [0.76-1.02]) from 122 patients.

Telomere length *vs.* tumor localization

TL in tumor tissues was shorter in the proximal colon ((diagnosis C18.0-C18.4), 0.87 [0.62-1.2]) than in the distal part of the colon (C18.5-C19, 1.03 [0.65-1.53], P=0.006) and rectum (C20, 1.27 [0.8-1.91], P<0.0001; **Figure III**). Adjacent mucosa in CRC patients with proximal tumor origin had shortest TL (1.11 [0.79-1.86]). However, the longer TL in the adjacent mucosa of distally located colon tumors (1.37 [0.89-2.42]) was not statistically significant (P=0.18) than that in proximal colon. Interestingly, TL in adjacent mucosa within

the proximal tumor origin was statistically significantly shorter than the TL in non-cancerous mucosa attributable to rectal tumors (1.50 [0.93-3.05], P=0.03).

Telomere length *vs.* microsatellite instability status

Tumor tissues from patients with MSI had statistically significantly shorter TL (n=61, 0.72 [0.60-1.04]) than the microsatellite stable (MSS) tumors (n=553, 0.99 [0.70-1.30], P=0.009). MSI tumors localized within the proximal part of the colon (n=51, 83.6% from all MSI patients) had statistically significantly shorter TL in tumor tissue (0.72 [0.61-1.07] than the MSS tumors in proximal colon (1.12 [0.75-1.62], P=0.001). MSS tumors in the proximal colon also had statistically significantly shorter TL than MSS tumors from the rectum (P<0.0001).

Telomere length *vs.* TNM status

Increased TL was observed in tumors with increased TNM stages. A similar trend was not observed in the adjacent mucosa. Patients with TNM stage I had statistically significantly shorter TL in tumor tissue (P=0.001; **Figure IV**) than in tumors with TNM stages II+III+IV. The TL ratio in colon TNM stage I was statistically significantly lower than TNM stages II+III+IV (P=0.001). Any difference in the TL ratio was not observed in TNM stages of rectal tumors.

Telomere length *vs.* tumor histology

TL was statistically significantly shorter in mucinous cancer tissue (n=80, 1.00 [0.66-1.44]) than in tubular cancer tissue (n=287, 1.46 [1.07-2.13], P<0.0001). TL in the adjacent mucosa of patients diagnosed to have mucinous tumors (n=77, 1.41 [0.92-2.14]) was also statistically significantly shorter than the adjacent mucosa of CRC patients with tubular carcinoma (n=281, 2.31 [1.54-3.98], P<0.0001).

Overall survival *vs.* telomere length ratio

Patients with a TL ratio higher than median (0.70) had a statistically significantly poorer OS than those patients with lower TL ratios (P=0.02; **Figure V**). We also analyzed the possible effect of TL differences on OS in association with MSS and MSI tumor characteristics. We did

not observe any significant difference in OS between CRC patients with MSI or MSS tumors (P=0.16).

Discussion

In this study, we observed that the telomeres were consistently shorter in a majority of tumor tissues than in adjacent mucosa. The proliferating activity of the rapidly growing tumors underlies this phenomenon as described in some studies carried out in smaller groups of CRC patient [15-17]. However, some studies have also shown contradictory results [18, 19]. We observed statistically significantly shorter TL in the tumors classified as TNM stage I than in tumors with advanced stages of the disease. One of the probable interpretations could be that the tumors with longer telomeres have increased possibility of advancing to higher disease stages than the tumors with short telomeres. We also demonstrated a prolonged OS in CRC patients with decreased TL ratio between tumor tissues and adjacent mucosa. A similar relationship between TL and OS, measured in 57 CRC patients, was observed earlier [15].

The determination of TL in paired primary tumors and metastatic lesions and paired liver samples and metastatic lesions showed excessive telomere erosion in the vast majority of metastases. Since all patients with distant metastases underwent various regimens of chemotherapy, short telomeres could be related to the treatment.

Studies investigating the relationship between MSI status and TL remain rare and inconsistent [16, 20]. Boardman *et al.* reported that MSS CIN- rectal cancers have statistically significantly longer telomeres than MSS CIN+ rectal samples [20]. We observed shorter telomeres in MSI tumor tissues than in MSS tumors. Previously studies [21] & [16], conducted on 55 & 118 CRC patients, respectively, support our observations of shorter TL in MSI CRC patients than in MSS patients. An explanation for pronounced shorter telomeres in MSI tumors may be the accumulation of unfixed deletions caused by DNA polymerase slippage events during DNA replication, as suggested by [22]. Such deletions could result in decreased levels of shelterin subunits at telomeres resulting in a disruption of telomere homeostasis [23]. We also observed that MSS tumors in the proximal colon had shorter telomeres than those arising in the rectum, which is in accordance with earlier reports [16]. While MSI status and MMR deficiency may only partly explain shorter telomeres in tumors arising at the proximal colon.

MSI tumors occur predominantly in in the proximal part of the colon [24-26]. Indeed, we found shorter TL in tumors arising from the proximal part of the colon than those in the distal colon and rectum. Those results highlight differences in the molecular carcinogenesis

between the colon and rectum [27]. The assumption that TL in malignant tissue is influenced by the site of tumor origin has been previously addressed by several authors, some of which are contradictory probable owing to limited sample sizes.

Mucinous histology can be counted as another characteristic of MSI tumors. Even though mucinous tumors are predominantly located in the proximal colon, the proportion of patients with mucinous histology in our cohort was not statistically significantly different in relation to other tumor origin sites. However, we found that patients with mucinous tumor histology had statistically significantly shorter TL in the tumor and adjacent mucosa, than the patients with tubular carcinoma.

Overall, TL was shorter in tumor tissues than in adjacent mucosa, in lower (initial) stages, in proximal colon and tumors with MSI instability. Furthermore, metastases originating from primary CRC tumors had shorter telomeres than the adjacent non-cancerous liver tissues. Finally, the smaller TL ratio between tumor tissues and adjacent mucosa, conferring to shorter telomeres in tumor, may represent a positive prognostic factor. The characteristics of TL in relation to the CRC heterogeneity emerges as an urgent predictive issue since the advent of therapeutically concepts based on targeting telomere or telomerase inhibition in order to overcome resistance, lack of drug sensitivity, toxicity etc. The understanding of TL in CRC with different clinicopathological features will be an important step.

Acknowledgements:

This study was supported by the Grant Agency of the Czech Republic (GA CR 18-09709S), the Internal Agency of the Ministry of Health of the Czech Republic (AZV 17-30920A, NV18-03-00199, AZV 15-27580A and NV19-09-00237), and by the Centrum of Clinical and Experimental liver surgery (UNCE/MED/006), the National Sustainability Program I (NPU I) Nr. LO1503 provided by the Ministry of Education Youth and Sports of the Czech Republic

References

1. Blackburn, E.H., C.W. Greider, and J.W. Szostak, *Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging.* Nat Med, 2006. **12**(10): p. 1133-8.
2. Heidenreich, B. and R. Kumar, *TERT promoter mutations in telomere biology.* Mutat Res, 2017. **771**: p. 15-31.
3. Kawanishi, S. and S. Oikawa, *Mechanism of telomere shortening by oxidative stress.* Ann N Y Acad Sci, 2004. **1019**: p. 278-84.
4. Savage, S.A., *Beginning at the ends: telomeres and human disease.* F1000Res, 2018. **7**.
5. Stewenius, Y., et al., *Structural and numerical chromosome changes in colon cancer develop through telomere-mediated anaphase bridges, not through mitotic multipolarity.* Proc Natl Acad Sci U S A, 2005. **102**(15): p. 5541-6.
6. Grady, W.M. and S.D. Markowitz, *The molecular pathogenesis of colorectal cancer and its potential application to colorectal cancer screening.* Dig Dis Sci, 2015. **60**(3): p. 762-72.
7. Aaltonen, L.A., et al., *Clues to the pathogenesis of familial colorectal cancer.* Science, 1993. **260**(5109): p. 812-6.
8. Cawthon, R.M., *Telomere length measurement by a novel monochrome multiplex quantitative PCR method.* Nucleic Acids Res, 2009. **37**(3): p. e21.
9. Rachakonda, S., et al., *Telomere length and survival in primary cutaneous melanoma patients.* Sci Rep, 2018. **8**(1): p. 10947.
10. O'Callaghan, N., et al., *A quantitative real-time PCR method for absolute telomere length.* Biotechniques, 2008. **44**(6): p. 807-9.
11. Kroupa, M., et al., *Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients.* Genes Chromosomes Cancer, 2018. **57**(2): p. 61-69.
12. Rachakonda, S., et al., *Telomere length, telomerase reverse transcriptase promoter mutations, and melanoma risk.* Genes Chromosomes Cancer, 2018. **57**(11): p. 564-572.
13. Srinivas, N., et al., *Telomere length, arsenic exposure and risk of basal cell carcinoma of skin.* Carcinogenesis, 2019.
14. Bustin, S.A., et al., *The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.* Clin Chem, 2009. **55**(4): p. 611-22.
15. Gertler, R., et al., *Telomere length and human telomerase reverse transcriptase expression as markers for progression and prognosis of colorectal carcinoma.* J Clin Oncol, 2004. **22**(10): p. 1807-14.
16. Rampazzo, E., et al., *Relationship between telomere shortening, genetic instability, and site of tumour origin in colorectal cancers.* Br J Cancer, 2010. **102**(8): p. 1300-5.
17. Suraweera, N., et al., *Relative telomere lengths in tumor and normal mucosa are related to disease progression and chromosome instability profiles in colorectal cancer.* Oncotarget, 2016. **7**(24): p. 36474-36488.
18. Katayama, S., et al., *Clinical usefulness of telomerase activity and telomere length in the preoperative diagnosis of gastric and colorectal cancer.* J Cancer Res Clin Oncol, 1999. **125**(7): p. 405-10.

19. O'Sullivan, J., et al., *Telomere length in the colon declines with age: a relation to colorectal cancer?* Cancer Epidemiol Biomarkers Prev, 2006. **15**(3): p. 573-7.
20. Boardman, L.A., et al., *Correlation of Chromosomal Instability, Telomere Length and Telomere Maintenance in Microsatellite Stable Rectal Cancer: A Molecular Subclass of Rectal Cancer.* Plos One, 2013. **8**(11).
21. Takagi, S., et al., *Relationship between microsatellite instability and telomere shortening in colorectal cancer.* Dis Colon Rectum, 2000. **43**(10 Suppl): p. S12-7.
22. Chatterjee, N. and G.C. Walker, *Mechanisms of DNA damage, repair, and mutagenesis.* Environ Mol Mutagen, 2017. **58**(5): p. 235-263.
23. Loayza, D. and T. De Lange, *POT1 as a terminal transducer of TRF1 telomere length control.* Nature, 2003. **423**(6943): p. 1013-8.
24. Walther, A., R. Houlston, and I. Tomlinson, *Association between chromosomal instability and prognosis in colorectal cancer: a meta-analysis.* Gut, 2008. **57**(7): p. 941-50.
25. Yamauchi, M., et al., *Colorectal cancer: a tale of two sides or a continuum?* Gut, 2012. **61**(6): p. 794-7.
26. Missiaglia, E., et al., *Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features.* Ann Oncol, 2014. **25**(10): p. 1995-2001.
27. Paschke, S., et al., *Are Colon and Rectal Cancer Two Different Tumor Entities? A Proposal to Abandon the Term Colorectal Cancer.* Int J Mol Sci, 2018. **19**(9).

	RTL ratio (RTL tumor / RTL adjacent mucosa)			age-adjusted	RTL tumor			age-adjusted	RTL adjacent mucosa			age-adjusted
		mean ± SD	median [IQR]	P value		mean ± SD	median [IQR]	P value		mean ± SD	median [IQR]	P value
	All				All				All			
Sex	n = 661*				n = 696				n = 677			
male	415	0.85 ± 0.56	0.75 [0.51 - 1.03]		438	1.27 ± 0.92	1.01 [0.69 - 1.54]		423	1.94 ± 1.73	1.30 [0.83 - 2.26]	
female	246	0.76 ± 0.40	0.72 [0.48 - 0.98]	0.036*	258	1:21 ± 0.91	0.99 [0.64 - 1.40]	0.38	254	1.94 ± 1.57	1.32 [0.86 - 2.42]	0.98
Smoking status	n = 635				n = 671				n = 650			
smokers	276	0.89 ± 0.58	0.78 [0.54 - 1.08]		293	1.29 ± 1.01	1.01 [0.70 - 1.49]		285	1.81 ± 1.51	1.25 [0.85 - 2.16]	
nonsmokers	359	0.76 ± 0.44	0.70 [0.46 - 0.95]	0.0051*	378	1.25 ± 0.86	1.02 [0.66 - 1.56]	0.99	365	2.09 ± 1.80	1.42 [0.87 - 2.58]	0.017*
Tumor stage	n = 634				n = 670				n = 650			
I (Ref)	88	0.67 ± 0.34	0.62 [0.39 - 0.90]		95	1.08 ± 0.78	0.86 [0.58 - 1.39]		91	2.10 ± 2.00	1.35 [0.76 - 2.10]	
II	224	0.84 ± 0.55	0.75 [0.49 - 1.03]		236	1.23 ± 0.93	0.99 [0.65 - 1.55]		229	1.91 ± 1.64	1.33 [0.79 - 2.47]	
III	200	0.82 ± 0.49	0.74 [0.51 - 1.00]		209	1.32 ± 0.93	1.10 [0.74 - 1.54]		206	1.98 ± 1.66	1.27 [0.89 - 2.42]	
IV	122	0.86 ± 0.51	0.76 [0.53 - 1.10]	0.0013*	130	1.38 ± 0.99	1.12 [0.72 - 1.73]	0.0012*	124	2.02 ± 1.60	1.49 [0.95 - 2.35]	0.77
I + II (Ref)	312	0.79 ± 0.51	0.71 [0.46 - 0.98]		331	1.19 ± 0.89	0.95 [0.63 - 1.47]		320	1.96 ± 1.75	1.34 [0.79 - 2.43]	
III + IV	322	0.84 ± 0.50	0.75 [0.52 - 1.03]	0.3	339	1.34 ± 0.95	1.10 [0.74 - 1.57]	0.095	330	2.00 ± 1.63	1.36 [0.91 - 2.39]	0.99
Tumor localisation	n = 656				n = 693				n = 672			
proximal (C18.0 - C18.4) (Ref)	222	0.78 ± 0.47	0.73 [0.46 - 0.98]		239	1.02 ± 0.67	0.87 [0.62 - 1.2]		231	1.73 ± 1.59	1.11 [0.79 - 1.86]	
distal (C18.5 - C19)	272	0.79 ± 0.42	0.71 [0.51 - 0.98]	0.70	286	1.32 ± 1.07	1.03 [0.65 - 1.53]	0.0057*	277	2.03 ± 1.78	1.37 [0.89 - 2.42]	0.18
rectum (C20)	162	0.92 ± 0.66	0.80 [0.54 - 1.11]	0.023*	168	1.46 ± 0.86	1.27 [0.80 - 1.91]	< 0.0001*	164	2.11 ± 1.57	1.50 [0.93 - 3.05]	0.031*
distal (C18.5 - C19) (Ref)	272	0.79 ± 0.42	0.71 [0.51 - 0.98]		286	1.32 ± 1.07	1.03 [0.65 - 1.53]		277	2.03 ± 1.78	1.37 [0.89 - 2.42]	
rectum (C20)	162	0.92 ± 0.66	0.80 [0.54 - 1.11]	0.017*	168	1.46 ± 0.86	1.27 [0.80 - 1.91]	0.25	164	2.11 ± 1.57	1.50 [0.93 - 3.05]	0.65
MSI status	n = 588				n = 614				n = 604			
stable	523	0.81 ± 0.49	0.73 [0.49 - 1.00]		553	1.34 ± 0.96	1.12 [0.75 - 1.62]		544	2.10 ± 1.72	1.44 [0.93 - 2.60]	
unstable	57	0.69 ± 0.36	0.67 [0.39 - 0.91]	0.098	61	0.90 ± 0.52	0.72 [0.61 - 1.07]	0.0012*	60	1.83 ± 1.72	1.39 [0.77 - 1.99]	0.33
Tumor histology	n = 353				n = 367				n = 358			
mucinous carcinoma	75	0.79 ± 0.54	0.65 [0.49 - 0.93]		80	1.16 ± 0.64	1.00 [0.66 - 1.44]		77	1.90 ± 1.55	1.41 [0.92 - 2.14]	
tubular carcinoma	278	0.74 ± 0.52	0.62 [0.41 - 0.91]	0.28	287	1.78 ± 1.10	1.46 [1.07 - 2.13]	< 0.0001*	281	3.00 ± 1.88	2.31 [1.54 - 3.98]	< 0.0001*
Tumor grade	n = 629				n = 662				n = 644			
low-grade	519	0.81 ± 0.52	0.72 [0.49 - 1.01]		546	1.29 ± 0.93	1.01 [0.70 - 1.57]		532	2.02 ± 1.68	1.39 [0.89 - 2.52]	
high-grade	110	0.88 ± 0.43	0.82 [0.58 - 1.10]	0.16	116	1.13 ± 0.88	1.0 [0.64 - 1.32]	0.13	112	1.68 ± 1.71	0.99 [0.72 - 1.87]	0.051

Table I: RTL values and clinicopathological characteristics of studied population

*Numbers of patients may not add up to 100% of available subjects because of RTL measurement failure or due to the lack of clinicopathological data

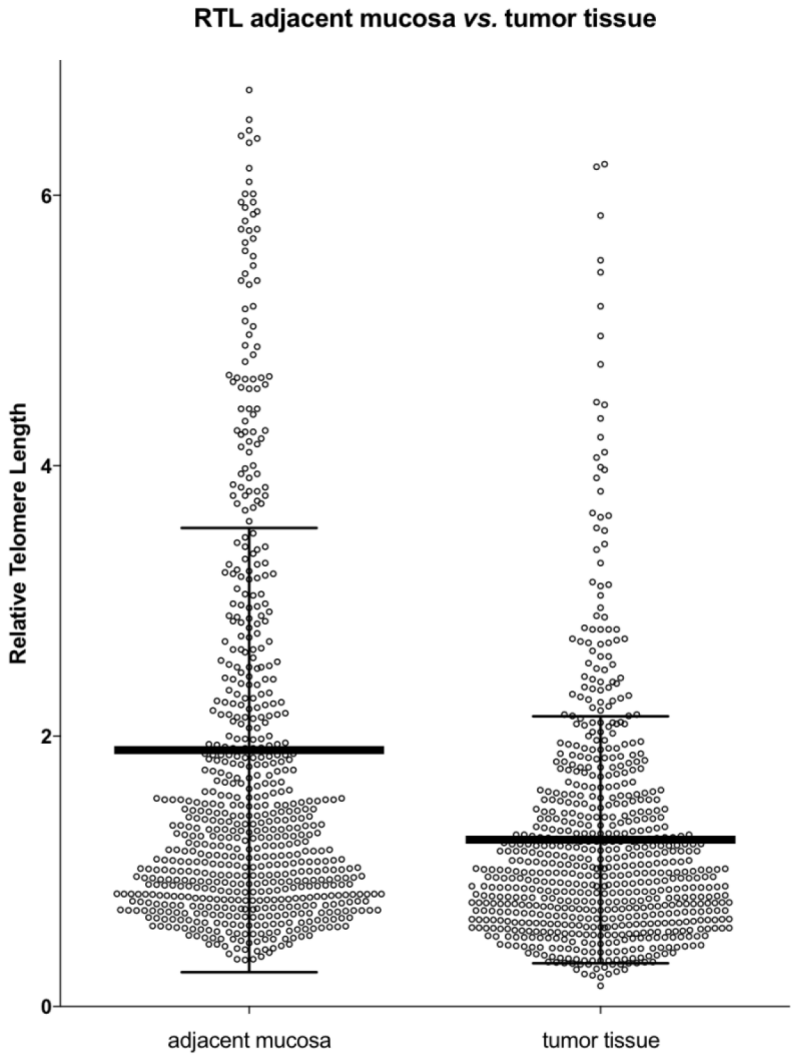


Figure I: Comparison of RTL in adjacent mucosa and tumor tissue (P < 0.0001)

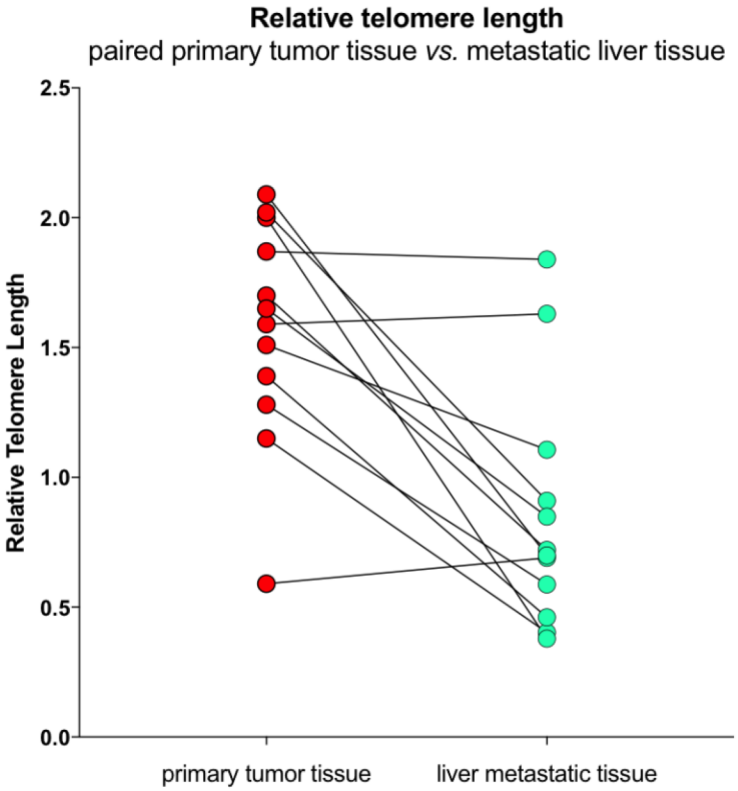


Figure II: Comparison of RTL in paired tumor and liver metastatic tissues (P < 0.0001)

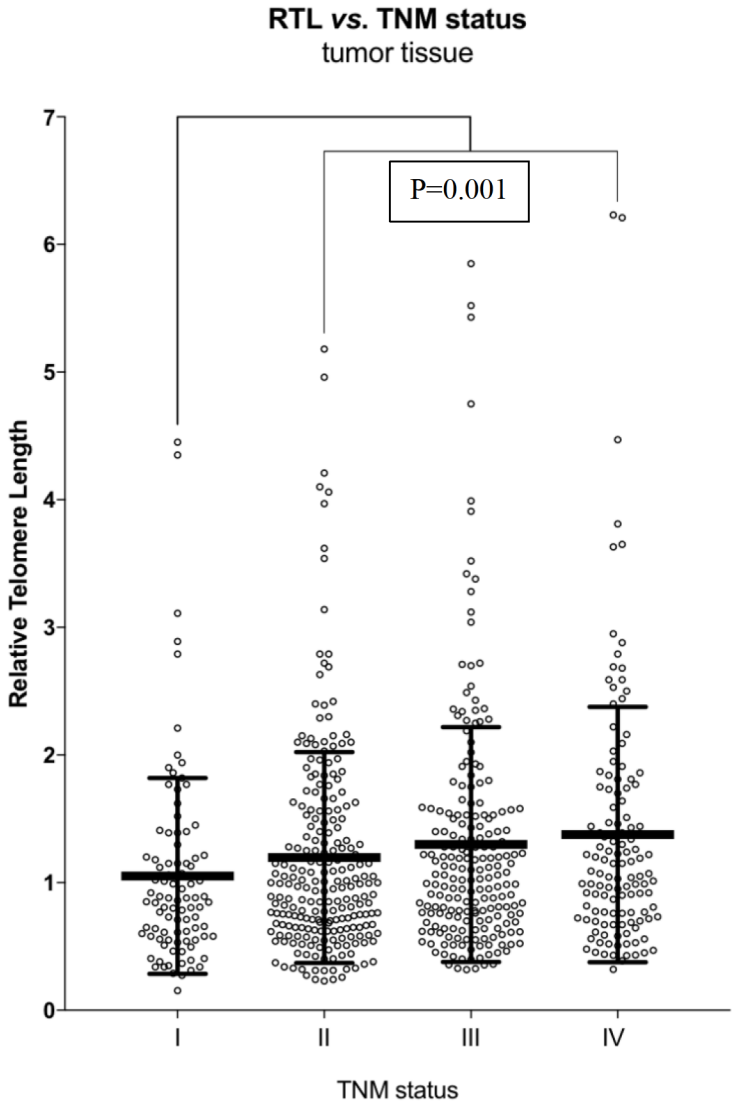


Figure III: Association between RTL in tumor tissue and TNM status. (P=0.001)

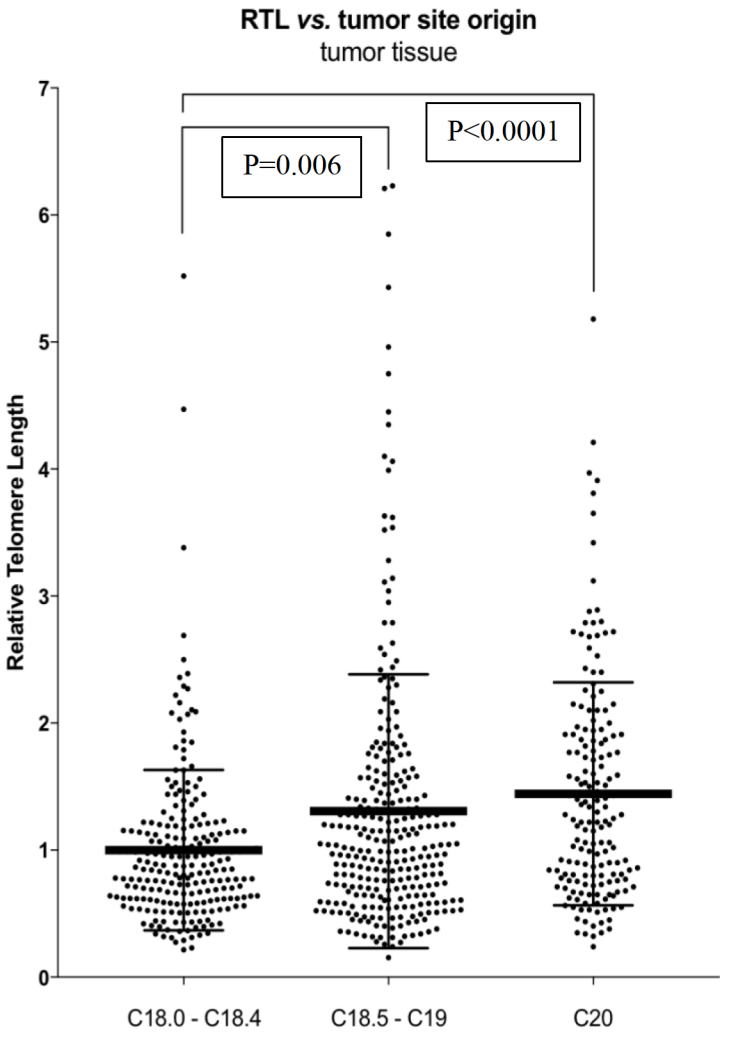


Figure IV: Association between RTL in tumor tissue and tumor site origin; proximal *vs.* distal site (P=0.006), proximal site *vs*. rectum (P<0.0001)

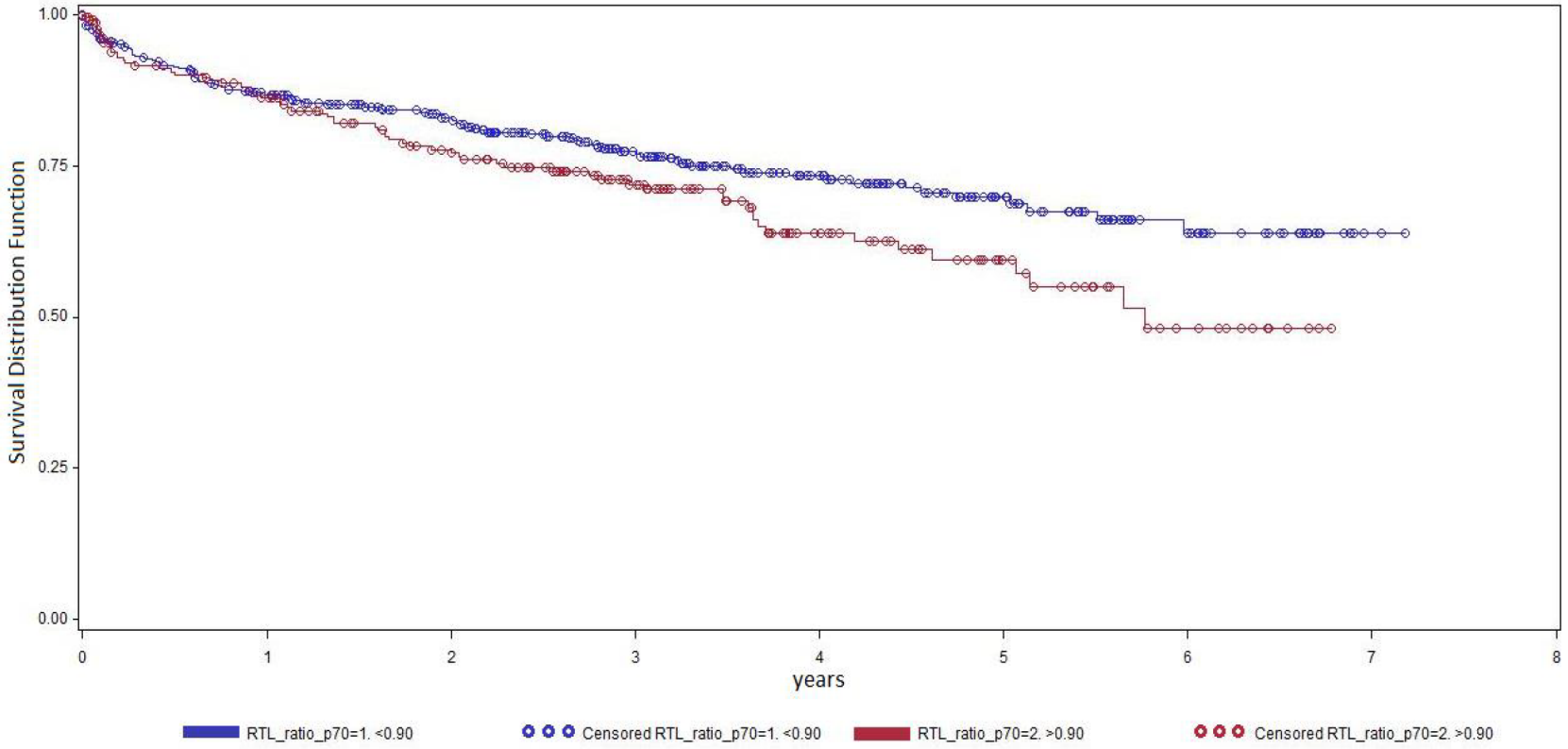


Figure V: Kaplan-Meier OS curves stratified for RTL ratio (RTL ratio cut-off=0.7, p=0.022)

Telg	5′ ACACTAAGGTTTGGGTTTGGGTTTGGGTTTGGGTTAGTGT 3′
Telc	5′ TGTTAGGTATCCCTATCCCTATCCCTATCCCTATCCCTAACA 3′
Albugcr2	5′ CGGCGGCGGGCGGCGCGGGCTGGGCGGCCATGCTTTTCAGCTCTGCAAGTC 3′
Albdgcr2	5′ GCCCGGCCCGCCGCGCCCGTCCCGCCGAGCATTAAGCTCTTTGGCAACGTAGGTTTC 3′

Supplementary Table I: Primers for RTL measurement

Manuscript VII

Genetic variation of acquired structural chromosomal aberrations.

Vodicka P, Musak L, Vodickova L, Vodenkova S, Catalano C, **Kroupa M**, Naccarati A, Polivkova Z, Vymetalkova V, Försti A, Hemminki K.

Mutat Res. 2018 Dec;836(Pt A):13-21. **IF: 1.996**

Mutat Res Gen Tox En 836 (2018) 13–21

Contents lists available at ScienceDirect

Mutat Res Gen Tox En

journal homepage: www.elsevier.com/locate/gentox

Minireview

Genetic variation of acquired structural chromosomal aberrations

Pavel Vodicka[a,b,c,*], Ludovit Musak[d], Ludmila Vodickova[a,b,c], Sona Vodenkova[a,b,e], Calogerina Catalano[f], Michal Kroupa[a,c], Alessio Naccarati[a,g], Zdena Polivkova[e], Veronika Vymetalkova[a,b,c], Asta Försti[f,h], Kari Hemminki[f,h]

[a] Department of Molecular Biology of Cancer, Institute of Experimental Medicine, Czech Academy of Sciences, Prague, 14220, Czech Republic
[b] Institute of Biology and Medical Genetics, First Faculty of Medicine, Charles University, Prague, 12800, Czech Republic
[c] Faculty of Medicine and Biomedical Center in Pilsen, Charles University, Pilsen, 30605, Czech Republic
[d] Biomedical Center Martin, Comenius University in Bratislava, Jessenius Faculty of Medicine, Martin, 03601, Slovakia
[e] Department of Medical Genetics, Third Faculty of Medicine, Charles University, Prague, 10000, Czech Republic
[f] Division of Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, D69120, Germany
[g] Italian Institute for Genomic Medicine (IIGM), Torino, 10126, Italy
[h] Center for Primary Health Care Research, Lund University, Malmö, 214 28, Sweden

ARTICLE INFO

Keywords:
Chromosomal aberrations
Genetics
Xenobiotic metabolizing enzymes
DNA repair
Mitotic checkpoints
Cyclin D1

ABSTRACT

Human malignancies are often hallmarked with genomic instability, which itself is also considered a causative event in malignant transformation. Genomic instability may manifest itself as genetic changes in the nucleotide sequence of DNA, or as structural or numerical changes of chromosomes. Unrepaired or insufficiently repaired DNA double-strand breaks, as well as telomere shortening, are important contributors in the formation of structural chromosomal aberrations (CAs). In the present review, we discuss potential mechanisms behind the formation of CAs and their relation to cancer. Based on our own studies, we also illustrate how inherited genetic variation may modify the frequency and types of CAs occurring in humans. Recently, we published a series of studies on variations in genes relevant to maintaining genomic integrity, such as those encoding xenobiotic-metabolising enzymes, DNA repair, the tumour suppressor TP53, the spindle assembly checkpoint, and cyclin D1 (CCND1). While individually genetic variation in these genes exerted small modulating effects, in interactions they were associated with CA frequencies in peripheral blood lymphocytes of healthy volunteers. Moreover, we observed opposite associations between the *CCND1* splice site polymorphism rs9344 G870A and the frequency of CAs compared to their association with translocation t(11,14). We discuss the functional consequences of the *CCND1* gene in interplay with DNA damage response and DNA repair during malignant transformation.

Our review summarizes existing evidence that gene variations in relevant cellular pathways modulate the frequency of CAs, predominantly in a complex interaction. More functional/mechanistic studies elucidating these observations are required. Several questions emerge, such as the role of CAs in malignancies with respect to a particular phenotype and heterogeneity, the formation of CAs during the process of malignant transformation, and the formation of CAs in individual types of lymphocytes in relation to the immune response.

1. Introduction

Human cancers are often associated with chromosomal instability, including both numerical and structural chromosomal aberrations (CAs) [1–3]; these are also considered causative events in malignant transformation [4]. Numerical chromosome variance refers to a multiple of haploid sets or any non-euploid deviation in chromosome number. Numerical aberrations are often incompatible with cell viability. On the contrary, structural CAs are commonly presented in viable human cells. Structural CAs may be a secondary consequence of malignant transformation or may directly trigger cancer onset, as in case of chronic myeloid leukaemia.

Structural CAs may be specific, such as translocations and inversions, or nonspecific, such as chromatid breaks, fragmented or missing parts of chromosomes, and fusion resulting in dicentric and ring chromosomes [5]. The former are often recurrent and they are analysed by molecular cytogenetic methods (sequencing, fluorescent *in situ* hybridization). The later are scored by classical cytogenetic techniques, able to recognise chromosome type aberrations (CSAs) and chromatid type aberrations (CTAs) according to morphological changes [6].

* Corresponding author.
E-mail address: pvodicka@biomed.cas.cz (P. Vodicka).

https://doi.org/10.1016/j.mrgentox.2018.05.014
Received 22 January 2018; Received in revised form 24 April 2018; Accepted 10 May 2018
Available online 19 May 2018

In the present review, we first discuss the potential mechanisms behind the formation of structural CAs and their relation to cancer. Later on the basis of available studies, we describe how inherited variations in genes relevant to important cellular pathways (such as those encoding xenobiotic-metabolising enzymes, DNA repair, the tumour suppressor TP53, spindle assembly checkpoint and cyclin D1) modify the frequency and types of CAs in humans. Genetic variation in relevant genes has been shown to modulate inter-individual variability in their ultimate functions. Moreover, the knowledge of the full effects of gene–gene interactions on chromosomal aberrations remains scarce.

1.1. Formation of structural CA and its relation to cancer risk

Unrepaired or insufficiently repaired DNA double-strand breaks (DSBs) as well as telomerase dysfunction are substantial players in the formation of structural CAs. Morphologically distinct types of CSAs and CTAs emerge depending on cell cycle stage and the type of DNA damage: whereas CSAs arise during G0 or G1 phase on a basis of DSBs induced by genotoxic damage, CTAs are formed due to insufficiently repaired DSBs during late S or G2 phase of the cell cycle [5,7,8]. Taken together in both mechanisms of CA formation, the repair of DNA damage plays a critical role [7], and, as was recently postulated, telomere shortening also acts as an important contributor to CA formation (Fig. 1). Telomeres are nucleoprotein structures composed of TTAGGG tandem repeats that cap ends of eukaryotic chromosomes. The telomere complex regulates a “cellular mitotic clock” and protects chromosomes against exonucleolytic degradation, DNA damage and chromosomal instability [9–11]. There is growing evidence of shorter telomeres being associated with the increased frequency of CAs in peripheral blood lymphocytes (PBLs), in particular with CSAs [12,13]. Telomere shortening, long hypothesized as being associated with cancer development [14,15], induces genomic and chromosomal instability [16,17]. Telomere becomes shorter at each round of replication, and critically shortened telomeres may be poorly end-capped and recognized as DSB by repair machinery [18]. These processes may underlie end-to-end chromosome fusions, the initiation of breakage-fusion-breakage cycles, and the development of numerical and structural CAs [19–22]. On the other hand, the reactivation of telomerase stabilizes telomeres and immortalizes premalignant cells, thus enabling cancer progression [23,24]. In this context, our recent report suggested that altered DSB repair, measured by sensitivity towards mutagen bleomycin in PBLs, occurs particularly in colorectal carcinogenesis. Irrespective of cancer type telomere shortening is associated with a decreased capacity to repair DSB [25].

Many human cancers and neoplastic cells exhibit chromosomal abnormalities [4]. In their review from 2007, Mitelman et al. summarized the role of translocations and related fusion genes in the initiation of cancer. Their catalogue included 358 gene fusions, of which involved 337 different genes. While gene fusions (i.e., the juxta positioning of two genes leading to the translation of a deregulated and/or chimeric protein) were initially identified; and their clinical value was recognized in malignant haematological disorders, they have become increasingly important also in solid tumors [26]. Mitelman et al. predicted in 2007 that fusion genes account for 20% of human cancer morbidity. In a more recent review from year 2015, they reported on a huge increase in identified fusion genes, amounting to 9000 during the covered 3 years [27]. Currently the number comprises 11,124 identified fusion genes (Mitelman database of chromosomal aberrations and gene fusions in cancer; https://cgap.nci.nih.gov/chromosomes/Mitelman).

The evidence linking nonspecific CAs with cancer is not as overwhelming as it is for specific, recurrent CAs. However, an increase in the frequency of CAs has been found in incident cancer patients thus closely linking CAs with cancer development [28,29]. Some of the CAs observed are also generated during cancer development, nevertheless CA frequency in PBLs is considered to be an early marker of cancer susceptibility, based on the hypothesis that genetic damage in PBLs reflects similar damage in other body cells undergoing carcinogenesis [30].

1.2. Genetic predisposition to CAs

There are only limited reports exploring the inherited genetic basis of acquired structural CAs. To achieve this goal a genome-wide association study (GWAS) may substantially contribute with an aim to finding novel genetic variants associated with CAs, and potentially with cancer risk, to elucidate the possible functional effects of these variants by *in silico* predictions.

However, through a Czech-Slovak-German collaboration, we recently published a series of studies related to DNA repair [47], metabolism [38] and mitotic checkpoint [72] (overviewed separately). We compared genotype frequencies between healthy volunteers whose total CA frequency was > 2% and those with lower frequency of CAs; for CSAs and CTAs the threshold was 1%. The numbers of volunteers genotyped ranged from 282 to 2200 depending on the SNP and DNA availability (Table 2). All SNPs were individually genotyped and multivariable logistic regression analyses adjusted for age, sex, occupational exposure and smoking; conducted for total CAs, CTAs, and CSAs. SNPs were considered individually, and in all possible pairwise combinations within the three sets of genes. Those studies are reviewed below, but they appear to share some overall features: individual SNPs had weak associations, but gene–gene interactions were common and often involved partners from different pathways.

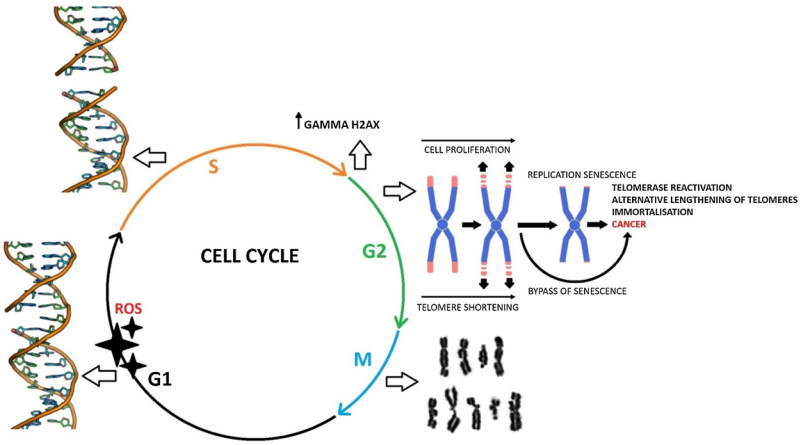


Fig. 1. Single-stranded DNA damage (a consequence of oxidative DNA damage in this particular case) if unrepaired gives rise to DSBs and CTAs (as assayed for by chemiluminescent measurement of γ-H2AX phosphorylation). Shortened telomeres may be poorly end-capped and recognized as DSB by repair machinery, resulting in the development of numerical and structural CAs.

Table 1
Background variables of age, smoking status, and occupational exposure ([34,45,70]) plus polymorphism rs1042522:G4C (Ex41119 G4C, Arg72Pro) in gene *TP53* in relation to chromosomal damage.

Variable	Persons	CAtot			CSA			CTA		
		Odds ratio	95% Confidence interval	P-value	Odds ratio	95% Confidence interval	P-value	Odds ratio	95% Confidence interval	P-value
CAtot (cases/control)	951/1245									
CTA (cases/control)	1042/1154									
CSA (cases/control)	983/1213									
Age (min, max, mean)	18, 88, 43	1.06[a]	1.00–1.13	0.07	1.07	1.00–1.13	0.04	0.95	0.90–1.01	0.11
Occupational environment (exposed/unexposed)	1207/989	2.36	1.97–2.83	< 0.01	1.73	1.45–2.06	< 0.01	1.64	1.38–1.96	< 0.01
Sex (M/F)	1171/1025	1.03	0.86–1.23	0.77	1.05	0.88–1.26	0.59	0.83	0.69–0.99	0.04
Smoker (S/NS)	614/1557	1.19	0.97–1.45	0.09	0.95	0.78–1.16	0.63	1.13	0.93–1.38	0.23
Arg72Arg[b]	278	–	–	–	–	–	–	–	–	–
Arg72Pro	269	0.91	0.64–1.29	0.60	1.14	0.81–1.62	0.45	1.11	0.79–1.58	0.52
Pro72Pro	46	1.03	0.54–1.96	0.93	0.82	0.92–1.60	0.55	1.42	0.76–2.66	0.27
Arg72Arg vs. Arg72Pro + Pro72Pro	–	0.93	0.66–1.30	0.66	1.03	0.73–1.44	0.88	1.16	0.84–1.61	0.37

[a] ORs for age were calculated for 10 year age difference.
[b] rs1042522:G4C (Ex41119 G4C, Arg72Pro).

Table 2
Polymorphisms in relevant genes of important cellular pathways (encoding xenobiotic-metabolising enzymes, DNA repair, the tumour suppressor TP53, spindle assembly checkpoint, and cyclin D1) evaluated in this study. The table additionally refers to the *in silico* predicted functional consequence of polymorphisms.

Genes	SNP ID	Amino Acid substitution	Alleles (major/minor)	NCBI	Chromosome	Location	Function*	Subjects with CAs
BER								
XRCC1	rs1799782	Arg194Trp	C > T	T > C	19q13.2	missense	deleterious	639
XRCC1	rs25489	Arg280His	G > A	G > A	19q13.2	missense	possibly deleterious	355
XRCC1	rs25487	Arg399Gln	G > A	A > G	19q13.2	missense	benign	1760
OGG1	rs1052133	Ser326Cys	C > G	A > G	3p26.2	missense	deleterious	1766
APEX1 (APE1)	rs1130409	Asp148Glu	T > G	T > G/A	14q11.2	missense	ambiguous	1299
NER								
XPA	rs1800975		G > A		9q22.3	5′UTR		1443
XPD	rs13181	Lys751Gln	G > A	C > A	19q13.3	missense	deleterious	1777
XPG	rs17655	Asp1104His	C > G	C > G	13q33	missense	deleterious	1756
XPC	rs2228001	Lys939Gln	A > C	C > A	3p25	missense	benign	1772
DSB								
XRCC2	rs3218536	Arg188His	G > A	G > A	7q36.1	missense	benign	813
XRCC3	rs861539	Thr241Met	C > T	C > T	14q32.3	missense	deleterious	1715
NBN (NBS1)	rs1805794	Glu185Gln	C > G	C > G	8q21	missense	benign	966
RAD54L	rs1048771	Ala730Ala	C > T	C > T/A	1p32	synonymous	splicing regulation	282
DDR								
TP53	rs1042522	Arg72Pro	G > C	C > G	17p13.1	missense	deleterious	593
CCND1	rs9344	Pro241Pro	G > A	G > A	11q13.3	synonymous	splicing regulation	730
XME								
CYP1B1/432	rs1056836	Leu432Val	C > G		2p22.2	missense	drug response	549
CYP1B1/453	rs1800440	Asn453Ser	A > G		2p22.2	missense	benign/likely benign	550
EPHX1	rs1051740	Tyr113His	T > C		1q42.12	missense	drug-response	1120
EPHX1	rs2234922	His139Arg	A > G		1q42.12	missense	drug-response	1120
NQO1	rs1800566	Pro187Ser	C > T		16q22.1	missense	pathogenic	1018
GSTP1	rs1695	Ile105Val	A > G		11q13.2	missense	drug-response	1493
GSTM1			allele deletion		1p13.3		deleterious	1499
GSTT1			allele deletion		22q11.23		deleterious	1489
SAC								
BUB1B	rs1801376	Gln349Arg	A > G		15q15.1	missense	benign	330
BUB3	rs3808960		G > T		10q26.13	promoter region	upstream variant 2KB	330
MAD2L1	rs903147		A > C		4q27	promoter region	upstream variant 2KB	330
CENPF	rs438034	Arg2943Gly	G > A		1q41	missense	missense	618
ESPL1	rs6580941		C > T		12q13.13	promoter region	intron variant, upstream variant 2KB	656
NEK2	rs701928		T > A		1q32.3	promoter region	intron variant, upstream variant 2KB	663
PTTG1	rs1862392		T > A		5q33.3	promoter region	upstream variant 2KB	729
ZWILCH	rs3087660		A > G		15q22.31	5′UTR	nc transcript variant, upstream variant 2KB, utr variant 5 prime	674
ZWINT	rs2241666	Arg187Gly	G > A		10q21.1	missense	intron variant, missense, nc transcript variant	662

*SIFT, PolyPhen-2-algorithms; SNP single nucleotide polymorphism; BER base excision repair pathway; NER nucleotide excision repair pathway; DSB double-strand break repair pathway; DDR DNA damage response pathway; SAC spindle assembly checkpoint genes XME xenobiotic-metabolising enzymes; Polymorphisms were studied on subjects with CAs; their numbers are shown in the right column.

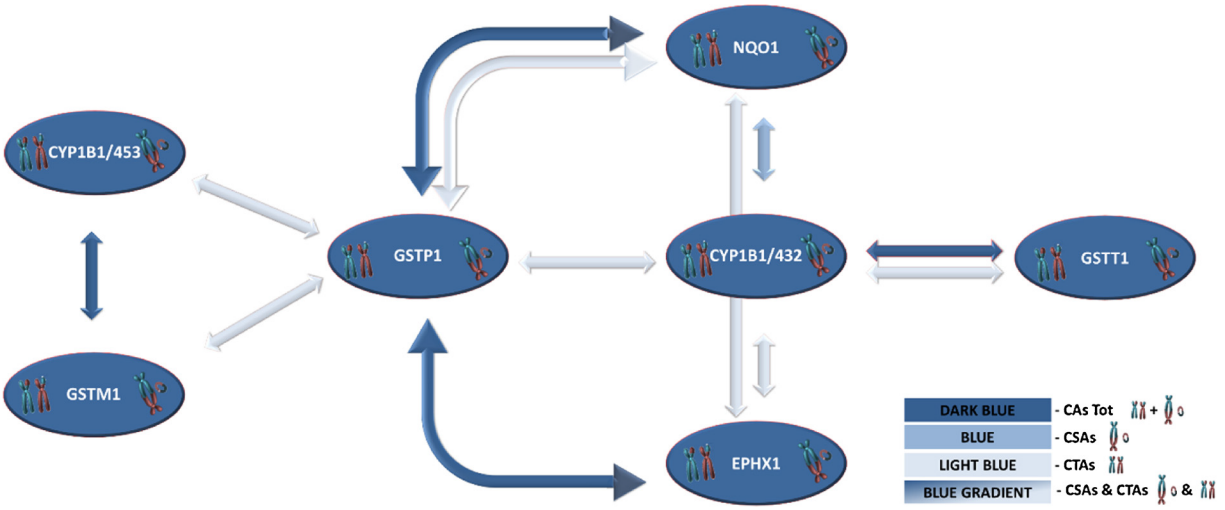


Fig. 2. Influence of metabolic gene polymorphisms on CAs in healthy volunteers [36]. The colour code indicates total CA (dark blue), CSA (blue), CTA (light blue), and CSA & CTA (blue gradient). The arrows indicate significant gene–gene interactions.

1.3. CAs and variants in genes encoding xenobiotic-metabolising enzymes (XMEs)

Only a limited number of reports have analysed effects of genetic predispositions on inter-individual variability in DNA and chromosomal damage [31–35]. In a recent study, we tested the associations of variants in genes encoding XMEs with CAs among the same cohort of volunteers described above [36]. We selected 5 genes for analysis: *cytochrome P450 1B1 (CYP1B1*, 2 separate polymorphisms), *epoxide hydrolase 1 (EPHX1), NAD(P)H:quinone oxidoreductase 1 (NQO1*), each coding for phase 1 enzymes, and *glutathione S-transferase P1 (GSTP1), glutathione S-transferases M1 (GSTM1), and T1 (GSTT1)*. All the 7 polymorphisms tested in our study have previously been shown to be functionally relevant [32–34,48]. Cytochromes P450 are the most important enzymes involved in the phase I biotransformation of various compounds. They catalyse a number of reactions modifying dietary, smoking, and environment-derived pre-carcinogens, and participate in the metabolism of endogenous compounds including hormones and bile acids. CYP1B1 is involved in activating polycyclic aromatic hydrocarbons (PAHs) or heterocyclic amines to reactive metabolites that cause DNA damage [37]. Epoxide hydrolase (EPHX1) converts reactive intermediates (epoxides) to less reactive, excretable diols. NAD(P)H:quinone oxidoreductase (NQO1) is a two-electron reductase involved in the biotransformation of quinones [39]. Glutathione S-transferases (GST), GSTM1, GSTP1, and GSTT1, belong to the most frequently studied XMEs in molecular epidemiology studies, playing a protective role against electrophilic chemicals. However, their outcomes in relation to CAs are often controversial: while Musak et al. observed significant association of CA frequency with *GSTT1-null* genotype, particularly in smoking tire plant workers [33], Rossi et al. did not record such an association [35]. In our analysis, we did not observe any modulating effect of GST polymorphisms on CA frequency, as assayed on 488 healthy individuals [40]. Our study describing the associations between functional polymorphisms in 6 genes encoding XME and CAs revealed that only *EPHX1* was individually associated with total CAs: high activity genotypes decreased CA frequency [33]. Tentative association of CAs with *EPHX1* polymorphisms has been shown by us earlier [33,37,38]. We also tested pairwise interactions between all 7 variants. A total of 12 nominally significant interactions were observed, 6 of them at $P < 0.01$ [36]. In most of the interactions, a *GST* variant was one partner of the pair (Fig. 2). The data provided evidence that variants in genes coding for metabolic enzymes (which individually have small effects) interact and are associated with CA frequencies in PBLs of healthy volunteers.

1.4. CAs and variants in DNA repair

Since human DNA is constantly exposed to physical (ultraviolet, ionizing radiation) and chemical (reactive oxygen species, alkylating, and aralkylating) damaging agents, the arising damage (particularly DSBs) may be critical for establishing persistent chromosomal alterations. CAs in PBLs reflect therefore inter-individual sensitivity to many genotoxic substances and serve as biomarkers of an early effect of genotoxic carcinogens and carcinogenic risk [29,41–45]. Efficient DNA repair machinery, comprising several distinct pathways, effectively maintains genomic integrity. Alterations in the DNA repair increase the vulnerability of the cells, resulting in an accumulation of mutations in the genome, and may ultimately lead to tumorigenesis [46].

Individual DNA repair capacity in response to DNA damage, effectively preventing an accumulation of CAs, is often modulated by the gene variants in different DNA repair pathways [47]. Indeed, we documented a correlation between gene variants involved in base excision repair (BER) and the corresponding BER repair capacity a decade ago [48]. However, single-nucleotide polymorphisms (SNPs) in non-coding regions and even changes in wobble bases that do not affect amino acid sequence, may be important as well [49].

The DNA repair genes that we studied encode proteins involved in BER (represented by *XRCC1, hOGG1 and APE1*), NER (*XPA*, *XPC*, *XPD* and *XPG*) and in double-strand break repair DSB (*XRCC2*, *XRCC3*, *NBS1* and *RAD54L*) [47]. Excision repair (both BER and NER) represents a crucial mechanism for a cells to preserve genomic integrity, since DSBs or permanent mutations can arise as a consequence of unrepaired or erroneously repaired single-stranded damage. BER removes small molecular lesions and oxidative DNA damage, in which XRCC1 protein acts as a coordinator of single-strand break repair proteins, hOGG1 is a glycosylase removing oxidative DNA adducts, and APE1 acts as endonuclease toward single-strand breaks. NER is a mechanism used by mammals for the elimination of bulky DNA lesions. Its importance is well documented on xeroderma pigmentosum or Cockayne syndrome disorders [50], associated with extreme UV light sensitivity and enormous skin cancer risk [51]. XPC (xeroderma pigmentosum complementation group C) protein participates in the recognition of lesions during global genomic-NER [52] – XPC enables the recruitment of two helicases – while XPB and XPD initiate the opening of the DNA around the lesion [53,54]. XPA and XPG proteins subsequently form a stable open structure and recruitment of ERCC1-XPF by XPA initiates the dual incision of the DNA 5′ and 3′ to the lesion by ERCC1-XPF and XPG, respectively [55,56]. The replication machinery subsequently fills the gap and seals the nick.

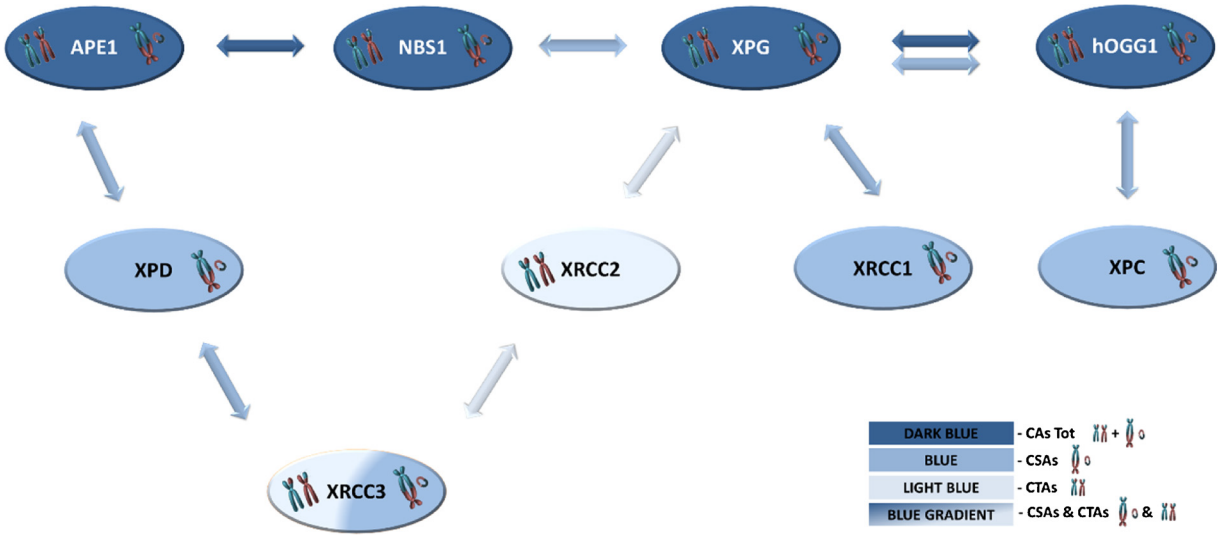


Fig. 3. Influence of DNA repair gene polymorphisms on CAs in healthy volunteers [47]. The colour code indicates total CA (dark blue), CSA (blue), CTA (light blue), and CSA & CTA (blue gradient). The arrows indicate significant gene–gene interactions.

DSB repair is a crucial mechanism for a cell to deal with the most serious DNA damage. Since DSBs have the potential to arrest cell cycle or even kill the cell, they have to be efficiently repaired to avoid such events. Different DNA repair mechanisms work in different stages of the cell cycle, namely homologous recombination (HR), non-homologous-end joining (NHEJ), and highly inaccurate microhomology-mediated end joining (MMEJ). When DSB occurs, a plethora of DNA repair proteins cooperate with each other to preserve genome integrity and functionality. NBS1 is one of the first proteins to accumulate at DSB sites. It creates a complex with MRE11 and RAD50 proteins, which recognizes the DNA damage site and promotes resection of its broken ends [57]. Interestingly, NBS1 is commonly overexpressed in different types of cancer, probably due to its participation in inaccurate MMEJ pathway [58,59]. Other key proteins such as XRCC2 and XRCC3 (the X-ray repair complementing defective repair in Chinese hamster cell 2 and 3, respectively) play an irreplaceable role during HR when an undamaged sister chromatid is invaded by single-stranded overhangs of the damaged strand to promote break repair. XRCC2 and XRCC3 (RAD51B, RAD51C and RAD51D) are paralogues of RAD51 recombinase. XRCC2 and XRCC3 create complexes with other RAD51 paralogues, which in turn facilitate recruitment and stabilisation of RAD51 at DSB sites [60,61]. The whole RAD51-mediated strand invasion is coordinated by RAD54L protein, which translocates along DSB sites and has diverse functions during HR, e.g. remodelling of chromatin structure, Rad51 filament assembly, and D-loop dissolution [62,63].

By analysing SNPs individually, we observed significantly lower CTA frequencies in association with the XPD Lys751Gln homozygous variant genotype (OR 0.64, P = 0.004). This association replicated with a similar significance level in a separate healthy population earlier [29]. An association of the heterozygous variant genotype in *RAD54L* with increased CSA frequency (OR 1.96, P = 0.03) was also found. By addressing gene–gene interactions, we discovered 14 interactions modulating CA frequencies, 9 modulating CTA and 12 modulating CSA frequencies with nominal significance. The interactions with strong statistical support are highlighted in Fig. 3. For total CAs, highly significant gene–gene interactions were observed mainly for genes participating in BER (*APE1, hOGG1, XRCC1*), NER (*XPC, XPD,XPG*), and DSB repair (*NBS1, XRCC2, XRCC3*). In general, gene–gene interactions always included pairs from two different pathways. Although individual variants in genes encoding DNA repair proteins modulated CAs only modestly, several gene–gene interactions were associated with CA frequencies suggesting that accumulation of CAs requires complex interplay between different DNA repair pathways.

1.5. CAs and variation in TP53 gene (DNA damage response pathway)

The tumour suppressor TP53 (MIM# 191170), located on chromosome 17p13.1, is known as 'the guardian of the genome' or 'the cellular gatekeeper of growth and division' owing to its diverse functions. The protein p53, a 393 amino acid long phosphoprotein, acts as a key regulator of cellular growth control and plays a central role in the induction of genes that are important in cell cycle arrest and apoptosis following DNA damage. Cell cycle arrest provides a time window for the cell to repair genomic damage before entering the critical stages of DNA synthesis and mitosis (previously reviewed [64]). The most frequently investigated *TP53* polymorphism is rs1042522, a G to C transversion in codon 72 of exon 4, which results in an amino acidic change from arginine to proline (*TP53* Arg72Pro). This rs1042522 polymorphism is located in a proline-rich region of the protein with importance for the growth suppression and apoptotic functions. Rs1042522 results in a structural change in the protein, giving rise to variants of distinct electrophoretic mobility; and the two isoforms of p53 due to polymorphism at codon 72 differ in biochemical and biological properties. Apparently, the *TP53* Arg72 form induces apoptosis more efficiently than the Pro72 form. The Arg72 variant, when in *cis*-form with certain tumour-derived mutations, might enhance tumour suppressive function owing to increased ability to activate p53 [64,65]. However, we did not observe any association between either chromosomal damage and rs1042522 *TP53* Arg72Pro (Tables 1 and 2) in 593 healthy individuals. Rs1042522 did not modulate the frequencies of CAs, CTAs, or CSAs, in all instances the dominant model of inheritance was tested.

1.6. CAs and variation in the major mitotic checkpoint genes

The fidelity of chromosome segregation during mitosis is ensured by the spindle assembly checkpoint (SAC). In case of unattached kinetochores to microtubules, the inhibition of anaphase promoting complex (APC or cyclosome) together with its adaptor binding partner Cdc20 prevents an occurrence of aneuploidy. The mitotic checkpoint complex, composed of Mad2, Bub1, Bub3, NEK2, ZWILCH and ZWINT proteins [66,67], is generated as a result of SAC activation. Mad2 forms a complex with the Cdc20 protein and therefore enables binding of

Bub1 and Bub3 to Cdc20 [68]. The human checkpoint kinase Bub1, and its adaptor protein Bub3 phosphorylate and subsequently inhibit the APC/Cdc20 complex during SAC signalling cascade [67,69]. When all chromosomes are properly attached in the mitotic spindle, MAD2 is released from the mitotic checkpoint complex, resulting in the activation of APC/C-CDC20, which in turn targets PTTG1 (securin) for degradation. ESPL1 (separase) gets activated and can cleave the cohesin complexes that keep the sister chromatids together [70]. Aberrant expression of genes involved in the mitotic checkpoint complex (NEK2) is related to carcinogenesis of various tumors [71].

In our study, we focused on genes constituting the mitotic checkpoint machinery in view of its role in maintaining chromosomal integrity with possible links to CAs [72]. We selected 9 variants in main checkpoint genes (*BUB1B, BUB3, MAD2L1, CENPF, ESPL1/separase, NEK2, PTTG1/securin, ZWILCH and ZWINT*) for a genotyping study in the above cohort of healthy individuals. The selected variants either caused a potentially deleterious amino acid change or were located at the core regulatory regions of the genes (Table 2). Genetic variation in individual genes exhibited a minor importance, with only a few nominally significant associations. Such data are consistent with the high conservation and selection pressure of the checkpoint system. However, many gene pairs were significantly associated with CAs and two pairs – *PTTG1-ZWILCH* and *PTTG1-ZWINT* – were associated with all types of CAs. *MAD2L1* and *PTTG1* were the most common partners in any of the two-way interactions (Fig. 4). These results suggest that interactions at the cohesin level (*PTTG1*) and kinetochore function (*ZWINT, ZWILCH* and *MAD2L1*) contribute to the frequency of CAs. The data imply that gene variants at different checkpoint functions appear to be required for the formation of CA.

1.7. Cyclin D1 and CAs

Random CAs may occur in normal cells, but they are a characteristic finding in advanced malignancies. They may be tumour-specific and contribute directly to malignant transformation through a rearrangement of specific underlying genes [4]. Many types of specific recurrent CAs, including deletions, amplifications, and translocations have been described for a large number of haematological malignancies, including lymphoid and myeloid leukaemias, multiple myeloma, and malignant lymphomas. The majority of haematological neoplasms are caused by clonal expansion of a single cell that has acquired a somatic mutation in one allele of a proto-oncogene. This mutated gene stimulates inappropriate cellular proliferation, leading to the development of cancer. Another mechanism of inappropriate stimulation of cellular proliferation is the misplacement of a proto-oncogene through CA under a strong promoter, and which may lead to the uncontrolled expression of this gene [73]. An example is translocation between chromosomes 11 and 14, abbreviated t(11;14). In t(11;14) the *cyclin D1 (CCND1)* gene from chromosome 11 is juxtaposed next to the promoter of the heavy chain immunoglobulin (*IgH*) gene, driving overexpression of cyclin D1 and stimulating cell division. It is found in several haematological malignancies, including mantle cell lymphomas in which every patient has this translocation. It is the most common translocation in multiple myeloma, found in 15% of all patients [74].

Our study on translocations in myeloma patients for the first time described a germline genetic basis of a translocation in humans: the t(11;14) was regulated by a splice site polymorphism rs9344 (archive dbSNP ID rs603965) in cyclin D1 [74]. The odds ratio (OR) between the risk allele G and the other allele A was about 2.0. The c.870G allele creates an optimal splice donor site at the intron 4/exon 5 boundary, resulting in the cyclin D1a transcript (Fig. 5). By contrast c.870A hinders splicing, allowing for read-through into intron 4 and the production of the variant cyclin D1b transcript. Although c.870A is preferentially associated with D1b production, the A allele is not fully penetrant [75,76]. Cyclin D1b has lost its capacity to associate with CDK4 and it does not enhance phosphorylation of Rb protein [77]. While cyclin D1b is a potent oncogene operating through aberrant kinase activity, cyclin D1a recruits RAD51 to local chromatin in response to DNA damage [78]. In our study, we showed that the levels of D1a and D1b transcripts correlated with the GG, GA, and AA genotypes in CD138-selected plasma cell [74]. D1a/D1b transcript ratios have been shown to shift in some cancers, suggesting that differential alternative splicing can influence tumourigenesis. It is intriguing that the G allele encoding the full length D1a would be associated with myeloma, because it has a low transforming capacity compared to D1b [79]. As the G allele is the risk factor for t(11;14) in myeloma, the D1a protein may promote the translocation, which in turn would drive cyclin D1 expression by the *IgH* promoter at the translocation locus. D1a may also

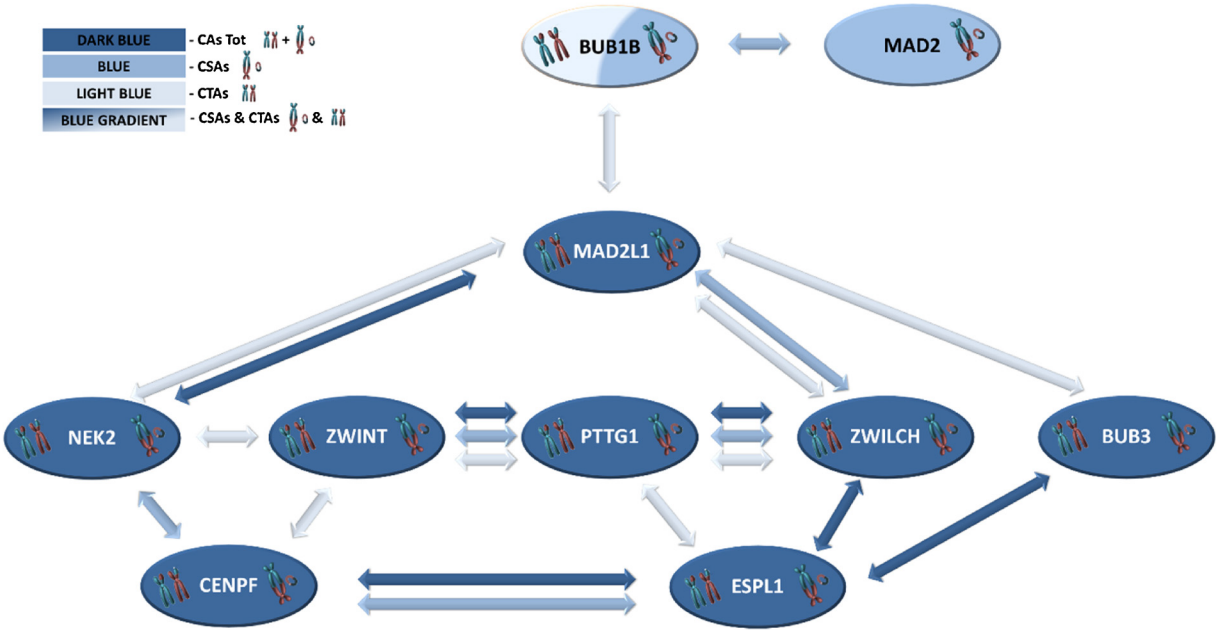


Fig. 4. Influence of mitotic checkpoint gene polymorphisms on CAs in healthy volunteers [72]. The colour code indicates total CA (dark blue), CSA (blue), CTA (light blue), and CSA & CTA (blue gradient). The arrows indicate significant gene–gene interactions.

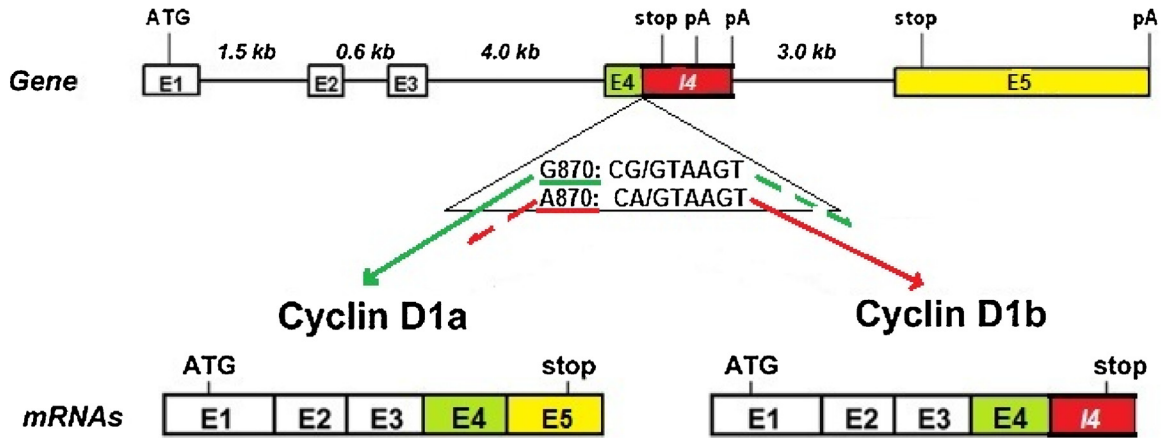


Fig. 5. Cyclin D1 participation in DSB repair by binding to RAD51 (a main recombinase involved in HR). The induction of the DNA damage response is mediated by cyclin D1a, associated with polymorphism rs9344 in *CCND1* (c.870G allele), whereas cyclin D1b (c.870A) lacks this activity.

act through separate pathways, promoting chromosomal aberrations because it is associated with double-strand break repair complexes of RAD51, BRCA1 and BRCA2 [78].

In myeloma the immunoglobulin gene translocations are generated by abnormal class switch recombination events occurring during the germinal centre stage of B cell development. Such abnormal events are usually present in all clonal myeloma cells, and are also detectable in monoclonal gammopathy of unknown significance (MGUS), a pre-malignant precursor of myeloma, and AL amyloidosis [80,81]. Normal class switch recombination involves a recognition step and double strand DNA cleavage steps, driven by activation induced deaminase. These strand breaks are then repaired using translocation prone pathways not involving classical non-homologous end joining (NHEJ) that can result in the formation of reciprocal translocation to other strand breaks located elsewhere in the genome [18]. *CCND1* polymorphism is also associated with t(11;14) in MGUS and AL amyloidosis, the G allele being the risk allele as in myeloma [74,82,83]. We also tested genetic associations with myeloma translocations other than t(11;14), but no positive results were found. This suggests that the impact of the genotype on t(11;14) is not entirely through the facilitation of errors in class switch recombination, nor can it be explained through errors in a general DNA repair mechanisms.

As t(11;14) is found in all mantle cell lymphomas, we also tested these cells for the CCND1 polymorphism, and found no effect [74]. However, in mantle cell lymphomas the translocation is mediated via abnormal VDJ recombination at the IgH locus. The breakpoints in mantle cell lymphomas are clustered, whereas the breakpoints in t (11;14) MM are heterogeneous spanning several hundred kb in the 5′ region of *CCND1* [84]. The absence of an association is consistent with a different etiological basis for t(11;14) myeloma and mantle cell lymphoma [18,85].

We also tested the possible impact of the *CCND1* genotype on non-specific CAs in a large cohort of persons for whom CAs were measured [86]. The results for total CAs showed a significant association with the AA genotype (OR 1.85). The OR was almost as high as it was for t (11;14) in myeloma, but to our surprise it was to the opposite direction. A was the risk allele for non-specific CAs, while G was the risk allele for translocation. How can these results be explained through known biology of the splice site variants? We can speculate that there are two key mechanisms. One is that allele G, encoding cyclin D1a isoform, may facilitate translocation through its DNA damage response functions. The other is that specificity for t(11;14), rather than other translocations, may depend on the juxtaposed *CCND1* gene while in other myeloma translocations other proto-oncogenes are juxtaposed at this locus. The transforming potential of isoform D1b may cause CAs in dormant lymphocytes.

To prove the above assumption we tested the association of CAs, CTAs, and CSAs with *CCND1* c.G870A polymorphism in an independent group of 255 sporadic cancer patients (mainly breast and colorectal) and controls. Moderately increased frequencies of chromosomal damage seemed to be associated with c.870A allele (OR 1.96 95% CI 0.94–4.06, P = 0.073 for CAs; OR 1.62 95%CI 0.90–2.94, P = 0.231 for CTAs; and OR 1.58 95% CI 0.77–3.21 for CSAs, P = 0.210, respectively). Although the data showed the same tendency as described above, the modulating effect was not statistically significant mainly due to a low number of observations.

2. Conclusions

Cellular DNA is continually attacked by numerous electrophilic compounds, generated or detoxified via XMEs. There is ample evidence that DNA repair pathways are crucial mechanisms for preserving of the genomic and chromosomal integrity. One source of variability in the function of XME and/or DNA repair is due to gene variants modulating activities of the particular proteins. Our review provides evidence that variants in the above genes may modulate the frequency of CAs. However, even in the conditions of an optimal DNA repair mechanism, CA may arise in a cell as a consequence of the improper mitotic division. Variants in major mitotic checkpoint genes, *TP53* and *cyclin D1* representing the DNA damage response pathway, may also contribute to the frequency of CAs in PBLs. Our results show that variants in genes encoding XMEs and DNA repair proteins as well as genes encoding key mitotic checkpoint proteins individually exerted small effects, while in interactions with each other they were associated with CA frequencies in PBLs of healthy volunteers. Obviously, the accumulation of CAs requires a complex interplay between genes involved in the maintenance of genomic integrity. However, more functional/mechanistic studies supporting these observations are required.

By assaying for the links between gene variants in the DNA damage response pathway and CAs in PBLs, we did not find any association with the common functional *TP53* Arg72Pro polymorphism. These data require rather cautious interpretation, since other *TP53* polymorphisms and haplotypes were not investigated. Opposite associations were obtained regarding the modulating effect of *CCND1* splice site polymorphism rs9344 G870A on CAs in healthy individuals and cancer patients compared to its association with t(11,14) translocation. Further study aimed at the disclosure of the functional consequences of the *CCND1* gene and its variants in interplay with the DNA damage response pathway and DNA repair on chromosomal instability during malignant transformation is warranted.

An increased frequency of CAs has also been associated with a decreased relative telomere length in healthy subjects. Notably, this

interesting association was not observed in patients with incident breast and colorectal cancers. The mechanisms underlying CA formation in association with telomere length in various pathologies deserves further attention.

In the context of the etiological role of suboptimal DNA processes in carcinogenesis, we recently suggested that altered DSB repair measured by a sensitivity towards mutagen in PBL occurs particularly in colorectal tumours. Irrespective of cancer type, telomere shortening is associated with a decreased capacity to repair DSB. There are several emerging questions, such as the role of CAs in malignancies with respect to a particular phenotype and heterogeneity, and the formation of CAs during the process of malignant transformation. Addressing the above points is unconditionally required prior to any translations into the disease prevention.

Acknowledgement

Support from grants No. 15-14789S (Czech Science Foundation, Czech Republic), 15-27580A (Ministry of Health, Czech Republic), and grant Biomedical Center Martin, Slovakia, ITMS 26220220187 (EU Commission) are acknowledged. The study was also partially supported by the projects Research program of Charles University PROGRES Q28 (Oncology) and “Centrum of clinical and experimental liver surgery”, UNCE/MED/006, and by the National Sustainability Program I (NPU I) Nr. LO1503 provided by the Ministry of Education Youth and Sports, Czech Republic.

References

[1] P.A. Futreal, L. Coin, M. Marshall, et al., A census of human cancer genes, Nat. Rev. Cancer 4 (3) (2004) 177–183.
[2] H. Rajagopalan, C. Lengauer, Aneuploidy and cancer, Nature 432 (7015) (2004) 338–341.
[3] R.A. Burrell, S.E. McClelland, D. Endesfelder, et al., Replication stress links structural and numerical cancer chromosomal instability, Nature 494 (7438) (2013) 492–496.
[4] F. Mitelman, B. Johansson, F. Mertens, The impact of translocations and gene fusions on cancer causation, Nat. Rev. Cancer 7 (4) (2007) 233–245.
[5] L.P. Bignold, Mechanisms of clastogen-induced chromosomal aberrations: a critical review and description of a model based on failures of tethering of DNA strand ends to strand-breaking enzymes, Mutat. Res. 681 (2–3) (2009) 271–298.
[6] L. Hagmar, U. Stromberg, H. Tinnerberg, Z. Mikoczy, Epidemiological evaluation of cytogenetic biomarkers as potential surrogate end-points for cancer, IARC Sci. Publ. 157 (2004) 207–215.
[7] A.T. Natarajan, F. Palitti, DNA repair and chromosomal alterations, Mutat. Res. 657 (1) (2008) 3–7.
[8] M. Durante, J.S. Bedford, D.J. Chen, et al., From DNA damage to chromosome aberrations: joining the break, Mutat. Res. 756 (1–2) (2013) 5–13.
[9] E.H. Blackburn, Switching and signaling at the telomere, Cell 106 (6) (2001) 661–673.
[10] J. Maciejowski, T. de Lange, Telomeres in cancer: tumour suppression and genome instability, Nat. Rev. Mol. Cell Biol. 18 (3) (2017) 175–186.
[11] B. Heidenreich, R. Kumar, TERT promoter mutations in telomere biology, Mutat. Res. 771 (2017) 15–31.
[12] H. Li, H.T. Hilmarsen, M.B. Hossain, et al., Telomere length and LINE1 methylation is associated with chromosomal aberrations in peripheral blood, Genes Chromosomes Cancer 52 (1) (2013) 1–10.
[13] K. Hemminki, S. Rachakonda, L. Musak, et al., Telomere length in circulating lymphocytes: association with chromosomal aberrations, Genes Chromosomes Cancer 54 (3) (2015) 194–196.
[14] K. Broberg, J. Bjork, K. Paulsson, M. Hoglund, M. Albin, Constitutional short telomeres are strong genetic susceptibility markers for bladder cancer, Carcinogenesis 26 (7) (2005) 1263–1271.
[15] J.S. Jang, Y.Y. Choi, W.K. Lee, et al., Telomere length and the risk of lung cancer, Cancer Sci. 99 (7) (2008) 1385–1389.
[16] R.A. DePinho, The age of cancer, Nature 408 (6809) (2000) 248–254.
[17] J.W. Shay, Role of telomeres and telomerase in aging and cancer, Cancer Discov. 6 (6) (2016) 584–593.
[18] M. Gostissa, F.W. Alt, R. Chiarle, Mechanisms that promote and suppress chromosomal translocations in lymphocytes, Annu. Rev. Immunol. 29 (2011) 319–350.
[19] R.S. Maser, R.A. DePinho, Connecting chromosomes, crisis, and cancer, Science (New York, NY) 297 (5581) (2002) 565–569.
[20] A.K. Meeker, J.L. Hicks, C.A. Iacobuzio-Donahue, et al., Telomere length abnormalities occur early in the initiation of epithelial carcinogenesis, Clin. Cancer Res. 10 (10) (2004) 3317–3326.
[21] J. Maciejowski, Y. Li, N. Bosco, P.J. Campbell, T. de Lange, Chromothripsis and kataegis induced by telomere crisis, Cell 163 (7) (2015) 1641–1654.
[22] M.J. Jones, P.V. Jallepalli, Chromothripsis: chromosomes in crisis, Dev. Cell 23 (5) (2012) 908–917.
[23] J.W. Shay, Molecular pathogenesis of aging and cancer: are telomeres and telomerase the connection? J. Clin. Pathol. 50 (10) (1997) 799–800.
[24] A.F. Vallarelli, P.S. Rachakonda, J. Andre, et al., TERT promoter mutations in melanoma render TERT expression dependent on MAPK pathway activation, Oncotarget 7 (33) (2016) 53127–53136.
[25] M. Kroupa, Z. Polivkova, S. Rachakonda, et al., Bleomycin-induced chromosomal damage and shortening of telomeres in peripheral blood lymphocytes of incident cancer patients, Genes Chromosomes Cancer 57 (2) (2018) 61–69.
[26] F. Mertens, C.R. Antonescu, F. Mitelman, Gene fusions in soft tissue tumors: recurrent and overlapping pathogenetic themes, Genes Chromosomes Cancer 55 (4) (2016) 291–310.
[27] F. Mertens, B. Johansson, T. Fioretos, F. Mitelman, The emerging complexity of gene fusions in cancer, Nat. Rev. Cancer 15 (6) (2015) 371–381.
[28] P. Vodicka, Z. Polivkova, S. Sytarova, et al., Chromosomal damage in peripheral blood lymphocytes of newly diagnosed cancer patients and healthy controls, Carcinogenesis 31 (7) (2010) 1238–1241.
[29] S. Vodenkova, Z. Polivkova, L. Musak, et al., Structural chromosomal aberrations as potential risk markers in incident cancer patients, Mutagenesis 30 (4) (2015) 557–563.
[30] P. Rossner, P. Boffetta, M. Ceppi, et al., Chromosomal aberrations in lymphocytes of healthy subjects and risk of cancer, Environ. Health Perspect. 113 (5) (2005) 517–520.
[31] P. Vodicka, R. Kumar, R. Stetina, et al., Genetic polymorphisms in DNA repair genes and possible links with DNA repair rates, chromosomal aberrations and single-strand breaks in DNA, Carcinogenesis 25 (5) (2004) 757–763.
[32] A. Naccarati, P. Soucek, R. Stetina, et al., Genetic polymorphisms and possible gene–gene interactions in metabolic and DNA repair genes: effects on DNA damage, Mutat. Res. 593 (1-2) (2006) 22–31.
[33] L. Musak, P. Soucek, L. Vodickova, et al., Chromosomal aberrations in tire plant workers and interaction with polymorphisms of biotransformation and DNA repair genes, Mutat. Res. 641 (1-2) (2008) 36–42.
[34] C.F. Skjelbred, M. Svendsen, V. Haugan, et al., Influence of GSTM1, GSTT1, GSTP1, NAT1, NAT2, EPHX1, MTR and MTHFR polymorphism on chromosomal aberration frequencies in human lymphocytes, Carcinogenesis 32 (3) (2011) 399–405.
[35] A.M. Rossi, I.L. Hansteen, C.F. Skjelbred, et al., Association between frequency of chromosomal aberrations and cancer risk is not influenced by genetic polymorphisms in GSTM1 and GSTT1, Environ. Health Perspect. 117 (2) (2009) 203–208.
[36] K. Hemminki, C. Frank, A. Forsti, et al., Metabolic gene variants associated with chromosomal aberrations in healthy humans, Genes Chromosomes Cancer 54 (4) (2015) 260–266.
[37] P. Vodicka, P. Soucek, A.D. Tates, et al., Association between genetic polymorphisms and biomarkers in styrene-exposed workers, Mutat. Res. 482 (1-2) (2001) 89–103.
[38] P. Vodicka, R. Kumar, R. Stetina, et al., Markers of individual susceptibility and DNA repair rate in workers exposed to xenobiotics in a tire plant, Environ. Mol. Mutagen. 44 (4) (2004) 283–292.
[39] I. Hlavata, D. Vrana, Z. Smerhovsky, et al., Association between exposure-relevant polymorphisms in CYP1B1, EPHX1, NQO1, GSTM1, GSTP1 and GSTT1 and risk of colorectal cancer in a Czech population, Oncol. Rep. 24 (5) (2010) 1347–1353.
[40] P. Vodicka, A. Naccarati, L. Vodickova, et al., Do GST polymorphisms modulate the frequency of chromosomal aberrations in healthy subjects? Environ. Health Perspect. 117 (9) (2009) A384–A385 (author reply A385).
[41] R.A. Mateuca, I. Decordier, M. Kirsch-Volders, Cytogenetic methods in human biomonitoring: principles and uses, Methods Mol. Biol. (Clifton, NJ) 817 (2012) 305–334.
[42] M. Dusinska, M. Barancokova, A. Kazimirova, et al., Does occupational exposure to mineral fibres cause DNA or chromosome damage? Mutat. Res. 553 (1-2) (2004) 103–110.
[43] M. Dusinska, A. Collins, A. Kazimirova, et al., Genotoxic effects of asbestos in humans, Mutat. Res. 553 (1–2) (2004) 91–102.
[44] A. Kazimirova, M. Barancokova, Z. Dzupinkova, L. Wsolova, M. Dusinska, Micronuclei and chromosomal aberrations, important markers of ageing: possible association with XPC and XPD polymorphisms, Mutat. Res. 661 (1–2) (2009) 35–40.
[45] L. Musak, Z. Smerhovsky, E. Halasova, et al., Chromosomal damage among medical staff occupationally exposed to volatile anesthetics, antineoplastic drugs, and formaldehyde, Scand. J. Work Environ. Health 39 (6) (2013) 618–630.
[46] R. Hakem, DNA-damage repair; the good, the bad, and the ugly, EMBO J. 27 (4) (2008) 589–605.
[47] P. Vodicka, L. Musak, C. Frank, et al., Interactions of DNA repair gene variants modulate chromosomal aberrations in healthy subjects, Carcinogenesis 36 (11) (2015) 1299–1306.
[48] P. Vodicka, R. Stetina, V. Polakova, et al., Association of DNA repair polymorphisms with DNA repair functional outcomes in healthy human subjects, Carcinogenesis 28 (3) (2007) 657–664.
[49] V. Simonelli, F. Mazzei, M. D'Errico, E. Dogliotti, Gene susceptibility to oxidative damage: from single nucleotide polymorphisms to function, Mutat. Res. 731 (1-2) (2012) 1–13.
[50] J.E. Cleaver, Defective repair replication of DNA in xeroderma pigmentosum, Nature 218 (5142) (1968) 652–656.
[51] J.J. DiGiovanna, K.H. Kraemer, Shining a light on xeroderma pigmentosum, J. Invest. Dermatol. 132 (3 Pt. (2)) (2012) 785–796.
[52] L. Nemzow, A. Lubin, L. Zhang, F.X.P.C. Gong, Going where no DNA damage sensor has gone before, DNA Repair (Amst). 36 (2015) 19–27.
[53] S. Benhamou, A. Sarasin, ERCC2/XPD gene polymorphisms and cancer risk,

Mutagenesis 17 (6) (2002) 463–469.
[54] M. Spies, Two steps forward, one step back: determining XPD helicase mechanism by single-molecule fluorescence and high-resolution optical tweezers, DNA Repair (Amst) 20 (2014) 58–70.
[55] O.D. Scharer, XPG: its products and biological roles, Adv. Exp. Med. Biol. 637 (2008) 83–92.
[56] N. Sugitani, R.M. Sivley, K.E. Perry, J.A. Capra, W.J. Chazin, XPA: a key scaffold for human nucleotide excision repair, DNA Repair (Amst). 44 (2016) 123–135.
[57] K. Komatsu, NBS1 and multiple regulations of DNA damage response, J. Radiat. Res. 57 (Suppl. 1) (2016) i11–i17.
[58] Y. Wang, M. Li, J. Long, et al., Clinical significance of increased expression of Nijmegen breakage syndrome gene (NBS1) in human primary liver cancer, Hepatol. Int. 8 (2) (2014) 250–259.
[59] D.S. Hsu, S.Y. Chang, C.J. Liu, et al., Identification of increased NBS1 expression as a prognostic marker of squamous cell carcinoma of the oral cavity, Cancer Sci. 101 (4) (2010) 1029–1037.
[60] N. Suwaki, K. Klare, M. Tarsounas, RAD51 paralogs: roles in DNA damage signalling, recombinational repair and tumorigenesis, Semin. Cell Dev. Biol. 22 (8) (2011) 898–905.
[61] J.F. Carvalho, R. Kanaar, Targeting homologous recombination-mediated DNA repair in cancer, Expert Opin. Ther. Targets 18 (4) (2014) 427–458.
[62] A.V. Mazin, O.M. Mazina, D.V. Bugreev, M.J. Rossi, Rad54, the motor of homologous recombination, DNA Repair (Amst) 9 (3) (2010) 286–302.
[63] S.J. Ceballos, W.D. Heyer, Functions of the Snf2/Swi2 family Rad54 motor protein in homologous recombination, Biochim. Biophys. Acta 1809 (9) (2011) 509–523.
[64] A. Naccarati, V. Polakova, B. Pardini, et al., Mutations and polymorphisms in TP53 gene–an overview on the role in colorectal cancer, Mutagenesis 27 (2) (2012) 211–218.
[65] X. Tian, S. Dai, J. Sun, S. Jiang, Y. Jiang, The association between the TP53 Arg72Pro polymorphism and colorectal cancer: an updated meta-analysis based on 32 studies, Oncotarget 8 (1) (2017) 1156–1165.
[66] J.M. Kasuboski, J.R. Bader, P.S. Vaughan, et al., Zwint-1 is a novel Aurora B substrate required for the assembly of a dynein-binding platform on kinetochores, Mol. Biol. Cell 22 (18) (2011) 3318–3330.
[67] L. Jia, B. Li, H. Yu, The Bub1-Plk1 kinase complex promotes spindle checkpoint signalling through Cdc20 phosphorylation, Nat. Commun. 7 (2016) 10818.
[68] M. Simonetta, R. Manzoni, R. Mosca, et al., The influence of catalysis on mad2 activation dynamics, PLoS Biol. 7 (1) (2009) e10.
[69] K. Overlack, T. Bange, F. Weissmann, et al., BubR1 promotes bub3-dependent APC/C inhibition during spindle assembly checkpoint signaling, Curr. Biol. 27 (19) (2017) 2915–2927 (e2917).
[70] S. Hauf, I.C. Waizenegger, J.M. Peters, Cohesin cleavage by separase required for anaphase and cytokinesis in human cells, Science (New York, NY) 293 (5533) (2001) 1320–1323.
[71] L. Fu, S. Liu, H. Wang, et al., Low expression of NEK2 is associated with hepatocellular carcinoma progression and poor prognosis, Cancer Biomark. 20 (1) (2017) 101–106.
[72] A. Forsti, C. Frank, B. Smolkova, et al., Genetic variation in the major mitotic checkpoint genes associated with chromosomal aberrations in healthy humans, Cancer Lett. 380 (2) (2016) 442–446.
[73] A.M. Rajan, S.V. Rajkumar, Interpretation of cytogenetic results in multiple myeloma for clinical practice, Blood Cancer J. 5 (2015) e365.
[74] N. Weinhold, D.C. Johnson, D. Chubb, et al., The CCND1 G870A polymorphism is a risk factor for t(11;14)(q13;q32) multiple myeloma, Nat. Genet. 45 (2013) 522–525.
[75] K.E. Knudsen, J.A. Diehl, C.A. Haiman, E.S. Knudsen, Cyclin D1: polymorphism, aberrant splicing and cancer risk, Oncogene 25 (2006) 1620–1628.
[76] E.K. Millar, J.L. Dean, C.M. McNeil, et al., Cyclin D1b protein expression in breast cancer is independent of cyclin D1a and associated with poor disease outcome, Oncogene 28 (2009) 1812–1820.
[77] C.J. Kim, K. Nishi, T. Isono, et al., Cyclin D1b variant promotes cell invasiveness independent of binding to CDK4 in human bladder cancer cells, Mol. Carcinog. 48 (10) (2009) 953–964.
[78] R.G. Pestell, New roles of cyclin D1, Am. J. Pathol. 183 (1) (2013) 3–9.
[79] N.A. Olshavsky, C.E. Comstock, M.J. Schiewer, et al., Identification of ASF/SF2 as a critical, allele-specific effector of the cyclin D1b oncogene, Cancer Res. 70 (2011) 3975–3984.
[80] W.J. Chng, O. Glebov, P.L. Bergsagel, W.M. Kuehl, Genetic events in the pathogenesis of multiple myeloma, Best Pract. Res. Clin. Haematol. 20 (2007) 571–596.
[81] J. Sawyer, The prognostic significance of cytogenetics and molecular profiling in multiple myeloma, Cancer Genet. 204 (2011) 3–12.
[82] M.I. da Silva Filho, A. Försti, N. Weinhold, et al., Genome-wide association study of immunoglobulin light chain amyloidosis in three patient cohorts: comparison to myeloma, Leukemia 31 (8) (2017) 1735–1742.
[83] I. Meziane, S. Huhn, M. Filho, et al., Genome-wide association study of clinical parameters in immunoglobulin light chain amyloidosis in three patient cohorts, Haematologica 102 (10) (2017) e411–e414.
[84] R. Fonseca, B. Barlogie, R. Bataille, et al., Genetics and cytogenetics of multiple myeloma: a workshop report, Cancer Res. 64 (2004) 1546–1558.
[85] P. Perez-Galan, M. Dreyling, A. Wiestner, Mantle cell lymphoma: biology, pathogenesis, and the molecular basis of treatment in the genomic era, Blood 117 (2011) 26–38.
[86] K. Hemminki, L. Musak, V. Vymetalkova, et al., Cyclin D1 splice site variant triggers chromosomal aberrations in healthy humans, Leukemia 28 (2014) 721–722.