A transgenerational mutational signature from ionizing radiation exposure.

Fabian Brand[1], Hannah Klinkhammer[1,3], Alexej Knaus[1], Manuel Holtgrewe[2], Leonie Weinhold[3], Dieter Beule[2], Kerstin Ludwig[4], Prachi Kothiyal[5], George Maxwell[5], Markus Noethen[4], Matthias Schmid[3], Karl Sperling[6], Peter Krawitz[1]

[1] Institute of Genomics Statistics and Bioinformatics, School of Medicine, University Hospital Bonn & University of Bonn, Germany
[2] Core Unit Bioinformatics, Berlin Institute of Health, Berlin, Germany
[3] Institute of Medical Biometry, Informatics and Epidemiology, University Hopsital Bonn, Bonn, Germany
[4] Institute of Human Genetics, School of Medicine, University Hospital Bonn & University of Bonn, Germany
[5] Inova Translational Medicine Institute, Inova Health System, Falls Church, VA, USA
[6] Institute of Medical and Human Genetics, Charitè-Universitaetsmedizin Berlin, Germany

Abstract

Effects of accidental radiation exposure on the human germline remain a topic of medical interest. Sequencing technology now allows the accurate analysis of the entire human genome in potentially exposed cohorts to search for possible signatures of ionizing radiation exposure. Clustered *de novo* mutations (cDNMs), that are multiple DNA lesions within 20 bp of each other, have been hypothesized to be a signature for paternal exposure to ionizing radiation. We analyzed whole genome trio data of 240 children and their likely exposed parents as well as 1,275 offspring from unexposed families. The cohort of exposed children consists of 130 offspring of Chernobyl cleanup workers (CRU, exposure range = 0-4,080 mSv) and 110 offspring of former soldiers from both german armies which have likely been irradiated during their service (Radar cohort, exposure range = 0-353 mSv). For the Radar cohort, a retrospective dosage estimation was conducted based on the service records from each soldier, but it remains a challenging task to retrospectively estimate received dose accurately. In agreement with previous epidemiological data, no transgenerational effects could be detected for the class of isolated *de novo* mutations, which cause the majority of genetic disorders. We found that the number of isolated *de novo* mutations increases by 2% per year of age of the father at conception, which is also in line with previous works.
On average, we observed 2.65 cDNMs per offspring in the CRU cohort, 1.48 in the Radar cohort and 0.88 in the control cohort, with a median of two clustered mutations per genome in offspring of irradiated fathers. This is a significant increase ($p < 0.005$) in the number of cDNMs compared to a set of age-matched controls. Furthermore, we show that the cDNM rates scale with paternal exposure to ionizing radiation ($p < 0.001$). Our findings corroborate that clustered *de novo* mutations represent a transgenerational biomarker for paternal exposure to ionizing radiation.

Introduction

Somatic mutational signatures can point to endogenous or exogenous mutagen exposure and have become an important biomarker to determine the cause and underlying processes in human cancers [1]. Exposure to ultraviolet light, for instance, results in pyrimidine dimers that can be removed by nucleotide excision repair on the transcribed strand. Consequently, a higher prevalence of C>T mutations on the untranscribed strand can be observed in e.g. malignant melanoma. Some signatures also result from anticancer drugs that are mutagens. For instance, systemic treatment with platinum-based drugs also affects the germline and can result in hypermutation in offspring [2].

While ultraviolet light is too weak to penetrate into cells of the germline, during oogenesis and spermatogenesis higher energy radiation could potentially induce DNA damage that is passed on to the next generation. Changes to human DNA in the form of chromosomal inversions and translocations have previously been found after long duration space flight onboard the International Space Station where astronauts were subjected to a dose of 146.34 mSv of high-linear energy transfer (LET) space radiation[3]. This environment has also been found to increase the cancer risk for astronauts following prolonged stays in orbit [4,5]. On earth, gamma rays with energies ranging from a few kiloelectronvolts (keV) to megaelectronvolts (MeV) occur naturally from various atmospheric interactions with cosmic particles, or from the decay of radioisotopes and accumulate to an annual whole body dose of 1-2 mSv [6]. Singular medical radiation exposure e.g. during computer tomography is in the same order of magnitude in the target region, but the gonads are usually shielded [6]. Unintentional, high exposure to man-made sources of gamma rays can happen during nuclear disasters or in certain military occupations (e.g. operating high powered radar devices) where stray radiation is emitted, and personnel is insufficiently protected [7]. However, since retrospective dosage estimations are difficult, biomarkers indicating prior radiation injury have been searched for decades with the respective methods available at that time.

While high-energy particles can cause double-strand breaks through direct impacts on the DNA, the primary damage is often a result from the generation of reactive oxygen species (ROS) following ionization of nearby water molecules in the cell [8–10]. Within a few helix turns of the DNA, this ionization can result in multiple single-strand breaks, abasic sites, and oxidized purine or pyrimidine bases. In contrast to damage caused by endogenous oxidative stress, clustered DNA lesions from ionizing radiation are less likely to be repaired due to their complexity and are thus more likely to persist after cell division [11]. Acute exposure to high doses (> 4Gy) of ionizing radiation has been studied extensively in mice, where an increase in *de novo* mutations within a few helix turns of DNA have been found to be associated with parental exposure to ionizing radiation, alongside structural changes to the genome [12,13]. Recently, Matsuda et al. demonstrated a dose-dependency of the increase of clustered DNA damage and structural changes to the genome by analyzing data from mouse hematopoietic stem cells following exposure to 3.8, 7.5 and 7.7 Gy [14]. Multiple damaged sites affecting nucleotides in close vicinity are thus a characteristic signature from ionizing radiation that could be passed on to the next generation when occurring in the germline [12].

Effects of ionizing radiation on human DNA

Until now, such studies have not been replicated with human data, mainly because of the lack of cohorts that are suited to trio whole genome sequencing analysis and have been exposed to elevated doses of ionizing radiation. One notable event that caused widespread irradiation in humans and animals is the Chernobyl nuclear accident. The first analyses after the Chernobyl

accident focused on mutation rates for minisatellites, microsatellites, and tandem repeat expansions and found a two-fold increase [15,16]. However, due to the small cohort size, potential confounders, and the small number of loci that could be investigated, the statistical significance of these results is still debated [17]. With the availability of high-throughput sequencing, further analyses focused on genome-wide screens of single nucleotide *de novo* mutations (DNMs). Yeager, *et al.* collected and sequenced a cohort of 130 offspring and their 105 parents that were exposed to ionizing radiation after the 1986 Chernobyl nuclear accident, where the parents were inhabitants of the town of Pripyat or liquidators guarding or cleaning up the accident site [18]. Doses prior to conception ranged from 0 to 4,080 mSv (Mean: 365.42 mSv, Std.: 684.55 mSv, Median: 29 mSv), but no elevated mutation rates for isolated DNMs were detected [18,19]. They also did not detect any enrichment of C>T mutations within 47 kb intervals, which has been hypothesized to be an indicator for hyper-mutability of single-strand intermediates during repair of double-strand breaks [20,21].

We reanalyzed the cohort of Yeager, *et al.* comprising 130 offspring, based on their data accessed under dbGAP accession number phs001163.v1.p1. Now, we recruited and analyzed an extended cohort of 110 offspring from 80 soldiers who served in German radar units [22]. This second cohort of exposed parents and their children was collected from former radar operators of both german armies (Bundeswehr and Nationale Volksarmee, NVA) for this study, which are known to have been subjected to high doses of ionizing radiation when servicing radar ground stations during live operations [7]. Even though the German government spent significant efforts in investigating health effects and occupational risks following long time service with radar units, and despite the fact that soldiers serving at unprotected radar installations have a higher risk to develop certain cancers, reliable data on the damage caused by the stray radiation from these devices is very limited [7,23]. Herein, we examine whether the signature of radiation-induced clustered DNA lesions is detectable in children born to parents with a high likelihood of exposure to ionizing radiation.

For the radar soldiers, the gonadal doses have been estimated by the radiation measurement facility (*Strahlenmessstelle*) of the Bundeswehr and yielded a lower mean exposure than in the Chernobyl (CRU) cohort (9.2 mSv vs 365.42 mSv, Figure 1b). Dosage estimations were based on sparse measurements taken during the active service of common radar devices or on emissions of devices reconstructed especially for retrospective assessments. Due to these complications and because some soldiers served in military roles that probably did not result in elevated levels of exposure, the dosage estimation remains inconclusive for the majority of soldiers (nexposed= 22, n no exposure = 55, nno documents=3, Supplementary Figure 1, 2, Supplementary Bericht S 209/20).

As a control cohort we chose the Inova study group consisting of 1,214 family trios and quadros, for which no exposure beyond naturally occurring ionizing radiation has been recorded. This cohort was also used to analyze the parental age effect on isolated DNMs in the germline [24,25]. All sequencing data were reprocessed using the NVIDIA Parabricks and Illumina DRAGEN toolkits for short-read mapping and variant calling on the GRCh37 reference genome, joint variant calling was performed on a total of 4,337 whole genomes using GLNexus v1.3.1 [26]. All data was subjected to a rigorous quality control pipeline, including checks for mean coverage, contamination with foreign DNA, the expected pedigree and more (Supplemental Material, Section 1.4.3), which resulted in the exclusion of 14 samples, mostly due to contamination of the saliva samples we received from these study participants. The resulting variant data was subsequently filtered for DNMs, with one of the criteria being that each *de novo* allele only occurs once in the entire call set (Supplemental Material, Section 1.4.2, Supplementary Table 2). In agreement with the literature, we could confirm that the number of isolated DNMs increased by 2% per year of paternal age (Figure 2, Supplementary Table 5) which translates to an accumulation of 1-2 mutations per year of

age of the father [24,27,28]. To account for the significant age effect and age differences between the parents in the three cohorts, we computed an age-matched subset of the control cohort which was used for the subsequent statistical analysis. This age-matching minimizes the age difference between the control or CRU and the Radar cohort (Supplemental Material, Section 1.6.1). For the age-matched subsets, we found a rate of (72.67 (18.15, median = 79) isolated DNMs per generation for the Inova cohort, 65.43 (13.57, median = 65) isolated DNMs for the CRU cohort, and 67.95 (17.25, median=64) isolated DNMs for the Radar cohort (Supplemental tables 1 and 3). These rates are comparable to the values reported in the original works and did not differ significantly in between the cohorts (Supplemental Material, Section 1.6.3) [18,24].

Analysis of clustered *de novo* mutations

We continued our analysis by filtering for loci with multiple lesions (clustered DNMs, cDNMs). These sites are defined as genomic regions where at least two *de novo* mutations occur within 20 bp distance of each other. If there are multiple successive pairs of DNMs that fulfill this criteria, they are always grouped into the same cluster, even if the first and last lesion in the cluster might exceed the 20 bp distance threshold. In total, we observed 1,989 clustered mutations in 1,515 offspring. The median number of clustered DNMs was two in offspring from the Chernobyl cohort, which is twice as many as in an age-matched subset of the unexposed controls. Using a negative binomial regression model we confirmed that the estimated number of cDNMs differ between either exposed and the control cohort (n = 110, p Radar= 0.005, pChernobyl< 1·10-3; Figure 3, Supplemental Table 6-7). These differences are more prominent when the Radar cohort is split into a subcohort of children with either exposed fathers (estimation > 0 mSv) or unexposed fathers, where the children of exposed parents show a higher number (exposed=1.72, not exposed = 1.39) of cDNMs on average (Supplemental Material, Section 1.6.8).

In contrast to isolated DNMs, the false positive rate is higher for clustered DNMs [29]. We, therefore, established criteria for the assessment of alignments from short read data in the IGV browser (Supplemental Material, Section 1.5). This was an iterative process involving validating different call sets of clusters from the Radar cohort, which is the only of the three cohorts where DNA for validation with secondary sequencing methods was available by Sanger and PacBio sequencing (Supplementary Figure 6). Our validation efforts resulted in an estimate for the positive predictive value (PPV), i.e. the probability of a discovered cDNM being true positive, of 0.23 (95% Clopper-Pearson confidence interval 0.17 - 0.30). It is noteworthy that in some individuals none of the detected cDNMs could be validated, including the outlier with 14 cDNMs shown in Figure 3. However, the negative binomial regression model is not substantially affected by this as it models the median count of cDNMs per offspring, which is robust against outliers. Additionally, simulations accounting for this PPV did not affect the significance of the test results (Supplemental Material, Section 1.6.6). All true positive cDNMs identified in the Radar cohort were analyzed with respect to their likely effect on diseases or the coding region in general, but none was found to have any implications with any genetic conditions in the study participants (Supplemental Material, Section 1.5.1).

We used read-based phasing of DNMs with informative variants to determine the parental gamete of origin [30]. The proportion of DNMs that can be phased varied between the three cohorts because of technical and stochastic reasons. The influencing factors are the distance between DNM and informative SNP, the coverage in that region, the length of the sequencing reads (100bp in the control cohort vs 150bp in both case cohorts), and the distribution of

fragment sizes. We did not observe changes in the number of isolated DNMs that were attributed to the paternal or maternal alleles between the three cohorts in this study (Chi-Squared-Test, Supplemental Material, Section 1.6.7). Since clustered DNMs can extend over several base pairs, not all lesions in a cluster can necessarily be phased. If read-based phasing suggested the paternal or maternal germline based on the information from at least one lesion we assumed this origin for the whole cluster. We did not observe any clusters with contradicting phasing information regarding the parental origin. In addition to read-based phasing, we also determined the parental origin with Sanger and PacBio long-read sequencing for a subset of cDNMs in the Radar cohort during our validation efforts. By resequencing the phase informative single nucleotide polymorphism (SNP) alongside each cluster, we validated the parental origin of 26 clusters in the Radar cohort, where 17 clusters of paternal and 9 of maternal origin were present. Due to the shorter read length of 100bp in the Inova cohort, a reliable estimate for this ratio in the population could not be computed.

In addition to a general increase in the amount of cDNMs per sample in the cohorts that were exposed to ionizing radiation, we found a positive association between the estimated dose and the number of cDNMs. Using a negative binomial regression model, we see a significant ($p < 0.001$) increase of clustered DNMs with paternal radiation exposure at $f(n) = 1.55 \cdot e0.0005n$ mutations per n mSv, when combining both exposed cohorts (Figure 4, Supplementary Table 12). However, we could not assert statistical significance for this model in the Radar cohort alone. The average number of lesions per cluster, which we term cluster size, is expected to increase with LET and the largest clusters that we observed in each of the three cohorts is eight (Inova), nine (Radar), and eleven (CRU) lesions respectively (Supplemental Figure 10). Analyzing the distribution of cluster sizes in all three cohorts did not yield a statistically significant shift (Supplemental Material, Section 1.6.9).

Discussion

The question of whether there is a transgenerational effect of ionizing radiation on the human genome has been studied ever since large human cohorts were subjected to increased doses of ionizing radiation. The early epidemiologic studies on the offspring of atomic bomb survivors now date back more than 70 years and the disadvantage of such analyses is that they require a readout on the phenotype-level such as malformations [31]. More recent studies indicate that larger quantities of environmental radiation lead to a higher incidence of cancer and birth defects, emphasizing that there are subtle effects of ionizing radiation on the human germline that have not been captured by earlier studies [32]. If an analysis is based on phenotypic data alone an increased rate of *de novo* mutations or other subtle changes may remain unobserved, since most *de novo* mutations are rare and occur in the non-coding part of the DNA. Therefore, the unique capabilities of whole genome sequencing can yield much deeper insights into the consequences of prolonged ionizing radiation exposure on the human genome than possible with earlier technologies. Yeager, *et al.* already studied isolated *de novo* mutations and clusters on the scale of kb, which are associated with the repair of double-strand breaks, and did not find a significant increase in mutation rates [18,24]. In mice and somatic cells however, clusters on the single-bp scale were previously implied as a consequence of the secondary effect of high-energy particles interacting with DNA [12–14]. We reanalyzed the cohort of Yeager et al., as well as an additional cohort with potential exposure to ionizing radiation, and found *de novo* mutation clusters to be significantly increased in

human DNA. Furthermore, our statistical models indicate that the number cDNMs in offspring increases with the dosage the paternal germline was exposed to.

The potential of transmission of radiation-induced DNA changes to the next generation is of particular concern for parents that were potentially subjected to higher doses of ionizing radiation than deemed safe. To address the concerns articulated by former radar soldiers, the German Ministry of Defense initiated this study with the goal of investigating the transgenerational effects of ionizing radiation in offspring of exposed parents. Subsequently, we recruited retired radar operators, but even though we took great care when assessing the inclusion criteria for each individual, the available information about exposition towards ionizing radiation was limited. In practice that meant, radar devices that were in active service more than 50 years ago had to be made operational again, to measure scattered radiation profiles [7]. In addition, the service hours and proximity to the radar device during operation and maintenance were derived from a generalized service manual of the German armies based on rank, position and mission of the soldier, while actual practice from memory protocols of the study participants differed, suggesting a potential source of errors. There is anecdotal evidence of higher ranking soldiers helping with maintenance work, however, this is not part of the official records. Thus, some of the individuals that we list as unexposed, might actually have been irradiated (Figure 3 and 4) and dosage values given for members of the Radar cohort are likely to be underestimated. For instance, we had one offspring in the Radar cohort with 14 clustered DNMs whose parents were unexposed according to the estimation based on official records, yet no other causes of hypermutation were known [2]. Despite the discussed inaccuracies of dosage estimation we proceeded with dose-effect estimations, and excluded this individual and other offspring of allegedly not exposed fathers from the negative binomial regression models (Figure 4).

Aside from difficulties in a-priori gauging potential radiation exposure scenarios for each soldier, there were several potential ascertainment biases in recruiting. First, individuals that were under the subjective impression that they had been exposed were more likely to participate (volunteer bias), and second, radar soldiers that were operating devices emitting the highest quantities of stray radiation, were in their eighties and had a high personal risk for diseases following their service (survivorship bias) [7]. However, also with these caveats, most of the statistical results indicate an influence of paternal radiation exposure on the number of cDNMs per offspring, even when considering the smaller and less exposed radar cohort alone. These findings are further supported by the increase of effect size and statistical significance in the more likely exposed subgroup.

In all exposed cohorts, we observed that the number of cDNMs increased by one to two per genome in children of fathers that were subjected to elevated doses of ionizing radiation for a longer time. To derive a clinical interpretation from these statistical results, we can compare the number and impact of cDNMs with the disease burden due to all DNMs. The total number of clusters exceeds that of the general population by 0.6 for the Radar and 1.77 for the CRU cohort. In addition to the increase in the total number of DNMs per sample due to radiation exposure, it is also hypothesized that the functional impact of cDNMs is larger, if they fall within coding regions of the human genome compared to isolated DNMs. This increased impact could lead to pathogenicity or even embryonic lethality, in cases where cDNMs affect important parts of the exome. However, given the low overall rate of excess cDNMs following paternal exposure to ionizing radiation and the portion of the genome that is protein coding, the likelihood that a disease in an offspring of exposed parents is triggered by a DNM cluster is minimal. Therefore, with the paternal age effect at roughly 1 additional DNM per year of age of the father, and the expected number of 60 to 80 DNMs per generation on average, the paternal exposure to low dose ionizing radiation contributes less to an individual's risk for genetic diseases then age [20,27,33]. Thus, the excess risk attributable to

DNMs that arose after paternal exposure to ionizing radiation is negligible compared to the base risk for genetic diseases. While the expected clinical consequences of each clustered or isolated DNM is of comparable order, the consequences of a double strand break that is incorrectly repaired, is usually more severe. Translocations are most likely an indirect consequence of double-strand breaks, and have been observed with increased frequency in irradiated mice as well as in offspring of Radar soldiers [14,22]. However, in contrast to cDNMs, comparing mutation rates for structural variants is more error-prone when cohorts were sequenced with different short read lengths and we have therefore refrained from assessing them statistically.

Further studies with longer read lengths, and ideally larger cohorts and more accurate radio-dosimetry are required to better characterize the disease risk and dose-response relationships exhibited by transgenerational signatures of prolonged paternal exposure to low dose ionizing radiation, like cDNMs. Furthermore, to determine the paternal to maternal ratio in cDNMs in the general population, it is necessary to sequence an appropriate cohort deeply and with greater read length than what was possible in this study. Since the number of clusters with paternal and maternal origin is expected to differ due to the accumulation of repair errors, an accurate measurement of the paternal to maternal cDNM ratio in the general population is necessary to assess the influence of radiation exposure accurately [34]. A further shift towards more paternally inherited clusters in exposed cohorts would provide additional evidence for the correlation between ionizing radiation exposure of the fathers and cDNM rates in the offspring. Targeting a statistical analysis of structural variants and translocations of the general population compared to exposed cohorts would also become feasible after subjecting samples to long read sequencing. It is also interesting to further investigate the potential impact of the linear energy transfer on the cluster size. Naturally, one would expect the damage to increase proportionally with the linear energy transfer of the ionizing radiation, leading to larger clusters, or an increased number of structural variants, when individuals are subjected to ionizing radiation with higher LET. This effect could also explain differences in the number and nature of the clusters between the two exposed cohorts, because the gamma ray spectrum of the ionizing radiation the two cohorts were exposed to differs and a difference in the LET is one of the consequences thereof [35].

While there are some remaining insecurities about the nature and quantity of the radiation involved, we could provide strong evidence for the existence of a transgenerational effect of prolonged paternal exposure to low-dose ionizing radiation on the human genome for the first time. Luckily for those affected families, the additional risk due to radiation induced cDNMs on the scale of single base-pairs is very low. In addition, this analysis gives rise to several further promising research avenues to characterize potential additional transgenerational signatures of ionizing radiation on the human genome, f.e. analyzing structural changes such as translocations, which are even more complicated to detect than cDNMs.

Figures

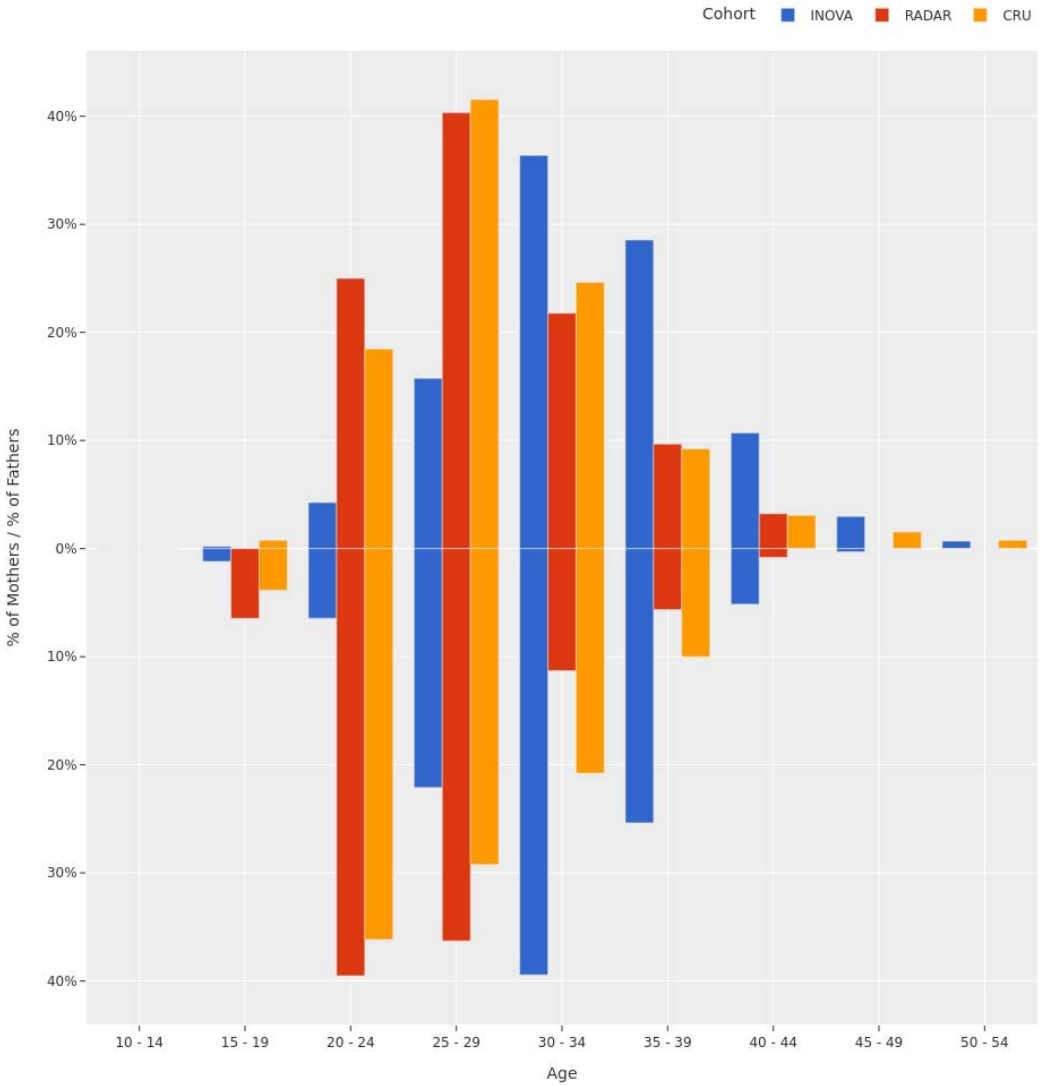
Figure 1: Study Cohorts

a) Age distribution of the three study cohorts. Due to the large difference in cohort size, Y-Axis is given in percentage of the total cohort size. Negative values on the Y-Axis show the distribution of maternal age, positive values that of paternal age. On average, fathers are >5 years older in the Inova cohort compared to the case cohorts and mothers are >5.5 years older on average.

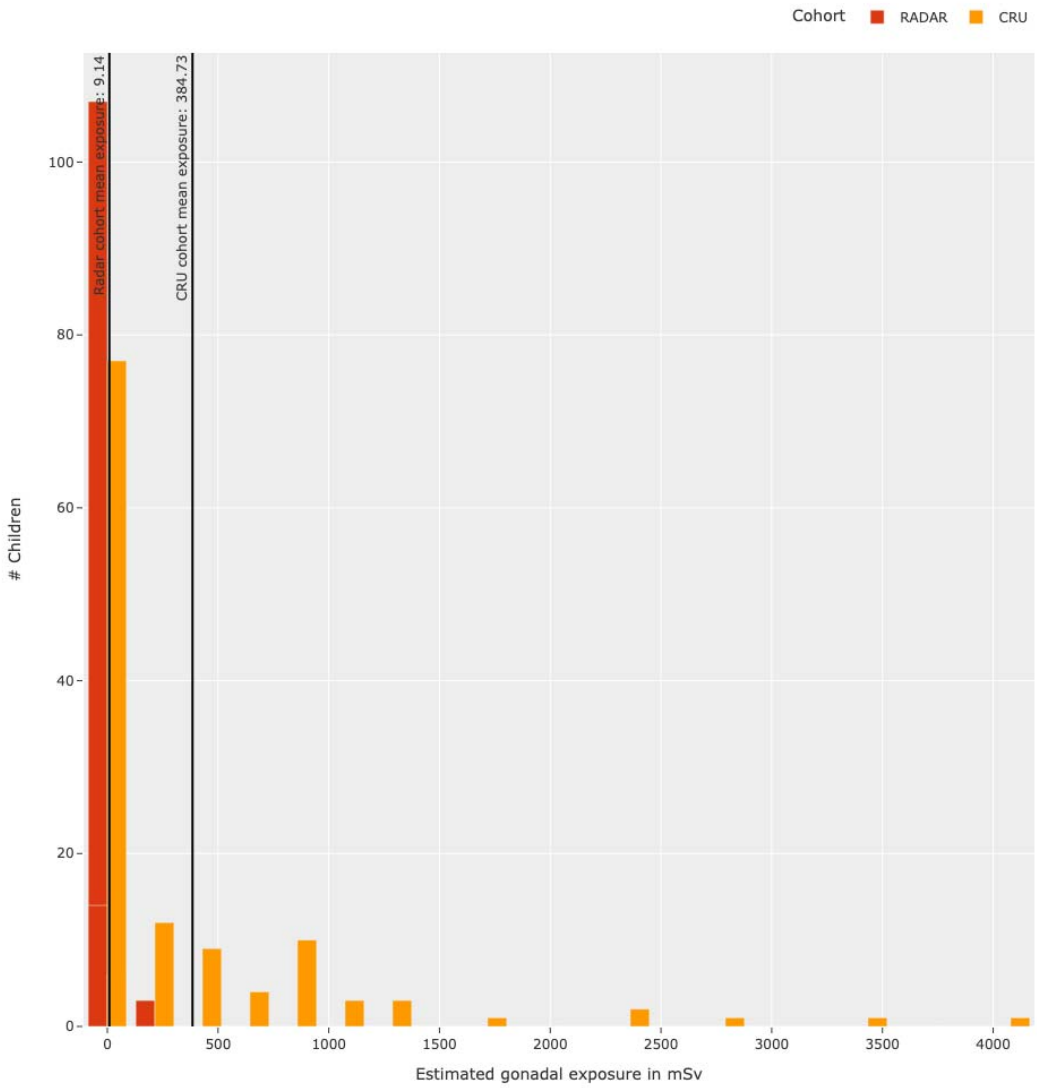
b) Distribution of paternal exposure for the Radar and CRU cohorts. The maximum exposure observed in the Radar cohort is 353 mSv, and 4,079 mSv in the CRU cohort. In the Radar cohort, 107 samples are recorded with an exposure estimation of exactly 0, meaning that their fathers are unlikely to have been exposed or no estimation could be made.

c) Distribution of sex of the children in each of the three cohorts.

a)



b)


Cohort
RADAR
CRU
Radar cohort mean exposure: 9.14
CRU cohort mean exposure: 384.73
Children
0
20
40
60
80
100
0
500
1000
1500
2000
2500
3000
3500
4000
Estimated gonadal exposure in mSv

c)

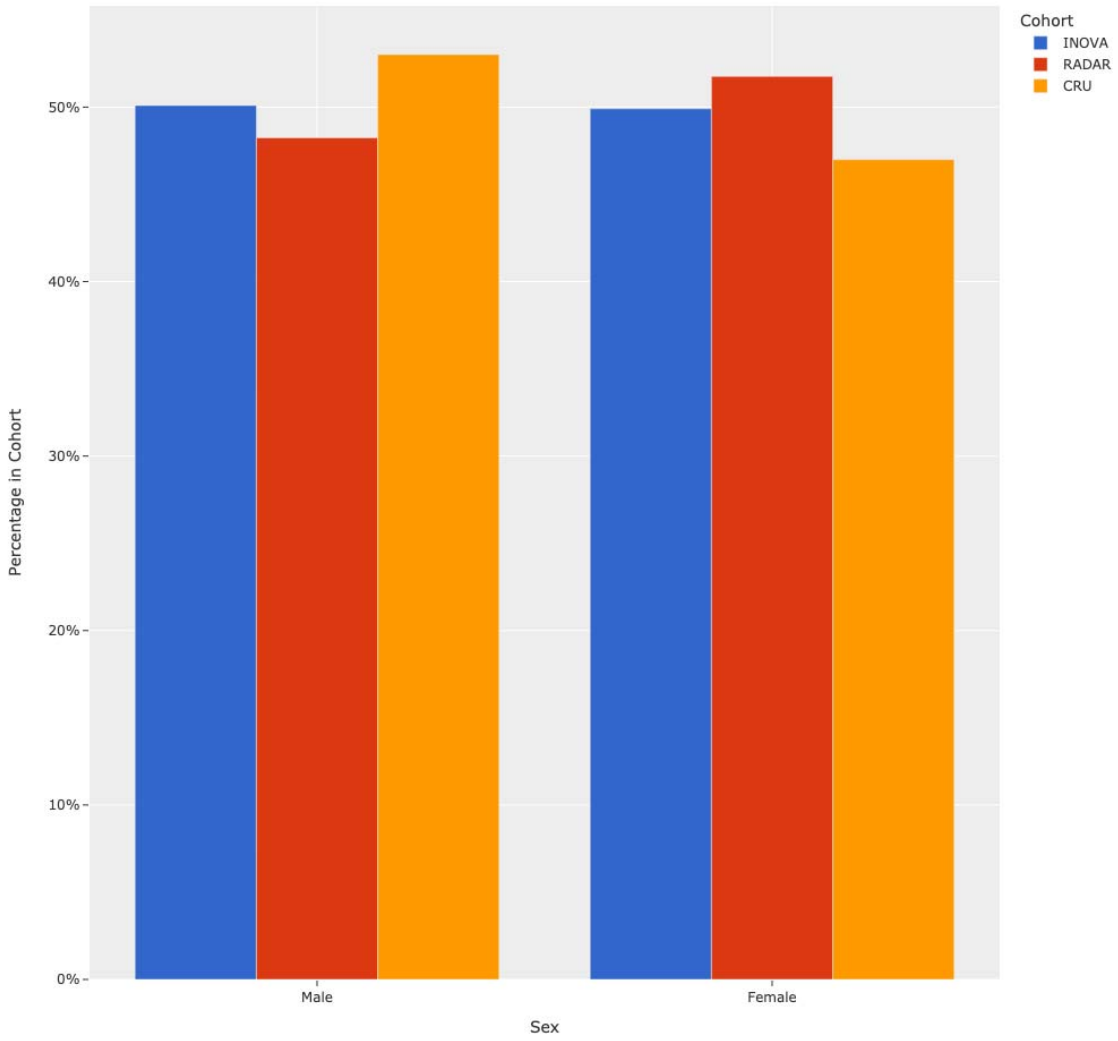

Cohort
INOVA
RADAR
CRU
50%
40%
30%
20%
10%
0%
Percentage in Cohort
Male
Female
Sex

Figure 2: Paternal Age Effect

Paternal age effects computed by a negative binomial regression (NBR) model estimating the number of DNMs given the paternal age at conception of the child for each cohort. When fitting this model, no age matching was applied to the data, so on average the parents are older in the Inova cohort. Nevertheless, the paternal age effect is roughly 2% for each of the cohorts, which results in an increase of ~1 DNM per year of age of the father.

Figure 3: Number of cDNMs per sample

Violin plot of the number of cDNMs per sample grouped by cohort. The width of the violin at each integer value of the y-axis visualizes the number of samples and their respective count of cDNM clusters. The box plot for each cohort is included inside the respective violin to display the quartile ranges and median number of cDNMs per sample in the given cohort. On average, we detected cDNMs in the Inova cohort, in the Radar cohort and in the CRU cohort on age matched data.

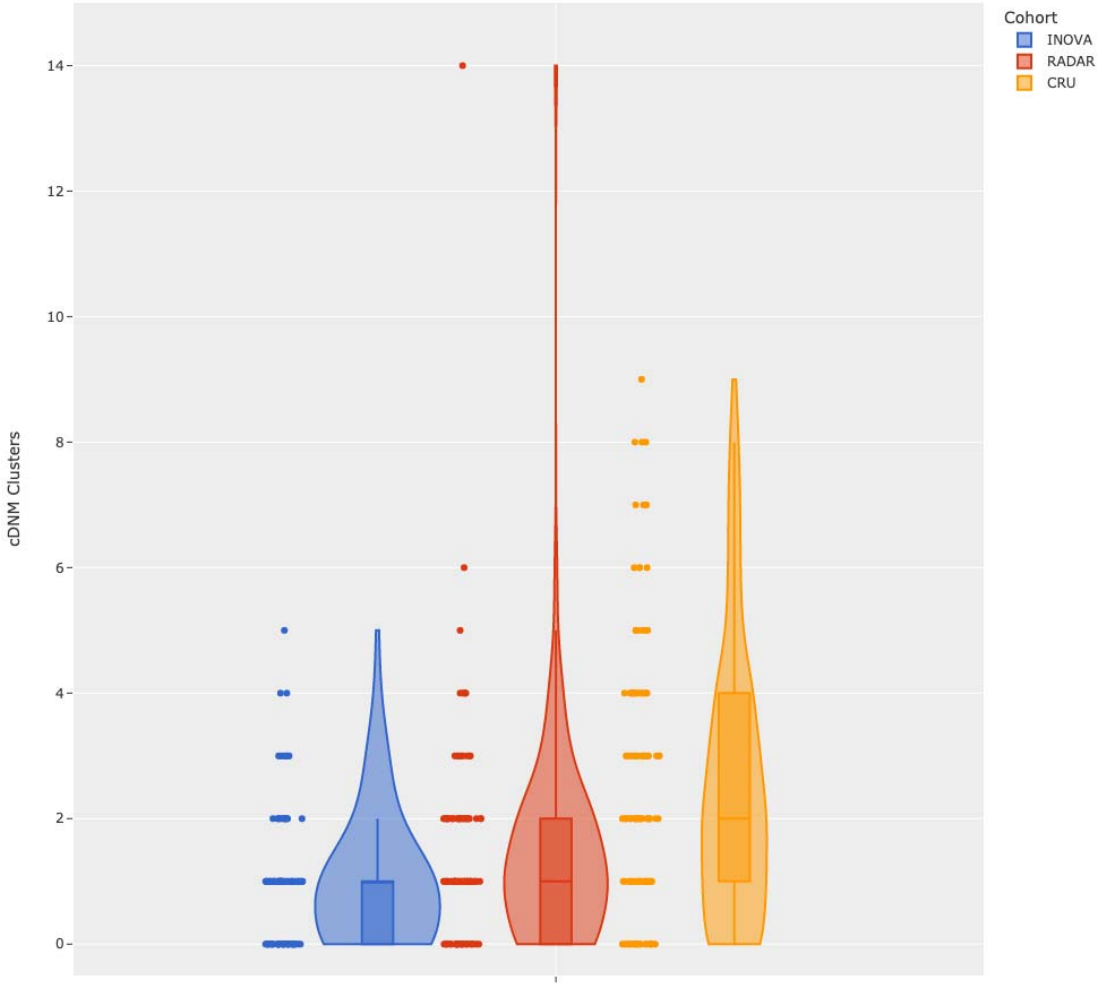
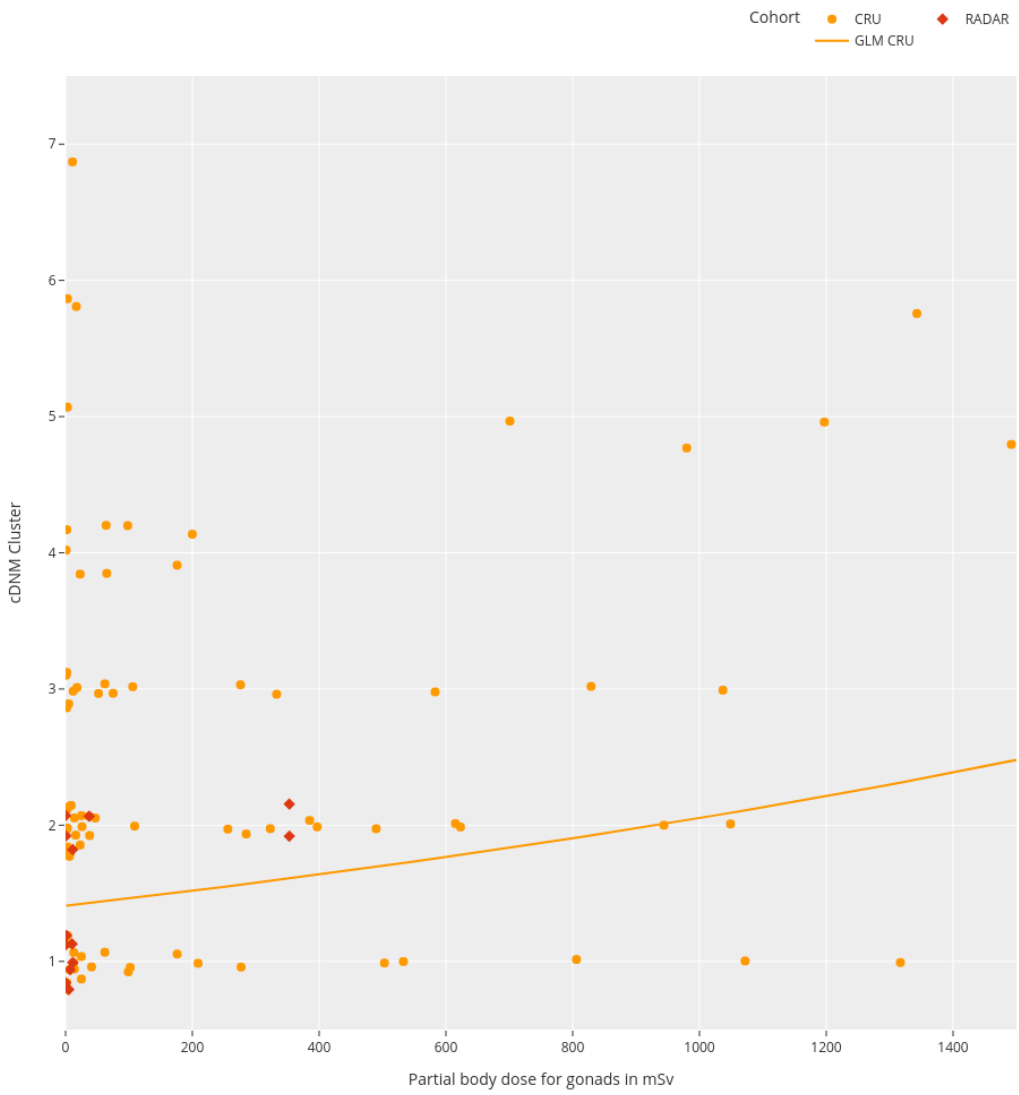


Figure 4: Number of cDNMs per mSv of paternal exposure

Estimation of the number of cDNMs depending on the paternal exposure in mSv as computed by a negative binomial regression (NBR) model with logarithmic link function modeling the number of clusters given the paternal exposure in mSv. Because the accuracy of the negative binomial regression model deteriorates rapidly with larger exposure estimates, the x-Axis has been cut off at 1.5k mSv, which means that five samples from the CRU cohort for which the estimation by the model is very inaccurate, are hidden. The fit shown in this image is conditional only on the number of mSv of paternal exposure, both exposed cohorts have been rolled into one supercohort for this analysis. Supplemental Figure 10 shows the analogous model also taking into account cohort membership for each sample. We only included the samples from the exposed subgroup of the radar cohort in this analysis, i.e. all offspring of fathers with an estimated exposure of >0 mSv. This restriction in the cohort leads to the exclusion of some children with many cDNMs, e.g. the outlier in the Radar cohort in Figure 3.



Data Availability

- Sequencing data can be accessed at EGA:
  - EGA: Study Accession EGAS00001007321
- Code can be accessed at GitHub and Zenodo:
  - GitHub: https://github.com/brand-fabian/radarstudy
  - Zenodo https://doi.org/10.5281/zenodo.8431077

Acknowledgement

The radar study was funded by the German Ministry of Defense and conducted by Deutsches Zentrum für Luft und Raumfahrt (DLR) under the supervision of an expert committee consisting of Michael Krawczak, Hajo Zeeb, Ivo Gut, André Reis and experts from the Bundeswehr (Reinhard Ullmann, Matthias Port). Dosages were estimated by Andreas Schirmer from Strahlenmessstelle der Bundeswehr. Recruitment was assisted by Dietmar Glaner, Josef Wiesner, and "Bund zur Unterstützung Radargeschädigter (BzUR e.V.)". *De novo* mutation rates in unexposed controls were also discussed with Hákon Jónsson from DeCode. Data was generated by the NGS core facility at Universitätsklinikum Bonn, which is part of the West German Genome Center (WGGC) and processed by the Core Unit for Bioinformatics in Bonn. Epigenetic analysis was performed by Sascha Tierling and Jörn Walter of University Saarland.

Ethics Declaration

All participants of the radarstudy were of legal age at time of recruitment and consented to the use of their data for research purposes. Ethics approval was granted by the "Ethikkommission der Medizinischen Fakultät Bonn" (Ethics committee of the medical faculty Bonn). All identifiers were pseudonymized prior to publication and the reidentification based on identifiers mentioned in the manuscript or supplemental material is only possible by members of the research group.

Author contributions:

Conceptualization: FB, HK, AK, MH, LW, DB, MN, MS, KS, PMK

Data curation: FB, AK

Formal analysis: FB, HK, LW

Funding acquisition: PMK, DB

Investigation: FB, HK, AK

Methodology: FB, HK, AK, LW, MS, PMK

Project administration: PMK

Resources: FB, AK, MH, KUL, PK, GM, DB, PMK

Supervision: DB, MN, MS, KS, PMK

Software: FB

Visualization: FB

Writing – original draft: FB, HK, AK, PMK

Writing – review & editing: FB, HK, AK, PMK

Bibliography

1. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
2. Kaplanis, J. *et al.* Genetic and chemotherapeutic influences on germline hypermutation. *Nature* **605**, 503–508 (2022).
3. Garrett-Bakelman, F. E. *et al.* The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science* **364**, (2019).
4. Cucinotta, F. A. Space Radiation Risks for Astronauts on Multiple International Space Station Missions. *PLoS ONE* **9**, e96099 (2014).
5. George, K., Rhone, J., Beitman, A. & Cucinotta, F. A. Cytogenetic damage in the blood lymphocytes of astronauts: Effects of repeat long-duration space missions. *Mutat. Res. Toxicol. Environ. Mutagen.* **756**, 165–169 (2013).
6. Radiation, U. N. S. C. on the E. of A. & others. *Sources and effects of ionizing radiation, united nations scientific committee on the effects of atomic radiation (UNSCEAR) 2008 report, volume I: Report to the general assembly, with scientific annexes A and B-sources*. (United Nations, 2010).
7. König, W. *et al.* Bericht der Expertenkommission zur Frage der Gefährdung durch Strahlung in früher Radareinrichtungen der Bundeswehr und der NVA

(Radarkommission)[Report of the expert group on the issue of the hazards of radiation in earlier types of radar devices and the German Radar Commission]. Berlin, German Radar Commission. (2003).

8. Frankenberg-Schwager, M. Induction, repair and biological relevance of radiation-induced DNA lesions in eukaryotic cells. *Radiat. Environ. Biophys.* **29**, 273–292 (1990).
9. Georgakilas, A. G., O'Neill, P. & Stewart, R. D. Induction and repair of clustered DNA lesions: what do we know so far? *Radiat. Res.* **180**, 100–109 (2013).
10. Sage, E. & Shikazono, N. Radiation-induced clustered DNA lesions: Repair and mutagenesis. *Free Radic. Biol. Med.* **107**, 125–135 (2017).
11. Sage, E. & Harrison, L. Clustered DNA lesion repair in eukaryotes: relevance to mutagenesis and cell survival. *Mutat. Res. Mol. Mech. Mutagen.* **711**, 123–133 (2011).
12. Adewoye, A. B., Lindsay, S. J., Dubrova, Y. E. & Hurles, M. E. The genome-wide effects of ionizing radiation on mutation induction in the mammalian germline. *Nat. Commun.* **6**, 6684 (2015).
13. Satoh, Y. *et al.* Characteristics of induced mutations in offspring derived from irradiated mouse spermatogonia and mature oocytes. *Sci. Rep.* **10**, 1–13 (2020).
14. Matsuda, Y. *et al.* Spectra and characteristics of somatic mutations induced by ionizing radiation in hematopoietic stem cells. *Proc. Natl. Acad. Sci.* **120**, e2216550120 (2023).
15. Dubrova, Y. E. *et al.* Human minisatellite mutation rate after the Chernobyl accident. *Nature* **380**, 683–686 (1996).
16. Slebos, R. J. *et al.* Mini-and microsatellite mutations in children from Chernobyl accident cleanup workers. *Mutat. Res. Toxicol. Environ. Mutagen.* **559**, 143–151 (2004).
17. Little, M. P., Goodhead, D. T., Bridges, B. A. & Bouffler, S. D. Evidence relevant to untargeted and transgenerational effects in the offspring of irradiated parents. *Mutat. Res. Mutat. Res.* **753**, 50–67 (2013).

18. Yeager, M. *et al.* Lack of transgenerational effects of ionizing radiation exposure from the Chernobyl accident. *Science* **372**, 725–729 (2021).
19. Chumak, V. *et al.* Estimation of radiation gonadal doses for the American–Ukrainian trio study of parental irradiation in Chornobyl cleanup workers and evacuees and germline mutations in their offspring. *J. Radiol. Prot.* **41**, 764–791 (2021).
20. Jónsson, H. *et al.* Parental influence on human germline de novo mutations in 1,548 trios from Iceland. *Nature* **549**, 519–522 (2017).
21. Roberts, S. A. & Gordenin, D. A. Hypermutation in human cancer genomes: footprints and mechanisms. *Nat. Rev. Cancer* **14**, 786–800 (2014).
22. Holtgrewe, M. *et al.* Multisite de novo mutations in human offspring after paternal exposure to ionizing radiation. *Sci. Rep.* **8**, 1–5 (2018).
23. Degrave, E., Meeusen, B., Grivegnée, A.-R., Boniol, M. & Autier, P. Causes of death among Belgian professional military radar operators: A 37-year retrospective cohort study. *Int. J. Cancer* **124**, 945–951 (2009).
24. Wong, W. S. *et al.* New observations on maternal age effect on germline de novo mutations. *Nat. Commun.* **7**, 10486 (2016).
25. Knijnenburg, T. A. *et al.* Genomic and molecular characterization of preterm birth. *Proc. Natl. Acad. Sci.* **116**, 5819–5827 (2019).
26. Lin, M. F. *et al.* GLnexus: joint variant calling for large cohort sequencing. *BioRxiv* 343970 (2018).
27. Kong, A. *et al.* Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
28. Sasani, T. A. *et al.* Large, three-generation CEPH families reveal post-zygotic mosaicism and variability in germline mutation accumulation. *bioRxiv* 552117 (2019).

29. Besenbacher, S. *et al*. Multi-nucleotide de novo Mutations in Humans. *PLOS Genet.* **12**, e1006315 (2016).

30. Belyeu, J. R., Sasani, T. A., Pedersen, B. S. & Quinlan, A. R. Unfazed: parent-of-origin detection for large and small de novo variants. *Bioinformatics* **37**, 4860–4861 (2021).

31. Yamada, M. *et al.* Congenital malformations and perinatal deaths among the children of atomic bomb survivors: a reappraisal. *Am. J. Epidemiol.* **190**, 2323–2333 (2021).

32. Zlobina, A. *et al.* Impact of Environmental Radiation on the Incidence of Cancer and Birth Defects in Regions with High Natural Radioactivity. *Int. J. Environ. Res. Public. Health* **19**, 8643 (2022).

33. Wood, K. A. & Goriely, A. The impact of paternal age on new mutations and disease in the next generation. *Fertil. Steril.* **118**, 1001–1012 (2022).

34. Wang, S., Meyer, D. H. & Schumacher, B. Inheritance of paternal DNA damage by histone-mediated repair restriction. *Nature* **613**, 365–374 (2023).

35. Semenenko, V. & Stewart, R. A fast Monte Carlo algorithm to simulate the spectrum of DNA damages formed by ionizing radiation. *Radiat. Res.* **161**, 451–457 (2004).