Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Gene–Radon Interaction Effects on Lung Cancer Risk: Validation of Susceptibility Loci Using Archived DNA from German Uranium Miners

Albert Rosenberger

arosenb@gwdg.de

University Medical Center, Georg-August-University Göttingen

Irina Bonzheim

University Hospital Tuebingen and Comprehensive Cancer Center South West

Falko Fend

University Hospital Tuebingen and Comprehensive Cancer Center South West

Beate Hochstrat

Federal Office for Radiation Protection

Georg Johnen

Institute of the Ruhr University Bochum (IPA)

Alexander Brik

Institute of the Ruhr University Bochum (IPA)

Peter Rozynek

Institute of the Ruhr University Bochum (IPA)

Heike Bickeböller

University Medical Center, Georg-August-University Göttingen

Simone Mörtl

Federal Office for Radiation Protection

Maria Gomolka

Federal Office for Radiation Protection

Research Article

Keywords: Arsenic, Lung Neoplasms, Uranium, Genotype, Radon, DNA, Miners, Occupational Exposure

Posted Date: July 27th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10446998/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: Competing interest reported. Mörtl Simone and Gomolka Maria are employees of the funder.

Version of Record: A version of this preprint was published at Genes and Environment on October 26th, 2026. See the published version at https://doi.org/10.1186/s41021-026-00365-z.

Abstract

Background

Radon exposure increases lung cancer risk, shown in studies on indoor exposure (low exposure) and uranium miners (high exposure). Genome-wide and candidate gene studies have identified DNA variants that interact with occupational radon to modify lung cancer risk.

Method

Highly fragmented DNA samples of 611 individuals were genotyped using next-generation sequencing. DNA variant associations could be assessed in 526 out of 611 WISMUT uranium miners with lung cancer. The relative risk of lung cancer per allele (marginal effect) as well as its dependence on radon exposure (gene-environment (GxE) interaction) could be estimated from the enrichment of one allele compared to the distribution of genotypes expected under the HWE assumption (Hardy-Weinberg equilibrium). Exposure to gamma radiation, long-lived radionuclides, fine dust, silicate dust and arsenic was considered in the same way.

Results

For markers from six (*UBE2U, CDKN2A, SIRT1, NAV2, ST8SIA2, TP53*) of the 10 genes examined, a significant modification of the genetic relative risk for lung cancer due to mining-specific radon exposure [in WLM] was observed.

Conclusion

Radon exposure is considered to be the primary cause of the observed GxE interaction for *UBE2U, CDKN2A* and *SIRT1*. This conclusion is supported by the function of these genes, as based on relevant literature. For another marker of the gene *SOX5*, a GxE interaction can presumably be attributed to exposure to arsenic. For the remaining genes, no single environmental factor (radon, dust or arsenic) can unambiguously be identified as sole cause of the GxE interaction.

Introduction

Radon-222 (Rn-222) is a radioactive gas formed naturally from uranium-238 decay. Its concentration depends on uranium content in rocks. (1) With a half-life of 3.8 days, it decays into short-lived solid radioactive progeny. As a gas, it diffuses through soil and accumulates in mines and homes. Humans are mainly exposed through inhalation. During decay into short-lived radioactive progeny products such as polonium-218 (Po-218) and polonium-214 (Po-214), it emits high-energy alpha particles, along with beta and gamma radiation, contributing to cumulative radiation exposure.

Alpha particles, the main exposure source, release high energy over short distances, strongly ionising and damaging DNA. Though poorly penetrating, their intense ionisation events cause severe direct damage to DNA and by ionizing water molecules also oxidative stress by generated free radicals. Inhaled radon and its progeny deposit on bronchial epithelium, where alpha particles primarily harm respiratory cells, initiating and promoting lung carcinogenesis through cytotoxic and genotoxic effects. These lead to far-reaching molecular DNA changes, such as double-strand breaks, mutations, deletions, and chromosomal rearrangements potentially initiating cancer development (2, 3).

Several studies have demonstrated an increased radon-induced risk of lung cancer (LC) in relation to indoor exposure (low dose) and occupational exposure in uranium miners (high dose) (1, 4–6).

In addition to a general genetic predisposition to lung cancer (7), a genome-wide study identified single nucleotide polymorphisms (SNPs) modifying the lung cancer risk in interaction with occupational radon exposure. The study included radon-exposed WISMUT miners, both with and without lung cancer (collected by the Federal Office for Radiation Protection, BfS),

as well as lung cancer patients and control subjects from the general population with no occupational exposure to radon (collected by the International Lung Cancer Consortium, ILCCO) (8). Radon gene interactions were observed at additional loci by other studies with candidate gene approaches and of much smaller sample size, e.g., for GSTM1 (9), IL-6 and *SIRT1* (10), or MGMT (11). However, none of these radon-specific associations/interactions observed in individual studies have yet been replicated.

The aim of this project was to validate and replicate the genome-wide interaction search and possibly the results of the other studies on radon gene interactions in an independent series of 611 DNA samples from uranium miners suffering from lung cancer. The samples from the German Uranium Miners' Biobank, which were provided by the Federal Office for Radiation Protection (BfS), consisted of highly fragmented DNA (find example figure of DNA in agarose gel in Additional file 1), isolated from archived formalin-fixed paraffin-embedded (FFPE) normal lung tissue (12, 13). As nucleic acid integrity decreases with age, and non-buffered formalin was used at the time of preservation, the isolated DNA quantity and quality limited strongly the number of markers amenable to genotyping.

Methods

Selection of markers for genotyping

Due to the limited amount of DNA available, the number of markers that could be genotyped was limited. Based on an extensive literature search, a list of 34 genes was compiled, evaluated and prioritised according to association, function and expression (14–17). These genes were either associated with radon-induced LC, known as lung cancer risk-genes or considered relevant to lung cancer pathogenesis. The prioritisation was factually justified, but to a certain extent subjective. Details are laid out in the Additional file 1. The following 10 genes (in the given order, listed in Table 1) were prioritised for validation: *FTO*, *UBE2U*, *CUBN*, *ST8SIA2*, *SOX5*, *TP53*, *CD163L1*, *CDKN2A*, *NAV2* and *SIRT1*. Two to four markers per gene were selected for typing, covering the affected regions (in total 29 markers). The selection was based on the marker-specific significance (p-value) of a radon x gene interaction (8), the linkage disequilibrium (LD) between the markers and the frequency of the rarer allele (minor allele frequency MAF, > 5%).

Table 1
External information on the markers selected to be genotyped

No.	marker	chr.	position	MAF	major allele	minor allele	gene	selection reason
1	rs7545208	1	64 665 046	33.9%	A	C	*UBE2U*	LD-Sig.
2	rs7526950	1	64 684 038	22.9%	C	T	*UBE2U*	LD-Sig.
3	rs10889478	1	64 804 890	48.1%	T	C	*UBE2U*	LD-Sig.
4	rs1908635	1	64 805 503	17.2%	C	A	*UBE2U*	significance
5	rs7023680	9	21 816 573	47.0%	A	G	*CDKN2A*	significance
6	rs3731217	9	21 984 661	14.3%	A	C	*CDKN2A*	significance
7	rs10965227	9	22 081 796	23.3%	A	G	*CDKN2A*	significance
8	rs73596224	10	17 005 956	6.6%	C	A	*CUBN*	significance
9	rs60918373	10	17 046 641	15.2%	T	A	*CUBN*	LD-Sig.
10	rs7095324	10	17 250 900	36.3%	G	T	*CUBN*	significance
11	rs10823112	10	69 670 816	6.0%	A	G	*SIRT1*	LD-Sig.
12	rs10762183	10	69 713 297	32.0%	A	C	*SIRT1*	LD-Sig.
13	rs7129281	11	19 265 226	26.8%	C	T	*NAV2*	significance
14	rs11025138	11	19 466 820	30.2%	T	C	*NAV2*	LD-Sig.
15	rs4756999	11	19 468 769	33.5%	C	A	*NAV2*	significance
16	rs7137426	12	7 382 236	37.0%	G	T	*CD163L1*	significance
17	rs7970379	12	7 451 788	5.0%	A	G	*CD163L1*	LD-Sig.
18	rs7302538	12	7 499 479	5.5%	C	T	*CD163L1*/ ACSM4	significance
19	rs1353793	12	24 190 988	45.0%	G	T	*SOX5*	LD-Sig.
20	rs7978583	12	24 244 724	46.5%	T	C	*SOX5*	significance
21	rs4777965	15	92 898 565	38.5%	T	C	*ST8SIA2*	LD-Sig.
22	rs1455777	15	92 988 741	49.9%	G	A	*ST8SIA2*	LD-Sig.
23	rs3784722	15	93 006 088	36.8%	A	G	*ST8SIA2*	significance
24	rs8061518	16	53 861 024	33.9%	A	G	*FTO*	LD
25	rs6499657	16	53 988 978	15.8%	G	C	*FTO*	significance
26	rs1111483	16	54 000 907	48.1%	A	C	*FTO*	LD
27	rs17886268	17	7 592 802	3.9%	C	T	*TP53*	significance
28	rs7141	17	7 614 601	36.6%	A	G	*TP53*	significance
29	rs7221056	17	7 711 126	33.2%	C	T	*TP53*	significance

Chr. chromosome; position on chromosome in [bp] according to Genome Reference Consortium Human Build 37 (GRCh37); MAF "minor allele frequency" according to SNPNexus (18); gene closest (assigned) gene according to SNPNexus. Selection reason: LD-Sig.= observed interaction and LD with most significant marker, LD = LD with most significant marker

Genotyping

Determination of the genotype by NGS

Genotypes were determined by next generation sequencing (NGS). To establish the NGS custom panel, 15 normal-tissue DNA samples with severe DNA fragmentation (100–200 base pairs), taken from the archive of the Institute of Pathology in Tübingen, were analysed.

A custom panel was designed for targeted analysis of 29 markers, optimized for highly degraded DNA with amplicon lengths up to 140 bp, with the Ampliseq Designer of Thermo Fisher Scientific (Waltham, MA, USA). Amplicon library preparation, quantification (LightCycler 480), and sequencing were performed manually and automatically using Ion Chef and Ion GeneStudio S5 Prime systems, following manufacturer protocols (AmpliSeq Library Kit v2.0, Ion 510/520/530 & 540 Kits). DNA concentrations in the WISMUT samples were measured via Qubit (Thermo Fisher Scientific). Genotyping was conducted with as little as 10 ng of DNA. DNA samples were mixed with AmpliSeq HiFi Mix and the primer pool. PCR amplification was carried out in a thermocycler (Eppendorf, Hamburg, Germany) under the following conditions: initial denaturation at 99°C for 2 min, followed by cycles of 99°C for 15 s and 60°C for 4 min. Following amplification, primer sequences were partially digested using FuPa reagent, and barcoded sequencing adapters (Ion Xpress Barcode Adapters, Thermo Fisher Scientific) were ligated to the amplicons. Libraries were subsequently purified with AMPure XP magnetic beads (Beckman Coulter, Brea, CA, USA) and quantified by qPCR on a LightCycler 480 Instrument (Roche, Basel, Switzerland) using the Ion Library TaqMan Quantitation Kit (Thermo Fisher Scientific). Quantified libraries were pooled and subjected to clonal amplification and enrichment on Ion Sphere particles using the Ion Chef system. Sequencing was then performed on an Ion GeneStudio S5 Prime instrument using Ion 530 or 540 chips. The feasibility of typing was tested in advance using another 35 pilot cases from the archive.

Initial genotyping of WISMUT samples showed poor DNA quality. To prevent amplification of fixation artefacts, PCR cycles were reduced from 25–26 to 22, and DNA amount was tripled if possible. Although more clear results were obtained, most samples remained insufficiently amplifiable. Thus, PCR cycles were increased back to 26, and genotyping repeated were possible. Therefore, NGS was performed with three different settings. A bioinformatics analysis (sequence generation, raw data processing, quality control, sample assignment, and file creation) was performed using Torrent Suite Software 5.16. Variant calling against the hg19 human genome was done with the variantCaller plugin.

A single genotype was coded by NGS with values between 0 and 100, where 0 stand for no copy of the reference allele (homozygous), 50 for 1 copy (heterozygous) and 100 for 2 copies (homozygous).

Translation of the genotype doses into genotype probabilities

Before testing the G×E interaction statistically, genotype doses were converted to genotype probabilities using the flexible parametric FLDSf function, transforming doses 0–50 and 100–50 to heterozygosity probabilities (0–1) under symmetric assumptions (19). This function was integrated into the G-model describing genotype distributions [g_0 g_1 g_2] among “cases” (details given in the Additional file 1). Parameters were estimated by maximizing likelihood under HWE, allowing deviations due to genetic relative risk for lung cancer (RR, by marker, case-only design) and excess heterozygosity (accounting for hidden relatedness, population structure, or genotyping errors). This was applied per NGS setting and marker. Of 87 theoretical clusters (29 markers x 3 NGS-settings), only 12 showed strong similarity (nearest centroid sorting via PROC FASTCLUS, SAS©).

Phenotyping

For each WISMUT miner, the BfS holds detailed occupational history data, including employment periods, work locations (sites and shafts), job tasks, work areas (underground mining, processing mills, open-cast mining, surface work), and absence periods. Individual radiation exposure (E) were estimated by others retrospectively, linking work tasks, places and durations with annual values for radon and decay products (working level months, WLM), external gamma radiation (mSv), and long-lived radionuclides (kBq·h/m³; LLR) in a Job Exposure Matrix (JEM) (20, 21). A second JEM covered exposure to fine dust, silica dust, and arsenic, expressed in dust years—defined as 1 mg/m³ (fine dust/silica dust) or 1 µg/m³ (arsenic) over 220 shifts of 8 hours each.

Statistical testing for genetic relative risk

As G-Model we considered the distribution of genotypes $[\, g_0 \quad g_1 \quad g_2 \,]$ among "cases", taking into account the minor allele frequency f (MAF), the genetic relative risk (RR, risk change per rare allele), with F allowing a deviation from g1 (excess heterozygosity). The distribution of genotypes in the source population is assumed to follow the HWE. Therefore, this is an HWE distortion test. Details are explained in the Additional file 1.

The genetic relative risk for lung cancer (RR) was estimated, independently of any occupational exposure, by mapping the posterior distribution of parameters of the G-model (likelihood x prior distributions of the parameters) within the Bayesian MCMC framework. The mean point estimate and the *95%-Highest Posterior Density Intervals* (95%-HDI) were derived from the a-posterior distributions. A parameter was labelled as significant in this report if $2x\ (1 - p_{direct}\ )$ ("probability of direction") was lower than 0.05 (22).

Statistical testing for gene-environment interaction

For testing gene-environment interaction (GxE), the parameter RR of the HWE distortion test was considered linearly dependent on occupational exposure (now labelled as GxE-model):

$$RR = e^{\alpha_{RR} + \beta_{RR} * E^{\tau}}$$

Exponentiating the linear relationship ensures that the RR does not take on negative values. The parameter τ enables a data-driven scale transformation of the exposure during the model estimation. If the parameter β_{RR} deviates from zero, an interaction between genotype and exposure can be inferred. The genetic RR is then modified by the exposure of interest. $e^{\alpha_{RR}}$ is equal to genetic relative risk RR from the G-Model in the case of no GxE-interaction ($\beta_{RR} = 0$). Otherwise $e^{\alpha_{RR}}$ is equal to genetic relative risk RR of unexposed individuals ($E = 0$; given the linearity assumption is correct).

Results

Description of the case series

All lung cancer biospecimens examined here were derived from formalin-fixed, paraffin-embedded (FFPE) lung tissue material from the pathological archive of the former Institute of Pathology of the Health Service of WISMUT SDAG, (23, 24) the world's fourth largest producer of uranium between 1946 and 1990. Biospecimens and phenotypes are stored today in the "German Uranium Miners' Biobank" (GUMB) (12).

The sample analysed consisted of 611 male miners born between 1899 and 1951 (see Table 2). This case series is neither representative nor related to the cohort of WISMUT miners. It includes miners with available biospecimens and data, and is only a small subgroup of all lung cancer cases which occurred among former WISMUT miners. The age at death of the 611 miners was between 37 and 88 years, on average 66 years (median 65 years). About one in three (36%) had also been diagnosed with silicosis. Adenocarcinomas (LUAD, 38%), small cell carcinomas (SCLC, 31%) and squamous cell carcinomas (SqCLC, 28%) were the most common subtypes. The proportions of LUAD and SqCLC correspond to the proportions among all lung cancer cases in Western Europe in 2020 (39% and 31%, respectively), while the proportion of SCLC is higher than in Western Europe (15%). Only one WISMUT miner (< 1%) was diagnosed with large cell carcinoma (LCLC; Western Europe: 7%) (25).

Table 2
Description of the case series of former WISMUT miners (1)

Category		N	%
All		611	100%
Smoking [y/n]	never smoker	19	3%
	ever smoker	128	21%
	unknown	464	76%
Silicosis [y/n]	no	392	64%
	yes	219	36%
Histology	LUAD	231	38%
	SCLC	189	31%
	SqCLC	170	28%
	others	19	3%
	LCLC	1	< 1%
	unknown	1	< 1%
ever smoker: miners who have reported smoking at least once during their period of employment;			
LUAD Adenocarcinomas, SCLC Small cell carcinoma, SqCLC Squamous cell carcinoma, LCLC large cell carcinomas			

Cumulative exposure to radon and its decay products, measured in Working-Level-Months (WLM), is considered as the primary exposure. The strength of exposure ranged between 0 and almost 3000 WLM, with a mean value of 727 WLM. However, the WLMs followed a bi-modal left-skewed distribution. One peak is observed at a very low level of exposure (< 50), another peak was observed at about 800 WLM. Around 25% of miners have accumulated more than 1000 WLM during their employment. This distribution was found independent of the age at death (see Table 3 and Fig. 1). The distribution of other occupational exposures is described in Table 3. These moderately to strongly correlate with each other: weakest - arsenic vs. silica dust (ρ = 0.47) and strongest - LLR vs. gamma radiation (ρ = 0.95). Age at death and year of birth are generally uncorrelated with these mining-specific phenotypes.

Table 3
Description of the occupational exposure

cumulative exposure to …	N	Minimum	25% Percentile	50% Percentile	Mean	75% Percentile	Maximum	s.d.
radon and its decay products [WLM]	611	0	177.8	663.1	727.2	1083.2	2947.4	590.4
LLR [kBq·h/m³]	611	0	1.3	4.4	8.8	12.2	81.4	10.9
gamma radiation[mSv]	611	0	8.3	32.7	71.9	96.7	576	92
fine dust [mg/m³·Jahre]	611	2.4	46.9	82.8	86.1	118.4	238.8	48.8
silica-dust [mg/m³·years]	611	0.4	8.2	15.1	15.8	22.1	51.4	9.2
arsenic [µg/m³·years]	596	0	0	31.9	95.1	126	1132.9	149.7
WLM Working-Level-Months, with weighting, s.d. standard deviation; LLR long-lived radionuclides								

No genotypes could be derived for n = 85 DNA samples from WISMUT miners (14% of 611). Therefore, data from 526 WISMUT miners were included in the statistical modelling.

G model: genetic relative risk

The genetic relative risks RR (risk change per rare allele), estimated by fitting the G-model, range from 0.80 (rs1455777-*ST8SIA2*) to 1.38 (rs10762183-*SIRT1*) (see Fig. 2). A significant increase in lung cancer risk per minor allele was observed for 9 of the 29 markers analysed, and a significant reduction was observed for one marker ($p < 0.05$). Three further markers with similarly high RRs (from the *TP53*, *SIRT1* and *FTO* genes) fell just short of the significance threshold (see S-Table 4).

The strongest genetic risk increase was observed for the marker rs10762183 (*SIRT1*) with a genetic RR = 1.38 (95%-HDI 1.17–1.61, $p < 0.0001$), the second strongest genetic risk increase per rare allele for the marker rs1353793 (*SOX5*) with an RR = 1.34 (95%-HDI 1.14–1.54, $p < 0.0001$). A comparable genetic risk increase of RR = 1.35 (95%-HDI 1.10–1.54, $p < 0.0073$) was observed for the marker rs73596224 (*CUBN*). Other markers with significant RRs can be assigned to the genes *CDKN2A*, *CUBN*, *FTO*, *NAV2*, *ST8SIA2* and *UBE2U*.

RR: genetic relative risk per rare allele

GxE model: interaction for radon exposure [in WLM]

A significant modification of the genetic RR by radon exposure [in WLM] was observed for seven of the 29 markers from 6 of the 10 genes analysed (see Fig. 3). The RR increases almost linearly by WLM. For six of these markers a significant genetic relative risk (RR) was found (see above). The 7th marker is rs7221056 (*TP53*). In the model the scale parameter τ was estimated as almost 1 for all markers, the exposure was not rescaled in estimating RR.

For all markers, a genetic RR of almost 1 was estimated at minimal exposure WLM = 0 (see Fig. 3). This means that the risk of lung cancer without being exposed does not depend on the genes investigated. The risk modification therefore appears to result exclusively from the interaction between genes and as modelled here, radon exposure (Fig. 4). This clear picture of "no genetic risk without occupational exposure" did not apply to the other exposures considered.

The two markers with the strongest positive interaction ($\beta_{RR} > 0$) are rs4756999 (*NAV2*, p=0.0002) and rs10762183 (*SIRT1*, p = 0.0014). The genetic RRs derived from α_{RR}, β_{RR} and τ at low (178 WLM), median (633 WLM) and high (1083 WLM; 75% percentile) exposure levels are about 1.05, 1.21 and 1.36, respectively (see Table 4). We also observed interaction with WLM for

the marker rs7545208 (*UBE2U*, p = 0.0187), but slightly less pronounced (RRs at low, median and high exposure: 1.04, 1.13 and 1.22).

Table 4
Estimated modification of RR by radon exposure [in WLM] (extract)

			0 WLM		178 WLM		663 WLM		1083 WLM	
	$\varvec{\beta}_{\varvec{R}}$	p-value	RR	95%-HDI	RR	95%-HDI	RR	95%-HDI	RR	95%-HDI
rs4756999 - *NAV2*	0.00030	0.0002	1.01	0.99–1.03	1.05	1.02–1.09	1.21	1.08–1.35	1.36	1.12–1.63
rs1455777 - *ST8SIA2*	-0.00018	0.0003	1.00	1.00-1.01	0.97	0.96–0.99	0.89	0.86–0.93	0.83	0.78–0.89
rs4777965 - *ST8SIA2*	0.00021	0.0046	1.01	1.00-1.04	1.06	1.02–1.10	1.19	1.06–1.34	1.32	1.09–1.57
rs10762183 - *SIRT1*	0.00024	0.0014	1.02	0.99–1.05	1.07	1.02–1.12	1.24	1.10–1.38	1.40	1.15–1.67
rs7545208 - *UBE2U*	0.00015	0.0187	1.01	0.99–1.03	1.04	1.01–1.08	1.13	1.01–1.26	1.22	1.00-1.45
rs7526950 - *UBE2U*	0.00008	0.3849	1.01	0.99–1.03	1.03	0.98–1.07	1.07	0.94–1.19	1.11	0.90–1.32
rs1353793 - *SOX5*	0.00017	0.0813	1.02	1.00-1.04	1.05	1.01–1.09	1.16	1.02–1.29	1.26	1.03–1.52
rs73596224 - *CUBN*	0.00017	0.2234	1.00	0.97–1.03	1.04	0.98–1.09	1.14	0.94–1.34	1.24	0.88–1.60
rs8061518 - *FTO*	0.00003	0.7404	1.02	1.00-1.04	1.03	1.00-1.07	1.07	0.95–1.20	1.11	0.91–1.33
RR genetic relative risk estimated from the GxE model; 95%-HDI 95% Highest Posterior Density Interval;										
178, 663 and 1083 WLM are the 25%, 50% and 75% percentile of the exposure in the collective;										
the table is sorted by $p_{directly}$ in ascending order, but markers of the same gene are kept together.										

Two markers with significant interaction effects (rs1455777 and rs4777965) belong to the gene *ST8SIA2*. While rs4777965 shows an interaction comparable to the above markers, β_{RR} is negative for rs1455777. This means that the risk of disease under radon exposure decreases with the number of rare alleles. The derived genetic RRs at low, median and high exposure are estimated to be 0.97 (= 1/1.03), 0.89 (= 1/1.12) and 0.83 (= 1/1.20).

For 4 markers with significant genetic RRs (rs7526950- *UBE2U*, rs73596224-*CUBN*, rs1353793-*SOX5* and rs8061518-*FTO*) no interaction with radon exposure (in WLM) could be shown to be significant.

RR genetic relative risk; the distribution of radiation exposures in WLM is shown in pale blue in the background. The minimum and the 25%, 50% and 75% percentiles are highlighted on the x-axis. RR = 1 all genotypes have the same genetic RR.

Horizontal curve gradients would indicate that the relative genetic risk does NOT change with the degree of exposure and therefore there is no GxE-interaction. In a pure GxE-interaction, the genetic relative risk should be 1 at zero exposure (starting point of the line).

GxE model: interaction for other exposures

Other than for radon exposure [in WLM], only four (out of the 29x5 = 145; 3%) investigated marker-exposure interactions were statistically significant (p < 5%). Two of these significant GxE interactions were observed in relation to arsenic, one each in relation to LLR or gamma radiation, but none in relation to fine dust or silica dust.

The scale parameter τ was estimated as almost 1 for radon and arsenic exposure for most markers. With regard to LLR radiation, τ ranged from 0.3 to 0.6, and with regard to gamma radiation, from 0.70 to 0.98, and with regard to fine dust or silica dust, it averaged 0.40 and 0.45, respectively. This underlines the importance of scale-adapted modelling for each exposure.

As can be seen in Fig. 4 (first line left panel) and Table 5, a GxE interaction of marker rs1353793 (*SOX5*) can be observed for radon almost (p = 0.0813) and for arsenic statistically significant (p = 0.0231), but not with any of the other exposure variables (horizontal lines indicate no GxE interaction). This, together with the correlation between the two exposure variables (r = 0.48), makes it difficult to regard either of them as the cause of the observed interaction. No significant interaction was observed for the other two markers of the *SOX5* gene. The situation, however, is similar with regard to the marker rs4777965 and rs1455777 (*ST8SIA2*; Fig. 4: first line right, second line left panel), whose interaction with arsenic (p = 0.0144; 0.0551) and with radon (p = 0.0046; p = 0.0003) was, statistically speaking, below or on the threshold of significance (Table 5).

Table 5
Significance of GxE interaction by exposure

Marker	Gene	radon and its decay products [WLM]	LLR [kBq·h/m³]	gamma radiation [mSv]	arsenic [µg/m³·years]	particulate matter [mg/m³·Jahre]	silica-dust [mg/m³·years]
rs7545208	*UBE2U*	**0.0187**	0.2697	0.4950	0.7034	0.6328	0.5785
rs7526950	*UBE2U*	0.3849	0.6903	0.4132	0.2284	0.8736	0.3191
rs10889478	*UBE2U*	0.6923	0.6918	0.3209	0.8911	0.5588	0.5625
rs1908635	*UBE2U*	0.6855	0.4752	0.8365	0.8360	0.3407	0.8203
rs7023680	*CDKN2A*	0.1515	0.0980	0.7233	0.4868	0.9287	0.6871
rs3731217	*CDKN2A*	0.3578	0.4368	0.3262	0.7665	0.2441	0.2320
rs10965227	*CDKN2A*	**0.0252**	0.4182	0.3781	0.0897	0.5800	0.3674
rs73596224	*CUBN*	0.2234	0.2804	0.2838	0.7519	0.5414	0.6404
rs60918373	*CUBN*	0.4681	0.8135	0.7697	0.6709	0.3753	0.3769
rs7095324	*CUBN*	0.4638	0.6074	0.7711	0.8652	0.4019	0.4736
rs10823112	*SIRT1*	0.1244	0.9675	0.6893	0.2862	0.1171	0.2268
rs10762183	*SIRT1*	**0.0014**	0.1031	0.2757	0.1455	0.2896	0.0955
rs7129281	*NAV2*	0.9352	0.2789	0.7028	0.8333	0.5258	0.1691
rs11025138	*NAV2*	0.3065	0.2973	0.1198	0.4753	0.3706	0.6067
rs4756999	*NAV2*	**0.0002**	0.1085	0.0712	0.8659	0.3298	0.4740
rs7137426	*CD163L1*	0.6735	0.1914	0.2736	0.8369	0.8089	0.9262
rs7970379	*CD163L1*	0.7757	0.6644	0.4348	0.8902	0.8329	0.3626
rs7302538	*CD163L1*/ACSM4	0.5323	0.4243	0.5861	0.3549	0.5892	0.6776
rs1353793	*SOX5*	0.0813	0.6878	0.4794	**0.0231**	0.8948	0.7339
rs7978583	*SOX5*	0.1811	0.6553	0.3021	0.7425	0.7732	0.6004
rs4777965	*ST8SIA2*	**0.0046**	0.2791	0.3009	**0.0144**	0.4123	0.2811
rs1455777	*ST8SIA2*	**0.0003**	0.2940	0.3742	0.0551	0.4457	0.0618
rs3784722	*ST8SIA2*	0.3993	**0.0191**	**0.0005**	0.0850	0.2331	0.1816
rs8061518	*FTO*	0.7404	0.3895	0.4751	0.3596	0.6355	0.3778
rs6499657	*FTO*	0.7741	0.6435	0.6587	0.3145	0.9741	0.9904
rs1111483	*FTO*	0.7604	0.6355	0.3752	0.4556	0.6169	0.6915
rs17886268	*TP53*	0.9532	0.5852	0.8059	0.8986	0.8802	0.4611
rs7141	*TP53*	0.1025	0.1052	0.0723	0.0907	0.2683	0.6897
rs7221056	*TP53*	**0.0303**	0.1427	0.2613	0.2761	0.3669	0.3049

[†] for rs7545208 a p-value of 0.0187 was observed considering WLM, when fixing τ at 1; LLR long-lived radionuclides

Although only significant for radon (p = 0.0002), a GxE interaction of the marker rs4756999 (*NAV2*; Fig. 4: 3rd line left panel) with other exposure variables (LLR and gamma radiation, fine dust, silica dust and arsenic) cannot be ruled out. The same appears for the markers rs10762183 (*SIRT1*), rs7221056 (*TP53*), and rs10965227 (*CDKN2A*; Fig. 4: 3rd line right and 4th line both panels) with a significant interaction with radon (p_{SIRT1}=0.0014, p_{TP53}=0.0303, p_{CDKN2A}=0.0252), however, no interaction to arsenic (all p > 0.1). Solely marker rs7545208 (*UBE2U*; Fig. 4: 2nd line right panel) shows no GxE interaction other than with radon (p = 0.0187; when τ is set to 1).

A relative genetic risk of RR = 1 all genotypes have the same genetic RR. Horizontal curve gradients would indicate that the relative genetic risk RR does NOT change with the degree of exposure and therefore there is no GxE interaction. In a pure GxE interaction, the genetic relative risk should be 1 at zero exposure (starting point of the line). The distribution of radon exposures in WLM is shown in pale blue in the background. For visual comparison, the exposures were scaled so that the quartiles are positioned identically. The minimum and the 25%, 50% and 75% percentiles are highlighted on the x-axis. The curves for each exposure were scaled so that the percentiles correspond graphically.

Discussion

Our aim was to validate/replicate previously reported radon-gene interactions in relation to lung cancer risk in an independent series of 611 DNA samples from uranium miners, all former employees of WISMUT AG in Germany. In reviewing the literature, we extracted 34 genes modifying the Radon-induced lung cancer risk.

The biological samples for studying genetic interactions with occupational radon exposure had been archived for a long time and isolated DNA from normal formalin-fixed paraffin-embedded lung tissue samples were therefore highly fragmented due to specific fixation procedures at that time. Despite the poor quality of the DNA, after optimization of procedures we were able to genotype 29 markers from 10 of the 34 literature-reported genes (*FTO*, *UBE2U*, *CUBN*, *ST8SIA2*, *SOX5*, *TP53*, *CD163L1*, *CDKN2A*, *NAV2*, and *SIRT1*).

Genotypes could be determined for the first time in this material by NGS in 526 of the 611 samples. However, the reliable translation of NGS results (genotype doses) into genotypes (or genotype probabilities) has proven to be a methodological challenge. Therefore, we first developed a flexible mathematical function for the isotonic conversion of doses into probabilities (FLDSf), specifically for this purpose (19). Subsequently, the genotype probabilities for each marker were estimated based on NGS reads, taking into account various NGS settings. We found that it was necessary to apply 12 different translations to the NGS reads to reliably fit statistical models to the uncertainties of the genetic data. When integrated into statistical models, we were able to detect significant modification in genetic RR due to radiation exposure [in WLM] for seven of the 29 markers from six of the ten genes examined.

Statistical significance, even in a replication study like this, should not be the only justification in postulating the existence of GxE interaction. The evidence for significant interactions between different genes and environmental exposures derived from the data is only corroborated by proven biological functional relationships within carcinogenesis or any known associations with lung cancer risk.

Markers in three genes show a clear dose-dependent modulating effect only in interaction with radon, *UBE2U*, an enzyme responsible for ubiquitination of proteins, *SIRT1*, a central epigenetic regulator and a genetic marker located in the *CDKN2A/CDKN2B/CDKN2B-AS1* gene cluster.

Genetic markers in the gene *UBE2U*, that modulate the risk for radon-induced lung cancer, were first discovered in our preceding genome-wide interaction study in Uranium miners embedded in a broader case control study of 15,077 cases and 13,522 controls of European ancestries (8). In the present study, the significant interaction with radon could be confirmed for the marker rs7545208, one of three investigated markers assigned to *UBE2U*. Ubiquitination affects proteins and important pathways in many ways. Pairing of E2-E3 enzymes regulating ubiquitination have been implicated in cancer initiation and progression or cancer suppression when cancer suppressors are degraded (26). *UBE2U* has been shown to have a central regulatory function to regulate tumor suppressor proteins involved in DNA double-strand break repair, cell cycle progression and checkpoint control.

The *UBE2U* gene is known to influence the response to DNA damage and other important proteins in double-strand break repair, e.g., via regulation of 53BP1. *UBE2U* interacts with the E3 ligase RNF17 to enforce DNA damage responses by promoting the accumulation of 53BP1 at ionizing radiation-induced double-strand breaks (DSBs) (27). Thus, mutations in the *UBE2U* gene may tip the balance toward genomic instability and oncogenic progression. Further studies are needed to elucidate the exact mechanism in lung cancer.

SIRT1 was the first gene identified predisposing for squamous cell carcinoma in the lung in interaction with radon in a cohort of former uranium miners from the Colorado Plateau (10). A significant association between *SIRT1* gene polymorphisms and human NSCLC was neither found in a Chinese study without radon exposure, nor in a GWAS on lung cancer or other lung diseases (28). However, the Chinese study was rather small and an increased OR of 1.7 was observed but did not reach significance. In our study, the interaction observed for the marker rs10762183 (*SIRT1*) is also likely to be attributable to the radiation component, confirming data from the US study. As a hub gene in cancer-related pathways, *SIRT1* plays a central role as an epigenetic regulator in a variety of cellular signalling mechanisms and triggers cell defence and cell survival in the event of radiation-induced damage (29, 30). Nonetheless, many aspects of the *SIRT1*–cancer relationship are ambiguous. However, *SIRT1* contributes to DNA repair after irradiation (10) as component of double-strand break (DSB) repair through homologous recombination (HR), non-homologous end joining (NHEJ), DNA damage response (DDR) and nucleotide excision repair (NER) (31). Although it is not yet clear whether *SIRT1* acts as a tumour promoter or tumour suppressor in relation to cancer in general (32), radon is assumed to be the interacting exposure.

A gene cluster at chromosome 9p21 encodes three tumour suppressor genes - *CDKN2A*, CDKN2B and the antisense non-coding RNA CDKN2B-AS1 (also known as ANRIL), and further MTAP, a gene encoding an enzyme for the salvage of both adenine and methionine. They are frequently mutated and co-deleted in a wide variety of tumours, including lung cancer (33). Two different markers located in CDKN2B-AS1 were linked to lung cancer in individuals of European decent. The SNP rs1333040 showed, amongst others, a genome-wide significant association with SqCLC (34). rs62560775 showed a suggestive significant association with LUAD (35). In addition, the SNP rs885518, located in MTAP, was found in the first node of a decision tree of genomic markers to identify lung cancer cases in never smokers (36). *CDKN2A* and – 2B are inhibitors of CDK4 kinase. Both genes are considered tumour suppressor genes with a central role in cell cycle regulation, including in lung cancer (35). The gene product of CDKN2B-AS1 is a functional RNA molecule and is, e.g., involved in the regulation of *CDKN2A* and – 2B (33). *CDKN2A* was recently labelled as hub gene for LUAD (37). With regard to occupational radiation exposure, epigenetic dysregulation of *CDKN2A* was observed in nuclear power plant workers in Mayak, Russia, who were exposed to alpha radiation (11). *CDKN2A* gene expression is regulated by methylation, and methylation rates have been reported to be affected by radon exposure in Chinese uranium miners (38). The GxE interaction observed for the marker rs10965227, located close to rs1333040 within CDKN2B-AS1, was most significant for radon, but its magnitude is comparable with that of other exposures. However, there is only little to moderate LD between the four mentioned markers (D': 0.003–0.742, European populations (39)). We therefore see strong evidence that this gene complex is a complex predictor of lung cancer, partly due to interactions with radon.

Gene-environment interaction due to arsenic occupational exposure

ST8SIA2 was first identified to influence lung cancer risk in interaction with radon in our previous GWAS study. *ST8SIA2* expression is linked to NSCLC recurrence risk and SCLC prognosis (40, 41), and its genomic markers associate with respiratory impairment (airflow obstruction (42)) and smoking cessation in Asians (43). Here, we observed an interaction not only due to radon exposure but also with other occupational exposure variables of the miners. Based on the pathway described below, we believe that the interaction with rs4777965 should be attributed to the risk of lung cancer from arsenic exposure rather than radiation exposure. Arsenic is classified as a Class 1 human carcinogen by the International Agency for Research on Cancer with strong evidence for lung cancer (44, 45). The observed interaction is in our study supported by recent functional data of *ST8SIA2*. *ST8SIA2* is a polysialyltransferase that synthesizes polysialic acid (PSA) on substrates like the neural cell adhesion molecule NCAM, enhancing tumour cell migration. It may promote tumorigenesis by shielding tumour cells from immune detection via a protective sialic acid layer. It has recently been discovered that arsenite increases the sialic acid content on the cell surface by inhibiting sialidase activity (46).

The same applies to *SOX5*, a transcription factor of the high mobility group box (SOX) family. It has been more strongly associated to arsenic exposure rather than radon exposure. We observed a G×E interaction with arsenic exposure at marker rs1353793 of *SOX5*. A link between arsenic exposure and tumour induction has not yet been described, although *SOX5* has been implicated in accelerating migration and invasion via epithelial-mesenchymal transition. *SOX5* has been described as an oncogene in many types of cancers, including lung cancer (47). However, four other markers of *SOX5* (rs1391789, rs12367235, rs1259887, rs1800973) have been associated with C-reactive protein (CRP) levels and are partly in strong LD with rs1353793 (48). CRP is considered a non-specific biomarker of acute inflammation. Arsenic, on the other hand, can alter proliferation status and expression of key pro-inflammatory cytokines (e.g. TNF-α, IL-1, IL-6). It has been proven that exposure to arsenic can trigger an inflammatory response and that lungs are particularly susceptible to arsenic (49–51). In summary, the observed interaction of *SOX5* markers with arsenic could be caused by genomic linkage to CRP (52).

Further observed GxE interactions should be viewed critically, although their significance has been reiterated.

We found *NAV2*, a member of the neuron navigator family, for the first time as a possible susceptibility gene modulating the radon-induced lung cancer risk. Our study also confirmed this previously found interaction. However, it was observed that the interaction was comparable in strength to the other exposures considered. *NAV2* is considered as a pro-tumorigenic factor for colorectal cancer. The Human Protein Atlas (16) identifies *NAV2* as a "multiplayer" due to a high "protein expression score" in many tissues/organs (especially the heart and lungs). Yet, little is known about possible functional pathways of members of the neuron navigator gene family. They may play a role in cell migration and tumour invasion (53, 54). The functional evidence to support the observed GxE interaction needs clearly to be further elaborated to strengthen the putative *NAV2* radon-induced lung cancer risk.

TP53 is a central hub gene in cancer progression. The coded protein acts as multifunctional tumour suppressor. This labelling of *TP53* seems obvious, as mutations in the *TP53* gene are the most common somatic changes in human cancer tissue that occur during carcinogenesis (55). Germline variants of the *TP53* gene are associated with the Li Fraumeni Syndrome. Humans bearing germline mutations are highly radiosensitive and cancer-prone to multiple sites, including lung cancer. However, rs7141, but not rs7221056, is also considered a weak proxy SNP for the *TP53*Arg72Pro mutation (rs1042522), which has been linked to lung cancer risk among non-Europeans (56). A higher proportion of heterozygotes at this marker was seen in 44 cases with radon-induced LC compared to 41 cases with non-radon-induced LC (57–59).

Apart from the GxE interactions replicated, we proved NGS to be an adequate technique for typing historical, highly fragmented DNA samples. However, due to uncertainties in genotyping results, adequate translation of genotype doses into genotype probabilities was necessary to account for this uncertainty. This was achieved by implementing a sophisticated model before performing data analysis.

Limitations

To estimate and test for GxE interaction using the case-only design employed here, we need to strictly assume independence between genotype G and environmental factor E (60). We consider this assumption to be met, as it is unlikely that a miner's genetic makeup would influence their assignment to a particular workplace (mine, above ground) and therefore exposure group. Distortions of genotype frequency due to admixing of different ethnic groups can be neglected since all 611 miners are of European descent. However, a distortion of the results due to the healthy worker effect or healthy worker survivor bias in the high exposure group cannot be ruled out in this sample of miners (61). We must also assume that the 14% of genotypes (from 85 miners) that could not be determined do not cause any distortion.

Conclusion

For markers from six of the 10 genes examined (*UBE2U*, *CDKN2A*, *SIRT1*, *NAV2*, *ST8SIA2*, *TP53*), a significant modification of the genetic relative risk for lung cancer due to mining-specific radon exposure [in WLM] was replicated/validated. Radon exposure is considered to be the primary cause of the observed GxE interaction for *UBE2U*, *CDKN2A* and *SIRT1*. This conclusion is based on

our findings and is supported by the functional relevance of the genes, as extracted from the literature. For another marker of gene *SOX5*, a GxE interaction can presumably be attributed to exposure to arsenic.

Additional information

Additional file 1.dox

Declarations

Ethics approval and consent to participate

This study is embedded in the German Uranium Miners' Wismut Biobank, which was conducted according to the guidelines of the Declaration of Helsinki and approved by the Bavarian Ethics Committee (Ethik-Kommission der Bayerischen Landesärztekammer No. 08082, 29 October 2008) and the Federal Commissioner for Data Protection and Freedom of Information (IV-508/020#0021,1 September 2008). Informed consent was obtained from all subjects involved in the study.

Consent for publication

Not applicable

Declarations Competing Interests

Mörtl Simone and Gomolka Maria are employees of the funder.

Funding

A.R. and H.B, were funded by the Bundesministerium für Umwelt, Klimaschutz, Naturschutz und nukleare Sicherheit (BMUKN, Germany) within project no. 3620S32271 and B.I and F.F. within project no. 3621S32220. The German Uranium Miners' Biobank was funded by the Bundesministerium für Umwelt, Klimaschutz, Naturschutz und nukleare Sicherheit (BMUKN, Germany) grant No. 3608S04532, 3610S10002, 3614S10015, 3611S10010, 3614S10013 and 3610S10001.

Author Contribution

RA made acquisition of the financial support, conceptualised the project, conducted the data analy-sis and drafted the manuscript. BI carried out genotyping and co-drafted the manuscript. FF made acquisition of the financial support, supervised genotyping, and reviewed the manuscript. HB was responsible data documentation in the Wismut biobank. JG made acquisition of financial support, supervised and coordinated the project to extract DNA and RNA to provide the material to the German Uranium Miners' Biobank BA carried out DNA extraction from the archived material RP carried out DNA extraction from the archived material BH made acquisition of the financial support and reviewed the manuscript. MS reviewed the manuscript. GM was responsible for the German Uranium Miners' Biobank at the BfS, initiated the project, organised biospecimens and data ex-change between the partners, supervised the project as an employed representative of the funder (BfS), and co-drafted the manuscript.

Acknowledgement

Many thanks to Felix Heinzl for valuable discussion on the applied statistical methods and the draft.

Data Availability

The phenotype data supporting the conclusions of this article are only available on request from the BfS server located MEDEORA databank. The genotype data are available on request from the STOREDB. (https://www.storedb.org/store_v3/study.jsp?studyId=1034).

Clinical trial number

not applicable

References

1. Zeeb H, Shannoun F. WHO Handbook on Indoor Radon: A Public Health Perspective [Internet]. Geneva: World Health Organization; 2009. Located at: 23762967. Verfügbar unter:. https://www.who.int/publications/i/item/9789241547673.
2. Cruz-Nova P, Trujillo-Nolasco M, Aranda-Lara L, Ferro-Flores G, Ocampo-García B. Radiobiological effect of alpha particles. The scientific basis of targeted alpha-particle therapy. Nuclear Med Biology 1 Juli. 2025;146–7. 10.1016/j.nucmedbio.2025.109044.
3. Riudavets M, Garcia de Herreros M, Besse B, Mezquita L. Radon and Lung Cancer: Current Trends and Future Perspectives. Cancers Januar. 2022;14(13):13. 10.3390/cancers14133142.
4. Martin-Gisbert L, Ruano-Ravina A, Varela-Lema L, Penabad M, Giraldo-Osorio A, Candal-Pedreira C. u. a. Lung cancer mortality attributable to residential radon: a systematic scoping review. J Expo Sci Environ Epidemiol Mai. 2023;33(3):368–76. 10.1038/s41370-022-00506. -w PubMed PMID: 36577801.
5. Liu Y, Xu Y, Xu W, He Z, Fu C, Du F. Radon and lung cancer: Current status and future prospects. Crit Reviews Oncology/Hematology 1 Juni. 2024;198:104363. 10.1016/j.critrevonc.2024.104363.
6. Kreuzer M, Sommer M, Deffner V, Bertke S, Demers PA, Kelly-Reif K. u. a. Lifetime excess absolute risk for lung cancer due to exposure to radon: results of the pooled uranium miners cohort study PUMA. Radiat Environ Biophys März. 2024;63(1):7–16. 10.1007/s00411-023-01049-. w PubMed PMID: 38172372; PubMed Central PMCID: PMC10920468.
7. Bosse Y, Amos CI. April. A Decade of GWAS Results in Lung Cancer. Cancer Epidemiol Biomarkers Prev. 2018;27(4):363–79. Located at: 28615365. 10.1158/1055-9965.EPI-16-0794
8. Rosenberger A, Hung RJ, Christiani DC, Caporaso NE, Liu G, Bojesen SE et al. Genetic modifiers of radon-induced lung cancer risk: a genome-wide interaction study in former uranium miners. Int Arch Occup Environ Health. 3. Juli 2018. Located at: 29971594. 10.1007/s00420-018-1334-3
9. Ruano-Ravina A, Pereyra MF, Castro MT, Perez-Rios M, Abal-Arca J, Barros-Dios JM. Genetic susceptibility, residential radon, and lung cancer in a radon prone area. J Thorac Oncol August. 2014;9(8):1073–80. 10.1097/JTO.0000000000000205. Located at: 24852519.
10. Leng S, Picchi MA, Liu Y, Thomas CL, Willis DG, Bernauer AM. u. a. Genetic variation in SIRT1 affects susceptibility of lung squamous cell carcinomas in former uranium miners from the Colorado plateau. Carcinog Mai. 2013;34(5):1044–50. 10.1093/carcin/bgt024. Located at: 23354305.
11. Belinsky SA, Klinge DM, Liechty KC, March TH, Kang T, Gilliland FD. u. a. Plutonium targets the p16 gene for inactivation by promoter hypermethylation in human lung adenocarcinoma. Carcinog Juni. 2004;25(6):1063–7. 10.1093/carcin/bgh096. Located at: 14742312.
12. Gomolka M, Bucher M, Duchrow L, Hochstrat B, Taeger D, Johnen G. u. a. The German Uranium Miners' Biobank—A Biobank for OMICs Radiation Research. Radiation März. 2022;2(1):1. 10.3390/radiation2010005.
13. Azimzadeh O, Gomolka M, Birschwilks M, Saigusa S, Grosche B, Moertl S. Advanced Omics and Radiobiological Tissue Archives: The Future in the Past. Appl Sci Januar. 2021;11(23):11108. 10.3390/app112311108.
14. Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C. u. a. The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 8 Januar. 2019;47(D1):D1005–12. 10.1093/nar/gky1120 . PubMed PMID: 30445434; PubMed Central PMCID: PMC6323933.

15. National Center for Biotechnology Information (NCBI)[Internet]. Bethesda (MD). dbGene [Internet]. Verfügbar unter: https://www.ncbi.nlm.nih.gov/gene
16. The Human Protein. Atlas [Internet]. [zitiert 5. Februar 2021]. Verfügbar unter: https://www.proteinatlas.org/
17. Cunningham F, Amode MR, Barrell D, Beal K, Billis K, Brent S. Ensembl 2015. Nucleic acids research. Januar 2015;43(Database issue):D662-9. Located at: 25352552. 10.1093/nar/gku1010
18. Oscanoa J, Sivapalan L, Gadaleta E, Dayem Ullah AZ, Lemoine NR, Chelala C. SNPnexus: a web server for functional annotation of human genome sequence variation (2020 update). Nucleic Acids Res 2 Juli. 2020;48(W1):W185–92. 10.1093/nar/gkaa420.
19. Rosenberger A, FLDS. A Flexible, Limited, Double-Sigmoid Transmission Function. J Probab Stat. 2026;2026(1):6076741. 10.1155/jpas/6076741.
20. Kreuzer M, Grosche B, Dufey F, Schnelzer M, Tschense A, Walsh L. The German Uranium Miners Cohort Study (Wismut cohort), 1946–2003 [Internet]. Oberschleissheim, February 2011: Bundesamt für Strahlenschutz. Verfügbar unter: German Federal Office of Radiation Prodection); 2011. p. 50. http://nbn-resolving.de/urn:nbn:de:0221–201102185211.
21. Dahmann D, Bauer HD, Stoyke G. Retrospective exposure assessment for respirable and inhalable dust, crystalline silica and arsenic in the former German uranium mines of SAG/SDAG Wismut. Int Arch Occup Environ Health 1 August. 2008;81(8):949–58. 10.1007/s00420-007-0287-8.
22. Makowski D, Ben-Shachar MS, Chen SHA, Lüdecke D. Indices of Effect Existence and Significance in the Bayesian Framework. Frontiers in Psychology [Internet]. 2019 [zitiert 1. November 2022];10. Verfügbar unter: https://www.frontiersin.org/articles/10.3389/fpsyg.2019.02767
23. Wiethege T, Wesch H, Wegener K, Müller KM, Mehlhorn J, Spiethoff A. German uranium miner study–pathological and molecular genetic findings. German Uranium Miner Study, Research Group Pathology. Radiat Res Dezember 1999;152(6 Suppl):S52-55. PubMed PMID: 10564937.
24. Taeger D, Johnen G, Wiethege T, Tapio S, Möhner M, Wesch H. u. a. Major histopathological patterns of lung cancer related to arsenic exposure in German uranium miners. Int Arch Occup Environ Health Juli. 2009;82(7):867–75. 10.1007/s00420-008-0386-1 . PubMed PMID: 19020892.
25. Zhang Y, Vaccarella S, Morgan E, Li M, Etxeberria J, Chokunonga E. u. a. Global variations in lung cancer incidence by histological subtype in 2020: a population-based study. Lancet Oncol 1 November. 2023;24(11):1206–18. 10.1016/S1470-2045(23)00444-8.
26. Chang SC, Zhang BX, Ding JL. März. E2-E3 ubiquitin enzyme pairing - partnership in provoking or mitigating cancers. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer. 1. 2022;1877(2):188679. 10.1016/j.bbcan.2022.188679
27. Guo Y, An L, Ng HM, Sy SM, Huen MS. An E2-guided E3 Screen Identifies the RNF17-UBE2U Pair as Regulator of the Radiosensitivity, Immunodeficiency, Dysmorphic Features, and Learning Difficulties (RIDDLE) Syndrome Protein RNF168. J Biol Chem 20 Januar. 2017;292(3):967–78. 10.1074/jbc.M116.758854. Located at: 27903633.
28. Lv Y, Lin S, Peng F. SIRT1 gene polymorphisms and risk of lung cancer. Cancer Manag Res. 2017;9:381–6. doi:10.2147/CMAR.S142677 PubMed PMID: 28919817; PubMed Central PMCID: PMC5590763.
29. Qin H, Zhang H, Zhang S, Zhu S, Wang H. Protective Effect of Sirt1 against Radiation-Induced Damage. Radiat Res 1 Dezember. 2021;196(6):647–57. 10.1667/RADE-20-00139.1 . PubMed PMID: 34459925.
30. Hashemi M, Zali N, Ghafarzadeh Dastjerdi SZ, Pakshad B, Aliahmadi M, Sharifi N. u. a. SIRT1 as a masterful reciprocal regulator of molecular mechanisms and signaling pathways involved in tumor growth and expansion. Clin Exp Med. 2025;25(1):225. 10.1007/s10238-025-01759-z . PubMed PMID: 40591009; PubMed Central PMCID: PMC12213863.
31. Mei Z, Zhang X, Yi J, Huang J, He J, Tao Y. Sirtuins in metabolism, DNA repair and cancer. J Experimental Clin Cancer Res 5 Dezember. 2016;35(1):182. 10.1186/s13046-016-0461-5.
32. Deng CX. SIRT1, is it a tumor promoter or tumor suppressor? Int J Biol Sci [Internet]. 2009;5(2):147–52. Located at: 19173036. Verfügbar unter:. https://www.ncbi.nlm.nih.gov/pubmed/19173036.
33. Cunnington MS, Koref MS, Mayosi BM, Burn J, Keavney B. Chromosome 9p21 SNPs Associated with Multiple Disease Phenotypes Correlate with ANRIL Expression. PLOS Genet 8 April. 2010;6(4):e1000899. 10.1371/journal.pgen.1000899.

34. Timofeeva MN, Hung RJ, Rafnar T, Christiani DC, Field JK, Bickeboller H. November, Influence of common genetic variation on lung cancer risk: meta-analysis of 14 900 cases and 29 485 controls. Human molecular genetics. 2012/08/18 Aufl. 15. 2012;21(22):4980–95. Located at: 22899653. 10.1093/hmg/dds334
35. Fehringer G, Kraft P, Pharoah PD, Eeles RA, Chatterjee N, Schumacher FR. u. a. Cross-Cancer Genome-Wide Analysis of Lung, Ovary, Breast, Prostate, and Colorectal Cancer Reveals Novel Pleiotropic Associations. Cancer Res 1 September. 2016;76(17):5103–14. 10.1158/0008-5472.CAN-15-2980 . PubMed PMID: 27197191; PubMed Central PMCID: PMC5010493.
36. Rosenberger A, Muttray N, Hung RJ, Christiani DC, Caporaso NE, Liu G. u. a. Gene–gene interaction of AhRwith and within the Wntcascade affects susceptibility to lung cancer. Eur J Med Res 31 Januar. 2022;27(1):14. 10.1186/s40001-022-00638-7.
37. Zeng C, Zhou Y, Ye W, Fang Z, Wang K. Exploration and validation of hub genes in lung adenocarcinoma based on bioinformatics analysis. Transl Cancer Res Oktober. 2022;11(10):3814–26. 10.21037/tcr-22-2225 . PubMed PMID: 36388051; PubMed Central PMCID: PMC9641136.
38. Su S, Jin Y, Zhang W, Yang L, Shen Y, Cao Y. u. a. Aberrant Promoter Methylation of p16-INK4a and O-6. J Occup Health. 2006;48(4):261–6. 10.1539/joh.48.261.
39. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. Bioinf 1 November. 2015;31(21):3555–7. 10.1093/bioinformatics/btv402 . PubMed PMID: 26139635; PubMed Central PMCID: PMC4626747.
40. Gong L, Zhou X, Yang J, Jiang Y, Yang H. Effects of the regulation of polysialyltransferase ST8SiaII on the invasiveness and metastasis of small cell lung cancer cells. Oncol Rep Januar. 2017;37(1):131–8. 10.3892/or.2016.5279 . PubMed PMID: 28004110.
41. Hao J, Zeltz C, Pintilie M, Li Q, Sakashita S, Wang T. u. a. Characterization of Distinct Populations of Carcinoma-Associated Fibroblasts from Non-Small Cell Lung Carcinoma Reveals a Role for ST8SIA2 in Cancer Cell Invasion. Neoplasia. 2019;21(5):482–93. .009 PubMed PMID: 30978569; PubMed Central PMCID: PMC6458340.
42. Wilk JB, Shrine NRG, Loehr LR, Zhao JH, Manichaikul A, Lopez LM. u. a. Genome-Wide Association Studies Identify CHRNA5/3 and HTR4 in the Development of Airflow Obstruction. Am J Respir Crit Care Med 1 Oktober. 2012;186(7):622–32. 10.1164/rccm.201202-0366OC . PubMed PMID: 22837378; PubMed Central PMCID: PMC3480517.
43. Argos M, Tong L, Pierce BL, Rakibuz-Zaman M, Ahmed A, Islam T. u. a. Genome-wide association study of smoking behaviors among Bangladeshi adults. J Med Genet Mai. 2014;51(5):327–33. 10.1136/jmedgenet-2013-102151 . PubMed PMID: 24665060; PubMed Central PMCID: PMC4126189.
44. Speer RM, Zhou X, Volk LB, Liu KJ, Hudson LG. Arsenic and cancer: Evidence and mechanisms. Adv Pharmacol. 2023;96:151–202. 10.1016/bs.apha. 2022.08.001 PubMed PMID: 36858772; PubMed Central PMCID: PMC10860672.
45. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Some drinking-water disinfectants and contaminants, including arsenic. IARC Monogr Eval Carcinog Risks Hum. 2004;84:1–477. PubMed PMID: 15645577; PubMed Central PMCID: PMC7682301.
46. Taguchi H, Fujishiro H, Sumi D. Arsenite increases sialic acid levels on the cellular surface through the inhibition of sialidase activity. Biochem Biophys Res Commun 20 Dezember. 2024;739:150973. 10.1016/j.bbrc.2024.150973 . PubMed PMID: 39541927.
47. Xue Jdi, Xiang W fang, Cai Mqin. Lv X yun. Biological functions and therapeutic potential of SRY related high mobility group box 5 in human cancer. Front Oncol. 21. Mai 2024;14. 10.3389/fonc.2024.1332148
48. The NHGRI-EBI. Catalog of human genome-wide association studies [publicly accessible database] [Internet]. 2020 [zitiert 22. Dezember 2020]. GWAS catalog. Verfügbar unter: https://www.ebi.ac.uk/gwas/
49. Celik I, Gallicchio L, Boyd K, Lam TK, Matanoski G, Tao X. u. a. Arsenic in drinking water and lung cancer: a systematic review. Environ Res September. 2008;108(1):48–55. 10.1016/j.envres.2008.04.001. Located at:.
50. Tutkun L, Gunduzoz M, Turksoy VA, Deniz S, Oztan O, Cetintepe SP. u. a. Arsenic-induced inflammation in workers. Mol Biol Rep 1 April. 2019;46(2):2371–8. 10.1007/s11033-019-04694-x.

51. Zhang Z, Pi R, Luo J, Liu J, Zhang A, Sun B. Association between arsenic exposure and inflammatory cytokines and C-reaction protein: A systematic review and meta-analysis. Med (Baltimore) 16 Dezember. 2022;101(50):e32352. 10.1097/MD.0000000000032352 . PubMed PMID: 36550845; PubMed Central PMCID: PMC9771207.
52. Solovieff N, Cotsapas C, Lee PH, Purcell SM, Smoller JW. Pleiotropy in complex traits: challenges and strategies. Nat Rev Genet Juli. 2013;14(7):483–95. 10.1038/nrg3461. Located at: 23752797.
53. Dwivedi A, Thippana M, Khammampalli S, Cholleti SN, Vindal V. Unraveling the gender-specific molecular landscape of lung squamous cell carcinoma progression. J Biomol Struct Dyn 3 Februar 2025;1–14. doi:10.1080/07391102.2025.2460069 PubMed PMID: 39895519.
54. Tan F, Zhu H, Tao Y, Yu N, Pei Q, Liu H. u. a. Neuron navigator 2 overexpression indicates poor prognosis of colorectal cancer and promotes invasion through the SSH1L/cofilin-1 pathway. J Experimental Clin Cancer Res 9 Oktober. 2015;34(1):117. 10.1186/s13046-015-0237-3.
55. Soussi T. The TP53 gene network in a postgenomic era. Hum Mutat Juni. 2014;35(6):641–2. 10.1002/humu.22562 . PubMed PMID: 24753184.
56. Jafrin S, Aziz MA, Anonna SN, Akter T, Naznin NE, Reza S. u. a. Association of TP53 Codon 72 Arg > Pro Polymorphism with Breast and Lung Cancer Risk in the South Asian Population: A Meta-Analysis. Asian Pac J Cancer Prev 1 Juni. 2020;21(6):1511–9. .1511 PubMed PMID: 32592343; PubMed Central PMCID: PMC7568897.
57. Bulgakova O, Kussainova A, Kakabayev A, Kausbekova A, Bersimbaev R. Association of polymorphism TP53 Arg72Pro with radon-induced lung cancer in the Kazakh population. Vavilovskii zhurnal genetiki i selektsii. 2019;23(5):594–9. 10.18699/VJ19.530.
58. Ren YW, Yin ZH, Wan Y, Guan P, Wu W, Li XL. u. a. P53 Arg72Pro and MDM2 SNP309 polymorphisms cooperate to increase lung adenocarcinoma risk in Chinese female non-smokers: a case control study. Asian Pac J Cancer Prev. 2013;14(9):5415–20. 10.7314/apjcp.2013.14. .9.5415 PubMed PMID: 24175836.
59. Mechanic LE, Bowman ED, Welsh JA, Khan MA, Hagiwara N, Enewold L. u. a. Common genetic variation in TP53 is associated with lung cancer risk and prognosis in African Americans and somatic mutations in lung tumors. Cancer Epidemiol Biomarkers Prev Februar. 2007;16(2):214–22. 10.1158/1055-9965.EPI-06-0790.
60. Lash TL, VanderWeele TJ, Haneuse S, Rothman KJ. Modern epidemiology. Fourth edition / Timothy L. Lash, Tyler J. VanderWeele, Sebastien Haneuse, Kenneth J. Rothman. Philadelphia: Wolters Kluwer; 2021.
61. Chowdhury R, Shah D, Payal AR. Healthy Worker Effect Phenomenon: Revisited with Emphasis on Statistical Methods – A Review. Indian J Occup Environ Med. 2017;21(1):2–8. 10.4103/ijoem.IJOEM_53_16 . PubMed PMID: 29391741; PubMed Central PMCID: PMC5763838.

Figures

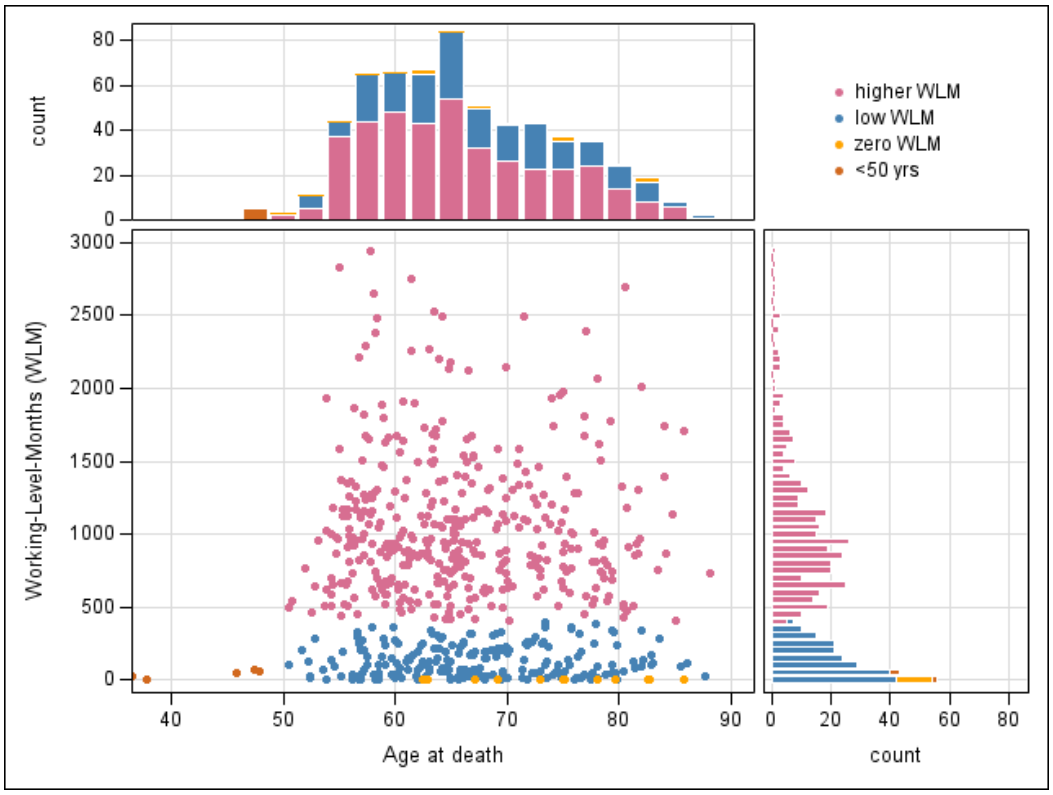


Figure 1

cumulative exposure to radon and its decay products [WLM] by age at death

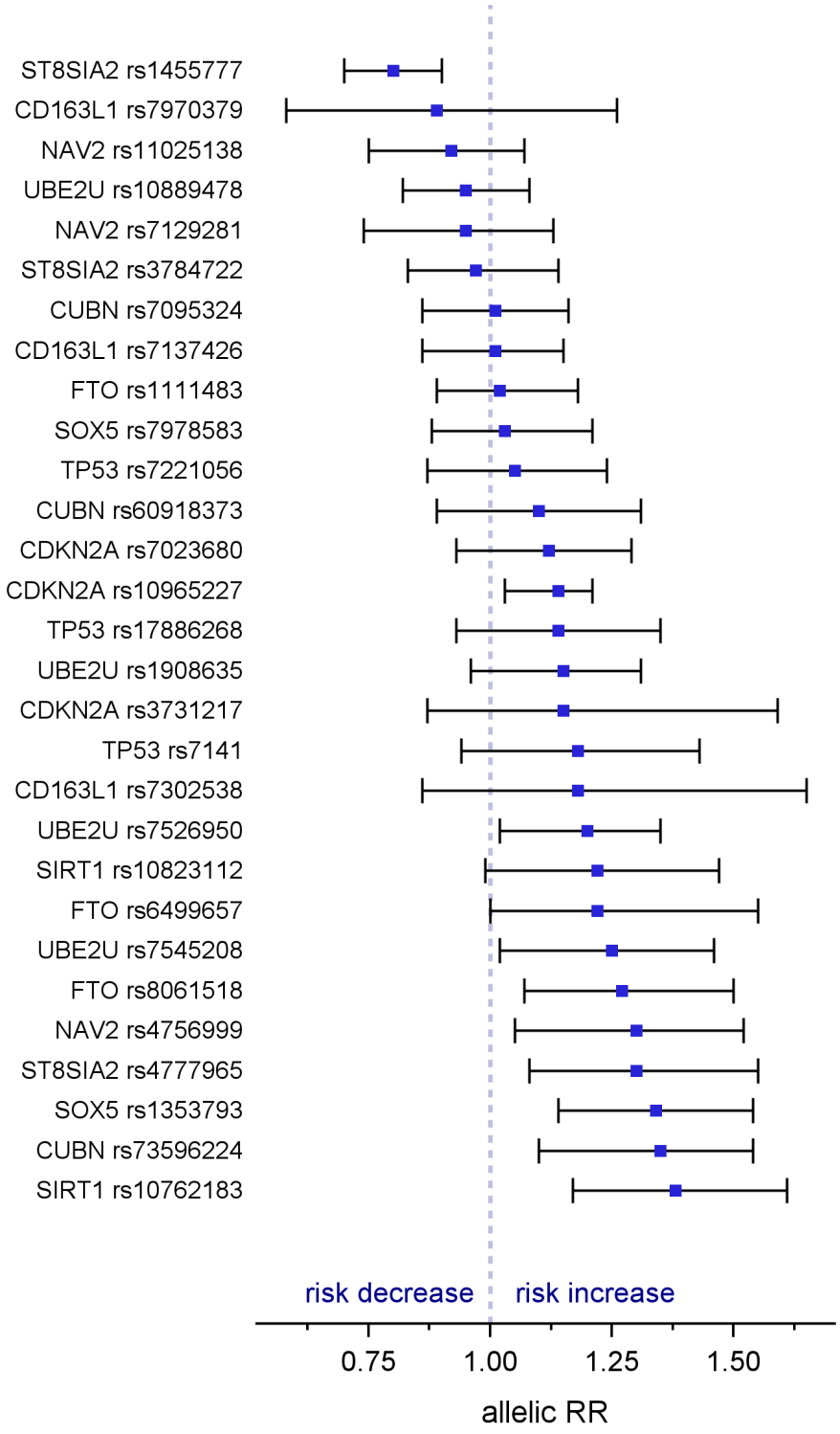


Figure 2

genetic relative risk (estimated by the G-model)

RR: genetic relative risk per rare allele

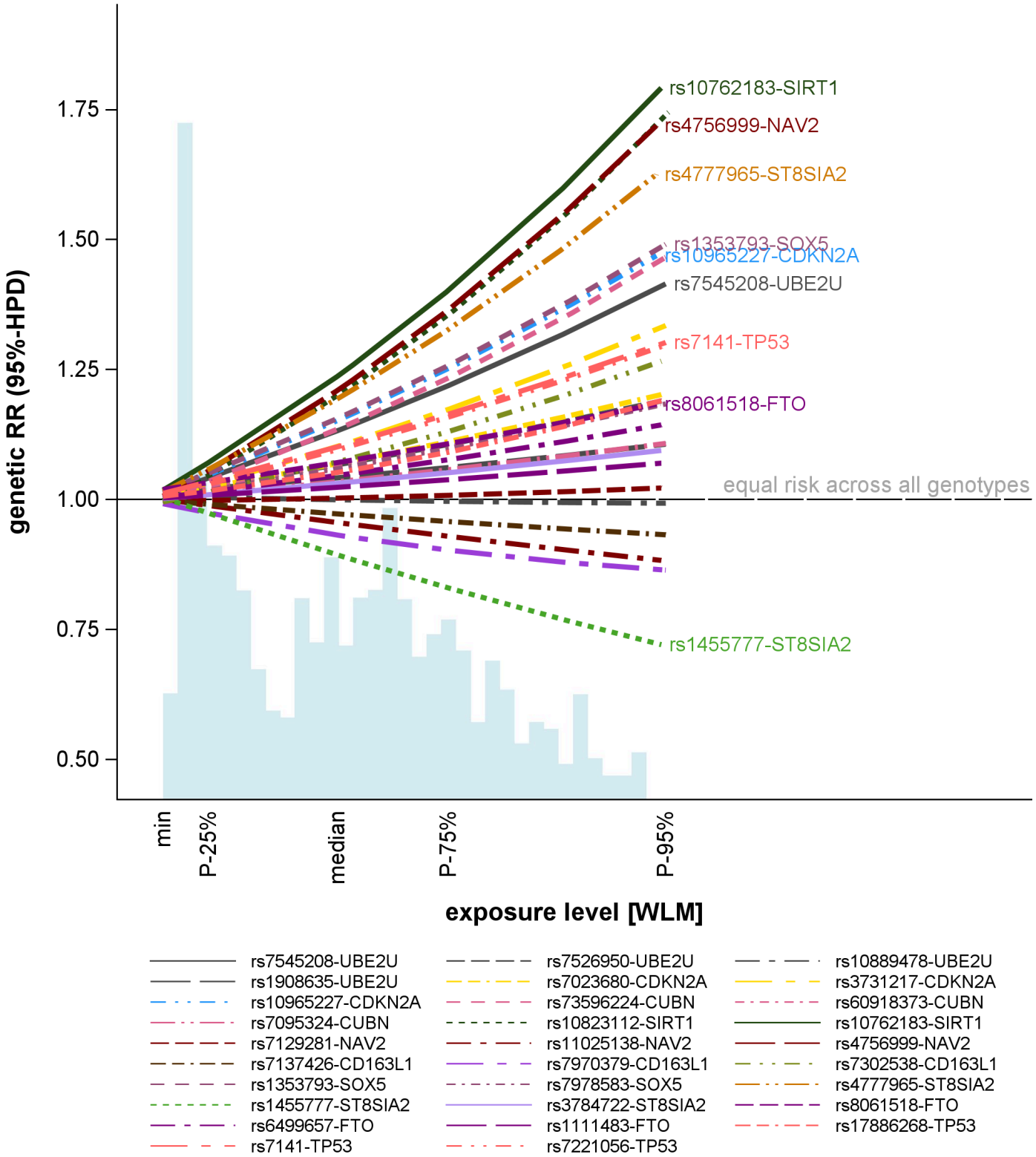


Figure 3

modification of the genetic relative risk by WLM (estimated by the GxE-model)

RR genetic relative risk; the distribution of radiation exposures in WLM is shown in pale blue in the background. The minimum and the 25%, 50% and 75% percentiles are highlighted on the x-axis. RR=1 all genotypes have the same genetic RR. Horizontal curve gradients would indicate that the relative genetic risk does NOT change with the degree of exposure and therefore there is no GxE-interaction. In a pure GxE-interaction, the genetic relative risk should be 1 at zero exposure (starting

point of the line).
Note: The line for rs10823112-*SIRT1*is hidden behind that of rs4756999-*NAV2*.

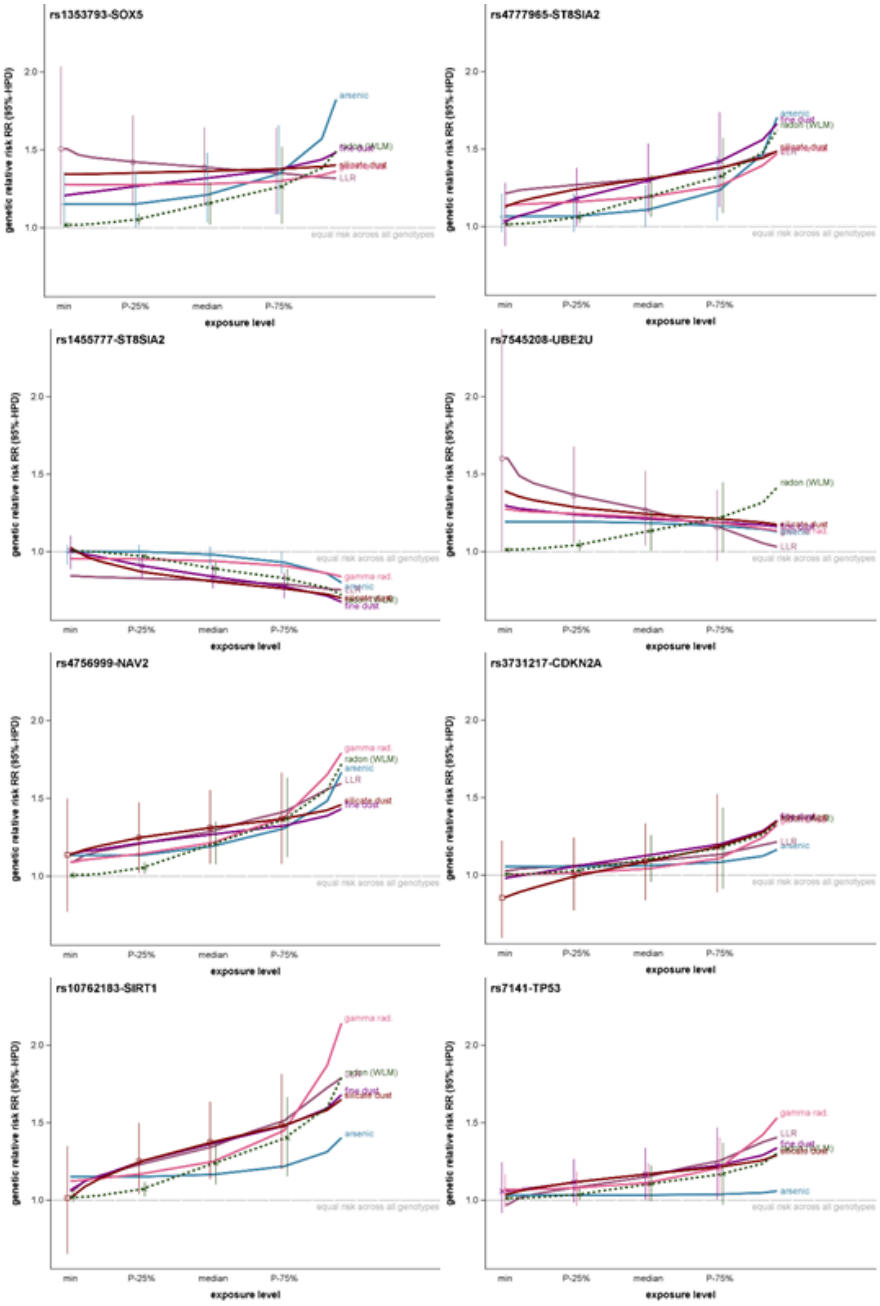


Figure 4

GxE-interaction with all exposures for selected markers

A relative genetic risk of RR=1 all genotypes have the same genetic RR. Horizontal curve gradients would indicate that the relative genetic risk RR does NOT change with the degree of exposure and therefore there is no GxE interaction. In a pure GxE interaction, the genetic relative risk should be 1 at zero exposure (starting point of the line). The distribution of radon exposures in WLM is shown in pale blue in the background. For visual comparison, the exposures were scaled so that the quartiles are positioned identically. The minimum and the 25%, 50% and 75% percentiles are highlighted on the x-axis. The curves for each exposure were scaled so that the percentiles correspond graphically.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx