Radiation-Induced Deletions in Mouse Spermatogonia are Usually Large (over 200 kb) and Contain Little Sequence Similarity at the Junctions

Authors: Kodaira, Mieko, Asakawa, Jun-ichi, and Nakamura, Nori

Source: Radiation Research, 187(6) : 722-731

Published By: Radiation Research Society

URL: https://doi.org/10.1667/RR14660.1

RADIATION RESEARCH **187**, 722–731 (2017)
0033-7587/17 $15.00

DOI: 10.1667/RR14660.1

Radiation-Induced Deletions in Mouse Spermatogonia are Usually Large (over 200 kb) and Contain Little Sequence Similarity at the Junctions

Mieko Kodaira,[1] Jun-ichi Asakawa and Nori Nakamura

Department of Molecular Biosciences, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima 732-0815, Japan

Kodaira, M., Asakawa, J. and Nakamura, N. Radiation-Induced Deletions in Mouse Spermatogonia are Usually Large (over 200 kb) and Contain Little Sequence Similarity at the Junctions. *Radiat. Res*. 187, 722–731 (2017).

Ionizing radiation can induce mutations, and the majority of radiation-induced mutations in mammalian cells are deletions. The most critical types of radiation-induced DNA damage are DNA double-strand breaks, and these breaks are repaired by either the homologous recombination (HR) pathway or the non-homologous end joining (NHEJ) pathway. The HR pathway is not as mutagenic as the NHEJ pathway, and it is expected that radiation-induced deletions would usually have little sequence similarity around the deletion junction points. Here we report sequence data from the regions around the rejoined junctions of 33 *de novo* copy-number mutations (27 deletions and 6 duplications) obtained from offspring sired by male mice that were irradiated at the spermatogonia stage and from nonirradiated controls. The results indicate that deletions can be classified into three major groups. In group 1, nine deletions were found to share long blocks of similar sequences (200–6,000 bp) at the junctions and the deletion size varied extensively (1 kb to 2 Mb) (e.g., illegitimate recombination). In group 2, five deletions shared short identical sequences (0–7 bp) at the junctions, and the deletion sizes were shorter than 200 kb (e.g., micro-homology-mediated repair). Additional three-deletion candidates of this group were also found but turned out to be inherited from mosaic parents. They are therefore not included in germline mutations. In group 3, twelve deletions shared little sequence similarity (only 0–2 bp) at the junctions (likely due to NHEJ repair) and deletion sizes were longer than 200 kb. Group 1 consisted of deletions found in both spontaneous and irradiated genomes and thus, were probably caused by spontaneous events during meiosis or DNA replication. Group 2 consisted mainly of deletions found in nonexposed genomes. Group 3 consisted primarily of deletions that occurred in the irradiated genomes. Among the duplications, we found no indication of any association with radiation exposures. These results indicate that large size (>200 kb) and little sequence similarity around the rejoined sites are likely to be a hallmark of radiation-induced deletions in mice. © 2017 by Radiation Research Society

INTRODUCTION

A series of large-scale epidemiological studies were conducted by the Radiation Effects Research Foundation (RERF) to monitor atomic bomb (A-bomb) survivors and to follow the radiation-induced health effects. The results clearly showed an increased risk of death from cancer and non-cancer diseases (*1*). To detect any radiation-induced effects in the next generation (transgenerational effects), various epidemiological and laboratory studies have been done to monitor the offspring of survivors; however, to date these studies have not had significant results. These studies include monitoring of the offspring of survivors for birth defects (*2*), chromosomal abnormalities (*3, 4*), electrophoretic studies of blood proteins (*5*) and mini- and microsatellite mutations (*6–8*), as well as epidemiological studies on mortality and cancer incidence (*9–11*). Currently, preliminary genome studies using blood samples of trios (samples from both of the parents and from the children), as well as longitudinal epidemiological studies, are ongoing. In addition, an F1 clinical study (FOCS) was recently initiated to examine the prevalence of multifactorial diseases in a cohort of 12,000 individuals. Here too, the results currently do not suggest the presence of any effects derived from parental radiation exposure (*12*), and do not correlate with any specific diseases or health conditions (i.e., diabetes, high blood pressure or hypercholesterolemia) (*13*).

In addition to studies of the children of A-bomb survivors, the offspring of childhood cancer survivors have also been evaluated. These patients' gonadal doses vary extensively depending on tumor site; for example, treatment of Wilms' tumor patients may lead to exposures of over 20 Gy in the ovaries. Nonetheless, even in these cases, no association has been observed between radiation exposures and an increased risk for chromosomal aberrations, Mendelian diseases or malformations in the offspring (*14–18*).

This information indicates that quantitative detection of genetic effects of radiation exposure in humans is an extremely difficult task. Therefore, we performed model experiments in mice to better understand the molecular characteristics of radiation-induced mutations. Historically, large-scale mouse studies conducted in the 1950s and 1960s utilized specific-locus tests. As the name indicates, 6 or 7

Editor's note. The online version of this article (DOI: 10.1667/RR14660.1) contains supplementary information that is available to all authorized users.

[1] Address for correspondence: Department of Molecular Biosciences, Radiation Effects Research Foundation, 5-2 Hijiyama Park, Minami-ku, Hiroshima, 732-0815, Japan; e-mail: kodaira@rerf.or.jp.

TABLE 1
Origin of 33 Copy Number Mutations Studied

Experiment no.	Method	Mouse strains		Radiation		No. of F1 mice screened		Number of copy number mutation				Ref.
								Amplification		Deletion		
		Sire	Dam	Type	Dose (Gy)	Control group	Exposed group	Control group	Exposed group	Control group	Exposed group	
1	RLGS	BALB/c	BALB/c	X rays	0, 3, 5	190	316	ND	ND	0	3	(*24*)
2	RLGS	B6C3	JF1	X rays	0, 4	502	505	ND	ND	1	5	(*26*)
3	aCGH	C57BL/6	C3H/He	γ rays	0, 4	100	100	2	4	5	8	(*27*)
4	aCGH	B6C3	JF1	X rays	4	ND	48	ND	0	ND	5	(*27*)
					Total	792	969	2	4	6	21	

Notes. RLGS = restriction landmark genome scanning or 2DE method; aCGH = array CGH; ND = not done.

specific genes were chosen as potential targets for mutagenesis, and these studies were an attempt to estimate mutation induction rates per locus per unit dose (*19, 20*). The observed mutation induction rates were found to vary considerably among the genes studied, and the mean rate is estimated as $1–2 \times 10^{-5}$/locus per Gy (*21*). In these studies, mutant alleles were regarded as large deletions when the animals bearing the mutant alleles under homozygous conditions were nonviable, and mutant alleles were regarded as small deletions or intragenic mutations if the homozygous animals were viable (*22*). These studies confirmed earlier results in fruit flies showing that radiation-induced mutations are primarily deletions (*21, 23*). However, detailed characterizations of the mutations were not available until the development of new molecular tools. For example, two-dimensional gel electrophoresis (2DE) method of ^{32}P-labeled DNA fragments at *Not*I sites, or the restriction landmark genome scanning (RLGS) method, can allow screening of approximately 1,000 sites in the genome for evidence of decreased copy numbers going from two to one copies (*24–26*). The more recent microarray-based comparative genomic hybridization (aCGH) method now enables us to screen over one million sites in the genome to detect deletions and duplications. Despite rapid progress in screening tools, however, mutation detection is still a laborious task, and the observed radiation effects often fail to reach a level of statistical significance (*27*).

Here we report sequence data from 33 *de novo* copy-number mutations (27 deletions and 6 duplications) detected in a series of studies that we performed previously. Because spontaneous mutations occur not only in mouse offspring of the control group, but also in the exposed group with nonirradiated maternal genomes or even in the irradiated paternal genome, the parental origins of the mutations were also determined. The overall results indicate that radiation-induced deletions are distinctive due to their large deletion sizes and very short- or no shared identical sequences around the rejoined junctions. These findings may be useful by allowing researchers to distinguish radiogenic deletions from spontaneous deletions.

MATERIALS AND METHODS

DNA Sources of Mutations

The 33 mutations studied here are *de novo* copy number variants (CNVs) detected in the offspring derived from irradiated sires using two genome-wide scanning techniques: the 2DE (i.e., RLGS) method and the aCGH method. Table 1 shows a summary of the four experiments.

Brief outlines of these mouse experiments are as follows. Purchased males were mated to nonirradiated females at 8 weeks of age. The males were then irradiated and mated again to the same females or new females after waiting for 8 weeks postirradiation (experiment 2 to experiment 4) or for 10 weeks postirradiation (experiment 1) so that offspring derived from irradiated spermatogonia could be obtained. High-molecular-weight DNA was extracted from spleens, livers and kidneys of F1 mice and their parents. For screening purposes, spleen DNA was used, whereas liver and kidney DNA was used to exclude apparent mosaic mutations caused by somatic mutations in early embryos. The procedures for mutant screenings are described by Asakawa *et al*. (*24, 26, 27*).

Copy Number Estimation

In experiments 1 and 2, in which the 2DE method was used for screening, mutations were detected as spots of ^{32}P-labeled DNA fragments, which had half the density of controls. In experiments 3 and 4, the aCGH method was used, in which the $\log_2$ ratio of the two fluorescent dyes used to label each genomic DNA fragment was calculated. Altered ratios at two or more adjacent probes on the genome were considered to indicate regions that were candidates for deletions or duplications. For confirmation of mutations, copy numbers were measured with quantitative PCR (qPCR) using primer sets designed to amplify the mutated regions. Sequence information for the primer sets is shown in Supplementary Table S1 (http://dx.doi.org/10.1667/RR14660.1.S1). The detailed qPCR protocols used to estimate copy number changes have been previously described elsewhere (*26*).

Rapid Mapping of Breakpoints with Array CGH

In experiments 1 and 2, deletions were screened with the 2DE method, but the deletion junctions were not determined. To determine the approximate localizations of the breakpoints, DNA samples bearing deletion mutations were subjected to aCGH tests (*27*). In experiments 3 and 4, deletions/duplications were detected with the aCGH method and therefore, the breakpoints were approximately localized (*27*). Using this initial information, the precise breakpoint locations were found by first determining the genomic regions in which the copy numbers had changed from two to one, or from one to two in the case of a deletion, and from two to three or three to two in the case of duplications. Breakpoints were mapped on the mouse

genome using probe information from the array UCSC Mouse (Mus musculus) Genome Browser Gateway NCBI37/mm9 (http://bit.ly/2opsQLs).

Amplification of Rejoined Junctions for Deletions and Duplications

To determine the sequences around the rejoined junctions of deletions or duplications, we amplified DNA segments that contained these junctions by using PCR. In the case of duplications, two primers were designed by assuming a simple, tandem duplication; namely, 5′ and 3′ primers were designed to extend in outward directions so that only DNA regions at the head-to-tail junction would be amplified (Supplementary Fig. S1A; http://dx.doi.org/10.1667/RR14660.1.S1). In the case of deletions, 5′ and 3′ primers were designed to amplify the regions inward towards the center of the deletion (Supplementary Fig. S1B). PCR-primer sets used to amplify the junctions are shown in Supplementary Table S2. Some experimental data describing sequences for the junctions amplified by PCR are shown in Supplementary Fig. S2. PCR conditions consisted of initial denaturation at 96°C for 3 min, followed by 29 or 30 cycles of denaturation at 96°C for 45 s, annealing at an optimized temperature between 60°C and 72°C for 45 s, and an extension at 72°C between 45 s to 2 min depending on the length of the amplified segments.

Molecular Analysis

Determination of the parental origins of the mutations. To determine the parental origins of the mutations, it was necessary to obtain parental single nucleotide polymorphism (SNP) information around the deleted or duplicated regions. For this purpose, we used two sources: http://www.informatics.jax.org (Jackson Laboratory, Bar Harbor, ME) and http://bit.ly/2nnjNcd (Sanger Institute, Cambridge, UK). We looked for SNPs within 20 kb in both the 5′ and 3′ directions from the junctions and amplified any region that contained the SNPs along with the rejoined junctions. The amplified products were sequenced and the results were compared with those from each parental mouse strain. Supplementary Table S3 (http://dx.doi.org/10.1667/RR14660.1.S) shows primer information for SNP genotyping and the PCR results.

Molecular characterization of rejoined sites. The amplified products were sequenced by using an ABI Prism 310 Genetic Analyzer (Applied Biosystems, Tokyo, Japan). Prior to aligning the amplified sequences on the reference genome, the sequences were screened to exclude interspersed repeat sequences (i.e., *LINE*s, *SINE*s or transposed elements, etc.) by using RepeatMasker (http://www.repeatmasker.org). Also segmental duplications, i.e., 1–200 kb nonrepeat sequence blocks, with base sequences that are in close agreement (>90%) with each other and which are located at different places in the genome (annotated on the reference genome, ''Segmental Dups'', in the UCSC Genome Browser), were also excluded using the UCSC Table Browser (http://bit.ly/2opC62b). Subsequently, amplified sequences not annotated by those programs were aligned on the mouse reference genome using ''blast'' methods within the NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi), UCSC ''mouse blat search'' (http://bit.ly/2nE5i5z), or with the homology search program, ''Nucleotide vs. Nucleotide Homology'' in GENETYX version 12 (GENETYX Corp., Tokyo, Japan). The alignments made it possible to determine base sequences at the junctions and sequence similarities between the 5′ and 3′ breakpoints.

Number of genes involved in a deletion or duplication. We used ''genes'' in Ensembl, which includes Ensembl and Havana genes, Ensembl short noncoding RNA genes, Merged Ensembl and Havana lincRNA genes. The assembly of Ensembl genes in this article is NCBI37/mm9, the April 2007 Mus musculus (strain C57BL/6J) high coverage assembly (http://bit.ly/2nnkPVG). Detailed information concerning the genes involved in the deletions is listed in Supplementary Table S4 (http://dx.doi.org/10.1667/RR14660.1.S1).

RESULTS

Characterization of Duplications

Among the six duplications examined, three occurred in the paternal genome (two in an irradiated genome and one in a nonirradiated genome), two in the maternal nonirradiated genome, and one for which parental origins could not be determined. These results do not indicate that they were radiation induced. Genomic locations and base sequence information at the junctions are given in Supplementary Table S5-1 (http://dx.doi.org/10.1667/RR14660.1.S1). Sequence data around the junctions showed similar sequence blocks varying from 0 to 1,500 bp. The duplications were all considered as duplicated in tandem after successful amplifications by PCR using primer pairs that were designed assuming head-to-tail tandem duplications (Supplementary Fig. S1-A). The results do not necessarily mean, however, that the duplicated segments are intact, and one duplication (CGH-CA1) was found to contain a deletion within the duplicated segment (for details, see Supplementary Fig. S3; http://dx.doi.org/10.1667/RR14660.1.S1).

Characterization of Deletions

Among the 27 deletions observed in the offspring, 15 occurred in paternal genomes (13 in irradiated genomes and two in nonirradiated genomes), six in maternal genomes (all of which were nonirradiated), and for six deletions, the parental origins could not be determined, but they occurred in the offspring of the exposed group (Table 2). In three deletions, junction sites could not be amplified due to the presence of repeat sequence blocks such as a TCR gene cluster and segmental duplications; thus, these deletion sizes are shown as ranges (CGH-ED3, -ED4 and -ED5). Junction sites for the remaining 24 deletions were successfully amplified. Among these, 18 produced information on base sequences around the rejoined junctions. From these results, it was possible to map a pair of breakpoints involved in each deletion on the normal genome, and subsequently examine sequence similarities at the junctions and estimate deletion sizes at the single nucleotide level (Table 2 and Supplementary Table S5-2; http://dx.doi.org/10.1667/RR14660.1.S1). It is important to note that breakpoints do not necessarily represent the location of the initial double-strand breaks (DSBs) caused by exposure to radiation. In the remaining six cases, the exact size of the deletions could not be determined due to the presence of repeat sequence blocks. For example, one (CGH-ED13) had a deletion and an insertion at the junction, and the deletion junction had LINE sequences (a group of transposable elements that exists as 100,000 copies in mammalian genomes). Three deletions (RLGS-CD1, CGH-CD1 and -ED7) had rejoined junctions within LINE sequences varying from 2 to 6 kb, and two deletions (CGH-CD6 and -ED15) had junctions within repeat sequence blocks, which are not interspersed (only two copies per haploid genome) and were approximately 200 and 2,000 bp, respectively. The deletion sizes of

TABLE 2
Summary of 27 Deletions

Parental origin	ID	Radiation dose	Deletion size (kb)	Similar sequence (bp)	Nature of similar sequence	No. of genes involved
Group 1						
P	RLGS-CD1	0 Gy	2,329–2,333	~2,000	LINE	9
P	CGH-ED3	4 Gy	735–982	>200	TCR, SD	87
P	CGH-CD1	0 Gy	137–150	~6,000	LINE	2
M	CGH-ED13	0 Gy	51–62	2	CT	1
				~250	LINE	
P	CGH-ED15	4 Gy	21–25	~2,000	Homologous sequence	2
M	CGH-CD6	0 Gy	1.5–1.9	~200	Homologous sequence	1
nd	CGH-ED4	4 Gy	53–68	>1,000	SD	1
nd	CGH-ED5	4 Gy	64–89	>1,000	SD	2
nd	CGH-ED7	4 Gy	12–25	~2,000	LINE	0
Group 2						
P	CGH-ED8	4 Gy	4.6	7	ATATC(T/–)C	0
M	CGH-ED14	0 Gy	47	6	AAG(–/C)TT	2
M	CGH-CD2	0 Gy	21	2	AG	2
M	CGH-CD4	0 Gy	13	4	TGCT	0
M	CGH-CD5	0 Gy	4.4	0		0
Group 3						
P	RLGS-ED7	4 Gy	13,032	1	T	59
P	RLGS-ED4	4 Gy	10,559	0		213
P	CGH-ED1	4 Gy	4,967	0		29
P	RLGS-ED5	4 Gy	4,696	1	G	70
P	RLGS-ED8	4 Gy	4,256	0		34
P	RLGS-ED6	4 Gy	1,933	0		27
P	CGH-ED9	4 Gy	593	1	C	5
P	CGH-ED10	4 Gy	568	0		9
P	CGH-ED6	4 Gy	323	1	G	1
nd	RLGS-ED2	3 Gy	2,235	1	A	5
nd	RLGS-ED1	3 Gy	1,965	0		10
nd	RLGS-ED3	5 Gy	359	2	CT	10
P	CGH-ED2	4 Gy	2,668	10	GAAAGT(–/C)TAT	6

Notes. P = paternal; M = maternal; nd = not determined.

these mutations are also expressed as ranges in Table 2, Supplementary Tables S5-3 and S5-4 and Supplementary Fig. S4. Detailed base sequence information at the junction sites is given in Supplementary Table S5.

Deletion Size and Shared Similar Sequences at the Junctions

From the deletion size and the length of the similar or identical sequences around the rejoined junctions, 21 *de novo* deletions, which were confirmed or suspected to be paternal in origin, are plotted in Fig. 1A. Among these, 13 were confirmed to have originated in the irradiated paternal genome, and two in the nonirradiated paternal genome. For the remaining six, the origin could not be determined because the same strain of mice was used for both parents, but they all occurred in the offspring of the irradiated group (i.e., they occurred in either the exposed paternal or nonexposed maternal genome). In the same way, confirmed maternal as well as possibly maternally derived deletions are plotted in Fig. 1B, which includes six deletions that occurred in nonirradiated maternal genomes, and six deletions that occurred in the offspring of the exposed group, but for which the parental origins could not be determined, as mentioned above.

Classification of Deletions

Each deletion was classified into one of three major groups according to the deletion size and the length of similar sequences around the two rejoined sites (breakpoints).

Group 1. Deletions in this group were easily recognized by two distinguishing characteristics: the presence of relatively long blocks of DNA bearing similar sequences in which the two junction points are mapped (similar sequence blocks of 200 bp to 6 kb) and large variations in the deletion size (approximately 2 kb to over 1 Mb) (Fig. 1A and B, Table 2). Nine deletions fell in this group: two in the exposed paternal genome, two in the nonirradiated paternal genome, two in the nonirradiated maternal genome and three with undetermined origins. One deletion (CGH-ED13) was a complex type, and consisted of an insertion and a deletion, and the latter occurred between two LINE blocks (Fig. 1B, see the open square). Possible mechanisms leading to the deletion are shown in Supplementary Fig. S4 (http://dx.doi.org/10.1667/RR14660.1.S1).

Group 2. This group of deletions consists of small sizes (below 200 kb long) and identical sequences of 0–7 bp around the junction points when mapped on the normal

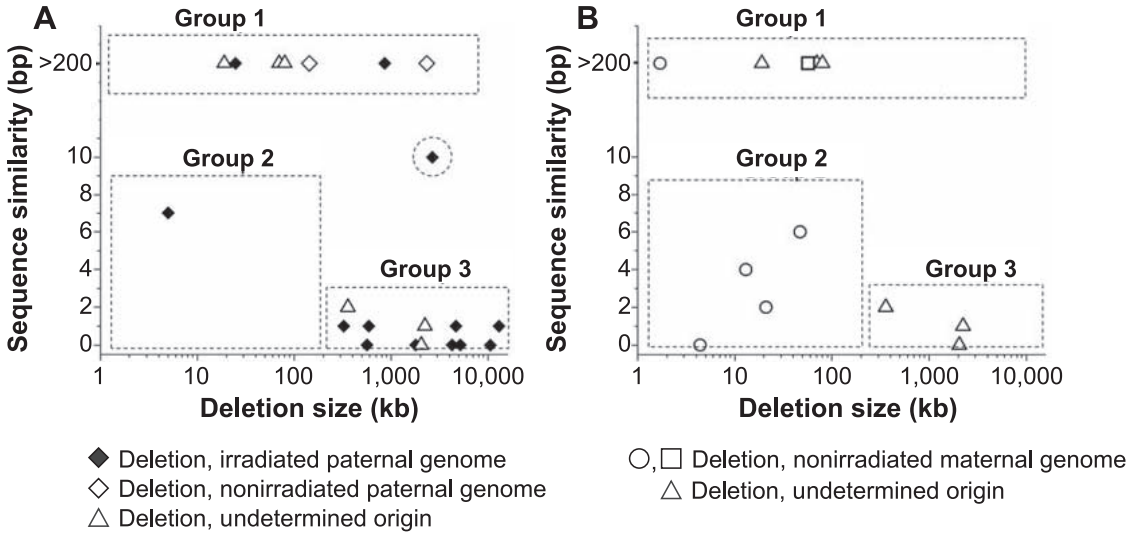

FIG. 1. *De novo* deletions were plotted according to deletion size (the X-axis) and length of similar or identical sequences around the junctions (the Y-axis). Panel A: Deletions occurring in irradiated (◆) or nonirradiated (◇) paternal genome. Panel B: Deletions occurring in nonirradiated maternal genome (○, □). Open triangles (△) indicate deletions for which parental origins could not be determined.

genome. The upper limit of the deletion size was tentatively set at 200 kb for two reasons: 1. Human data on polymorphic copy number changes (spontaneously occurred sometime in the past) showed that deletions were generally smaller than 200 kb (*28*) (see also Discussion); and 2. Other mouse deletions to be classified as group 3 (see below) clustered at the right bottom corner in Fig. 1A. They mainly consisted of deletions that occurred in irradiated paternal genomes, and the smallest one was 323 kb long. A total of five deletions fell in with group 2; one was found in the irradiated paternal genome and four were found in the nonirradiated maternal genome. Most of the deletions appear to be simple ones in which the two rejoined sequences are mapped in two blocks of DNA bearing short identical sequences. It is noted further that one deletion, which had no sequence similarity at the junctions (i.e., from NHEJ repair), accompanied an insertion (CGH-CD5 had an insertion of a 5 kb LTR segment; Supplementary Table S5-2).

Group 3. Deletions in this group are characterized by a large size (from 300 kb up to over 10 Mb) and little sequence similarity (only 0–2 bp) around the junction sites. Twelve deletions fell in with this group, of which nine were derived from the irradiated paternal genome. The remaining three were found in the offspring of the exposed group but the parental origins could not be determined (Fig. 1A and B).

One exceptional case, which did not fit in any of the three groups, was deletion CGH-ED2. This deletion was made up of a long deletion (approximately 2,600 kb) and relatively long identical sequence blocks of 10 bases around the junction sites (enclosed with a broken circle in Fig. 1A). It was tentatively classified as a member of group 3 (Table 2), but might be a member of group 1.

Mean Number of Genes Involved in a Deletion or Duplication

Figure 2 shows plots of the estimated number of genes involved in a deletion or duplication versus deletion or duplication size, respectively. In this context, a gene is defined in a broader sense: it includes not only ordinary protein-coding genes but also pseudogenes, noncoding

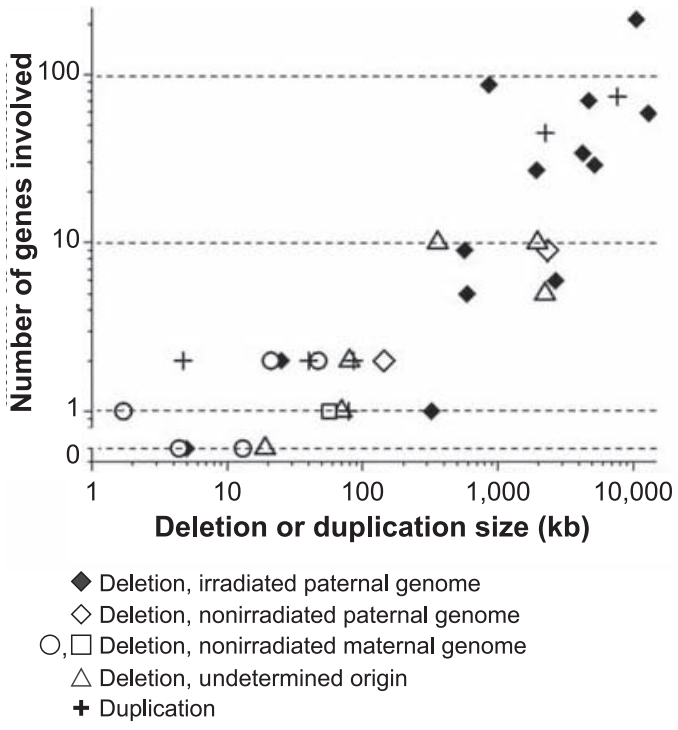

FIG. 2. Estimated number of genes involved in a deletion (the symbols are the same as those given Fig. 1 legend) or in a duplication (+).

RNA genes and lincRNA genes, as described in the Materials and Methods section. Briefly, a 10–100 kb deletion contained 0–2 genes, a 1,000-kb deletion could have involved 10–100 genes, and a 10 Mb deletion had lost approximately 100 genes (Supplementary Table S4; http://dx.doi.org/10.1667/RR14660.1.S1). It is impressive to find offspring bearing large deletions encompassing over 1–10 Mb but without any apparent phenotypic effect; i.e., mutant RLGS-ED4 lost 213 genes, CGH-ED3 lost 87 genes and RLGS-ED7 lost 59 genes. However, it is worth noting that the deletion RLGS-ED4 included a cluster of olfactory genes in chromosome 9 and deletion CGH-ED3 lost a cluster of TCR-Vα genes. These two gene clusters are expressed only in a limited number of organ(s) and the deletions are maintained under hemizygous conditions (i.e., one copy still remains), which would explain the apparent lack of phenotypic effect at 3 weeks of age when the animals were sacrificed (*27*). It is also important to note that the results shown in Fig. 2 imply that the number of genes lost in a deletion (various symbols) and the number of genes gained by a duplication (+ symbols) per unit length of affected DNA appear to be similar to each other.

Mutations Originated from Mosaic Parents

During the course of the study, five more mutation candidates were detected by the screening, however, they turned out to be inherited from mosaic parents after deletion-specific PCR tests on parental spleen or other somatic cells [two duplications and three deletions, Supplementary Table S6 and Fig. S5 (http://dx.doi.org/10.1667/RR14660.1.S1); these mutations are not included in Figs. 1 and 2]. Among the three deletions, one was found in the irradiated paternal genome (but occurred prior to the irradiation: 87 kb) and two in nonirradiated maternal genomes (63 and 155 kb). All three deletions had characteristics of group 2 deletions.

DISCUSSION

Types of Deletions and the Molecular Mechanisms Involved

In this study, we found that deletion mutations could be classified into three groups according to the size of the deletion and the length of the similar or identical sequences at the junctions.

Group 1 deletions. There are two characteristics of this group: the joined junctions contained long blocks of similar sequences (over 200 bp), and deletion sizes were highly variable (approximately 2 kb to over 2 Mb). They were found in both irradiated and nonirradiated genomes (Fig. 1A). There are two possible molecular mechanisms to consider. One possible mechanism is non-homologous (i.e., illegitimate) recombination between two homologous chromosomes, which is a unique characteristic of meiosis. Another possible mechanism involves intrachromosomal deletions caused by unequal sister chromatid exchanges (USCEs) or intrastrand recombinations. They may occur at any time during mitotic cell cycles in spermatogonia or progenitor cells, as well as during meiosis. In either event, these are termed nonallelic homologous recombination (NAHR), which is suggested to require >200 bp of similar sequence blocks (*29*). Although the possibility can be raised that exposure to radiation has the potential to increase the frequency of meiotic recombination through the induction of DSBs in DNA, which are required for the initiation of meiotic recombination, several lines of evidence do not support this hypothesis. Specifically, radiation is reported not to increase the frequency of meiotic recombination in mice (*30*), and the F1 mice studied here were not designed to be derived from irradiated meiotic cells, but from more immature spermatogonia cells. Furthermore, group 1 deletions were found in both nonirradiated and irradiated male genomes (Fig. 1A), as well as in nonirradiated female genomes (Fig. 1B). In short, group 1 deletions cannot be considered as a hallmark of radiation-induced deletions. This result in turn indicates that the deletion size alone would not be sufficient to classify a deletion as radiogenic or not, and further information on the sequence homology at the deletion junction is important.

Group 2 deletions. This group of deletions is represented by a relatively short deletion size (<200 kb) and the presence of short blocks of identical sequences (microhomology of 0–7 bp) around the junction. These deletions are likely to be spontaneous for a number of reasons. First, deletions in this group are common among deletions that occurred in nonirradiated female genomes (4 out of 4 nongroup-1 deletions) while they are rare among those that occurred in the irradiated paternal genome (only one case out of 14 non-group-1 deletions) (Table 2). Second, the presence of short similar sequences at the deletion junction appears as an indication of an S-phase event, and S phase is not a sensitive cell cycle stage for mutation induction by radiation (*31*). Third, the largest deletion in group 2 that we could observe was 155 kb (CGH-ED11), which is one of the three deletions inherited from mosaic parents (i.e., not a germline event) and thus is spontaneous. Fourth, 302 polymorphic deletions (selected size of >400 bp), which were detected in the genomes of three human individuals, fell in with this area, as shown in Fig. 3 (*28*, *32*).

Group 3 deletions. Compared to the group 2 deletions, group 3 deletions are characterized by their large size (>300 kb) and shorter stretches of identical sequences (microhomology of 0–2 bp) around the junctions. These are most likely induced through the NHEJ repair pathway, which is believed to be the most common mechanism involved in DSB repair in cells in a resting stage (G_0 or G_1). The NHEJ mechanism may also be active in cycling cells along with homology-related repair (HR), which is apparently less mutagenic (*33*). Thus, it is not surprising that nine deletions out of 11 nongroup-1 deletions that occurred in the exposed paternal genomes fell in with this group. These considerations strongly indicate that group 3

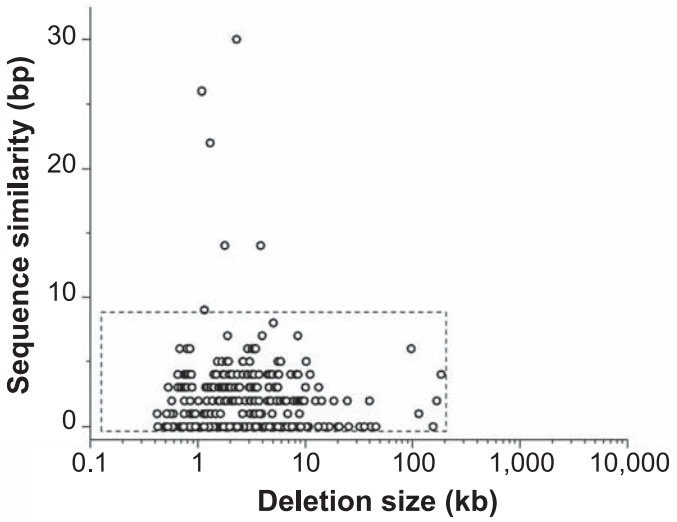

FIG. 3. Plots of 302 polymorphic deletions (selected size of >400 bp) detected in the genomes of three human individuals. Data were obtained from ref. (*28*). The rectangle defined by broken lines represents the area for the group 2 deletions.

deletions are primarily radiation-induced deletions. In addition, there were three deletions whose parental origins were not determined (Fig. 2 and Table 2) because these offspring were born to parents that were both from the same BALB/c strain. Although parental origins of the three deletions could not be determined, they appear to be radiation-induced since the deletion sizes are all >300 kb and shared little identical sequences at the junctions, which fits the hallmarks of radiation-induced deletions. Further evidence showing that cells at the G_1 stage are most sensitive to radiation for the induction of mutations in the *HPRT* gene (*31*) supports the interpretation that group 3 deletions having the fingerprints of NHEJ repair are likely to be radiogenic. In support of this interpretation, in cultured human fibroblasts, 12 out of 14 radiation-induced deletions shared only 0–2 bp sequence homologies at the junctions (*34*).

It may be worth asking why group 2 deletions (most likely to be spontaneous deletions) bore short blocks of similar sequences at the deletion junctions and had deletion sizes that appeared distinctively smaller compared to those that were radiation induced (group 3 deletions). Regarding the first point, the presence of short stretches of similar sequences indicates that deletions of this group occurred primarily in the S phase of the cell cycle. When DNA replication is temporally stalled during S phase, growing single-strand DNA unwinds and the 3′ end switches the template (single-strand migration) to a nascent strand of the sister chromatids so that the damaged site may be bypassed, which may give rise to sister chromatid exchanges (SCEs). In this situation, if illegitimate short blocks of similar sequences were utilized, deletion mutations are expected to occur. An alternative pathway is to use an error-prone DNA polymerase(s) to forcibly replicate the damaged site, which tends to induce base-change mutations. However, our study was not designed to detect such mutations. Regarding the second point (shorter deletion sizes), for the nascent DNA to switch templates, the target short stretches of similar sequences must be in close proximity to the damaged sites; i.e., within a distance shorter than the length of the unwound single-strand DNA. It seems that the target sequences are often found within the same DNA segment that is under replication (i.e., within the same replicon or the same cluster of replicons). In short, this may be a reflection of the fact that the mean replicon size is of the order of 100 kb (*35, 36*). Evidence that ionizing radiation is not effective in inducing SCEs in G_0 lymphocytes (*37*) may also support the notion that the group 2 deletions are not radiogenic.

Comparisons with Other Genomic Data

Adewoye *et al*. recently reported aCGH results from studies of deletion mutations detected in the offspring of male mice irradiated at post- and pre-meiotic germ cells stages (*38*). Since the sequence information at the junctions is not reported, it is not possible to plot the deletions in a two-dimensional format as shown in Fig. 1, but only by deletion size. The deletion size is distributed as two clusters: one was below 100 kb (8 deletions that were confirmed as derived from irradiated genomes plus 4 deletions whose origins were not defined), and the other was over 1,000 kb (5 deletions all derived from irradiated genomes). There was an apparent lack of deletions between 100 and 1,000 kb long. The smaller size group comprises our group 2 and a part of the group 1 deletions, and the larger size group includes our group 3 and a part of the group 1 deletions. When 13 deletions detected in our experiment 3 (100 offspring of the control group and 100 offspring of the irradiated group) were classified according to deletion size only (ignoring the parental origins), there were 8 deletions below 100 kb and 5 deletions over 100 kb {Supplementary Tables S5-2 and S5-3; http://dx.doi.org/10.1667/RR14660.1.S1 [see also Table 2 in ref. (*27*)]}. The corresponding numbers from Adewoye *et al*. are 12 and 5 in 93 offspring from the control group and 169 offspring from the irradiated group, respectively. Therefore, the two sets of data are in reasonable agreement.

In another study, Behjati *et al*. examined 12 radiation-associated second malignancies using the whole genome sequencing (WGS) method and reported mutational signatures of ionizing radiation; namely, radiation-associated cancers had on average 201 excess indels (size 1–100 bp), which often accompanied microhomology at the junctions (*39*). While it is well established that radiation-induced deletions are often quite large (*21–23, 40–43*), only short indels were reported in that study. This is probably attributable to the WGS technique because it is not best suited to detect large deletions (*44*), which is also supported by the observation that large inversions were significantly elevated (up to 100 Mb) but deletions, counterparts of

inversions, of that size were not (*39*). A combination of CGH and WGS techniques would be required in future studies to better understand a radiation fingerprint in somatic and germ cells.

Why do Radiation-Induced Deletions Tend to Exceed 100 kb?

Among the deletion mutations that were confirmed as derived from irradiated paternal genomes, those that fell into group 2 (less than 100 kb) were rare; there was only one deletion out of 13 (Fig. 1A). Furthermore, since this deletion in group 2 shared identical sequences of 7 bp at the junctions compared to ordinary 0–2 bp commonly observed among deletions in group 3, it seems likely that this group 2 deletion is a spontaneous deletion, which occurred in the irradiated genome. Although one may raise the possibility that the current microarray CGH method might not be sensitive enough to detect small deletions and therefore, the current results are biased toward the selective detection of larger deletions, we hypothesize that the current CGH data are not extensively biased. Specifically, our previously utilized DNA-2DE (or RLGS) method (experiments 1 and 2 in Table 1) detects deletion mutations as half-density spots of ^{32}P-labeled DNA fragments in autoradiograms, and therefore its detection efficiency is not influenced by deletion size. In fact, deletion mutations that occurred in microsatellite sequences were detected by the method (*24*). Nonetheless, no deletions smaller than 100 kb were found among 9 deletions (mutant ID marked with ''RLGS'': Table 2). These considerations led us to hypothesize that deletions smaller than 100 kb are not frequently induced by radiation. This may be reasonable if one considers that genomic DNA is organized into DNA loop domains that are approximately 100 kb long (*45, 46*). Then, if two DSBs are induced by a single track of an electron within a single loop domain (several tens of kb apart on stretched DNA), the breaks may often be physically too far apart to be able to be rejoined with each other. In contrast, if the two DSBs occurred in two different loop domains but are located in close spatial proximity, they can join and result in large structural alterations leading to the formation of intrachromosomal events (deletions or inversions) as well as interchromosomal events (dicentrics or translocations). Future information describing gross chromatin structures within a so-called chromosomal domain in interphase nuclei (*45*) would help our understanding of the basic mechanisms involved in radiation-induced deletions.

Mosaic Mutations

Three deletion candidates were observed but turned out to be inherited from mosaic parents. Obviously, the mutations occurred at an early stage of embryogenesis in the parents, and therefore they are not appropriate to be included in germline mutations. Nonetheless, they provide useful information that is not otherwise available. Because they all fell in the defined area for group 2 deletions in Fig. 1, setting the borderline between group 2 and group 3 deletions at approximately 200 kb appears to be satisfactory. In addition, the presence of mosaic parents, especially in the control group, has been recognized to be important in estimating the genetic risks from radiation because such parents give rise to mutant offspring as a cluster, which makes the interpretation of the results complicated (*47*). Furthermore, information on mosaicism is important in genetic counseling because mosaic parents are at higher risk of giving rise to recurrent affected offspring compared to parents with germline mutations. In this regard, Campbell *et al*. screened 100 families with children affected by genomic deletions that were previously ascertained by SNP and aCGH platforms, and looked for the presence of mosaic parents. They found four parental mosaics (4/100) (*48*). The corresponding numbers in our study are three parental mosaics in addition to 14 most likely spontaneous *de novo* deletions (9 deletions in group 1 plus 5 deletions in group 2), or 3/17. The large difference in the fraction of mosaic cases between the two studies may be attributed to different approaches in selecting the deletions (genomic screening vs. clinical screening), as well as to the difference in the mean deletion size. Namely, in our study, deletions larger than 1,000 kb were rare among spontaneous deletions, and there was no deletion that showed an apparent sign of detrimental effect. In contrast, in the Campbell study, deletions were selected by the presence of clinical symptoms and the deletion size varied widely (100–10,000 kb). The results seem to indicate that negative selective forces are operating *in vivo* to remove deletion-bearing somatic cells, especially in blood cells, which divide most frequently. In other words, potential contributions of mosaic parents in a family with affected children may be even higher than is currently detected.

Application of the Current Results to Possible Human Cases

Another important aspect of the Campbell study mentioned above is that shared identical sequences at the junctions of deletions in human genomic disease patients varied from 0 to 5 bp irrespective of the deletion size (100–10,000 kb, Supplementary Fig. S6; http://dx.doi.org/10.1667/RR14660.1.S1). This result indicates that if the current findings are to be applied to possible human cases, deletions that fall into groups 1 and 2 can be classified as most likely derived from spontaneous events. In contrast, deletions which fall into group 3 are expected to consist of a mixture of spontaneous and radiation-induced ones, and the length of identical sequences at the junctions may help us distinguish one from the other; namely, the longer the sequence (e.g., $\geq$3 bp), the higher the probability of spontaneous origins. Further collections of radiation-induced deletions in mice would help to improve our understanding of the nature of radiation-induced deletions in germ cells.

SUPPLEMENTARY INFORMATION

Table S1. Primer information for copy number estimation.

Table S2. Primer information for amplification of rejoined-junctions.

Table S3. Estimation of parental origins of mutations by SNP analyses.

Table S4. Genes involved in each mutation.

Table S5. Molecular characteristics at rejoined junctions of 33 *de novo* copy number mutations.

Table S6. Molecular characteristics of mutations derived from mosaic parents (inherited mutations).

Fig. S1. Design strategy of PCR-primer set.

Fig. S2. Some experimental data for the junctions amplified by PCR.

Fig. S3. A complex duplication (CGH-CA1).

Fig. S4. Mutation process of one complex deletion (CGH-ED13) in Table S5-4.

Fig. S5. Five mutations originated from mosaic parents.

Fig. S6. Summary data of genome disease patients (*48*) and mouse data of group 3 deletions.

ACKNOWLEDGMENTS

We thank A. Miura, T. Tsuji, M. Imanaka, Y. Nakamoto, J. Kitamura, E. Nishikori and J. Kaneko for their excellent technical assistance and Dr. L. Kapp for his careful reading of the manuscript. The Radiation Effects Research Foundation (Hiroshima and Nagasaki, Japan) is a public interest foundation funded by the Japanese Ministry of Health, Labor and Welfare and the U.S. Department of Energy. This work was supported by RERF Research Protocols RP 7-85 and RP 1-10. A part of this work was supported by JSPS KAKENHI (grant no. 20312033 to JA). The views of the authors do not necessarily reflect those of the two governments.

Received: October 21, 2016; accepted: March 2, 2017; published online: April 18, 2017

REFERENCES

1. Ozasa K, Shimizu Y, Suyama A, Kasagi F, Soda M, Grant EJ, et al. Studies of the mortality of atomic bomb survivors, Report 14, 1950–2003: an overview of cancer and noncancer diseases. Radiat Res 2012; 177:229–43. Erratum: Radiat Res 2013; 179:e40–1.
2. Otake M, Schull WJ, Neel JV. The effects of parental exposure to the atomic bombings of Hiroshima and Nagasaki on congenital malformations, stillbirths and early mortality among their children: a reanalysis. Radiat Res 1990; 122:1–11.
3. Awa AA, Honda T, Neriishi S, Sofuni T, Shimba H, Ohtaki K, et al. Cytogenetic study of the offspring of atomic bomb survivors. In: Obe G, Baseler A, Editors. Cytogenetics: basic and applied aspects. Berlin: Springer-Verlag; 1987. p. 166–183.
4. Neel JV, Schull WJ. The children of atomic bomb survivors; a genetic study. Washington, DC: National Academy Press; 1991.
5. Neel JV, Satoh C, Goriki, K, Asakawa J, Fujita M, Takahashi N, et al. Search for mutations altering protein charge and/or function in children of atomic bomb survivors: final report. Am J Hum Genet 1988; 42:663–76.
6. Kodaira M, Satoh C, Hiyama K, Toyama K. Lack of effects of atomic bomb radiation on genetic instability of tandem-repetitive elements in human germ cells. Am J Hum Genet 1995; 57:1275–83.
7. Kodaira M, Izumi S, Takahashi N, Nakamura N. No evidence of radiation effect on mutation rates at hypervariable minisatellite loci in the germ cells of atomic bomb survivors. Radiat Res 2004; 162:350–6.
8. Kodaira M, Ryo H, Kamada N, Furukawa K, Takahashi N, Nakajima H, et al. No evidence of increased mutation rates at microsatellite loci in offspring of A-bomb survivors. Radiat Res 2010; 173:205–13.
9. Izumi S, Suyama A, Koyama K. Radiation-related mortality among offspring of atomic bomb survivors: a half-century of follow-up. Int J Cancer 2003; 107:292–7.
10. Izumi S, Koyama K, Soda M, Suyama A. Cancer incidence in children and young adults did not increase relative to parental exposure to atomic bombs. Br J Cancer 2003; 89:1709–13.
11. Grant EJ, Furukawa K, Sakata R, Sugiyama H, Sadakane A, Takahashi I, et al. Risk of death among children of atomic bomb survivors after 62 years of follow-up: a cohort study. Lancet Oncol 2015; 16:1316–23.
12. Fujiwara S, Suyama A, Cologne JB, Akahoshi M, Yamada M, Suzuki G, et al. Prevalence of adult-onset multifactorial disease among offspring of atomic bomb survivors. Radiat Res 2008; 170:451–7.
13. Tatsukawa Y, Cologne JB, Hsu WL, Yamada M, Ohishi W, Hida A, et al. Radiation risk of individual multifactorial diseases in offspring of the atomic-bomb survivors: a clinical health study. J Radiol Prot 2013; 33:281–93.
14. Winther JF, Boice JD Jr, Frederiksen K, Bautz A, Mulvihill JJ, Stovall M, et al. Radiotherapy for childhood cancer and risk for congenital malformations in offspring: a population-based cohort study. Clin Genet 2009; 75:50–6.
15. Winther JF, Olsen JH, Wu H, Shyr Y, Mulvihill JJ, Stovall M, et al. Genetic disease in the children of Danish survivors of childhood and adolescent cancer. J Clin Oncol 2012; 30:27–33.
16. Signorello LB, Mulvihill JJ, Green DM, Munro HM, Stovall M, Weathers RE, et al. Stillbirth and neonatal death in relation to radiation exposure before conception: a retrospective cohort study. Lancet 2010; 376:624–30.
17. Green DM, Sklar CA, Boice JD Jr, Mulvihill JJ, Whitton JA, Stovall M, et al. Ovarian failure and reproductive outcomes after childhood cancer treatment: results from the Childhood Cancer Survivor Study. J Clin Oncol 2009; 27:2374–81.
18. Green DM, Lange JM, Peabody EM, Grigorieva NN, Peterson SM, Kalapurakal JA, et al. Pregnancy outcome after treatment for Wilms tumor: a report from the national Wilms tumor long-term follow-up study. J Clin Oncol 2010; 28:2824–30.
19. Russell WL, Russell LB, Kelly EM. Radiation dose and mutation frequency. Science 1958; 128:1546–50.
20. Russell WL, Kelly EM. Mutation frequencies in male mice and the estimation of genetic hazards of radiation in men. Proc Natl Acad Sci U S A 1982; 79:542–4.
21. United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Hereditary effects of radiation. New York: United Nations; 2001.
22. Searle AG. Mutation induction in mice. Adv Radiat Biol 1974; 4:131–207.
23. Sankaranarayanan K. Genetic effects of ionizing radiation in multicellular eukaryotes and the assessment of genetic radiation hazard in man. Amsterdam: Elsevier Biomedical Press; 1982.
24. Asakawa J, Kuick R, Kodaira M, Nakamura N, Katayama H, Pierce D, et al. A genome scanning approach to assess the genetic effects of radiation in mice and humans. Radiat Res 2004; 161:380–90.
25. Asakawa J, Nakamura N, Katayama H, Cullings HM. Estimation of mutation induction rates in AT-rich sequences using a genome

scanning approach after X irradiation of mouse spermatogonia. Radiat Res 2007; 168:158–67.

26. Asakawa J, Kodaira M, Cullings HM, Katayama H, Nakamura N. The genetic risk in mice from radiation: an estimate of the mutation induction rate per genome. Radiat Res 2013; 179:293–303.

27. Asakawa J, Kodaira M, Miura A, Tsuji T, Nakamoto Y, Imanaka M, et al. Genome-wide deletion screening with the array-CGH method in mouse offspring derived from irradiated spermatogonia indicates that mutagenic responses are highly variable among genes. Radiat Res 2016; 186:568–76.

28. Conrad DF, Bird C, Blackburne B, Lindsay S, Mamanova L, Lee C, et al. Mutation spectrum revealed by breakpoint sequencing of human germline CNVs. Nat Genet 2010; 42:385–91.

29. Inoue K, Lupski JR. Molecular mechanisms for genomic disorders. Annu Rev Genomics Hum Genet 2002; 3:199–242.

30. Barber R, Plumb M, Smith AG, Cesar CE, Boulton E, Jeffreys AJ, et al. No correlation between germline mutation at repeat DNA and meiotic crossover in male mice exposed to X-rays or cisplatin. Mutat Res 2000; 457:79–91.

31. Burki JH. Ionizing radiation-induced 6-thioguanine-resistant clones in synchronous CHO Cells. Radiat Res 1980; 81:76–84.

32. Mills RE, Walter K, Stewart C, Handsaker RE, Chen K, Alkan C, et al. Mapping copy number variation by population-scale genome sequencing. Nature 2011; 470:59–65.

33. Lieber MR. The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 2010; 79:181–211.

34. Arlt MF, Rajendran S, Birkeland SR, Wilson TE, Glover TW. Copy number variants are produced in response to low-dose ionizing radiation in cultured cells. Environ Mol Mutagen 2014; 55:103–13.

35. Berezney R, Dubey DD, Huberman JA. Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma 2000; 108:471–84.

36. Chagin VO, Casas-Delucchi CS, Reinhart M, Schermelleh L, Markaki Y, Maiser A, et al. 4D Visualization of replication foci in mammalian cells corresponding to individual replicons. Nat Commun 2016; 7:11231.

37. Littlefield LG, Colyer SP, Joiner EE, DuFrain RJ. Sister chromatid exchanges in human lymphocytes exposed to ionizing radiation during G0. Radiati Res 1979; 78:514–21.

38. Adewoye AB, Lindsay SJ, Dubrova YE, Hurles ME. The genome-wide effects of ionizing radiation on mutation induction in the mammalian germline. Nat Commun 2015; 6:6684.

39. Behjati S, Gundem G, Wedge DC, Roberts ND, Tarpey PS, Cooke SL, et al. Mutational signatures of ionizing radiation in second malignancies. Nat Commun 2016; 7:12605.

40. Aghamohammadi SZ, Morris T, Stevens DL, Thacker J. Rapid screening for deletion mutations in the hprt gene using the polymerase chain reaction: X-ray and alpha-particle mutant spectra. Mutat Res 1992; 269:1–7.

41. Park MS, Hanks T, Jaberaboansari A, Chen DJ. Molecular analysis of gamma-ray-induced mutations at the hprt locus in primary human skin fibroblasts by multiplex polymerase chain reaction. Radiat Res 1995; 141:11–8.

42. Schwartz JL, Jordan R, Sun J, Ma H, Hsieb AW. Dose-dependent changes in the spectrum of mutations induced by ionizing radiation. Radiat Res 2000; 153:312–7.

43. Bao CY, Ma AH, Evans HH, Horng MF, Mencl J, Hui TE, et al. Molecular analysis of hypoxanthine phosphoribosyltransferase gene deletions induced by alpha- and X-radiation in human lymphoblastoid cells. Mutat Res 1995; 326:1–15.

44. Alioto TS, Buchhalter I, Derdak S, Hutter B, Eldridge MD, Hovig E, et al. A comprehensive assessment of somatic mutation detection in cancer using whole-genome sequencing. Nat Commun 2015 9; 6:10001.

45. Cremer T, Cremer M. Chromosome territories. Cold Spring Harb Perspect Biol 2010; 2:a003889.

46. Kantidze OL, Razin SV. Chromatin loops, illegitimate recombination, and genome evolution. Bioessays 2009; 31:278–86.

47. Selby. PB. Major impact of gonadal mosaicism on hereditary risk estimation, origin of hereditary diseases, and evolution. Genetica 1998; 102–103:445–62.

48. Campbell IM, Yuan B, Robberecht C, Pfundt R, Szafranski P, McEntagart ME, et al. Parental somatic mosaicism is underrecognized and influences recurrence risk of genomic disorders. Am J Hum Genet 2014; 95:173–82.