



## OPEN ACCESS

## EDITED BY

Hesham M. H. Zakaly,  
Ural Federal University, Russia

## REVIEWED BY

Hussam Askar,  
Al Azhar University, Egypt  
Svetlana Chulkova,  
Russian Cancer Research Center NN Blokhin,  
Russia

## \*CORRESPONDENCE

Sandeep K. Singhal  
✉ Sandeep.Singhal@und.edu

RECEIVED 04 November 2023

ACCEPTED 27 February 2024

PUBLISHED 22 March 2024

## CITATION

Ruprecht NA, Singhal S, Sens D and Singhal SK  
(2024) Translating genetic findings to  
epigenetics: identifying the mechanisms  
associated with aging after high-radiation  
exposure on earth and in space.  
*Front. Public Health* 12:1333222.  
doi: 10.3389/fpubh.2024.1333222

## COPYRIGHT

© 2024 Ruprecht, Singhal, Sens and Singhal.  
This is an open-access article distributed  
under the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other forums is  
permitted, provided the original author(s) and  
the copyright owner(s) are credited and that  
the original publication in this journal is cited,  
in accordance with accepted academic  
practice. No use, distribution or reproduction  
is permitted which does not comply with  
these terms.

# Translating genetic findings to epigenetics: identifying the mechanisms associated with aging after high-radiation exposure on earth and in space

Nathan A. Ruprecht <sup>1</sup>, Sonalika Singhal <sup>2</sup>, Donald Sens <sup>2</sup>  
and Sandeep K. Singhal <sup>1,2\*</sup>

<sup>1</sup>Department of Biomedical Engineering, University of North Dakota, Grand Forks, ND, United States,

<sup>2</sup>Department of Pathology, University of North Dakota, Grand Forks, ND, United States

**Purpose:** Exposure to radiation is a health concern within and beyond the Earth's atmosphere for aircrew and astronauts in their respective austere environments. The biological effects of radiation exposure from a multiomics standpoint are relatively unexplored and stand to shed light on tailored monitoring and treatment for those in these career fields. To establish a reference variable for genetic damage, biological age seems to be closely associated with the effect of radiation. Following a genetic-based study, this study explores the epigenetic landscape of radiation exposure along with its associative effects on aging processes.

**Methods:** We imported the results of the genetics-based study that was a secondary analysis of five publicly available datasets (noted as Data1). The overlap of these genes with new data involving methylation data from two datasets (noted as Data2) following similar secondary analysis procedures is the basis of this study. We performed the standard statistical analysis on these datasets along with supervised and unsupervised learning to create preranked gene lists used for functional analysis in Ingenuity Pathway Analysis (IPA).

**Results:** There were 664 genes of interest from Data1 and 577 genes from Data2. There were 40 statistically significant methylation probes within 500 base pairs of the gene's transcription start site and 10 probes within 100 base pairs, which are discussed in depth. IPA yielded 21 significant pathways involving metabolism, cellular development, cell death, and diseases. Compared to gold standards for gestational age, we observed relatively low error and standard deviation using newly identified biomarkers.

**Conclusion:** We have identified 17 methylated genes that exhibited particular interest and potential in future studies. This study suggests that there are common trends in oxidative stress, cell development, and metabolism that indicate an association between aging processes and the effects of ionizing radiation exposure.

## KEYWORDS

bioinformatics, biological aging, epigenomics, ionizing radiation, rad-age

# 1 Introduction

Biological aging is an inevitable process characterized by the progressive decline of cellular and physiological functions, leading to increased vulnerability to age-related diseases and mortality. This complex phenomenon results from dynamic interactions between genetic, epigenetic, and environmental factors. In recent years, there has been an interest in deciphering the molecular mechanisms that underlie accelerated aging induced by exposure to ionizing radiation (1–3). Ionizing radiation, encompassing X-rays and gamma rays, carries sufficient energy to ionize atoms and molecules, generating highly reactive free radicals and causing cellular damage. The biological repercussions of ionizing radiation are extensively studied due to its pervasive use in medical diagnostics and radiation therapy, as well as the potential risk of exposure during nuclear accidents or space exploration. As such, a central mechanism through which ionizing radiation expedites biological aging is by inducing DNA damage (4). Radiation-induced double-strand breaks (DSBs) and oxidative DNA damage trigger intricate repair processes, leading to the accrual of genetic mutations and genomic instability. These genetic alterations are closely associated with cellular senescence, a hallmark of aging, and are implicated in the elevated risk of cancer and other age-related disorders (5). A review article thoroughly explores the dependent association between radiation damage and biological aging theory, leveraging modern technology to revisit this significantly complex topic (6).

Understanding the influence of ionizing radiation on biological aging necessitates consideration of the hallmarks of aging, an ensemble of interconnected cellular and molecular processes that collectively contribute to aging (7). These hallmarks include genomic instability, telomere attrition, epigenetic alterations, proteostasis loss, cellular senescence, mitochondrial dysfunction, and altered intercellular communication. Epigenetic modifications have recently gained substantial attention for their pivotal roles in orchestrating these hallmarks (8). Estimating biological age, in contrast to chronological age, constitutes a crucial strategy for evaluating the consequences of ionizing radiation on aging processes. Biological age offers a more precise representation of an individual's physiological state and susceptibility to age-related diseases. Gold standard methods for estimating biological age rely on DNA methylation-based clocks, such as the Horvath and Hannum clocks, which leverage epigenetic modifications as reliable aging biomarkers (9, 10). These clocks have been validated across diverse tissues and cell types, demonstrating their ability to predict health outcomes and mortality rates.

Computational biology has become an indispensable tool for unraveling the relationship between ionizing radiation, epigenetic modifications, and biological aging. The secondary analysis of existing multiomics datasets has proven to be a valid approach for investigating the epigenetic effects of radiation on aging (11, 12). These datasets typically encompass genetic, epigenetic, and transcriptomic profiles of individuals exposed to radiation, providing a rich resource for in-depth exploration. The secondary analysis has allowed researchers to harness existing resources and uncover novel insights by applying advanced data science and bioinformatics techniques (13). It can facilitate

the identification of radiation-responsive biomarkers, epigenetic modifications, and gene expression changes associated with accelerated aging (14). Moreover, it can enable the construction of gene regulatory networks, shedding light on the regulatory mechanisms that underlie these effects. No matter the purpose or goals, researchers can access several valuable resources for data acquisition. Large-scale collaborative initiatives, such as The Cancer Genome Atlas (TCGA) and the Gene Expression Omnibus (GEO), offer comprehensive multiomics datasets, encompassing DNA methylation, gene expression, and clinical information for diverse cohorts (15, 16). These resources empower researchers to explore the epigenetic and genetic landscape of radiation-exposed individuals and their aging phenotypes.

Epigenetic alterations, including DNA methylation changes, histone modifications, and non-coding RNA expression, have emerged as central players in mediating the effects of ionizing radiation on biological aging. Radiation-induced epigenetic modifications can directly influence the expression of genes involved in DNA repair, cellular senescence, and oxidative stress responses (17–19). These epigenetic changes may contribute to the aging phenotype observed in radiation-exposed individuals. Moreover, radiation-induced epigenetic modifications have been implicated in the regulation of telomere length, another hallmark of aging (20). Telomeres, protective caps at the ends of chromosomes, shorten with each cell division, and their length is a crucial determinant of cellular senescence. Radiation-induced changes in DNA methylation and histone modifications can influence telomere maintenance, potentially accelerating the aging process.

The epigenetic association between ionizing radiation and biological aging represents an opportunity to revisit with modern computational methods and multiomic techniques. By harnessing the power of existing datasets and advanced analytical tools, researchers can uncover the intricate mechanisms that drive radiation-induced aging, providing valuable insights for both biomedical engineering and personalized medicine. This research not only contributes to our fundamental understanding of aging processes but also holds the promise of identifying potential therapeutic interventions to mitigate the adverse effects of radiation exposure on human health. In this study, we use the findings from a genomic-only study on a radiation-age association to explore the epigenetic landscape of the effects of ionizing radiation. We specifically examine high-dose (2Gy) exposure to identify differentiating probes and understand the underlying methylation and gene expression changes. This potentially identifies genes of interest to target as control mechanisms in studying radiosensitivity or biological age.

## 2 Materials and methods

### 2.1 Data characteristics

The epigenetic aspect serves as a continuation of a multiomic exploration into the biological effects of radiation exposure. The genetics foundation that we carry over is noted as Data1, which is shown in red in Figure 1. Specifically, Data1-1 ( $n = 91$ ) include dose analysis with sex and age as clinical factors while Data1-2 ( $n$

= 75) includes sex as a secondary clinical factor (21). GSE21240, GSE23515 (22), and GSE20173 (23) (collectively referred to as Data1-1) included a total of 163 originally collected samples (of which 91 are utilized in this analysis). The samples were obtained with a chronological age ranging from 21 to 64 years, who were exposed to various levels of ionizing radiation, including 0 Gy (controls), 0.1Gy, 0.5Gy, and 2Gy. GSE21240 was a study of peripheral blood mononuclear cells (PBMCs) collected from 6 individuals, following 2 different blood preservation methods, performing RNA extraction immediately or 3 h after an ex vivo exposure to 0.5 Gy of Cesium-137 gamma rays for one minute. In total, 48 samples were analyzed GSE23515 studied peripheral blood cells from 24 different donors (95 samples in total, as one sample was lost) exposed ex vivo to 0 Gy (controls), 0.1 Gy, 0.5 Gy, and 2 Gy at a dose rate of 0.82Gy/min from Cesium-137 gamma radiation to study radioactive responses between sex and smoking behavior. GSE20173 analyzed the miRNA expression profile of peripheral blood lymphocytes incubated for 4 and 24 h in normal gravity (1g) and in modeled microgravity after irradiation with 0.2 and 2 Gy of gamma rays (5 participants across 4 conditions). In GSE44201 study, we then performed time-dependent analysis on another dataset that focused on genetic effects of radiation sometime after exposure (noted as Data1-2;  $n = 75$ ) (24). The original study was an analysis of human peripheral blood, collected 5 healthy donors, that was exposed to varying levels of gamma-ray radiation and evaluated up to 48 h after exposure. For the overlap with our data studies, we only focused on time since exposed to 2 Gy. We also identify the significant gene changes with respect to sex and age. This genetics-only analysis of Data1-1/2 was used to set a foundation and identified an intersection with significantly methylated genes in this portion.

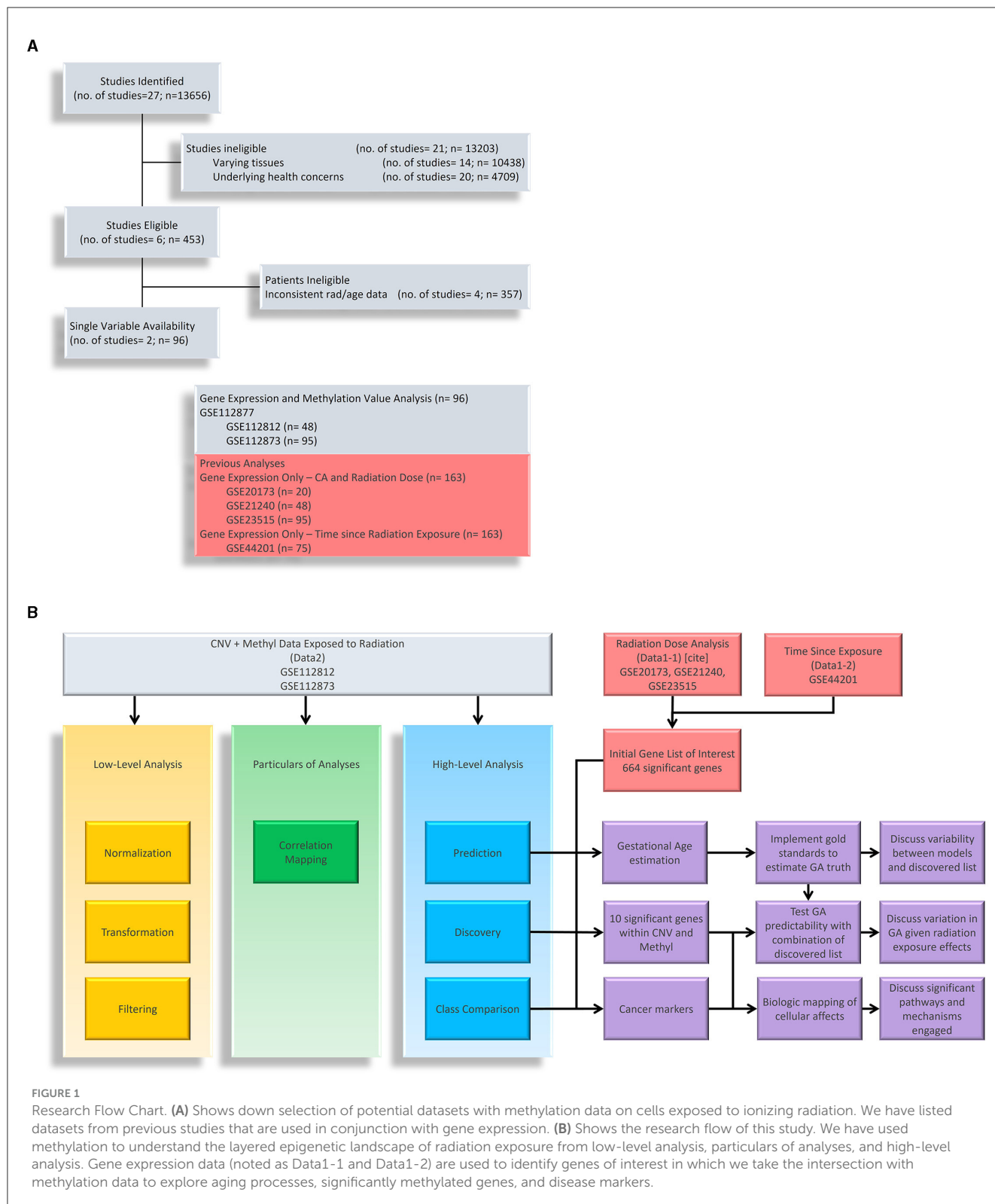
At its core, this study aimed to explore epigenetic changes in samples exposed to radiation levels that overlap with our previous study, and we primarily explored Gene Expression Omnibus (GEO) and NASA's GeneLab repositories for datasets. Figure 1 shows a breakdown of potential studies for secondary analysis (light grey) and how we utilize chosen sets (shown in yellow, green, blue, and purple boxes for respective purposes). In Figure 1A, we focused on identifying datasets with methylation data and initially found 13,656 samples. Narrowing down on clinical factors and holding variables constant, we removed those with varying tissues (brain, skin, lung, etc.), underlying health concerns that may bias expression values (cancer), and those that did not overlap with the radiation levels of the previous study (the only common dose level became 2Gy). We ended with two studies that belonged to a parent dataset, meaning that two strains aimed toward the same goal: GSE112812 (Strain 1;  $n = 48$ ) and GSE112873 (Strain 2;  $n=48$ ) falling under GSE112877 (25). This study analyzed delayed genetic and epigenetic radiation effects that may trigger radiation-induced carcinogenesis. The researchers explored cloned descendants of fetal fibroblasts irradiated to a single dose of 2 Gy of X-ray and observed copy number variation (CNV) and methylation changes in genes. The combined dataset is referred to as Data2 ( $n=96$ ) in which we downloaded the raw data of both strains separately so that they could be processed to remove the batch effect of strains. The demographics of these datasets are summarized in Table 1. Again, our focus is on the samples exposed to 0 or 2 Gy to associate with age with interest in

clinical factors such as sex. The gestational age of Data2 was determined using R's methylClock v1.2.1 package with further elaboration provided later for conducting our estimation analysis (26).

## 2.2 Recreating epigenetic workflows

With regards to Figure 1B, this subsection covers the low-level analysis (yellow box) and particulars of analysis (green box). These two, in addition to the outline of high-level analysis (blue box only) represent a recreation of methodology created particularly for epigenetic understanding (27). It is worth noting that, for data in GeneLab, fastq was the most common data form used to generate coverage files. While we did not use human data from this repository, mice data seemed promising. To explore it, methylation data were processed using nf-core/methylseq v2.3.0 (doi: 10.5281/zenodo.1343417) of the nf-core collection of workflows (28). The pipelines were executed using Nextflow v22.10.4 (29) with the below bash command. Unfortunately, several reports failed quality assurance as well as a very small number of probes relative to datasets found on GEO. Additionally, there was a lack of overlapping annotated probes to combine data from different repositories. of the nf-core collection of workflows (28). The pipelines were executed using Nextflow v22.10.4 (29) with the below bash command. Unfortunately, several reports failed quality assurance as well as a very small number of probes relative to datasets found on GEO. Additionally, there was a lack of overlapping annotated probes to combine data from different repositories.

For Data2 processing and normalization, we used R v4.2 library minfi v1.42 to turn raw idat files into usable count data (30). Recreating processing checks and techniques, we plot and evaluate a histogram of M-values, beta values, and log2 transformed gene expression with a distribution curve overlaid. This is to check reasonable density plots for each before moving forward to plotting the dataset's standard deviation and MA plots. MA plots are a visualization of log fold change ( $y - x$ ) vs. average intensity ( $(x + y)/2$ ) of two populations  $x$  and  $y$ . For our data, we used respective irradiated samples as  $x$  and non-irradiated as  $y$ . MA plots help observe the assumption that the majority of gene expression or methylation is insignificantly changing and should be along the zero line and can be an indication of what one can expect later in the analysis, such as how many probes are indeed at a higher fold change. Finally, for low-level analysis, we split the CNV into quantiles of low, medium, and high expression. We take the corresponding probe in methylation data to plot the histogram/density plot of the average beta value along with the distance to the transcription start site (TSS). Probes of interest are those typically within 500 bp of TSS in the promoter region. Correlation mapping involves analyzing the correlation between gene expression probes and methylation. Since we do not expect CNV to be fruitful as it is used for mutations, we calculate the correlation matrix between the methylation probes of Data2 and the corresponding gene expression probes of Data1. Here, an assumption is that significant findings are methylation probes that are heavily anti-correlated with their respective gene expression



probes. Bringing together these recent observations, we find that observations of extreme interest are those that have high, inverse relationship in correlation, within 500 bp of the TSS, along with significant *p*-values/log FC between irradiated and non-irradiated populations.

### 2.3 Statistical analysis

Finally, the high-level analysis emerges, which quickly becomes complex and is further elaborated upon in [Figure 1B](#). The methylated data, represented by the blue box, is integrated with the

TABLE 1 Demographics of the five datasets we use to analyze the association between radiation exposure and aging processes.

		Gene Exp. (Data1)				Methyl & CNV		Total
		Data1-1		Data1-2	Data2			
		GSE20173	GSE21240	GSE23515	GSE44201	GSE112812	GSE112873	
Sex	Male	-	40	48	15	-	-	103
	Female	-	8	47	15	-	-	70
Chronological age (yrs)	Minimum	-	21	21	20	-	-	-
	Maximum	-	64	45	53	-	-	-
Biological age (wks)	Minimum	-	-	-	-	12.09	12.41	-
	Maximum	-	-	-	-	42.77	41.47	-
Radiation Dose (Gy)	Control ( $D_T = 0$ )	10	24	24	15	13	22	108
	High Rad ( $D_T = 2$ )	10	0	23	15	35	26	109
Time since exposure (hrs)	6	-	-	-	10	-	-	10
	24	-	-	-	10	-	-	10
	48	-	-	-	10	-	-	10
	Total	20	48	95	30	48	48	289

genetics-only radiation analysis (red) to become the purple portion of the flow chart and is the bulk of our results and discussion section.

To begin, we find the overlap of the genes of interest from Data1 with those in Data2 with regards to radiation-only analysis, age-only analysis, rad-age interaction, and sex-only analysis. We aim to determine the significant genes categorized under each analysis type. Additionally, we seek to provide a holistic comparison across all categories. Since we are exploring the epigenetic landscape of radiation effects on aging, we specifically analyze a PCA plot of radiation groups with gene expression and methylation data to observe the overlap or differentiation between control and high radiation-exposed groups. Along with just the number of overlapping genes, we use the Circos v0.4.10 library in R to create a circular plot for an initial understanding of expression levels, clustering, and trends when analyzing expression and methylation (31). We perform Student t-tests on Data2 to create volcano plots to best visualize p-value and log<sub>2</sub> fold change significance. We then look at the overlap between significantly changing M-values and gene expression values to evaluate potential control mechanisms. We create another circos plot and heatmap to visualize values and trends.

## 2.4 Functional and clinical relevance

We used k-means clustering and identified 5 groups of interest (also shown on the circos plot and discussed later). We then create preranked gene lists for enrichment analysis divided into these 5 groups. Data are then analyzed using IPA (QIAGEN Inc., <https://digitalinsights.qiagen.com/IPA>) (32). We used both p-values and

log<sub>2</sub> fold change between radiation groups to run Individual Pathway Analysis (IPA) to flag significantly changing and linked biological pathways and diseases associated with our findings. As previously mentioned, we use methyClock v1.2.1 to evaluate identified markers with aging (26). Since Data2 are cloned fetal cells, we implement methylated gestational age (GA) estimation techniques found in this library listed as Bohlin, Epic, Knight, Mayne, Lee CPC, Lee reference RPC, and Lee RPC. Using all available probes of Data2, we get the estimated GA from these techniques to serve as our reference or "truth" value. Notably, we recognize that these are the estimations and, truth is relative. This is a limiting assumption as we develop a rad-age association. We then use our identified, significant rad-age methylated probes to predict sample GA and analyze the error rates compared to these established techniques. We plot the errors in weeks using root mean squared (RMS) error and discuss variation.

## 3 Results

### 3.1 Understanding the epigenetic landscape

Our analysis begins with understanding the epigenetic landscape of the data we are using. With specifics outlined in the methodology section, we present results covered in the "low-level analysis" and "particulars of analysis" portions represented by yellow and green boxes, respectively, of Figure 1 where we recreate epigenetic workflows to be implemented for our purposes (27). Overall, this is to gain a detailed understanding of the data and how it pertains to the problem at hand.

We use the R *minfi* library to process the raw *idat* files from GEO. We used the *preprocessFunnorm* function on the red-green channel set object to generate functionally normalized ratio set data. *Funnorm* extends the idea of quantile normalization from the same library, and we use *getBeta*, *getM*, and *getCN* functions to extract beta-values, M-values, and copy number variant data, respectively. Similar to the study we replicated, we show a histogram and density plot of the data to check for expected shape and intensity shown in [Supplementary Figure S1](#). Since we are focusing on epigenetic and genetic relations, we do not use the CNV data for downstream analysis. We show standard deviation blots of beta-values and gene expression along with MA plots for M-values and gene expression in [Supplementary Figure S2](#) where we see expected results of the majority of probes having insignificant fold change. Finally, in [Supplementary Figure S3](#), we again mirror our reference on exploring beta values for different quantiles of copy number variations and the corresponding distance from the gene's transcription start site (TSS). Our primary observation centers around the significant decrease in hypermethylated genes with the near constant scatter of expected distance from the TSS.

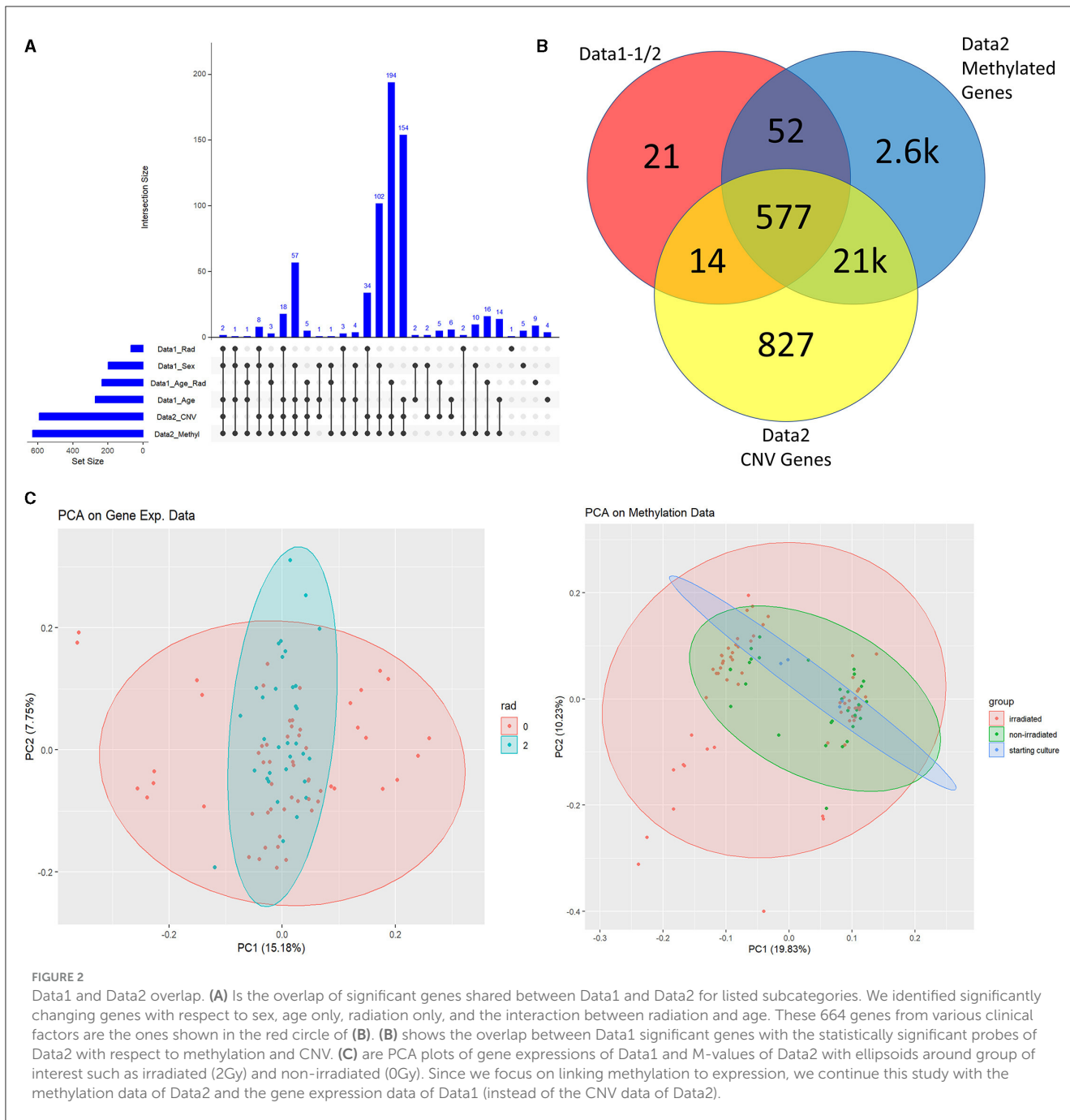
## 3.2 Statistical and high-level analysis

Understandably, most of our findings deal with the high-level analysis of epigenetic data indicated by the blue and purple sections of [Figure 1](#). We started by importing all t-test results from Data1-1 and Data1-2 (collectively referred to as Data1) where we extracted genes that were significantly different between populations of sex, age, radiation exposure (only 0 Gy vs. 2 Gy), and the interaction test between the same radiation exposure levels with age (referred to using the R notation of *rad \* age* or in plots with the underscore in place of the asterisk). We examined the overlap between this compiled list of 664 significant genes of interest from Data1 and those present in Data2. This alternate form of a Venn diagram using the R package *UpSetR* is shown in [Figure 2A](#) where vertically connected dots represent the overlap in the label category. This shows the breakdown of overlap by the t-test while a summarized gene list is shown in [Figure 2B](#) where we see 577 of the 664 genes are also in Data2. As mentioned, since CNV data are not necessarily relevant toward understanding the epigenetic landscape or control mechanisms we continue with the overlap between gene expression levels of Data1 with corresponding methylated levels in Data2 to compare trends. This analysis leads to [Figure 2C](#) where we see unsupervised clustering via principal component analysis (PCA) of irradiated groups of gene expression and methylation data, revealing a broad overlap between the two groups. A list of all genes and their overlapping groups (used to create [Figure 2A](#)) are presented in [Supplementary Table 1](#).

We conducted Student's t-test on all Data1 probes which included those found in Data2 to potentially identify overlapping genetic/epigenetic markers to eventually control. [Figure 3A](#) shows a familiar plot structure by showing a beta histogram for different quantiles of gene expression and the corresponding distance from the gene's transcription start site. With less probes of interest, we are able to see a bit more resolution trying to see trends. Of note with beta values, we identify an increase in gene

expression corresponded to a lower percentage of methylated genes as somewhat indicated with the hypermethylated hump pulling left as we go across the plots. Meanwhile, the scatterplots of methylation probe TSS distance seemed to increase in variability with increased gene expression. [Figure 3B](#) represents those results with volcano plots of the Data1 gene expression and Data2 methylation data independently evaluated. There were no genes of significant fold change, so we picked an artificial threshold for plotting. Of the methylation probes that had a  $p\text{-value} < 0.05$  and  $\log_2\text{FC}(\text{radiation}) > 2$ , 40 probes were within 500 base pairs of the gene's transcription start site and 10 within 100 base pairs. While we performed the analysis on both gene expression and methylation data for completion, we again proceed with evaluating methylated genes. [Figures 3C, D](#) show the methylation results, plotted with corresponding gene expression values from Data1. [Figure 3C](#) is a fold change heatmap showing the aforementioned 10 methylation probes matched to the closest gene symbol along with the gene expression  $\log_2\text{FC}$ . Next to the colored fold-change values is a list detailing the distance to the TSS for that methylated gene. [Figure 3D](#) shows those same 10 probes in a circos plot with the methylation values on the outer ring and corresponding gene expression (regardless of significance) on the inner ring. Here, it is easier to see the relationship between gene and epigenetic changes. Raw t-test findings (to include the change and  $p$ -values that were plotted here) are presented in [Supplementary Table 2](#).

Next, we focused on overlaying methylation levels with gene expression without digging into the genes themselves. While the previous paragraph/figure explores all significant probes of Data2 corresponding to their gene expression, this section starts with the significant gene list of Data1 (664 genes of interest mentioned) and explores the available methylation values. We used k-means clustering of methylation values to identify five groups of similar expressions. [Figure 4A](#) shows a circos plot of these clusters along with corresponding gene expression. The outer ring, representing the methylation values that are ordered by clusters, we see one group of hypomethylation, one group of hypermethylation, one of seemingly unmethylated, and two where there is some type of changing inverse relationship. The inner ring, as mentioned, is the gene expression of the corresponding probe to the plotted methylation values. While there might seem some correlation, it is not ordered and much harder to tell with the naked eye. While we dig into specifics more, this is merely a visualization step to see that there is indeed some trend to be explored. Since we are interested in the understanding landscape of radiation effects, [Table 2](#) is a summary of these five clusters when looking at irradiated vs. non-irradiated samples. Due to the changing relationship, the probes of clusters 1 and 2 are pulled for correlation plotting. Specifically looking for positive or negative correlation between methylated genes and expression, [Figure 4B](#) shows the top 1% of most absolute correlated values. For example, upon reviewing this plot, as will be discussed in the next section, the methylation of *DDX3Y* (far right column) demonstrates a significant positive-correlation (blue) with the expression of 15 genes, such as *HBD*, *BRPF3*, *ZNF177*, *KCM1B*, *RGS9*, *FAM132A*, *RSC1A1*, *SLPI*, *KIAA0586*, *ETV7*, *FEM1A*, *MXRA7*, *TIMD4*, *CAND1*, and *DFNB31* while anti-correlated (red) with the expression of 20 genes, such as *SEL1L1*, *GGT8P*, *AK4*, *SCAMP4*, *BCL2L1*, *SASH1*, *SLC10A7*,

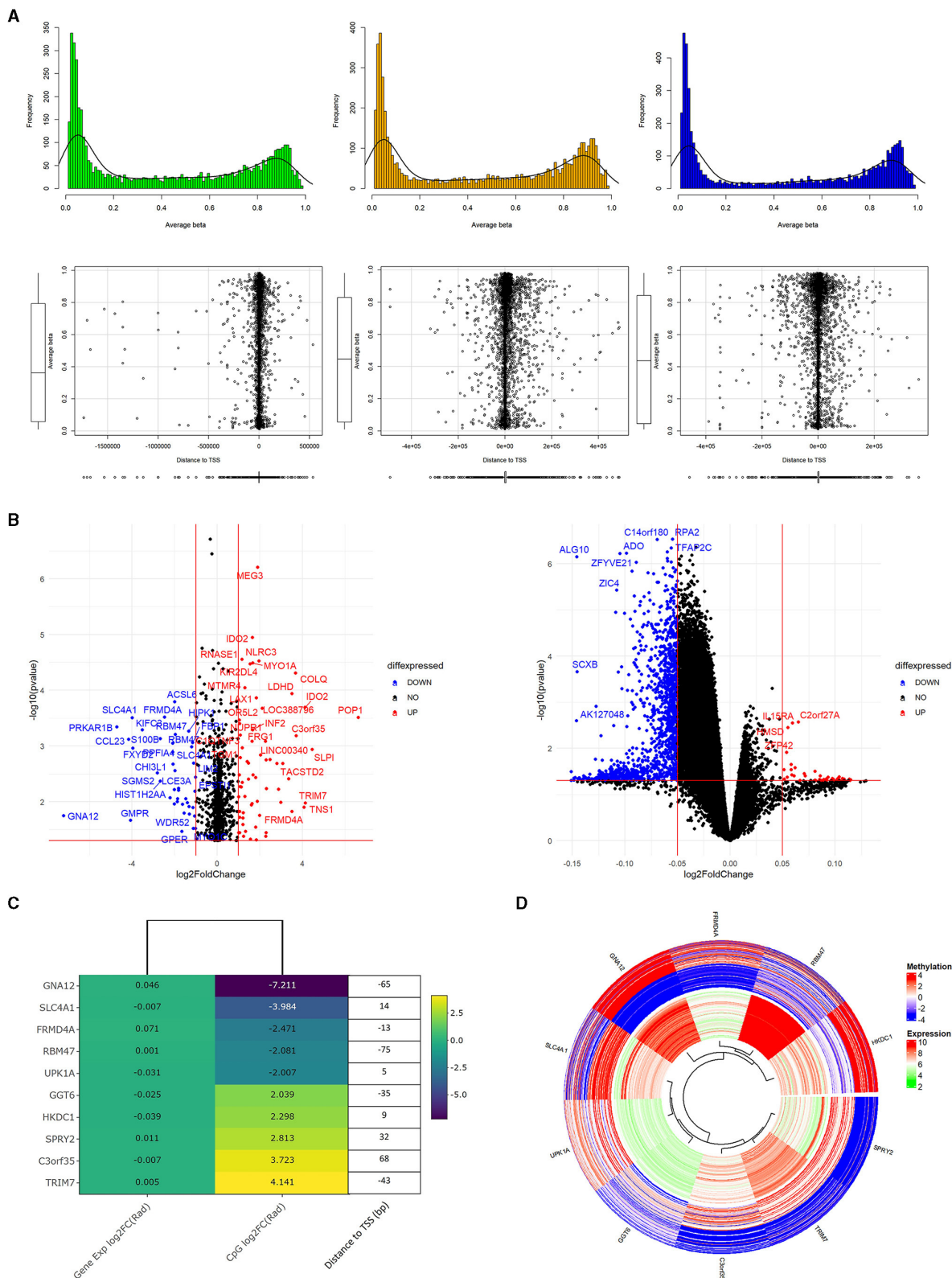


*POLDIP2, GARS, ORM1, KIAA2018, GPER, LAMA5, VAMP4, NAALADL1, TSPYL2, MCTP2, PPAP2B, GAL3ST4, and SOX4.* The raw correlation values can be found in [Supplementary Table 3](#).

### 3.3 Functional and clinical relevance

Overlapping statistical findings with functional analysis, we create preranked lists based on p-values and log2FC(radiation) from the five k-means clusters of the original genes (611 found within these data of the 664 genes of interest). These preranked lists were analyzed using Ingenuity Pathway Analysis (IPA) for

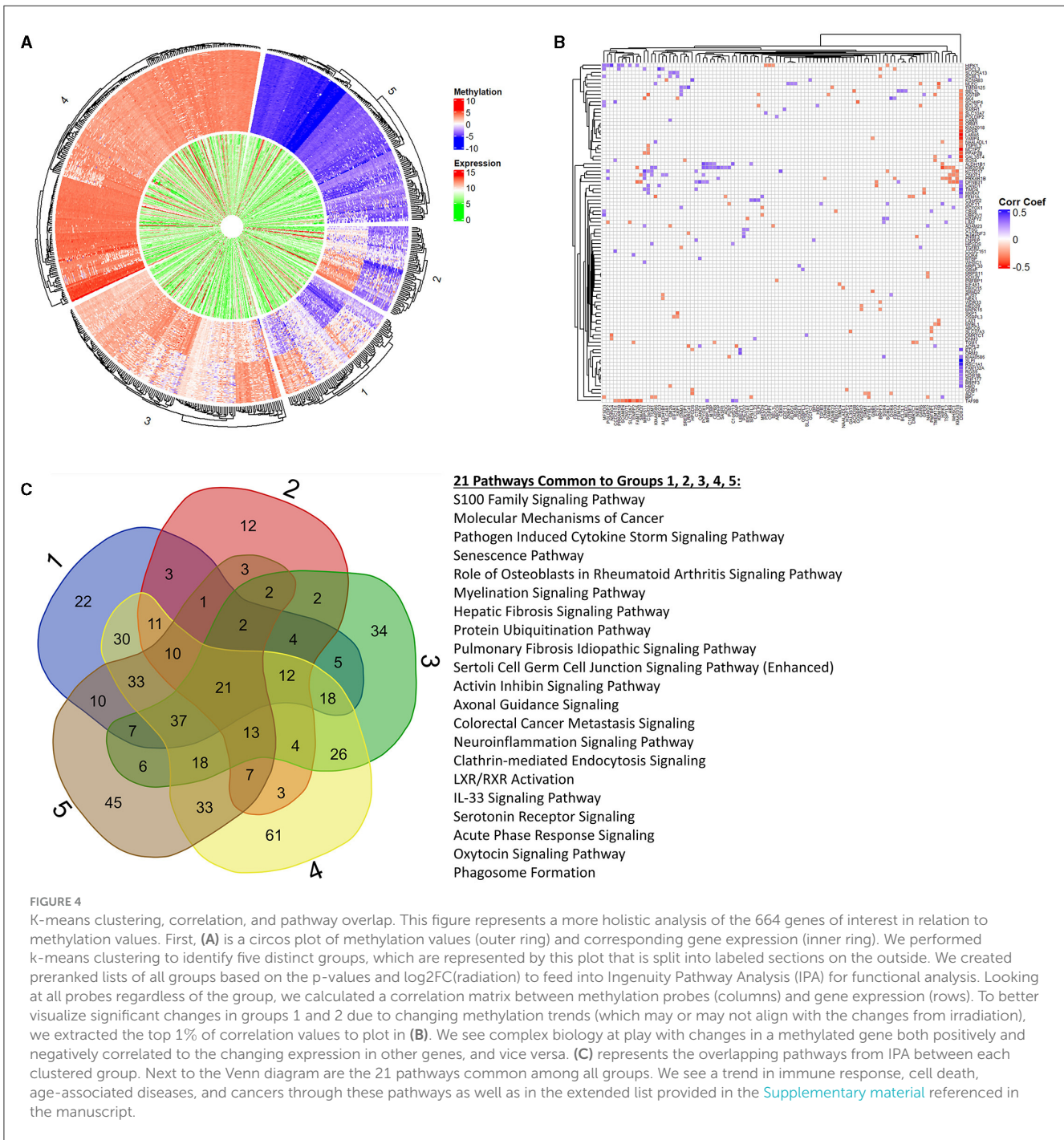
identifying significant pathways and diseases. [Figure 4C](#) shows a Venn diagram of output pathways with those intersecting 21 among all groups specified. Of particular note are the senescence, myelination signaling, neuroinflammation, and IL-33 signaling pathways overlapping in all groups, which clearly ties to aging processes. Seeing such pathways that would be tied to old age in fetal cells supplements a rad-age association with the epigenetic background. Looking specifically at cluster groups 1 and 2 (inverse relationship of M-values), we see pathways involved in the metabolism (*PFKFB4* signaling and fatty acid  $\alpha$ -oxidation), cellular development (Hippo signaling and putrescine biosynthesis/degradation), cell death (MYC mediated apoptosis and induction of apoptosis by HIV1), and disease (*THO1* in



**FIGURE 3** Methylation analysis supplementing genetic findings. **(A)** is a plot of the processed Data2 sites of interest from Data (previous figure). We separate Data2 into three quantiles of gene expression to view average beta trends (top) as well as distance of the methylation probes to the corresponding gene's transcription start site (bottom). Of note is the general similar shape in average beta distribution and there is increased variance in TSS distance with increased gene expression. **(B)** are volcano plots of  $p$ -value vs.  $\log_2\text{FC}$  of radiation only analysis between methylation (left) values and gene expression (right). These are only the probes that correspond to genes of interest from the original 664 list. The methylation values here are those with  $p < 0.05$ ,  $-\log_2\text{FC}(\text{radiation}) > 2$ . Additionally, 40 of these probes also had a  $-TSS$  distance  $< 500\text{bp}$  and 10 probes with a distance  $< 500\text{bp}$  (*Continued*)



**FIGURE 3 (Continued)**  
 < 100bp. Looking specifically at these 10 probes are the focus of (C, D). (C) shows a heatmap of the 10 methylation probes with the corresponding gene expression colored to log2FC(radiation) of the respective dataset. Added onto the right of the heatmap is the basepair distance to the TSS from the annotation file for the methylation probe. (D) is methylation (out ring) and gene expression (inner ring) of genes from the heatmap. Specifically, the study highlights the varying methylation rates while somewhat consistent gene expression.



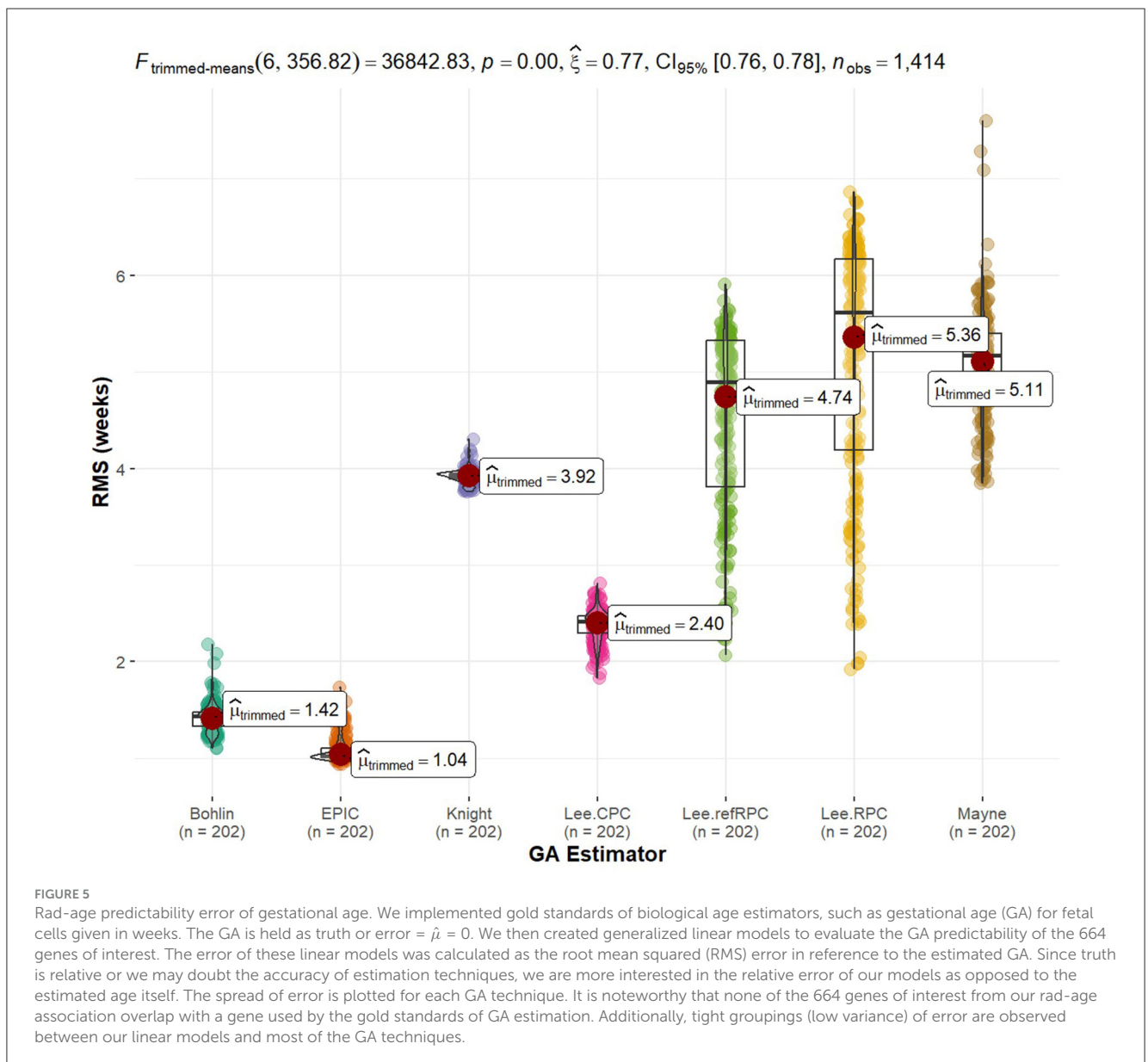
Alzheimer’s disease, Huntington’s disease, and colorectal cancer metastasis) to name a few. These intersecting pathways, and others of significant value to a rad-age association, are discussed more in the next section. The labeled k-means groups for each probe is listed in [Supplementary Table 4](#) while all IPA results for these groups

(molecules, pathways, overlapping groups, etc.) are presented in [Supplementary Table 5](#) and [6](#).

Tying back to the radioactive effects on biological aging, we implemented standard gestational age (GA) estimators in R as our reference value so we could explore the predictability of our

TABLE 2 Sample split by k-means cluster group on hyper- or hypo-methylated probes.

Cluster group	Irradiated		Non-irradiated	
	Hypomethylated	Hypermethylated	Hypomethylated	Hypermethylated
1	28	33	22	9
2	35	26	9	22
3	0	61	0	31
4	0	61	0	31
5	61	0	31	0



statistically significant methylation probes. Using a 60/40 split between training and testing data, we fit a generalized linear model to the data and calculated the RMS error to each of the GA references. The GA of each sample in itself was not of interest to us because truth is relative and there is currently no way to tell how

close to truth they are. Instead, we are interested in how close these methylated genes were to the standards. Figure 5 shows boxplots of error with models listed along the x-axis and RMS error to the standards along the y-axis (33). An interesting note is that none of these statistically significant probes when comparing radiation

were used to implement the gold standards. We were truly using new potential markers to estimate age. Additionally, it is interesting to see tight groupings (little variation in error) on non-Lee models and there is a low relative error in these models. The train/test split of samples can be found in [Supplementary Table 7](#) while the statistically significant probe list was previously mentioned in [Supplementary Table 8](#).

## 4 Discussion

Our first thing of note is the difference in the quantile separation of beta values and TSS plots in [Supplementary Figure S3](#) and [Figure 3A](#) compared to the findings of Singhal et al. (27). In our data, we have relatively unchanged histogram shape with increasing gene expression in dealing with radiation exposure with spikes in both hypo- and hyper-methylated regions. However, Singhal et al. had primarily hypomethylated genes in medium and high gene expression with almost a noise-like histogram of beta values among low expressed genes. This seems to be our first characterization of expectations when dealing with the epigenetic landscape of those exposed to ionizing radiation. With regards to significant, low absolute distance from the TSS of  $v$ , given our focus on gene expression instead of mutation rates, we look at the regions of [-500, 500] and [-100, 100]. In the highly expressed genes, 35.8% (1895 of 5289) and 13.3% (704 of 5,289) of the methylation probes were in these regions. There were 33.8% and 12.1% (out of 5223 probes) in these regions among medium gene expression. And lastly, 35.7% and 12.3% (out of 4,810 probes) in these regions respectively amongst low gene expression. Additionally, since we brought up a potential shift in hypermethylation, we looked at the frequency of beta value counts in that region of the histogram. High gene expression saw a peak count of 147 at 93% methylation, medium gene expression peaks at 124 counts each of 92–93% methylation, and low gene expression peaks at 95 counts each of 92–93% methylation. This can help research in recognizing this epigenetic landscape with a trend of increased hypermethylation of genes with increased gene expression when exposed to ionizing radiation while maintaining a ratio of distances to the gene's transcription start site.

Looking at the 10 genes highlighted from [Figures 3B, C](#), the first to draw attention is Gene Protein Subunit Alpha 12 (*GNA12*) which is predicted to enable the GTPase activity and involved in regulating TOR signaling. This is arguably the more significant finding because the GTPase activity with another gene has been shown to promote radioresistance in cells, and TOR signaling is directly linked to aging processes (34–37). The Solute Carrier Family 4 Member 1 (*SLC4A1*) gene encodes a protein part of the anion exchanger in the red blood cell plasma member where it is involved in CO<sub>2</sub> transport from tissues to lungs. This becomes an interesting finding in the perspective of radiation being a known causal factor of cancers in high proliferating cells (such as in blood forming organs) as well as CO<sub>2</sub> (38, 39). FERM Domain Containing 4A (*FRMD4A*) encodes a protein that regulates epithelial cell polarity. While associated with nicotine dependence, this has more relevance with a rad-age association because of its connection with Alzheimer's disease (40). RNA Binding Motif Protein 47 (*RBM47*) enables RNA binding activity and is predicted to act upstream of cytidine to uridine editing as well as upstream of hematopoietic

progenitor cell differentiation. Since C-to-U editing often involves the hydrolytic breakdown of amino acids for energy following excess protein intake, this is yet another gene found to be tied to energy in some regard (41). Additionally, a repeat of some sense is seeing reference again to blood cells. With observed changes from radiation exposure on a gene involved in hematopoietic cell differentiation, tied to one of the hallmarks of aging is stem cell exhaustion, *RBM47* is creating a loose corroboration to a rad-age association (42). Gamma-Glutamyl transferase 6 (*GGT6*) belongs to the GGT family that is a membrane-bound extracellular enzyme that is key to glutathione homeostasis as it provides synthesis substrates. This is indirectly significant with regards to a rad-age association because glutathione is a powerful antioxidant produced by the liver to deal with free radicals, peroxides, and heavy metals (43). Following connections with the liver, Hexokinase Domain Containing 1 (*HKDC1*) encodes a protein involved in glucose metabolism. Reduced expression may be associated with gestational diabetes while high expression may be associated with poor prognosis in hepatocellular carcinoma (44). *C3orf35* is an alias and previous HGNC symbol for the *APRG1* tumor suppressor candidate which is an RNA gene with little known properties or associations beyond a study in 2005 that it may suppress tumor growth in breast cancer (45). The protein from Sprouty RTK Signaling Antagonist 2 (*SPRY2*) is involved in the non-cell autonomous inhibitory effect on fibroblast growth factor two signaling. It is noteworthy that a mutation in this gene has been found to inhibit MAPK pathway which is involved in tumorigenesis (46). Uroplakin 1A (*UPK1A*) codes a cell-surface protein that may play a role in normal bladder epithelial physiology and possible tumor suppression (47). Tripartite Motif (TRIM), containing 7, encodes a protein that may participate in glycogen synthesis as well as both tumor-promoting and tumor-suppressing functions in innate immunity (48). Finally, while we present all results for viewer's understanding and interpretation, we noted that it was interesting to see the methylation of *DDX3Y* both positively and negatively correlated with the expression of many distinct genes. The protein from this gene is considered to be involved in ATP binding, hydrolysis, and RNA binding (49–51). Mutations in this gene result in male infertility being on the Y chromosome. While there is connection to aging processes, the continued trend of energy molecules is an exciting find with this dataset and approach.

Shifting to IPA results while continuing the discussion of significant genes, we see five genes that are recurrently found in the most significant pathways of cluster groups 1 and 2. *BCL2L1* codes a protein in the BCL-2 family that acts as an anti- or pro-apoptotic regulator located at the outer mitochondrial membrane. More significantly to our cause, it regulates the outer mitochondrial membrane channel (VDAC) potential and thus controls the production of reactive oxygen species (52). Again, continuing a theme of energy, ATPase sarcoplasmic/endoplasmic reticulum CA2+ Transporting 3 (*ATP2A3*) encodes one of the SERCA CA2+ ATPases located in muscle cells. This enzyme catalyzes the hydrolysis of ATP and is involved in calcium sequestration associated with muscular contraction and loosely tied to radioresistance and aging separately (53, 54). G protein subunit beta 1 (*GNB1*) codes a beta subunit, which regulates alpha subunits of nucleotide-binding proteins that integrate signals between receptors and effector proteins. One such case is when

beta and gamma chains are required for the GTPase activity which we saw relevant with *GNAI2* (one of the 10 methylation probes within 100bp of its gene's TSS). SMAD family member 2 (*SMAD2*) encodes a protein that mediates the signal of the transforming growth factor (TGF)-beta, thereby regulating processes, such as cell proliferation, apoptosis, and differentiation, as well as associated with cardiovascular diseases - all of which give confidence in linking effects of radiation and aging (55–57). Aldehyde Dehydrogenase 1 family member B1 (*ALDH1B1*) is the second enzyme of the major oxidate pathway of alcohol metabolism. The review of the literature connects downregulation of *ALDH1B1* to numerous cancers including colorectal, pancreatic, liver, prostate, lung, brain, and breast cancers to name a few (58–64). While *ALDH1B1* may indirectly or loosely support a rad-age association, given its ties to cancer progression and radioresistance, this enzyme also perks our interest once more on being yet another metabolic constituent.

Understandably with this dataset, cellular and embryonic development pathways are highlighted as significant. We cannot know if that's due to the radiation effects on the epigenome or because of the cell types used in this study. While briefly mentioning the pathways, we can revisit knowing significant genes/molecules that play a role. Energy and metabolism was an interesting trend compared to others in that it is not directly called out in response to radiation and aging. It could be underneath mitochondrial dysfunction and deregulated nutrient-sensing for hallmarks of aging but not considered obvious. Fatty acid oxidation comes close and was highlighted in groups 2, 3, 4, and 5 primarily due to the significance of *ALDH1B1* that was just discussed. From groups 1, 3, 4, and 5 is *PFKFB4* signaling which is another significant pathway involved in metabolism. Although the gene itself was not in our gene list, we wanted to highlight its function. The protein encoded by this gene is highly expressed in cancer cells and is induced by hypoxia (65). It forces the cell to increase the amount of energy (i.e., ATP) production beyond its typical constraints (66). Overlapping with cancer pathology is the Hippo signaling pathway from groups 2 and 3 (also sharing T-Cell exhaustion signaling) which modulates the proliferation, differentiation, and survival of cells (67). Another pathway without the gene itself is *MYC* mediated apoptosis from groups 1 and 4 (along with a necroptosis and pyroptosis signaling pathway). While the gene itself plays a role in cell growth, proliferation, differentiation, and apoptosis, the pathway focuses on inducing apoptosis within a cell when survival factors are missing (68). As may be expected, a number of cancer pathways were flagged as significant from our rad-age associated gene list. Colorectal and gastrointestinal cancers were recurring while breast, ovarian, lung, prostate, and bladder cancers were also found from multiple cluster analyses.

## 5 Conclusion

In this study, we utilized genetic findings by exploring previously identified genes of interest to investigate the epigenetic landscape of ionizing radiation and its relationship to aging processes. Our primary source of results stem from conducting the secondary analysis on two publicly available datasets of fetal fibroblasts exposed to 2 Gy of radiation. The Student t-test was performed to create preranked gene lists based on the p-values and

log<sub>2</sub>FC(radiation) for functional analysis in IPA where significant pathways and diseases were identified. Upon close examination of k-means clustering groups, the correlation between methylation and gene expression, and molecules that emphasized pathway results, we discussed 17 methylated genes that showed particular interest and potential in future studies: *GNAI2*, *SLC4A1*, *FRMD4A*, *RBM47*, *GGT6*, *HKDC1*, *APRG1*, *SPRY2*, *UPK1A*, *TRIM7*, *DDX3Y*, *BCL2L1*, *ATP2A3*, *GNB1*, *SMAD2*, *ALDH1B1*, and *PFKFB4*. Many other findings (such as 40 methylation probes within 500 base pairs of their gene's transcription start site) in addition to raw pathway analysis are provided in the [Supplementary material](#) for the reader's use. We found common trends in oxidative stress, cell development/growth/death, immune response, and (in an unforeseen manner) metabolism/energy without direct links to mitochondrial dysfunction.

## Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

## Author contributions

NR: Writing – review & editing, Software, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. SS: Formal analysis, Writing – review & editing, Software. DS: Writing – review & editing, Supervision, Project administration, Funding acquisition. SKS: Conceptualization, Formal analysis, Investigation, Project administration, Supervision, Writing – review & editing.

## Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpubh.2024.1333222/full#supplementary-material>

## References

- Uwineza A, Kalligeraki AA, Hamada N, Jarrin M, Quinlan RA. Cataractogenic load-A concept to study the contribution of ionizing radiation to accelerated aging in the eye lens. *Mutation Res/Rev Mutat Res.* (2019) 779:68–81. doi: 10.1016/j.mrrev.2019.02.004
- Richardson RB. Ionizing radiation and aging: rejuvenating an old idea. *Aging (Albany NY).* (2009) 1:887. doi: 10.18632/aging.100081
- Hernandez L, Terradas M, Camps J, Martin M, Tusell L, Genesca A. Aging and radiation: bad companions. *Aging Cell.* (2015) 14:153–61. doi: 10.1111/acel.12306
- Sharma NK, Sharma R, Mathur D, Sharad S, Minhas G, Bhatia K, et al. Role of ionizing radiation in neurodegenerative diseases. *Front Aging Neurosci.* (2018) 10:134. doi: 10.3389/fnagi.2018.00134
- Lidzbarsky G, Gutman D, Shekhdem HA, Sharvit L, Atzmon G. Genomic instabilities, cellular senescence, and aging: *in vitro*, *in vivo* and aging-like human syndromes. *Front Med.* (2018) 5:104. doi: 10.3389/fmed.2018.00104
- Ruprecht NA, Singhal S, Schaefer K, Panda O, Sens D, Singhal SK, et al. Review: multi-omics approach to studying the association between ionizing radiation effects on biological aging. *Biology.* (2024) 13:98. doi: 10.3390/biology13020098
- López-Otin C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. *Cell.* (2013) 153:1194–17. doi: 10.1016/j.cell.2013.05.039
- Brunet A, Berger SL. Epigenetics of aging and aging-related disease. *J Gerontol Series A.* (2014) 69:S17–20. doi: 10.1093/gerona/glu042
- Horvath S, Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. *Nat Rev Genet.* (2018) 19:371–84. doi: 10.1038/s41576-018-0004-3
- Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sadda S, et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol Cell.* (2013) 49:359–67. doi: 10.1016/j.molcel.2012.10.016
- Windle PE. Secondary data analysis: is it useful and valid? *J Peri Anesthesia Nursing.* (2010) 25:322–4. doi: 10.1016/j.jopan.2010.07.005
- Johnston MP. Secondary data analysis: a method of which the time has come. *Qual Quant Methods Librar.* (2014) 3:619–26.
- Huang Y, Tzeng JY, Maguire R, Hoyo C, Allen T. The association between neuraxial anesthesia and the development of childhood asthma—a secondary analysis of the newborn epigenetics study cohort. *Curr Med Res Opin.* (2020) 36:1025–32. doi: 10.1080/03007995.2020.1747417
- Kabacik S, Mackay A, Tamber N, Manning G, Finnon P, Paillier F, et al. Gene expression following ionising radiation: identification of biomarkers for dose estimation and prediction of individual response. *Int J Radiat Biol.* (2011) 87:115–29. doi: 10.3109/09553002.2010.519424
- Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. *Nucleic Acids Res.* (2012) 41:D991–5. doi: 10.1093/nar/gks1193
- Tomczak K, Czerwińska P, Wiznerowicz M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemp Oncol/Współczesna Onkologia.* (2015) 2015:68–77. doi: 10.5114/wo.2014.47136
- Wei J, Wang B, Wang H, Meng L, Zhao Q, Li X, et al. Radiation-induced normal tissue damage: oxidative stress and epigenetic mechanisms. *Oxidat Med Cell Longev.* (2019) 2019:3010342. doi: 10.1155/2019/3010342
- Belli M, Tabocchini MA. Ionizing radiation-induced epigenetic modifications and their relevance to radiation protection. *Int J Mol Sci.* (2020) 21:5993. doi: 10.3390/ijms21175993
- Isermann A, Mann C, Rube CE. Histone variant H2A J marks persistent DNA damage and triggers the senary phenotype in radiation-induced senescence. *Int J Mol Sci.* (2020) 21:9130. doi: 10.3390/ijms21239130
- Ilyenko I, Lyaskivska O, Bazyka D. Analysis of relative telomere length and apoptosis in humans exposed to ionising radiation. *Exp Oncol.* (2011).
- Ruprecht NA, Singhal S, Schaefer K, Gill JS, Bansal B, Sens D, et al. Establishing a genomic radiation-age association for space exploration supplements lung disease differentiation. *Front Public Health.* (2023) 11:1161124. doi: 10.3389/fpubh.2023.1161124
- Paul S, Amundson SA. Gene expression signatures of radiation exposure in peripheral white blood cells of smokers and non-smokers. *Int J Radiat Biol.* (2011) 87:791–801. doi: 10.3109/09553002.2011.568574
- Girardi C, De Pittà C, Casara S, Sales G, Lanfranchi G, Celotti L, et al. Analysis of miRNA and mRNA expression profiles highlights alterations in ionizing radiation response of human lymphocytes under modeled microgravity. *PLoS ONE.* (2012) 7:e31293. doi: 10.1371/journal.pone.0031293
- Paul S, Smilenov LB, Amundson SA. Widespread decreased expression of immune function genes in human peripheral blood following radiation exposure. *Radiat Res.* (2013) 180:575–83. doi: 10.1667/RR13343.1
- Flunkert J, Maierhofer A, Dittrich M, Müller T, Horvath S, Nanda I, et al. Genetic and epigenetic changes in clonal descendants of irradiated human fibroblasts. *Exp Cell Res.* (2018) 370:322–32. doi: 10.1016/j.yexcr.2018.06.034
- Pelegi-Sisó D, de Prado P, Ronkainen J, Bustamante M, González JR. methylclock: a Bioconductor package to estimate DNA methylation age. *Bioinformatics.* (2021) 37:1759–60. doi: 10.1093/bioinformatics/btaa825
- Singhal SK, Usmani N, Michiels S, Metzger-Filho O, Saini KS, Kovalchuk O, et al. Towards understanding the breast cancer epigenome: a comparison of genome-wide DNA methylation and gene expression data. *Oncotarget.* (2016) 7:3002. doi: 10.18632/oncotarget.6503
- Ewels PA, Peltzer A, Fillinger S, Patel H, Alneberg J, Wilm A, et al. The nf-core framework for community-curated bioinformatics pipelines. *Nat Biotechnol.* (2020) 38:276–8. doi: 10.1038/s41587-020-0439-x
- Di Tommaso P, Chatzou M, Barja PP, Palumbo E, Notredame C. Nextflow enables reproducible computational workflows. *Nat Biotechnol.* (2017) 35:316–9. doi: 10.1038/nbt.3820
- Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. Minfi: a flexible and comprehensive Bioconductor package for the analysis of Infinium DNA methylation microarrays. *Bioinformatics.* (2014) 30:1363–9. doi: 10.1093/bioinformatics/btu049
- Krzywinski M, Schein J, Birol I, Connors J, Gascoyne R, Horsman D, et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* (2009) 19:1639–45. doi: 10.1101/gr.092759.109
- Krämer A, Green J, Pollard Jr J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics.* (2014) 30:523–30. doi: 10.1093/bioinformatics/btt703
- Patil I. Visualizations with statistical details: The'ggstatsplot'approach. *J Open Sour Softw.* (2021) 6:3167. doi: 10.21105/joss.03167
- Hein AL, Post CM, Sheinin YM, Lakshmanan I, Natarajan A, Enke CA, et al. RAC1 GTPase promotes the survival of breast cancer cells in response to hyper-fractionated radiation treatment. *Oncogene.* (2016) 35:6319–29. doi: 10.1038/onc.2016.163
- Magalhaes YT, Boell VK, Cardella GD, Forti FL. Downregulation of the Rho GTPase pathway abrogates resistance to ionizing radiation in wild-type p53 glioblastoma by suppressing DNA repair mechanisms. *Cell Death & Disease.* (2023) 14:283. doi: 10.1038/s41419-023-05812-1
- Kaeberlein M, Kennedy BK. Hot topics in aging research: protein translation and TOR signaling, 2010. *Aging Cell.* (2011) 10:185–90. doi: 10.1111/j.1474-9726.2010.00665.x
- Blackwell TK, Sewell AK, Wu Z, Han M, TOR signaling in *Caenorhabditis elegans* development, metabolism, and aging. *Genetics.* (2019) 213:329–60. doi: 10.1534/genetics.119.302504
- Pross M, Schulz H, Flechsig A, Manger T, Halangk W, Augustin W, et al. Oxidative stress in lung tissue induced by CO<sub>2</sub> pneumoperitoneum in the rat. *Surg Endosc.* (2000) 14:1180–4. doi: 10.1007/s004640000189
- Sia J, Szymd R, Hau E, Gee HE. Molecular mechanisms of radiation-induced cancer cell death: a primer. *Front Cell Dev Biol.* (2020) 8:41. doi: 10.3389/fcell.2020.00041
- Lambert JC, Grenier-Boley B, Harold D, Zelenika D, Chouraki V, Kamatani Y, et al. Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. *Mol Psychiatry.* (2013) 18:461–70. doi: 10.1038/mp.2012.14
- Blanc V, Xie Y, Kennedy S, Riordan JD, Rubin DC, Madison BB, et al. Apobec1 complementation factor (A1CF) and RBM47 interact in tissue-specific regulation of C to U RNA editing in mouse intestine and liver. *Rna.* (2019) 25:70–81. doi: 10.1261/rna.068395.118
- Yeganeh M, Seyedjafari E, Kamrani FA, Ghaemi N. RNA-binding protein Rbm47 binds to Nanog in mouse embryonic stem cells. *Mol Biol Rep.* (2013) 40:4391–6. doi: 10.1007/s11033-013-2528-0
- Timme-Laragy AR, Goldstone JV, Imhoff BR, Stegeman JJ, Hahn ME, Hansen JM. Glutathione redox dynamics and expression of glutathione-related genes in the developing embryo. *Free Rad Biol Med.* (2013) 65:89–101. doi: 10.1016/j.freeradbiomed.2013.06.011
- Jiang H, Yao Q, An Y, Fan L, Wang J, Li H. Baicalin suppresses the progression of Type 2 diabetes-induced liver tumor through regulating METTL3/m6A/HKDC1 axis and downstream p-JAK2/STAT1/cleaved Capase3 pathway. *Phytomedicine.* (2022) 94:153823. doi: 10.1016/j.phymed.2021.153823
- Leris A, Roberts T, Jiang WG, Newbold R, Mokbel K. Evidence for a tumour suppressive function of APRG1 in breast cancer. *Breast Cancer Res Treat.* (2005) 93:97–100. doi: 10.1007/s10549-005-4169-z

46. Milillo A, La Carpia F, Costanzi S, D'urbano V, Martini M, Lanuti P, et al. A SPRY2 mutation leading to MAPK/ERK pathway inhibition is associated with an autosomal dominant form of IgA nephropathy. *Eur J Human Genet.* (2015) 23:1673–8. doi: 10.1038/ejhg.2015.52
47. Du F, Guo T, Cao C. Restoration of UPK1A-AS1 expression suppresses cell proliferation, migration, and invasion in esophageal squamous cell carcinoma cells partially by sponging microRNA-1248. *Cancer Manage Res.* (2020) 2020:2653–62. doi: 10.2147/CMAR.S239418
48. Zhou C, Zhang Z, Zhu X, Qian G, Zhou Y, Sun Y, et al. N6-Methyladenosine modification of the TRIM7 positively regulates tumorigenesis and chemoresistance in osteosarcoma through ubiquitination of BRMS1. *EBioMedicine.* (2020) 59:102955. doi: 10.1016/j.ebiom.2020.102955
49. Shen H, Yanas A, Owens MC, Zhang C, Fritsch C, Fare CM, et al. Sexually dimorphic RNA helicases DDX3X and DDX3Y differentially regulate RNA metabolism through phase separation. *Mol Cell.* (2022) 82:2588–603. doi: 10.1016/j.molcel.2022.04.022
50. Kotov AA, Olenkina OM, Godneeva BK, Adashev VE, Olenina LV. Progress in understanding the molecular functions of DDX3Y (DBY) in male germ cell development and maintenance. *Biosci Trends.* (2017) 11:46–53. doi: 10.5582/bst.2016.01216
51. Li G, Wang C, Guan X, Bai Y, Feng Y, Wei W, et al. Age-related DNA methylation on Y chromosome and their associations with total mortality among Chinese males. *Aging Cell.* (2022) 21:e13563. doi: 10.1111/accel.13563
52. Boise LH, González-García M, Postema CE, Ding L, Lindsten T, Turka LA, et al. bcl-x, a bcl-2-related gene that functions as a dominant regulator of apoptotic cell death. *Cell.* (1993) 74:597–608. doi: 10.1016/0092-8674(93)90508-N
53. Funahashi Y, Hernandez SL, Das I, Ahn A, Huang J, Vorontchikhina M, et al. A notch1 ectodomain construct inhibits endothelial notch signaling, tumor growth, and angiogenesis. *Cancer Res.* (2008) 68:4727–35. doi: 10.1158/0008-5472.CAN-07-6499
54. Li G, Petkova TD, Laritsky E, Kessler N, Baker MS, Zhu S, et al. Early postnatal overnutrition accelerates aging-associated epigenetic drift in pancreatic islets. *Environm Epigenet.* (2019) 5:dvz015. doi: 10.1093/eep/dvz015
55. Wang X, Ma W, Han S, Meng Z, Zhao L, Yin Y, et al. TGF- $\beta$  participates choroid neovascularization through Smad2/3-VEGF/TNF- $\alpha$  signaling in mice with Laser-induced wet age-related macular degeneration. *Sci Rep.* (2017) 7:9672. doi: 10.1038/s41598-017-10124-4
56. Yin X, Tian W, Wang L, Wang J, Zhang S, Cao J, et al. Radiation quality-dependence of bystander effect in unirradiated fibroblasts is associated with TGF- $\beta$ 1-Smad2 pathway and miR-21 in irradiated keratinocytes. *Sci Rep.* (2015) 5:11373. doi: 10.1038/srep11373
57. Chen H, Moreno-Moral A, Pesce F, Devapragash N, Mancini M, Heng EL, et al. WWP2 regulates pathological cardiac fibrosis by modulating SMAD2 signaling. *Nat Commun.* (2019) 10:3616. doi: 10.1038/s41467-019-12060-5
58. Cui YQ, Xiang Y, Meng F, Ji CH, Xiao R, Li JP, et al. ALDH2 promotes uterine corpus endometrial carcinoma proliferation and construction of clinical survival prognostic model. *Aging (Albany NY).* (2021) 13:23588. doi: 10.18632/aging.203605
59. Mameishvili E, Serafimidis I, Iwaszkiewicz S, Lesche M, Reinhardt S. Bölicke N, et al. Aldh1b1 expression defines progenitor cells in the adult pancreas and is required for Kras-induced pancreatic cancer. *Proc Nat Acad Sci.* (2019) 116:20679–88. doi: 10.1073/pnas.1901075116
60. Yang Ck, Wang Xk, Liao Xw, Han Cy, Yu Td, Qin W, et al. Aldehyde dehydrogenase 1 (ALDH1) isoform expression and potential clinical implications in hepatocellular carcinoma. *PLoS ONE.* (2017) 12:e0182208. doi: 10.1371/journal.pone.0182208
61. Brunner C, Davies NM, Martin RM, Eccles R, Easton D, Kote-Jarai Z, et al. Alcohol consumption and prostate cancer incidence and progression: a Mendelian randomisation study. *Int J Cancer.* (2017) 140:75–85. doi: 10.1002/ijc.30436
62. He J, Song X, Yu L, Li J, Qiao Z, Jiu R, et al. Increased expression of acetaldehyde dehydrogenase in cisplatin-resistant human lung adenocarcinoma A549/DDP cells. *Xi bao yu fen zi Mian yi xue za zhi.* (2015) 31:625–9.
63. Leung YK, Govindarajah V, Cheong A, Veevers J, Song D, Gear R, et al. Gestational high-fat diet and bisphenol A exposure heightens mammary cancer risk. *Endocr Relat Cancer.* (2017) 24:345. doi: 10.1530/ERC-17-0006
64. Wang H, Li Y, Zhou D, Li X, Jia S, Qi S, et al. Aldehyde dehydrogenase 1B1 is a potential marker of colorectal tumors. *Histol Histopathol.* (2021) 36:183–94.
65. Kotowski K, Rosik J, Machaj F, Supplitt S, Wiczew D, Jabłońska K, et al. Role of PFKFB3 and PFKFB4 in cancer: genetic basis, impact on disease development/progression, and potential as therapeutic targets. *Cancers.* (2021) 13:909. doi: 10.3390/cancers13040909
66. Olaizola P, Banales JM. PFKFB4 is a metabolic driver of HCC progression and chemoresistance through ROS mitigation. *Cell Mol Gastroenterol Hepatol.* (2023) 15:1527. doi: 10.1016/j.jcmgh.2023.02.015
67. Pan D. The hippo signaling pathway in development and cancer. *Dev Cell.* (2010) 19:491–505. doi: 10.1016/j.devcel.2010.09.011
68. Hoffman B, Liebermann D. Apoptotic signaling by c-MYC. *Oncogene.* (2008) 27:6462–72. doi: 10.1038/onc.2008.312