



Molecular karyotyping and gene expression analysis in childhood cancer patients

Galetzka Danuta¹ · Müller Tobias² · Dittrich Marcus² · Endres Miriam¹ · Kartal Nergiz¹ · Sinizyn Olesja¹ · Rapp Steffen^{3,4} · Zeller Tanja^{5,6} · Müller Christian^{5,6} · Hankeln Thomas³ · Scholz-Kreisel Peter⁷ · Chorzempa Heather^{8,9} · Mirsch Johanna¹⁰ · Poplawski Alicia⁷ · Rossmann Heidi¹¹ · Spix Claudia¹² · Haaf Thomas⁹ · Prawitt Dirk⁸ · Marron Manuela¹³ · Schmidberger Heinz¹

Received: 6 November 2019 / Revised: 20 April 2020 / Accepted: 8 June 2020 / Published online: 23 June 2020
© The Author(s) 2020

Abstract

The genetic etiology of sporadic childhood cancer cases remains unclear. We recruited a cohort of 20 patients who survived a childhood malignancy and then developed a second primary cancer (2N), and 20 carefully matched patients who survived a childhood cancer without developing a second malignancy (1N). Twenty matched cancer-free (0N) and additional 1000 (0N) GHS participants served as controls. Aiming to identify new candidate loci for cancer predisposition, we compared the genome-wide DNA copy number variations (CNV) with the RNA-expression data obtained after in vitro irradiation of primary fibroblasts. In 2N patients, we detected a total of 142 genes affected by CNV. A total of 53 genes of these were not altered in controls. Six genes (*POLR3F*, *SEC23B*, *ZNF133*, *C16orf45*, *RRN3*, and *NTANI*) that we found to be overexpressed after irradiation were also duplicated in the genome of the 2N patients. For the 1N collective, 185 genes were affected by CNV and 38 of these genes were not altered in controls. Five genes (*ZCWPW2*, *SYNCRIP*, *DHX30*, *DHRS4L2*, and *THSD1*) were located in duplicated genomic regions and exhibited altered RNA expression after irradiation. One gene (*ABCC6*) was partially duplicated in one 1N and one 2N patient. Analysis of methylation levels of *THSD1* and *GSTT2* genes which were detected in duplicated regions and are frequently aberrantly methylated in cancer showed no changes in patient's fibroblasts. In summary, we describe rare and radiation-sensitive genes affected by CNV in childhood sporadic cancer cases, which may have an impact on cancer development.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00109-020-01937-4>) contains supplementary material, which is available to authorized users.

✉ Galetzka Danuta
Danuta.Galetzka@unimedizin-mainz.de; <http://www.unimedizin-mainz.de/radioonkologie-und-strahlentherapie>

¹ Department of Radiation Oncology and Radiation Therapy, University Medical Centre, Johannes Gutenberg University Mainz, Obere Zahlbacher Str. 63, 55131 Mainz, Germany

² Bioinformatics Department, Julius Maximilians University, Würzburg, Germany

³ Institute of Organismal and Molecular Evolution, Molecular Genetics and Genome Analysis, Johannes Gutenberg University, Mainz, Germany

⁴ Preventive Cardiology and Preventive Medicine, Center for Cardiology, University Medical Centre, Mainz, Germany

⁵ University Heart & Vascular Center, Clinic for Cardiology, University Medical Centre Hamburg-Eppendorf, Hamburg, Germany

⁶ Dzhk (Deutsches Zentrum für Herz-Kreislauf-Forschung), Standort Hamburg, Lübeck, Kiel, Hamburg, Germany

⁷ Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Centre, Mainz, Germany

⁸ Center for Pediatrics and Adolescent Medicine, University Medical Centre, Mainz, Germany

⁹ Institute of Human Genetics, Julius Maximilians University, Würzburg, Germany

¹⁰ Radiation Biology and DNA Repair, University of Technology, Darmstadt, Germany

¹¹ Institute of Clinical Chemistry and Laboratory Medicine, University Medical Centre, Mainz, Germany

¹² German Childhood Cancer Registry, Institute of Medical Biostatistics, Epidemiology and Informatics, University Medical Centre, Mainz, Germany

¹³ Leibniz Institute for Prevention Research and Epidemiology – BIPS, Bremen, Germany

Key messages

- Rare CNV's may have an impact on cancer development in sporadic, non-familial, non-syndromic childhood cancer cases.
- In our cohort, each patient displayed a unique pattern of cancer-related gene CNVs, and only few cases shared similar CNV.
- Genes that are transcriptionally regulated after radiation can be located in CNVs in cancer patients and controls.
- *THSD1* and *GSTT2* methylation is not altered by CNV.

Keywords Childhood cancer · Copy number variation · Radiation · Gene expression · Primary secondary cancer

Introduction

Determining genetic risk factors for cancer is a major goal of medical research. The increase of knowledge about genetic risk factors aims to improve cancer diagnostics and, due to therapeutic advances, contribute to increased overall survival for pediatric cancer. As pediatric cancer survivors reach adulthood, the development of secondary malignancies becomes a significant issue for these patients. Treatment of the primary neoplasm with chemotherapy (systemic therapy) and/or radiotherapy has been described as a risk factor for second neoplasms after childhood cancer [1]. As only a small percentage of the treated children suffer from a second neoplasm, other factors are likely to be involved [2]. A predisposition for the occurrence of a second neoplasm in childhood might be a pre-existing somatic genetic variation responsible for, or associated with, DNA-repair, cell cycle control, and other genes crucial for tumor development [3]. Genetic variation among other modifications may manifest as SNP (single nucleotide polymorphism, mutation) or/and chromosomal copy number variation (CNV). Additive, epigenetic modifications, like aberrant methylation may also lead to tumor development.

CNVs may harbor genes and/or regulatory regions that could contribute to complex diseases such as cancer, whose development is triggered and orchestrated by the interaction of many genes. Two consecutive chromosomal aberrations can be described in neoplasias: Primary somatic variations as initiating events and secondary aberrations that are acquired during transformation toward cancer [4]. Typically, chromosomal abnormalities that accumulate during tumor evolution lead to an unbalanced genome. On the other hand, balanced chromosomal alterations are often associated with cytogenetically cryptic deletions or duplications in the breakpoint regions [5]. These alterations can have a direct effect on transcript levels and thus gene expression [6]. So far there have been few studies on primary fibroblasts of cancer patients to study underlying predisposing genomic variations and associated gene expression changes. Fibroblasts of breast and thyroid cancer patients were almost always found to have defective DNA repair and/or cell cycle regulation [7]. Abnormal gene expression in the somatic cells of unaffected parents of retinoblastoma patients is also consistent with an inherited predisposition to cancer development [8].

Radiation and chemotherapeutic agents do not mechanistically distinguish a tumor cell from healthy tissue and the application of these genotoxic agents may be another source of acquired CNVs. Frequent induction of chromosomal aberrations after irradiation has been reported by Massenkeil et al. [9] in skin fibroblasts in vivo.

To protect the healthy tissue from damage, it is especially important to understand the molecular mechanisms involved in the cellular response to radiation. Several attempts have been undertaken to analyze the transcriptional effects after irradiation of different tissues or cells, with different cell culture conditions, doses, and time points. Furthermore, it remains unclear whether DNA duplications or deletions are associated with the formation of epigenetic alterations, such as DNA methylation, which could play a role in cancer development [10].

To identify genomic susceptibility factors for primary and secondary cancer formation in childhood, we compared molecular cytogenetic profiles by SNP array analysis in primary fibroblasts of childhood cancer survivors (1N) and carefully matched patients with second cancer (2N), alongside with cancer-free controls (0N). We determined the gene expression profile in primary fibroblasts in vitro after X-ray treatment and correlated it with the genes located within the deletions and duplications detected by SNP array analyses. Specifically, we tested the hypothesis that the occurrence of secondary cancer is associated with modifications in the expression of cell cycle control and DNA repair pathways. Finally, we analyzed the methylation patterns in the putative promoter regions of two candidate cancer-relevant genes, which resided within CNV regions and displayed differential expression after irradiation.

Material and methods

Patient collective

This study was approved by the Ethics Committee of the Medical Association of Rhineland-Palatinate (no. 837.440.03 (4102) and no. 837.262.12(8363-F)). With the help of the German Childhood Cancer Registry, 20 individuals who survived a childhood malignancy and then developed a second primary cancer (2N) and 20 carefully matched (first tumor, manifestation age, sex) individuals who survived a childhood cancer without developing a second malignancy

(1N) were recruited for the KiKme study (Cancer in Childhood and Molecular Epidemiology). Twenty matched patients (sex and age) without cancer from the Department of Accident Surgery and Orthopedics in Mainz Germany served as controls (0N). Written informed consent to use fibroblasts for research purposes was obtained after genetic counseling for all participating patients. No patient had intellectual disability or any other severe mental disease based on the clinical impression and personal history.

The numbering of the patients does not represent the recruiting number and was chosen randomly. Skin biopsies were taken at the earliest 2 years after the last cancer therapy. Eleven patients suffered from acute lymphatic or myeloid leukemia, 5 patients from Hodgkin or Burkitt lymphoma, and 4 patients from other solid tumors as primary malignancy. The second cancers in the 2N group included myelodysplastic syndrome, lymphoma, thyroid cancer, and other solid tumors. All patients were followed up from primary cancer diagnosis to the time when they were recruited. With the exception of one patient (1N), all patients received chemotherapy, radiotherapy, or combination therapy. Six patients received allogeneic bone marrow transplantation. Clinical data of the participating patients are shown in Table 1.

Although it was reported that 7–8% of children affected by cancer carry an unambiguous predisposing germline variant, predominantly within *TP53* and *BRCA2* [11], no proven pathogenic germline variants in *TP53*, *BRCA1*, or *BRCA2* were identified using Sanger sequencing applying the ACMG criteria, in our cohort (Mutation Databases: (<http://p53.iarc.fr/>; <https://www.ncbi.nlm.nih.gov/clinvar/> and <https://www.lovd.nl/>)). In one case, an *RB1* oncogenic splice mutation was detected. This patient was excluded from further analysis. The remaining patients did not fit the criteria of an inherited childhood cancer syndrome [12].

Currently, there are several attempts to characterize the common CNVs which have no impact on disease. The 1000 Genomes project (<http://www.1000genomes.org>), the genome of the Netherlands project (<http://www.nlgenome.nl>), and the Toronto Database of Genomic Variants (<http://dgv.tcag.ca/>) are examples of these attempts. Since the boundaries of the variants are not well defined, there could be an over-estimation in the actual size of the variants. In addition, some ethnic aspects may contribute to the prevalence of a specific CNV. Therefore, we compared the cancer patients' CNV with the data of matched 0N controls and 1000 0N cases without cancer, diabetes, obesity, dyslipidemia, and stroke from the Gutenberg Heart Study (GHS), which had the advantage that the samples were analyzed in the same laboratory with the same technique and the participants' samples came from similar ethnic backgrounds. The aim was to detect genes which were affected in the cancer patients but not in the controls, which may therefore be considered as rare, putative, and predisposing variations. The GHS is a community-based prospective, observational single-center cohort study in the Rhein-Main-Region in western Mid-Germany. The GHS has been approved by the local ethics committee and by the local and federal data safety commissioners. The primary aim of the GHS study is to evaluate and improve cardiovascular risk stratification.

Cell culture and experimental procedure

Primary fibroblasts from skin biopsies were cultured in DMEM (Invitrogen, Karlsruhe, Germany) and were supplemented with 15% fetal bovine serum (FBS) (Biochrom, Berlin, Germany), 1% vitamins, and 1% antibiotics (Pen/Strep) (Biochrom, Berlin, Germany) at 37 °C and 5% CO₂. All experiments using the primary fibroblasts were performed with growth-arrested cells in the G0/G1 stage in 10-cm cell culture dishes. Confluency of the cells was achieved by

Table 1 Clinical details of the 2N and 1N cancer patient groups

Tumors	2N		1N	
	First	Second	First	Second
Age at diagnosis of cancer (year)	7.2 ± 4.3		7.5 ± 4.4	
Tumor entity				
Acute lymphatic/myeloid leukemia	11	1	9	-
Hodgkin or Burkitt lymphoma	5	5	5	-
Solid tumors	4	14	6	-
Age at fibroblast collection (year)	27.1 ± 3.7		26.3 ± 4.4	
Treatment				
CT-RT-	1		4	
CT+RT-	3		2	
CT-RT+	3		4	
CT+RT+	13		10	

CT, chemotherapy; RT, radiation therapy including radioiodine therapy for papillary thyroid cancer

contact inhibition and subsequent cultivation for 2 weeks. Over 90% of the cells were in the G0/G1 stage of the cell cycle which was confirmed by FACS (flow cytometric cell cycle analysis). For comparisons of 0N, 1N, and 2N patients, fibroblasts with similar passages 9 (± 2) were used. Cells were exposed to X-rays with a D3150 X-Ray Therapy System (Gulmay Ltd., Surrey, UK) at 140 kV and a dose rate of 3.62 Gray (Gy)/min at room temperature. Sham-irradiated cells were kept at the same conditions in the radiation device control room. Cells were exposed to single doses ranging from 2 to 8 Gy and were returned to the incubator. Cells were harvested by a brief treatment with trypsin/EDTA (Biochrom, Berlin, Germany) and washed with PBS ($-\text{Mg}/-\text{Cl}$) at 15 min, 2 h, and 24 h after irradiation. Resulting pellets were stored at $-80\text{ }^{\circ}\text{C}$ until DNA or RNA preparation. Cell lines MCF7 (ATCC, Manassas, VA, USA), ZR-75-1, EFO27, and T47D (ATCC, Manassas, VA, USA) were cultivated in RPMI1640 (Gibco) supplemented with 10% FBS, 2.5% HEPES buffer (Sigma), and 1% antibiotics (Pen/Strep) (Life Technologies). The A549 (ATCC, Manassas, VA, USA) cell line was cultured in DMEM modified with 10% FBS.

SNP array (molecular karyotype analysis)

Molecular karyotyping (SNP array) was performed using DNA (isolated with the NucleoSpin tissue kit Macherey-Nagel, Germany) from untreated primary fibroblasts in passage 5 (2N, 1N, 0N). High-resolution screening for microdeletions and duplications was performed with the Affymetrix GeneChip Genome-Wide Human SNP array 6.0 and the GeneChip Genome-Wide SNP Sty Assay Kit 5.0/6.0, following the protocol developed by the manufacturer (Affymetrix, Santa Clara, CA, USA). Data calculation was performed with Affymetrix Genotyping Console 4.2.0.26 and Chromosome Analysis Suite 3.1.0.15. The segment filters for gains and losses were set at a minimum of 5 markers and 20 kb. All samples passed the QC control filters (MAPD < 0.25 , SNPQC > 15.00 , Waviness SD < 0.12).

RNA-sequencing, data analysis, and statistics

Total RNAs were prepared from treated and untreated fibroblast cultures using the Nucleo Spin RNA Plus Kit from Macherey-Nagel. The RNA integrity was assessed with a Bioanalyzer 2100 (Agilent RNA 6000 Nano Kit, Agilent Technologies, Santa Clara, USA). One microgram of total RNA (QuBit, Thermo Fisher Scientific) RIN ≥ 8 was used for library construction using the TruSeq RNA Sample Prep Kit v2 (Set A and B, Illumina) following the manufacturer's instruction. RNA-Seq libraries were pooled, cBot clustered, and sequenced on a HiSeq2500 instrument (Illumina) in high-output mode. Reads with a length of 50 nucleotides were generated using TruSeq SR (single read) Cluster Kit v3 (Illumina) and

TruSeq SBS Kit v3 (Illumina). Data was generated by RTA Version 1.8.4 (real-time analysis) and converted to FASTQ format using bcl2fastq Version 1.8.4. (Illumina).

Raw reads were cleaned from adapter sequences using Trimmomatic. Cleaned reads were aligned to the human reference genome (GRCh38) using STAR. Expression per gene expressed as the number of aligned reads per gene was quantified using FeatureCounts. Data analysis was performed using R with 51 samples varying in the dose of applied radiation (0 Gy, 2 Gy, 5 Gy, 8 Gy) and time post-irradiation (15 min, 2 h, 24 h) for analysis. Genes with less than 10 counts in 4 samples were discarded. Data were normalized for sequencing depth using the EdgeR package. Transformation to log2 counts per million was performed via the Voom method, implemented in the limma-package. Differential gene expression dependent on dose and time points was detected using linear models implemented in the limma-package. Genes with an adjusted p value smaller than 0.05 were flagged as significant for further analyses. p values were adjusted for false discovery rate (FDR) (Benjamini-Hochberg procedure).

Quantitative real-time PCR for gene expression and copy number variation

Total RNAs were prepared from treated and untreated fibroblast cultures using the Nucleo Spin RNA Plus Kit from Macherey-Nagel. Two micrograms of the RNA samples was reversely transcribed into cDNA using the SuperScript IV First-Strand random hexamer Synthesis System (Invitrogen). Genomic DNA was isolated with the NucleoSpin tissue kit (Macherey-Nagel, Germany). Forward and reverse primers (Exon-spanning for gene expression) were designed with the Primer-Blast program (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). *RRN18S* and *TBP* for gene expression and *HEM3* and *RFC3* for copy number calculations were used as endogenous control genes (Online Resource Table: Primer sequences). Each 10- μl reaction volume contained 25 ng cDNA or DNA template in 5 μl Sybr-Green Master Mix (Roche), 2 μl RNase-free PCR graded water (Roche), and 1 μl each of forward and reverse primer (10 μM). All reactions were performed in triplicate and in two stages, with one cycle of 95 $^{\circ}\text{C}$ for 10 min (first stage) and 45 cycles of 94 $^{\circ}\text{C}$ for 10 s, (TM-primer) $^{\circ}\text{C}$ for 10 s, and 72 $^{\circ}\text{C}$ for 10 s (second stage) using the LightCycler 480II Roche. Amplification qualities were assayed using melting curves and agarose gel analysis. The qPCR amplification efficiency was calculated using the LingReg program and the CT values were corrected using the mean amplification efficiency. Relative quantification was carried out with the $\Delta\Delta\text{CT}$ method using the two endogenous control genes and the control 0 Gy or 0N probands for calibration. Statistical analyses were conducted using the unpaired t test. Expression changes with p value < 0.05 were considered significant.

Bisulfite pyrosequencing

Genomic DNA was isolated with the NucleoSpin tissue kit (Macherey-Nagel, Germany). Bisulfite conversion of 0.2 µg DNA was performed with the EpiTect Bisulfite Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. PCR and sequencing primers for the analyzed genes were designed with PyroMark Assay Design 2.0 software (Qiagen) (online Resource Table: Primer sequences). The 25 µl PCR reactions consisted of 2.5 µl 10x PCR buffer, 20 mM MgCl₂, 0.5 µl dNTP mix (10 mM), 1 µl of each forward and reverse primer (10 µM), 0.2 µl FastStart Taq DNA Polymerase (5 U/µl) (Roche Diagnostics, Mannheim, Germany), 18.8 µl PCR-grade water, and 1 µl (~100 ng) bisulfite-converted template DNA. PCR amplifications were performed with an initial denaturation step at 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 55 °C for 30 s and 72 °C for 45 s and a final extension step at 72 °C for 5 min. Bisulfite pyrosequencing was performed on a PyroMark Q96 MD Pyrosequencing System using the PyroMark Gold Q96 CDT Reagent Kit (Qiagen) and 0.5 µl of sequencing primers (10 mM). Data analysis was performed with the Pyro Q-CpG software (Qiagen).

FISH analysis

Metaphase chromosome spreads of the patients were prepared from primary mitotic fibroblasts. BAC clones (RP11-139D07 for the 2N4 and RP11-327M19 for the 2N7 patient) were selected from the Wellcome Trust Sanger Institute Ensembl contigs and obtained from the Resource Center Primary Database of the German Human Genome Project and ResGen (Invitrogen). Genomic BAC DNAs were labeled with Tetramethyl-rhodamine-5-dUTP* (Roche) or 25 nmol Fluorescein-12-dUTP* (Roche) by standard nick-translation and FISH-mapped on metaphase chromosomes. Control BAC clones were chosen for 16q terminal or 2q terminal chromosome areas. Images were generated using the Leica microscope CTR MIC and Software CW4000.

Results

Molecular karyotype analysis (SNP array) of 2N, 1N, and 0N controls

The concept of the study was to detect genes that were affected in the cancer patients (2N, 1N) but not in the controls (0N and 1000 GHS), which may, therefore, be considered as a rare, putative, and predisposing variation. We detected rare germline CNVs in eighteen 2N and sixteen 1N patients. In some cases, the detected aberrations in the SNP array analysis overlapped

between patients and controls. For the final compendium of putative pathogenic aberrations, we selected only genes that were not affected in the control section, but the annotation of the aberration reflects the complete duplicated or deleted CNV area. Altogether we detected 142 affected genes in 2N patients of which 53 genes were not altered in controls (matched 0N and 1000 GHS). For the 1N collective, there were 185 genes affected by CNVs of which 38 genes were uniquely altered in the 1N cancer patients. Interestingly, 22 genes in 2N patients within CNV and eighteen genes in 1N have previously been described to be associated with tumor development, growth, apoptosis, and chromosomal stability, or as differentially expressed in cancer. (TCGA data base: <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>). Only one gene (*ABCC6*) was partially duplicated in two matched patients (1N5 and 2N4) and duplicated in five out of 1000 controls (2N; arr[hg19] 16p13.11(15,048,755-16,295,900)x3 and 1N; arr[hg19] 16p13.11(16,294,705-16,798,651)x3). Both patients suffered from leukemia and patient 2N4 later developed a slow-growing brain tumor. Each 2N patient displayed a unique CNV pattern which was not seen in other patients of the 2N group. Altogether we detected sixteen heterozygous and one homozygous duplication, as well as eleven heterozygous and one homozygous deletion in the 2N group. The homozygous deletion affected the *TPTE2P3* gene, which is classified as a pseudogene with expression restricted to the testis [13]. The findings for the 1N patient group were similar to the observed ones in 2N patients. Here we also detected unique CNV patterns, with the exception of three regions: 19q13.42(54,716,827-54,741,307)x3, 22q11.21(21,567,218-21,845,282)x3, and 14q11.2(24,431,136-24,499,742)x1 which were duplicated in more than one 1N case. The duplication in chromosome 19q13.42 contains the *LILRB3* gene and occurred in two leukemia cancer patients, whereas the duplication in 22q11.21 encompasses five genes (*HIC2*, *PI4KAP2* (pseudogene), *POM121L8P* (pseudogene), *RIMBP3B*, and *RIMBP3C*). The patients carrying this aberration suffered from leukemia and solid tumors. The aberration 14q11.2(24,431,136-24,499,742)x1 contains the *DHRS4L2* gene, which is downregulated after radiation and occurred in our patients with leukemia and solid tumor. In total, we detected thirteen heterozygous, one homozygous duplicated, and twelve deleted heterozygous regions in the 1N patient cohort.

The CNVs did not always affect a whole gene. We detected intronic deletions in *IGSF21*, *NCK1*, and *MCU*, and intronic duplications in the *RBFOX3*, *COL11A*, *SORCS1*, *FMNL2*, and *NLGN1* genes. These sites harbor transcription

factor binding regions. In total, seven pseudogenes and eight microRNAs were affected. Six long coding (LINCO) and two anti-sense RNAs (AS) were affected by CNVs in cancer patients, but not in controls (see Tables 2 and 3 and for more detailed information Online Resource 1, Table: 1N CNV; 2N CNV).

qPCR and FISH analysis to confirm CNV

To confirm the results of the SNP array analyses, we chose exemplary regions for verification by qPCR and FISH analysis. In all explored cases, the qPCR result confirmed the duplications detected by the SNP array and the hybridization using specific probes (FISH) indicates tandem duplication in both cases (2N4; 2N7) analyzed. In Fig. 1, the duplication in 2N4 is shown. Since the qPCR technique is suitable for screening for aberrations even in a mosaic state, further verification

was conducted using qPCR. Case 2N12 displayed a deletion in chromosome 2q32.1. qPCR analysis with specific primers for this region suggests a heterozygous deletion, whereas the duplication in 19q13.41 in case 2N9 may be a mosaic (Fig. 2).

Analysis of common CNVs and transcription factor binding sites in controls (0N and 1000 CGH) and patient's gene free areas

Our analysis of frequently altered regions in the human genome (Online Resource 1, Table: Suppl. common CNV) revealed the presence of transcription factor binding sites either within gene-loci (e.g., the *DUSP22* gene) or within previously described immune response regulating areas (IGK, immunoglobulin kappa locus; IGH immunoglobulin heavy locus, etc.). Our hypothesis is that there may be several areas altered in the patients which do not contain genes, but enhancer-

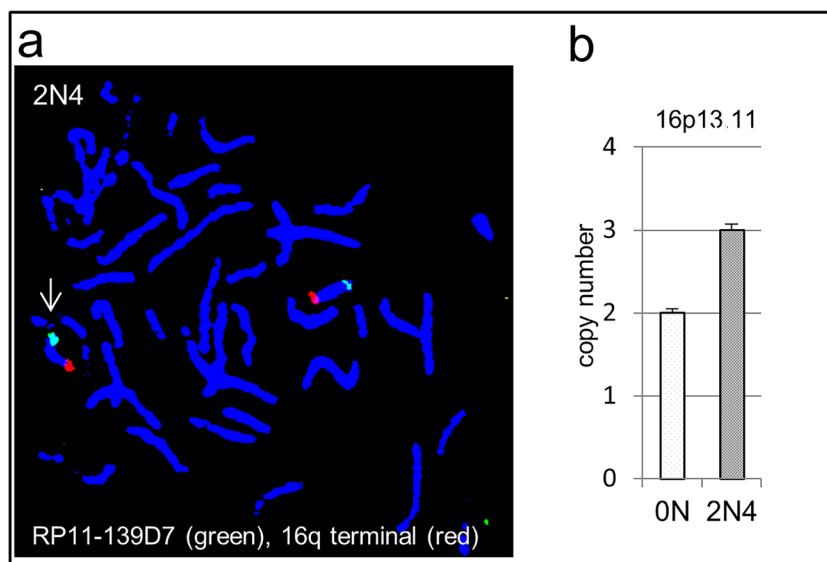
Table 2 Rare germline CNV's detected in 16 of 20 patients suffering from one cancer (1N). Genes indicated in bold are cancer related (Pubmed 2019; see further information in Online Resource 1)

Patients 1N	Type	Chromosome position [hg19]	Number of genes/probes	Genes	1000 GHS controls
1	Del	4q22.3(98,338,184-98,360,773)x1	1/16	<i>STPG2-AS1</i>	0
2	Del	4q31.21(141,590,313-141,612,737)x1	1/25	<i>TBC1D9</i>	0
3	Dupl	3p24.1(28,355,385-28,522,017)x3	3/65	<i>AZI2, CMC1, ZCWPW2</i>	0
4	Dupl	16p13.11(16,294,705-16,798,651)x3	5/30	<i>PKD1P1 (pseudogene), MIR3179-2, MIR3670-2, MIR3180-2, MIR6511A2</i>	0
5	Del	8q11.21q11.23(50,048,130-53,211,910)x1	3/2006	<i>PCMTD1, PXDNL, ST18</i>	0
6	Dupl	9p24.1(5,221,817-5,408,358)x3	2/145	<i>PLGRKT, INSL4</i>	0
7	Del	17p13.3(513-38,924)x1	1/19	<i>DOC2B</i>	0
	Del	4q22.1(88,176,912-88,295,008)x1	2/61	<i>HSD17B11, HSD17B13</i>	0
8	Dupl	10q25.1(108,746,879-108,777,331)x4	1/25	<i>SORCSI (intron)</i>	0
	Del	11q14.3(89,438,613-89,461,619)x1	1/19	<i>TRIM77</i>	2
	Del	16q23.1(78,057,102-78,117,280)x1	1/48	<i>CLEC3A</i>	2
9	Dupl	6q14.3(86,136,432-86,346,699)x3	3/87	<i>NT5E, SNX14, SYNCRIP</i>	0
	Del	10q22.1(74,515,548-74,560,343)x1	1/23	<i>MCU (intron)</i>	0
	Del	19p12(22,159,002-22,179,901)x1	1/10	<i>ZNF208</i>	0
	Del	6p24.3(10,468,665-10,535,651)x1	1/78	<i>GCNT2</i>	5
10	Dupl	3q26.31(173,247,466-173,289,668)x3	1/23	<i>NLGNI (intron)</i>	2
11	Dupl	3p21.31(47,799,742-47,849,936)x3	2/50	<i>DHX30, SMARCC1</i>	0
	Dupl	21q22.12(37,481,942-37,617,188)x3	3/96	<i>CBR3-AS1, CBR3, DOPEY2</i>	0
12	Del	2p21(45,312,361-45,504,006)x1	1/170	<i>LINC01121</i>	0
13	Dupl	13q14.3(52,923,759-53,053,990)x3	3/47	<i>CKAP2, THSD1, VPS36</i>	0
	Dupl	2q23.3(153,212,500-153,273,886)x3	1/54	<i>FMNL2 (intron)</i>	0
14	Dupl	6q16.2q16.3(100,277,839-100,643,681)x3	1/204	<i>MCHR2</i>	0
15	Del	19q13.12(35,774,555-35,795,979)x1	2/14	<i>HAMP, MAG</i>	0
5, 3	Dupl	19q13.42(54,716,827-54,741,307)x3	1/26	<i>LILRB3</i>	2
16, 14	Dupl	22q11.21(21,567,218-21,845,282)x3	5/35	<i>HIC2, PI4KAP2 (pseudogene), POMI21L8P (pseudogene), RIMBP3B, RIMBP3C</i>	4
4, 5	Del	14q11.2(24,431,136-24,499,742)x1	1/49	<i>DHRS4L2</i>	0

Table 3 Rare germline CNV's detected in 16 of 20 patients with two independent cancers (2N). Genes indicated in bold are cancer related (Pubmed 2019; see further information in Online Resource 1)

Patients 2N	Type	Chromosome position [hg19]	Number of genes/probes	Genes	1000 GHS controls
1	Del	7q11.21(63,526,244-63,784,883)x1	5/159	ZNF727, ZNF735, ZNF679, ZNF736, ZNF83	0
2	Dupl	Xp11.23(47,859,614-47,960,125)x3	2/69	ZNF182, ZNF630	6
3	Dupl	1p36.13(17,594,883-17,622,897)x3	1/26	PADI3	3
3	Dupl	16q23.1(75,539,395-75,575,997)x4	2/20	CHST5, TMEM231	7
4	Dupl	16p13.11(15,048,755-16,295,900)x3	10/682	MPV17L, ABCCI , FOPNL , MYH11 , NDEI , NTANI , <i>Corif45</i> , RRN3 , KIAA0430 , MIR484	0
	Del	1p36.13(18,578,074-18,601,898)x1	1/26	IGSF21 (intron)	0
	Del	3p21.31(46,763,773-46,859,820)x1	1/56	PRSS45	1
	Del	3q22.3(136,596,428-136,633,070)x1	1/17	NCK1 (intron)	0
	Del	7p22.1(6,776,023-6,919,827)x1	1/32	PMS2CL (pseudogene)	0
	Dupl	4q31.1(139,830,003-139,974,328)x3	2/77	NOCT , LOC105377448	0
	Del	2q21.1(132,395,403-132,425,713)x1	1/17	LINC01087	0
5	Dupl	17q25.3(77,363,500-77,389,101)x3	1/20	RBFOX3 (intron)	6
6	Dupl	9p13.3p13.2(36,229,595-36,728,544)x3	3/235	MELK , RNF38 , GNE	0
	Dupl	12q13.12(50,619,763-50,670,065)x3	2/24	LIMAI , MIR1293	0
7	Dupl	2q11.1(95,415,055-95,777,315)x3	4/108	ANKRD20A8P (pseudogene), MAL , MRP55 , TEKT4	0
8	Dupl	2q13(112,118,311-112,475,467)x3	1/129	MIR4435-2HG	0
9	Dupl	15q25.1(80,104,017-80,175,045)x3	1/67	MTHFS	0
	Dupl	19q13.41(52,974,191-53,127,459)x3	5/102	ZNF83, ZNF137P (pseudogene), ZNF578 , ZNF701, ZNF808	0
10	Dupl	19q13.32(46,693,400-46,782,606)x3	2/41	<i>IGFL1</i> , LOC93429	0
11	Dupl	9q34.3(138,154,269-138,309,335)x3	1/114	<i>C9orf62</i>	0
12	Dupl	1q24.2(167,549,279-167,602,533)x3	1/46	RCSDI	0
	Del	2q32.1(186,847,769-186,953,324)x1	1/71	LINC01473	0
13	Dupl	Xp22.2(13,089,235-13,621,642)x3	4/370	EGFL6 , ATXN3L , LINC01203 , <i>GS1-600G8.3</i>	3
14	Del	16p13.3(6,797,814-6,835,414)x1	1/65	RBFOX1	7
	Del	13q14.3(53,071,139-53,145,646)x0	1/7	<i>TPTE2P3</i> (pseudogene)	1
	Del	9q33.1(118,454,941-118,598,371)x1	1/107	<i>LOC101928775</i>	0
	Del	17q22(50,562,850-50,858,462)x1	1/220	LINC01982	0
15	Dupl	20p11.23(18,149,194-18,500,918)x3	7/266	DZANK1 , CSR2BP , POLR3F , SEC23B , RBBP9 , ZNF133 , MIR3192	0
	Dupl	11q25(134,347,973-134,719,556)x3	1/346	<i>LOC283177</i>	4
	Del	18q22.3(69,388,518-69,587,022)x1	1/158	LINC01899	0
16	Del	7q11.23(76,179,701-76,303,498)x1	2/46	POMZP3 , <i>LOC100133091</i>	0

Fig. 1 FISH analysis of duplications in 16p13.11 (2N4) and 2q11.1 (2N7). Possible tandem duplication in one 2N cancer patient (with brighter signals than on corresponding chromosomes indicated with white arrow) (a). Copy number qPCR analysis may indicate duplications in a mosaic state (b)



transcription binding- or CpG-sites that might be important for the regulation of genes outside of CNVs. On these grounds, we conducted an analysis on parts in the genome which are classified to date as gene free (data base UCSC <https://genome.ucsc.edu/cgi-bin/hgGateway>) and were found to be homozygously or heterozygously deleted mainly in the tumor patients (Online Resource 1 Table: Gene free CNV 2N and 1N). We did not detect any homozygously deleted sites, whereas twelve regions were heterozygously deleted in 2N cases and six in 1N tumor cases. Eight areas were duplicated (six in 2N and two in 1N). Only four of these regions did not contain any transcription factor binding site. Other regions like 5q21.2(103,509,767-103,534,114)x1,

harboring *MYC*, *RAD21*, or *SMC3* binding sites, suggest some involvement in the regulation of DNA repair or growth control. There were no significant differences or overlaps between 1N and 2N cases. None of the detected areas contained CpG islands (<https://genome.ucsc.edu/>).

Gene expression in cells of 0N patients after irradiation

As stated by other researchers, low background radiation and therapeutic radiation treatments are important inducers of cancer and secondary independent cancers. To estimate the influence of CNVs on gene regulation after irradiation, we

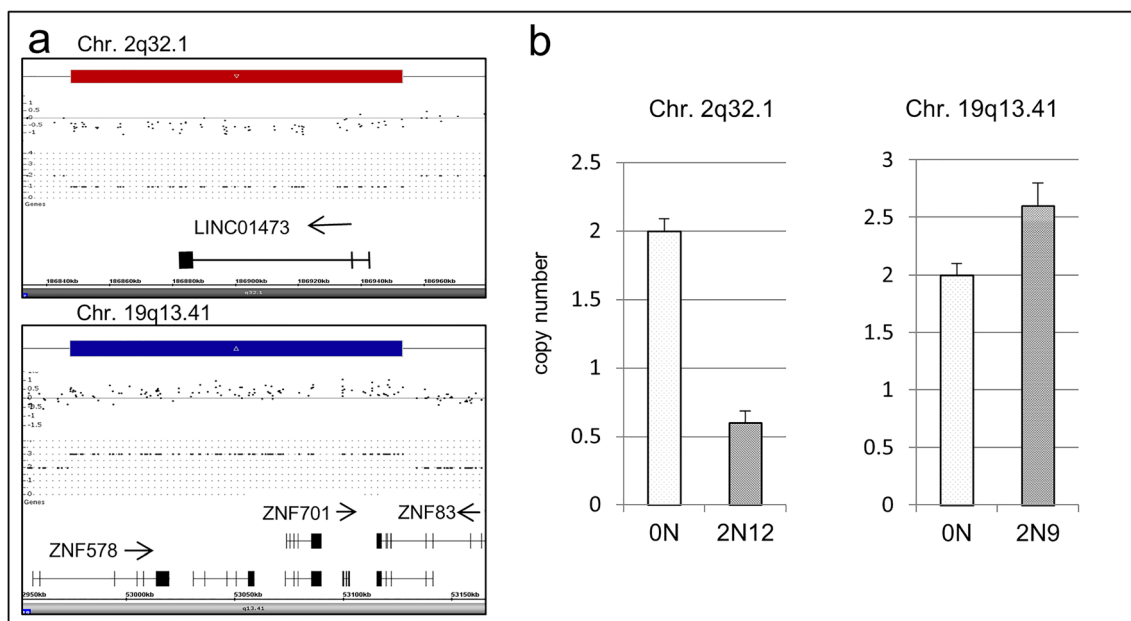


Fig. 2 SNP array copy number analysis (a) and verification using a qPCR approach (b). Case 2N12 shows a deletion in chromosome 2q32.1 and case 2N9 shows a duplication in 19q13.41. qPCR analyses with specific primers show a heterozygous/homozygous deletion

designed a study to analyze collateral radiation damage, aiming to detect genes that are transcriptionally altered after irradiation and affected in patients by CNVs.

To compare the gene expression with the genomic gains and losses of former tumor patients, it was necessary to generate a comparable gene expression data set. There are several studies published on gene expression after gamma radiation in human primary fibroblasts, mostly performed using chip technology and examining the expression in 2D and generally at 80% confluency. As G0/G1 is probably the predominant cell cycle stage in collaterally irradiated tissues, we designed our experiments in cell cycle arrested cells, to omit mitotic gene expression and DNA repair in dividing cells, as previously published with skin fibroblasts and neonatal foreskin cell lines [14, 15]. To ensure a wide spectrum of the gene induction after radiation, we used three independent 0N fibroblast cell lines and extracted the RNA after 15 min (early response), 2 h (mid response), and 24 h (late response). The radiation doses were chosen either to be therapeutically relevant (2 Gy) or experimental (5 Gy and 8 Gy). Using the entire data set, without regard for the differences in radiation dose and time, we detected 21,459 dysregulated genes (p value < 0.05) post-radiation. After stratifying the results for false discovery rate (FDR) < 0.05, we found 2619 genes to be altered in their transcription rate (Online Resource 2). Considering the post-radiation time, the gene expression varied strongly between 15 min, 2 h, and 24 h. After 15 min, we detected only one regulated gene (*ANAPC4*) (FDR < 0.05) in comparison with gene induction after 2 h (FDR < 0.05; 1472 regulated genes) and 24 h (FDR < 0.05; 1567 regulated genes).

For the verification of the RNA-seq experiments, we chose ten representative genes (*APG1*, *CDKN1A*, *CSRNP1*, *FAM111B*, *FBXO22*, *KRT17*, *MDM2*, *MYBL2*, *RAD54L*, and *THSD1*) for further characterization using qPCR. Among them are known marker genes that have already been described to be regulated upon radiation, such as *CDKN1A*, *RAD54L*, and *KRT17* (Fig. 3).

As stated by Christmann and Kaina [16], mammalian cells express DNA repair genes at a detectable basal level and just a slight upregulation or downregulation may significantly ameliorate the repair capacity of the cell. By convention, an expression change of ± 1.5 –2-fold is considered to be biologically relevant. In our experience, the calculation of the fold changes depends on the platform used to generate the data and the bioinformatics normalization approach. Therefore, we used a data set based on the FDR value < 0.05 with 2619 genes (Online Resource 2), with no regard to time and dose (duplicate genes were removed) for comparison with the molecular karyotypes of the patients.

Irradiation sensitive genes affected by CNVs in cancer patients

To analyze the impact of radiation-induced genes in unique patient-related CNVs, we compared the SNP array data with

the gene expression signatures obtained after irradiation. Among the 2N patients, we detected six genes (*POLR3F*, *SEC23B*, *ZNF133*, *C16orf45*, *RRN3*, and *NTANI*) which were overexpressed after irradiation and were duplicated in the genome of ALL patients with the second independent cancer being either meningioma or thyroid carcinoma. None of these genes has been described to promote cancer, but *ZNF133* has been identified to be overexpressed in osteosarcoma [17]. Among the 1N patients, we detected five genes (*ZCWPW*, *SYNCR1*, *DHX3*, *DHRS4L2*, and *THSD1*) that were differentially regulated after irradiation and were located in duplicated regions. We analyzed the expression profile of *THSD1* in three independent controls (0N) and six patient-derived fibroblast cell lines (three 1N and three 2N) using qPCR and detected highly variable expression changes after radiation among controls as well as cancer patients (1N, 2N). We could not establish a clear connection between the duplication of *THSD1* and an increased expression before and after irradiation (Fig. 4).

Other than the previously mentioned genes, we also detected radiation-sensitive genes within common CNVs. The copy number of the *DUSP22* gene is highly variable among individuals and, surprisingly, this gene changes its expression upon radiation treatment, probably contributing to individual response upon therapy. To make a classifying rating, we created a term called aberration frequency, which estimates a value for the incidence of a given aberration in a given cohort. Some alterations in the genome such as *FAM86B1* or *GOLGA8A* and duplication in the intron of *PTPRN2* were up to three times more frequent in our cancer cases than in controls. A compilation of the results is given in Online Resource 1, Table X-ray response. The highest response upon radiation (fold change) was calculated for *GOLGA8A*. The deletion occurs at the 5' end of the retained intron transcript variant 2 non-coding RNA. qPCR examination of the copy number status in matched 1N8/0N11 and 2N8/0N11 samples revealed a loss for the 1N8 case, whereas in 2N8, the loss was heterozygous. We analyzed the expression of *GOLGA8A* before and after radiation in the corresponding matching samples. The 0N11 control shows an increase of expression after 2 h proportional to the radiation intensity, whereas the cancer patient samples 1N8 and 2N8 show a diminished response after irradiation (Fig. 5).

Methylation analysis of duplicated genes

Gene expression is also modulated by methylation. In a previous study, we did not find global methylation changes in normal fetal fibroblasts 1–72 h after irradiation, neither in genic (promoters, 5' UTRs, first exons, gene bodies, and 3' UTRs) nor in intergenic regions [18]. To analyze the possibility of methylation changes upon altered DNA content in fibroblasts of our cancer patients, we chose two genes that

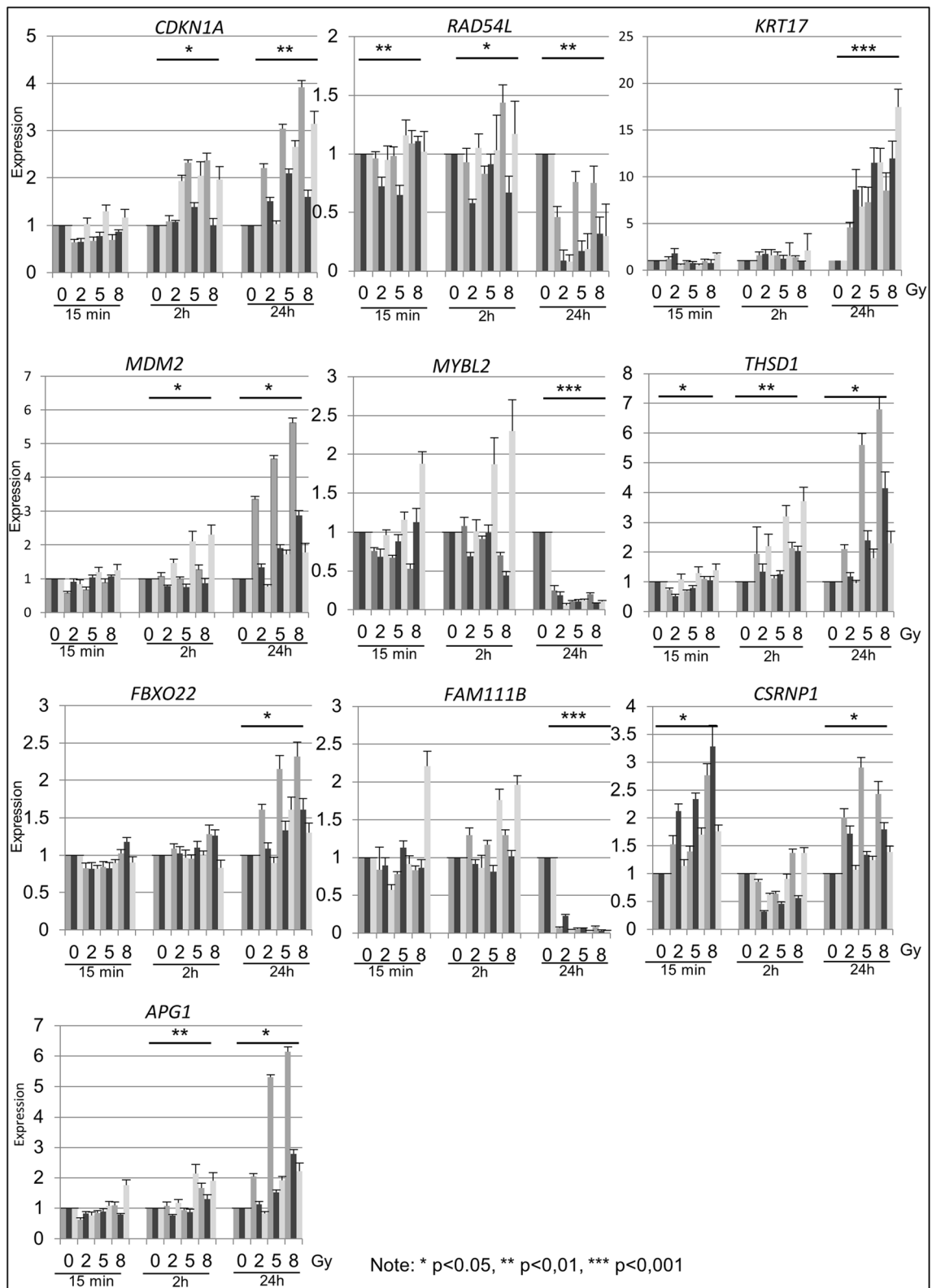


Fig. 3 QPCR experiments in three independent 0N cell lines. The analysis of ten genes was performed with irradiated (2, 5, 8 Gy) and control (0 Gy) primary fibroblast cell lines at 15 min, 2 h, and 24 h post-treatment. Significant differences in expression are indicated by asterisk

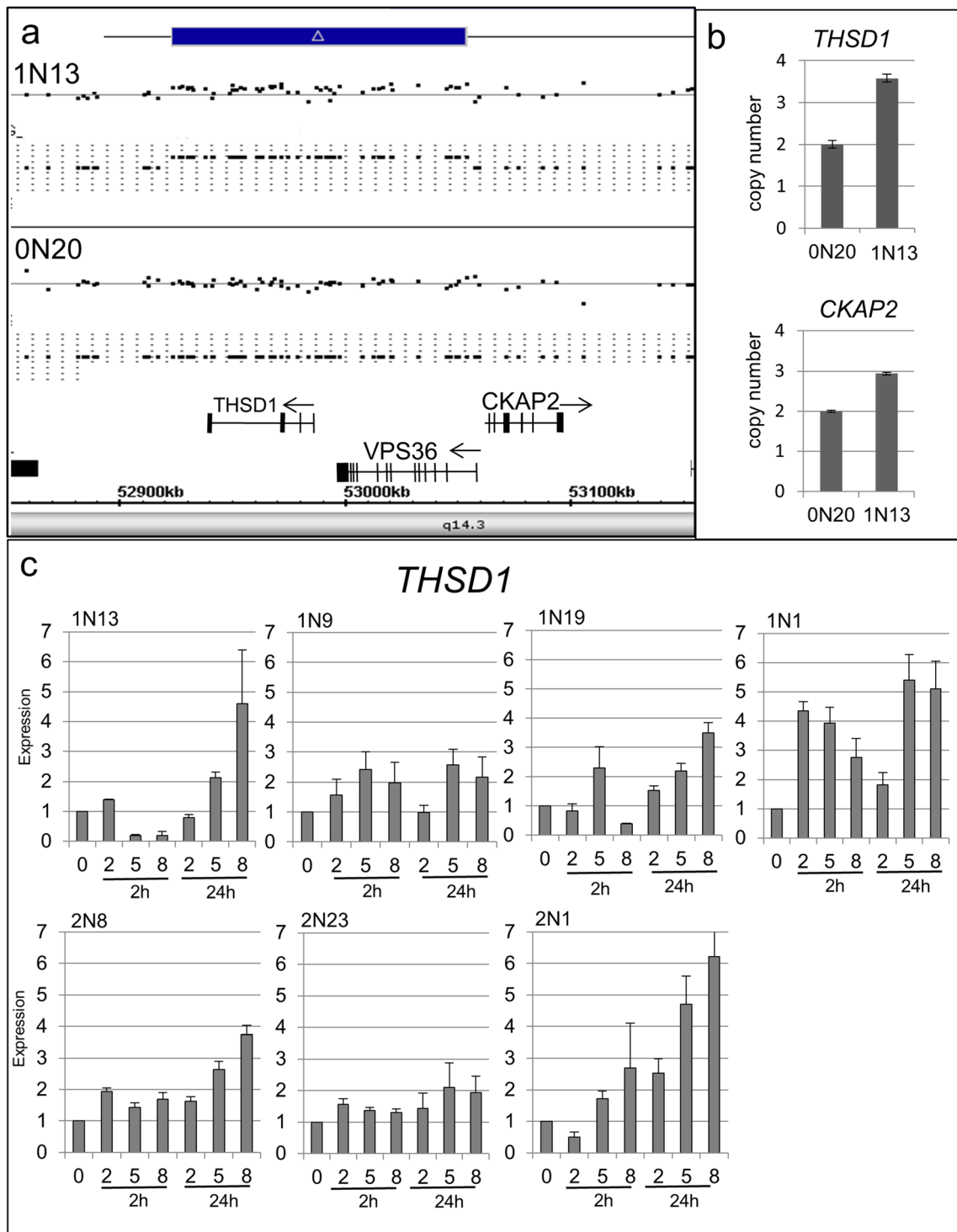


Fig. 4 Analysis of the duplication and gene expression of the *THSD1* gene. The scheme shows duplication in 13q14.3 including the *THSD1* gene (a). qPCR copy number analysis of this region reveals a heterozygous duplication of the DNA fragment in 1N13 (b). RNA

expression analysis using qPCR in index patient 1N13 and three independent 2N and 1N cell lines. The analysis was performed with irradiated (2, 5, 8 Gy) and control (0 Gy) cells at 2 h and 24 h post-treatment (c)

presented with a CNV and were also differentially methylated in several cancer cell lines. *GSTT2* is deleted in patient 2N7 (arr[hg19] 22q11.23(24,283,003-24,330,206)x1) and in 58 participants of the control GHS collective. In comparison with

the hypermethylation of the *GSTT2* CpG island, consisting of six CpGs, in the A549, MCF7, and EFO27 cancer cell lines, the patient’s sample was hypomethylated similar to the matched control sample 0N7 and the FancD1 fibroblast line.

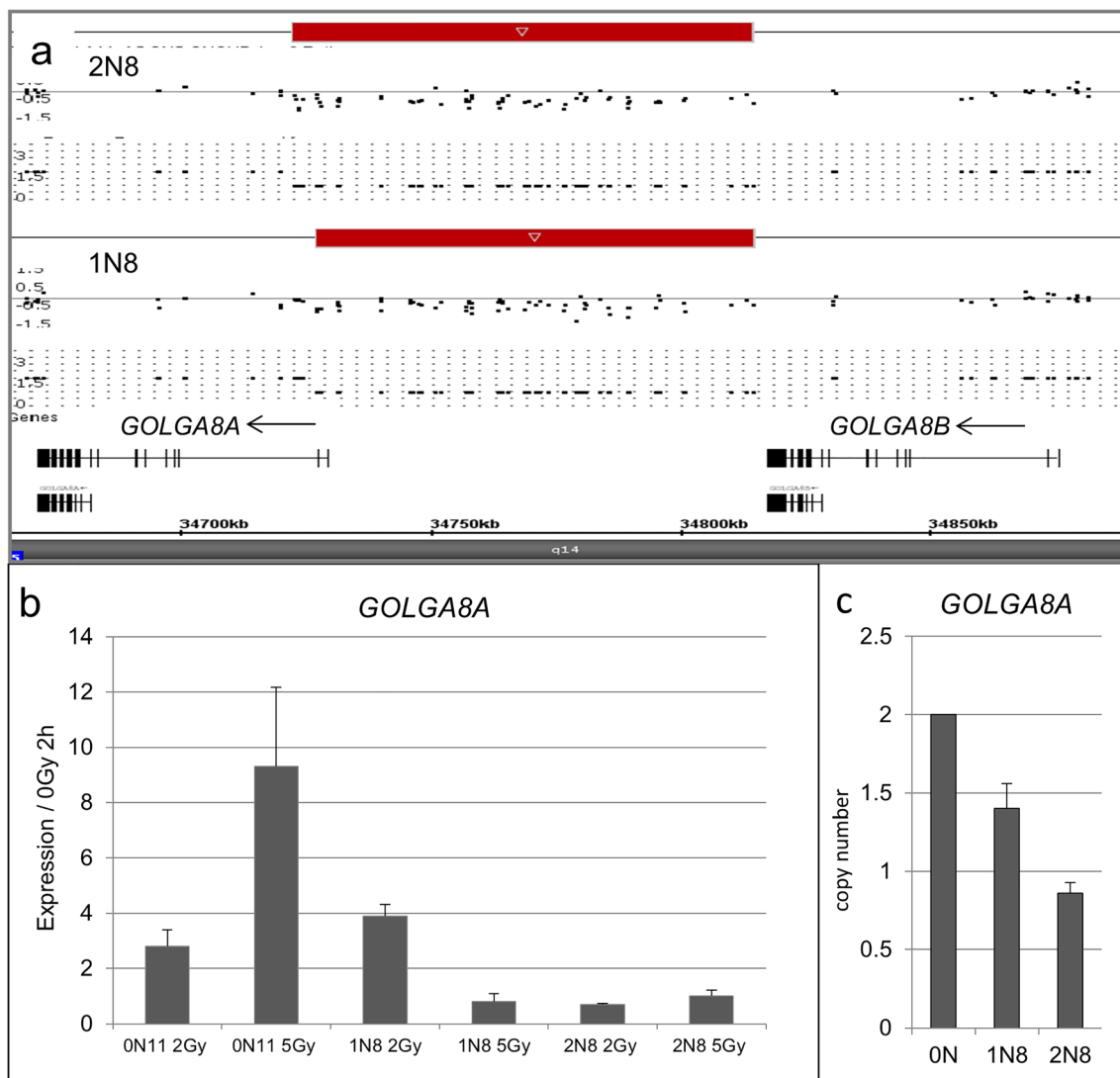


Fig. 5 Detail analysis of a deletion in Chr.15q14 of two matched patients. Genomic deletion in 15q14, including the *GOLGA8A* gene (a). RNA expression analysis using qPCR in 1N8, 2N8, and corresponding 0N11 patients. The analysis was performed with irradiated (2, 5 Gy) and control

(0 Gy) cells at 2 h post-treatment (b). qPCR copy number analysis of this region reveals heterozygous deletions of the DNA fragment in a heterozygous state (c)

The analysis of the duplicated *THSD1* gene promoter with ten CpGs showed similar results to the *GSTT2* gene. In the case 1N13 and the matched samples 2N19 and 0N20, the values matched the values of normative samples in contrast to the hypermethylation in the two cancer cell lines ZR-75-1 and T47D (Fig. 6).

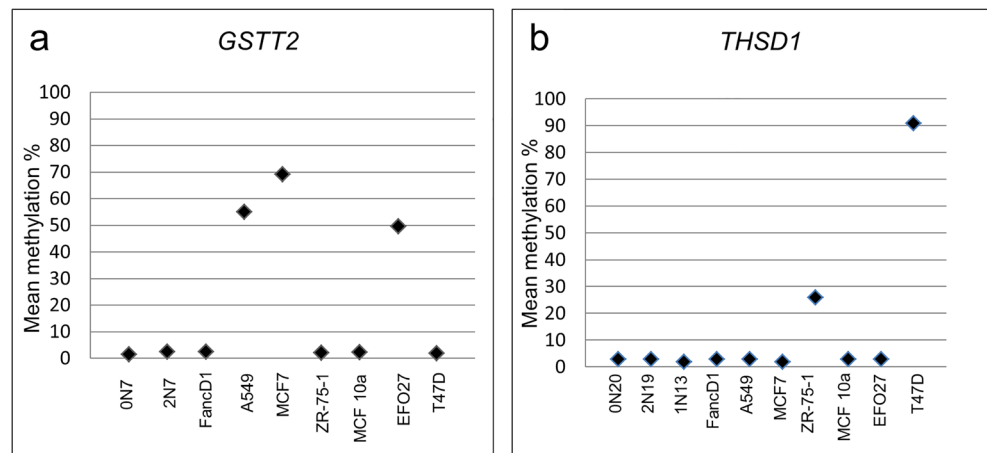
Discussion

Copy number variation in cancer patients

In this study, we focused on rare CNVs, which may have an impact on cancer predisposition and recurrent cancer incidents in sporadic, non-familial, non-syndromic childhood cancer cases. There have been some studies on rare copy number

aberrations (CNA) in hereditary cancer predisposition syndromes [19–22], whereas studies in sporadic cancer cases are sparse, or the cancer family history is not stated [10, 23, 24]. The possibility of CNVs occurring due to radiotherapy cannot be excluded but is unlikely because the fibroblasts in our study were usually obtained several years after the last treatment and the number of aberrant cells decreases over time and after 13 months post-therapy, normal karyotypes are prevalent [9]. The findings of acquired CNVs in leukemia patients from other studies [25, 26] did not match with the detected CNVs in our study and make them unlikely to be a secondary event. In addition, none of the CNVs in our cohort matched the reported de novo induced CNVs in clonal descendants of irradiated human fibroblasts [27]. Nevertheless, pre-existing somatic mutations acquired prior to treatment may be selected during chemotherapy/radiation and may lead to therapy-

Fig. 6 Methylation analysis of *THSD1* and *GSTT2* genes. Depicted are the mean methylation values of six CpGs for *GSTT2* and ten CpGs for *THSD1*. The tumor cell lines ZR-75-1 and T47D display hypermethylation of *THSD1* in contrast to 0N20, 2N19, 1N13, and other samples (a). The A549, MCF7, and EFO27 tumor cell lines show hypermethylation of *GSTT2* compared with 0N7, 2N7, and the remaining samples (b)



related secondary cancer [2, 28]. The aim, therefore, was to detect CNVs that could harbor genes that function as modifiers of cancer risk rather than being innocuous [29]. Since six of our patients received allogeneic bone marrow transplants, their blood DNA would represent the donor's profile. We, therefore, did not use EBV-transformed lymphoblasts, but primary fibroblasts, which constitute a homogenous cell population with intact cell cycle and DNA repair checkpoints. To date, few childhood cancer-predisposing mutations are known [30]. The patients of our collective showed no germline mutations in high-penetrant cancer predisposition genes like *TP53* or *BRCA1/BRCA2*. None of the patients classified for a genetic cancer (predisposition) syndrome.

In our study, 75% of the former cancer patients displayed unique CNVs and very few were shared between the 1N and 2N group. Only one gene (*ABCC6*) was partially duplicated in two patients and a similar duplication was reported in an adult cancer patient by Villacis et al. [31]. A deletion of the *LINC01473* gene present in a 2N patient was reported to be deleted in a childhood cancer patient by Krepischi [23]. In 2N patients, we saw aberrations, e.g., in Chr.16p13.11, that harbor at least four genes (*ABCC1*, *FOPNL*, *MYH11*, *KIAA0430*) involved in cancer development or that are present in rare CNVs in cancer patients. Another duplicated region, 9p13.3p13.2, includes the genes *MELK*, *RNF38*, and *GNE*. The *MELK* gene is involved in cell proliferation and apoptosis whereas multiple losses of *RNF38* were detected in CML samples. Another gene that is duplicated in this region is *GNE*, which has been reported to be overexpressed in cancer (see Online Resource 1 for detailed information). Loss of it is important for the induction of apoptosis. At this point, we cannot completely exclude the possibility that the detected aberrations may have some impact on primary and secondary tumor development in 2N patients.

Similar findings were made for 1N patients. We detected a heterozygous loss of the *ST18* gene in 1N patient who suffered from leukemia, which has been previously described in two papers [32, 33]. The finding in a 1N patient of duplicated *CKAP2*,

THSD1, and *VPS36* genes in 13q14.3 is very interesting because *CKAP2* is responsible for spindle bipolarity and chromosome stability and may represent a new factor contributing to eye tissue cancer development. As the patient was cured by surgical treatment only, the possible therapy-related changes can be excluded in this case. Altogether we would consider aberrations in *SNX14*, *SYNCRIP*, *CBR3-AS1*, and *CKAP2* as putatively responsible for tumor development or being at least an important passenger aberration in the respective 1N patients.

We verified a subset of the CNV results by qPCR and FISH analysis. Some of the CNV regions are duplicated in tandem mode and some cases may present as mosaic duplications, which conforms with the theory stated by Hu et al., [19]. As we showed in previous work, mosaic CNVs also occur in other cancer patients [34]. The absence of selective pressure might preclude the phenotypic manifestations of the minor mosaic population in a phenotypically normal individual [35], but this does not exclude the possibility of progression or change in the cell microenvironment toward cancer [36].

Some conditions might not be associated with a specific gene dosage, but rather the simple presence of a structural change at a given position in the human genome. It was stated before that intragenic regions may have an impact on cancer risk. A deleted intergenic locus may contain an enhancer which modulates breast cancer risk [37] or intergenic regions may harbor novel transcripts [38]. We, therefore, conducted a survey to find altered areas with annotated transcription binding sites in our patient collective. We included, in our study, the analysis of to date “gene free” regions and for all conspicuous areas, the presence of transcription factor binding sites, which to our knowledge, has not been done previously. These detected structural changes in cancer patient's genomes may cause perturbation in particular pathways regardless of gene dosage. The importance of the detected “gene free” loci will be the subject of further surveys.

There are certain limitations of our study. Firstly, we were unable to collect samples from relatives to evaluate the familial status of the aberration as well as the corresponding tumor

material. Also, the involvement of mutations in genes unknown to date cannot be excluded. Thirdly, there is evidence that distinct tumor classes exist, one driven by mutations and the other driven by CNVs. The question if the detected CNVs may represent driver or passenger mutations cannot be answered at this stage. Due to the chosen technique, only duplications and deletions are described in this study, while balanced structural changes remain undetected. Each patient displayed an almost unique pattern of aberrations which is consistent with the idea of the multi-causation of cancer and findings of spontaneous abnormalities in normal human fibroblasts from patients with Li-Fraumeni cancer syndrome [39] and chromosomal changes in non-cancerous breast tissue for breast carcinoma patients [40]. Interestingly, an abnormal gene expression in fibroblasts was detected in patients with Gorlin syndrome (GS), which is a hereditary disorder with tumorigenicity, caused by constitutive hyperactivity of hedgehog signaling. The hyper-activated hedgehog signaling contributes to low miR-196a-5p expression and high MAP3K1 expression in human fibroblasts and mouse cells [41]. The described CNVs in this study may consequently deregulate gene expression and important pathways. Although it was not possible to connect the occurrence of particular CNVs and 2N or 1N cancer incidence, our findings may inspire new insights in cancer pathways regulation. To our knowledge, this is the first time that such an investigation in childhood cancer survival patients prevalent in ALL and matched triplet design and additional 1000 well-chosen controls was performed. Further studies are on the way with increased patient numbers to corroborate our findings.

Radiation sensitive genes inside copy number variations

We generated an extensive gene list, which represents the dose- and time-dependent response upon irradiation damage to define sensitive genes within CNVs. qPCR analysis and former studies with radiation-induced transcriptional responses performed on quiescent fibroblasts support our results. We detected altered regions harboring genes, which respond upon irradiation and were not described to date as radiation-sensitive, but some of the detected genes were described in cancer. One gene (*ZNF133*), which responded upon irradiation, was detected in a patient of the 2N collective and has been identified as being overexpressed in osteosarcoma [17]. More genes described in cancer were detected in 1N patients. The gene *ZCWPW2* is a histone-modifying enzyme [42] and was downregulated after radiation in controls. The gene locus was found to be duplicated in a patient who suffered from Hodgkin's lymphoma. *SYNCRIP*, an RNA-binding protein that controls the myeloid leukemia stem cell program [43], was found to be overexpressed after radiation and was duplicated in a patient with sarcoma. *DHX30*, which was found to be frequently mutated in childhood

AML [44], was duplicated in a patient with ALL. The *THSD1* gene is often mutated in cancer [45] and was found to be duplicated in a patient with unilateral retinoblastoma (1N13), lacking an *RBI* mutation. None of the findings may explain the proneness to secondary cancers in 2N participants of this study. One of the most likely reasons for this result is the limited number of patients. Additional studies by analyzing larger cohorts may thus uncover more responsible gene sites. We show exemplarily that gene expression may depend upon copy number alterations but there are also exceptions. *THSD1* proved to be expressed in some individuals after irradiation on a copy number independent status. This is not unusual, because individual gene expression response after genotoxic agents in fibroblasts has already been described. Such findings complicate future studies and make intensive studies on RNA and protein level necessary. In some frequency, genomic alterations in gene copy numbers were seen also in controls, which may indicate a certain individual plasticity upon damage. Genes like *DUSP22*, which is deleted in some cases of cutaneous anaplastic large T cell lymphomas, may have more than one physiological substrate and the regulation of specific signaling cascades by this enzyme may be cell-type and context-specific. Recently, *DUSP22* was described to contribute to inflammatory bowel disease [46]. Surprisingly, its copy number varies in fibroblasts in 0N cases as well as in our patients. Another gene that was affected in controls and more frequently in our patients may play a role in tumor tissue. Significant biallelic deletions of *GOLGA8A* have been described in gastrointestinal tumors and pancreatic ductal adenocarcinomas. The detected downregulation in our patients correlates with the deletion status of the region, as shown also by Wang et al., [47].

It was shown that CNVs may be associated with aberrant methylation and have an impact on tumor prognosis [48]. Thus, to compensate for the addition or loss of genetic material, affected genes may be fine-tuned by methylation [10, 49]. In our two analyzed gene loci (*THSD1* and *GSTT2*), no aberrant methylation in patients was detected in contrast to hypermethylation in the analyzed tumor cell lines. Nevertheless, additional methylation surveys should be conducted in further studies.

To our knowledge, this is the first study that uses primary fibroblasts of childhood sporadic cancer cases. In conclusion, although we did not detect a consistent overall candidate gene, we describe potential vulnerable sites or rare CNVs in our collective, which may contribute to tumor development. Furthermore, we detected genes sensitive to radiation treatment that are transcriptionally altered by CNVs. As we detected aberrations seen also by other researchers, it is worthwhile to conduct further investigations in a larger collective and extensive study to address expression, cellular localization, putative deletion, and overexpression of genes to determine the impact of a given aberration on maintaining genome stability.

Acknowledgments Open Access funding provided by Projekt DEAL. We would like to thank the patients participating in this study. Our clinical staff also deserves special thanks. We are also grateful to Ursula Disque-Kaiser and Martina Hermanns for technical assistance and Heiko Karle for support during the radiation experiments. We thank Prof. Detlev Schindler (Institute of Human Genetics, Julius Maximilians University, Würzburg, Germany) for providing the FANCD1 cell line.

Authors' contributions All authors contributed to the study conception and design. Material preparation, data collection, and analysis were performed by Galetzka Danuta, Müller Tobias, Diederich Markus, Endres Miriam, Katal Nergiz, Sinizyn Olesja, Rapp Steffen, Hankeln Thomas, Prawitt Dirk, Scholz-Kreisel Peter, Mirsch Johanna, Poplawski Alicia, and Rossman Heidi. Additive data sets (1000 GHS controls) were provided by Zeller Tanja and Müller Christian. The first draft of the manuscript was written by Galetzka Danuta, Prawitt Dirk, and Chorzempa Heather and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding information This study was financially supported by the BMBF German ministry for education and science, Grants: 02NUK016A and 02NUK042A to DG and 02NUK042B to TH.

Data availability All data generated or analyzed during this study are included in this published article (and its Online Resources).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethics approval and consent to participate and for publication This study was approved by the Ethics Committee of the Medical Association of Rhineland-Palatinate in Mainz, Germany: Genkik: no. 837.440.03 (4102) DG, Kikme: no. 837.103.04 (4261) MM, Isabel: no. 837.262.12(8363-F) TH, and GHS: 837.394.17 (11228) HR.

Informed consent Written informed consent to use fibroblasts for research purposes was obtained after genetic counseling for all participating patients.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Kaatsch P, Reinisch I, Spix C, Berthold F, Janka-Schaub G, Mergenthaler A, Michaelis J, Blettner M (2009) Case-control study on the therapy of childhood cancer and the occurrence of second malignant neoplasms in Germany. *Cancer Causes Control* 20:965–980
- McNerney ME, Godley LA, Le Beau MM (2017) Therapy-related myeloid neoplasms: when genetics and environment collide. *Nat Rev Cancer* 17:513–527
- Whitworth J, Hoffman J, Chapman C, Ong KR, Lalloo F, Evans DG, Maher ER (2015) A clinical and genetic analysis of multiple primary cancer referrals to genetics services. *Eur J Hum Genet* 23: 581–587
- Mitelman F, Johansson B, Mertens F (2007) The impact of translocations and gene fusions on cancer causation. *Nat Rev Cancer* 7: 233–245
- Kolomietz E, Al-Maghrabi J, Brennan S, Karaskova J, Minkin S, Lipton J, Squire JA (2001) Primary chromosomal rearrangements of leukemia are frequently accompanied by extensive submicroscopic deletions and may lead to altered prognosis. *Blood*. 97: 3581–3588
- Stranger BE, Forrest MS, Dunning M, Ingle CE, Beazley C, Thorne N, Redon R, Bird CP, de Grassi A, Lee C, Tyler-Smith C, Carter N, Scherer SW, Tavare S, Deloukas P, Hurles ME, Dermitzakis ET (2007) Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science*. 315:848–853
- Al-Khodairy FM et al (2004) Defective repair of UV-induced DNA damage in cultured primary skin fibroblasts from Saudi thyroid cancer patients. *Asian Pac J Cancer*:139–143
- Chuang EY, Chen X, Tsai M-H, Yan H, Li C-Y, Mitchell JB, Nagasawa H, Wilson PF, Peng Y, Fitzek MM, Bedford JS, Little JB (2006) Abnormal gene expression profiles in unaffected parents of patients with hereditary-type retinoblastoma. *Cancer Res* 66: 3428–3433
- Massenkeil G, Zscheschang P, Thiel G, Hemmati PG, Budach V, Dörken B, Pross J, Arnold R (2015) Frequent induction of chromosomal aberrations in in vivo skin fibroblasts after allogeneic stem cell transplantation: hints to chromosomal instability after irradiation. *Radiat Oncol* 10:266
- Kuiper RP, Ligtenberg MJL, Hoogerbrugge N, van Geurts Kessel A (2010) Germline copy number variation and cancer risk. *Curr Opin Genet Dev* 20:282–289
- Kuhlen M, Taebner J, Brozou T, Wiczorek D, Siebert R, Borkhardt A (2019) Family-based germline sequencing in children with cancer. *Oncogene*. 38:1367–1380
- Ripperger T, Bielack SS, Borkhardt A, Brecht IB, Burkhardt B, Calaminus B et al (2017) Childhood cancer predisposition syndromes—a concise review and recommendations by the Cancer Predisposition Working Group of the Society for Pediatric Oncology and Hematology. *Am J Med Genet Part A* 173A:1017–1037
- Fagerberg L, Hallström BM, Oksvold P, Kampf C, Djureinovic D, Odeberg J, Habuka M, Tahmasebpoor S, Danielsson A, Edlund K, Asplund A, Sjöstedt E, Lundberg E, Szijarto CAK, Skogs M, Takanen JO, Berling H, Tegel H, Mulder J, Nilsson P, Schwenk JM, Lindskog C, Danielsson F, Mardinoglu A, Sivertsson Å, von Feilitzen K, Forsberg M, Zvalnen M, Olsson IM, Navani S, Huss M, Nielsen J, Ponten F, Uhlén M (2014) Analysis of the human tissue-specific expression by genome-wide integration of transcriptomics and antibody-based proteomics. *Mol Cell Proteomics* 13:397–406
- Ding L-H, Shingyoji M, Chen F, Hwang J-J, Burma S, Lee C, Cheng JF, Chen DJ (2005) Gene expression profiles of normal human fibroblasts after exposure to ionizing radiation: a comparative study of low and high doses. *Radiat Res* 164:17–26
- Warters RL, Packard AT, Kramer GF, Gaffney DK, Moos PJ (2009) Differential gene expression in primary human skin keratinocytes and fibroblasts in response to ionizing radiation. *Radiat Res* 172:82–95
- Christmann M, Kaina B (2013) Transcriptional regulation of human DNA repair genes following genotoxic stress: trigger

- mechanisms, inducible responses and genotoxic adaptation. *Nucleic Acids Res* 41:8403–8420
17. Li Y, Liang Q, Wen Y-q, Chen L-l, Wang L-t, Liu Y-l, et al. Comparative proteomics analysis of human osteosarcomas and benign tumor of bone. *Cancer Genet Cytogenet* 2010;198:97–106. doi:
 18. Maierhofer A, Flunkert J, Dittrich M, Müller T, Schindler D, Nanda I, Haaf T (2017) Analysis of global DNA methylation changes in primary human fibroblasts in the early phase following X-ray irradiation. *PLoS One* 12:e0177442
 19. Hu L, Yao X, Huang H, Guo Z, Cheng X, Xu Y, Shen Y, Xu B, Li D (2018) Clinical significance of germline copy number variation in susceptibility of human diseases. *J Genet Genomics* 45:3–12
 20. Gu X, Coates PJ, Boldrup L, Wang L, Krejci A, Hupp T, Fahraeus R, Norberg-Spaak L, Sgaramella N, Wilms T, Nylander K (2019) Copy number variation: a prognostic marker for young patients with squamous cell carcinoma of the oral tongue. *J Oral Pathol Med* 48:24–30
 21. Brown JR, Hanna M, Tesar B, Pochet N, Vartanov A, Fernandes SM, Werner L, Ash M, Roden CA, MacConaill L, Hainz U, Longtine J, Wang YE, Correll M, van de Peer Y, Regev A, Wu C, Neuberger D, Freedman AS (2012) Germline copy number variation associated with Mendelian inheritance of CLL in two families. *Leukemia*. 26:1710–1713
 22. Kayser K, Degenhardt F, Holzapfel S, Horpaopan S, Peters S, Spier I, Morak M, Vangala D, Rahner N, von Knebel-Doerberitz M, Schackert HK, Engel C, Büttner R, Wijnen J, Doerks T, Bork P, Moebus S, Herms S, Fischer S, Hoffmann P, Aretz S, Steinke-Lange V (2018) Copy number variation analysis and targeted NGS in 77 families with suspected Lynch syndrome reveals novel potential causative genes. *Int J Cancer* 143:2800–2813
 23. Krepischi ACV, Capelli LP, Silva AG, de Araújo ÉSS, Pearson PL, Heck B, da Costa CML, de Camargo B, Rosenberg C (2014) Large germline copy number variations as predisposing factor in childhood neoplasms. *Future Oncol* 10:1627–1633
 24. Walker LC, Marquart L, Pearson JF, Wiggins GAR, O'Mara TA, Parsons MT et al (2017) Evaluation of copy-number variants as modifiers of breast and ovarian cancer risk for BRCA1 pathogenic variant carriers. *Eur J Hum Genet* 25:432–438
 25. Walter MJ, Payton JE, Ries RE, Shannon WD, Deshmukh H, Zhao Y, Baty J, Heath S, Westervelt P, Watson MA, Tomasson MH, Nagarajan R, O'Gara BP, Bloomfield CD, Mrozek K, Selzer RR, Richmond TA, Kitzman J, Geoghegan J, Eis PS, Maupin R, Fulton RS, McLellan M, Wilson RK, Mardis ER, Link DC, Graubert TA, DiPersio JF, Ley TJ (2009) Acquired copy number alterations in adult acute myeloid leukemia genomes. *Proc Natl Acad Sci U S A* 106:12950–12955
 26. Papaemmanuil E, Rapado I, Li Y, Potter NE, Wedge DC, Tubio J, Alexandrov LB, van Loo P, Cooke SL, Marshall J, Martincorena I, Hinton J, Gundem G, van Delft FW, Nik-Zainal S, Jones DR, Ramakrishna M, Tittley I, Stebbings L, Leroy C, Menzies A, Gamble J, Robinson B, Mudie L, Raine K, O'Meara S, Teague JW, Butler AP, Cazzaniga G, Biondi A, Zuna J, Kempski H, Muschen M, Ford AM, Stratton MR, Greaves M, Campbell PJ (2014) RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet* 46:116–125
 27. Flunkert J, Maierhofer A, Dittrich M, Müller T, Horvath S, Nanda I, Haaf T (2018) Genetic and epigenetic changes in clonal descendants of irradiated human fibroblasts. *Exp Cell Res* 370:322–332
 28. Klimek VM, Tray NJ (2016) Therapy-related myeloid neoplasms: what's in a name? *Curr Opin Hematol* 23:161–166
 29. Krepischi ACV, Pearson PL, Rosenberg C (2012) Germline copy number variations and cancer predisposition. *Future Oncol* 8:441–450
 30. Sweet-Cordero EA, Biegel JA (2019) The genomic landscape of pediatric cancers: implications for diagnosis and treatment. *Science*. 363:1170–1175
 31. Villacis RAR, Basso TR, Canto LM, Pinheiro M, Santiago KM, Giacomazzi J, de Paula CAA, Carraro DM, Ashton-Prolla P, Achatz MI, Rogatto SR (2017) Rare germline alterations in cancer-related genes associated with the risk of multiple primary tumor development. *J Mol Med* 95:523–533
 32. Jandrig B, Seitz S, Hinzmann B, Arnold W, Micheel B, Koelble K, Siebert R, Schwartz A, Ruecker K, Schlag PM, Scherneck S, Rosenthal A (2004) ST18 is a breast cancer tumor suppressor gene at human chromosome 8q11.2. *Oncogene*. 23:9295–9302
 33. Job B, Bernheim A, Beau-Faller M, Camilleri-Broët S, Girard P, Hofman P, Mazières J, Toujani S, Lacroix L, Laffaire J, Dessen P, Fouret P, for the LG Investigators (2010) Genomic aberrations in lung adenocarcinoma in never smokers. *PLoS One* 5:e15145
 34. Galetzka D, Hansmann T, El Hajj N, Weis E, Imrscher B, Ludwig M et al (2012) Monozygotic twins discordant for constitutive BRCA1 promoter methylation, childhood cancer and secondary cancer. *Epigenetics*. 7:47–54
 35. Youssoufian H, Pyeritz RE (2002) Mechanisms and consequences of somatic mosaicism in humans. *Nat Rev Genet* 3:748–758
 36. Etzold A, Galetzka D, Weis E, Bartsch O, Haaf T, Spix C, Itzel T, Schweiger S, Strand D, Strand S, Zechner U (2016) CAF-like state in primary skin fibroblasts with constitutional BRCA1 epimutation sheds new light on tumor suppressor deficiency-related changes in healthy tissue. *Epigenetics*. 11:120–131
 37. Wyszynski A, Hong C-C, Lam K, Michailidou K, Lytle C, Yao S, Zhang Y, Bolla MK, Wang Q, Dennis J, Hopper JL, Southey MC, Schmidt MK, Broeks A, Muir K, Lophatananon A, Fasching PA, Beckmann MW, Peto J, dos-Santos-Silva I, Sawyer EJ, Tomlinson I, Burwinkel B, Marme F, Guénel P, Truong T, Bojesen SE, Nordestgaard BG, González-Neira A, Benitez J, Neuhausen SL, Brenner H, Dieffenbach AK, Meindl A, Schmutzler RK, Brauch H, The GENICA Network, Nevanlinna H, Khan S, Matsuo K, Ito H, Dörk T, Bogdanova NV, Lindblom A, Margolin S, Mannermaa A, Kosma VM, kConFab Investigators, Australian Ovarian Cancer Study Group, Wu AH, van den Berg D, Lambrechts D, Wildiers H, Chang-Claude J, Rudolph A, Radice P, Peterlongo P, Couch FJ, Olson JE, Giles GG, Milne RL, Haiman CA, Henderson BE, Dumont M, Teo SH, Wong TY, Kristensen V, Zheng W, Long J, Winqvist R, Pylkäs K, Andrusis IL, Knight JA, Devilee P, Seynaeve C, Garcia-Closas M, Figueroa J, Klevebring D, Czene K, Hooning MJ, van den Ouweland AMW, Darabi H, Shu XO, Gao YT, Cox A, Blot W, Signorello LB, Shah M, Kang D, Choi JY, Hartman M, Miao H, Hamann U, Jakubowska A, Lubinski J, Sangrajrang S, McKay J, Toland AE, Yannoukakis D, Shen CY, Wu PE, Swerdlow A, Orr N, Simard J, Pharoah PDP, Dunning AM, Chenevix-Trench G, Hall P, Bandera E, Amos C, Ambrosone C, Easton DF, Cole MD (2016) An intergenic risk locus containing an enhancer deletion in 2q35 modulates breast cancer risk by deregulating IGFBP5 expression. *Hum Mol Genet* 25:3863–3876
 38. Bartonicek N, Clark MB, Quek XC, Torpy JR, Pritchard AL, Maag JLV, Gloss BS, Crawford J, Taft RJ, Hayward NK, Montgomery GW, Mattick JS, Mercer TR, Dinger ME (2017) Intergenic disease-associated regions are abundant in novel transcripts. *Genome Biol* 18:241
 39. Bischoff FZ, Yim SO, Pathak S, Grant G, Siciliano MJ, Giovannella BC et al (1990) Spontaneous abnormalities in normal fibroblasts from patients with Li-Fraumeni cancer syndrome: aneuploidy and immortalization. *Cancer Res* 50:7979–7984
 40. Steinarsdottir M, Gudmundsson IH, Jonasson JG, Olafsdottir EJ, Eyfjörd JE, Ogmundsdottir HM (2011) Cytogenetic polyclonality of breast carcinomas: association with clinico-pathological characteristics and outcome. *Genes Chromosomes Cancer* 50:930–939

41. Shiohama T, Fujii K, Miyashita T, Takatani T, Ikehara H, Uchikawa H, Motojima T, Uchida T, Shimojo N (2019) MicroRNAs profiling in fibroblasts derived from patients with Gorlin syndrome. *J Hum Genet* 64:757–765
42. Liu Y, Tempel W, Zhang Q, Liang X, Loppnau P, Qin S, Min J (2016) Family-wide characterization of histone binding abilities of human CW domain-containing proteins. *J Biol Chem* 291:9000–9013
43. Vu LP, Prieto C, Amin EM, Chhangawala S, Krivtsov A, Calvo-Vidal MN, Chou T, Chow A, Minuesa G, Park SM, Barlowe TS, Taggart J, Tivnan P, Deering RP, Chu LP, Kwon JA, Meydan C, Perales-Paton J, Arshi A, Gönen M, Famulare C, Patel M, Paietta E, Tallman MS, Lu Y, Glass J, Garret-Bakelman FE, Melnick A, Levine R, al-Shahrour F, Järås M, Hacohen N, Hwang A, Garippa R, Lengner CJ, Armstrong SA, Cerchietti L, Cowley GS, Root D, Doench J, Leslie C, Ebert BL, Kharas MG (2017) Functional screen of MSI2 interactors identifies an essential role for SYNCRIP in myeloid leukemia stem cells. *Nat Genet* 49:866–875
44. Farrar JE, Schuback HL, Ries RE, Wai D, Hampton OA, Trevino LR, Alonzo TA, Guidry Auvil JM, Davidsen TM, Gesuwan P, Hermida L, Muzny DM, Dewal N, Rustagi N, Lewis LR, Gamis AS, Wheeler DA, Smith MA, Gerhard DS, Meshinchi S (2016) Genomic profiling of pediatric acute myeloid leukemia reveals a changing mutational landscape from disease diagnosis to relapse. *Cancer Res* 76:2197–2205
45. Liu H, Li F, Zhu Y, Li T, Huang H, Lin T et al (2016) Whole-exome sequencing to identify somatic mutations in peritoneal metastatic gastric adenocarcinoma: a preliminary study. *Oncotarget* 7: 43894–43906
46. Zhou R, Chang Y, Liu J, Chen M, Wang H, Huang M, Liu S, Wang X, Zhao Q (2017) JNK pathway-associated phosphatase/DUSP22 suppresses CD4+ T-cell activation and Th1/Th17-cell differentiation and negatively correlates with clinical activity in inflammatory bowel disease. *Front Immunol* 8:781
47. Wang J, Dumartin L, Mafficini A, Ulug P, Sangaralingam A, Alamiry NA, Radon TP, Salvia R, Lawlor RT, Lemoine NR, Scarpa A, Chelala C, Crnogorac-Jurcevic T (2017) Splice variants as novel targets in pancreatic ductal adenocarcinoma. *Sci Rep* 7: 2980
48. Woo HG, Choi J-H, Yoon S, Jee BA, Cho EJ, Lee J-H, Yu SJ, Yoon JH, Yi NJ, Lee KW, Suh KS, Kim YJ (2017) Integrative analysis of genomic and epigenomic regulation of the transcriptome in liver cancer. *Nat Commun* 8:839
49. El Hajj N, Dittrich M, Böck J, Kraus T FJ, Nanda I, Müller T et al (2016) Epigenetic dysregulation in the developing Down syndrome cortex. *Epigenetics*. 11:563–578

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.