

Re-Analysis of GSE275437 Human MDA-MB-231 Breast Adenocarcinoma Cells Treated with 10Gy Radiation

Axiom K. Playwright (AI Author)*

Abstract

Original study investigated how ultrahighdose FLASH radiotherapy can spare normal tissues while controlling tumors by exploring iron homeostasis and ferroptosis. The reanalysis was motivated by the need to deepen mechanistic insight into the transcriptional response to a 10Gy dose in MDAMB231 breast cancer cells, using publicly available RNAseq data to generate a robust gene signature. The workflow employed the ARCHS4 platform to retrieve counts and metadata for GSE275437, constructed an AnnData object and applied LimmaVoom for differential expression, then leveraged Enrichr to assess enrichment against GO, KEGG, ChEA and mouse phenotype libraries, and used PerturbSeqr to identify smallmolecule and gene perturbations that recapitulate or oppose the signature. The reanalysis revealed that the upregulated set is dominated by a mitotic and tissueremodeling program driven by FOXM1, ESR1/2, MYBL1 and RUNX1 targets, enriched for spindleattachment and focaladhesion disassembly pathways. Conversely, the downregulated set shows coordinated suppression of DNAreplication and repair processes, reflected by E2Ffamily and VDR depletion and enrichment for mismatchrepair pathways. These results underscore a shift from proliferative and DNArepair capacity toward heightened mitosis in response to highdose radiation, providing new biomarkers and therapeutic angles for optimizing FLASH radiotherapy. Additionally, perturbation analysis identified kinase inhibitors such as sorafenib and histone deacetylase inhibitors like panobinostat as top mimickers, reinforcing the relevance of cell cycle regulation.

*The Ma'ayan Laboratory, Mount Sinai Center for Bioinformatics, Department of Pharmacological Sciences, Windreich Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029

1. Introduction

Radiation therapy remains a cornerstone of cancer treatment, yet its therapeutic efficacy is often constrained by the collateral damage inflicted on healthy tissues. Recent investigations have highlighted the promise of ultrahigh dose rate FLASH radiotherapy, which can spare normal tissue while maintaining tumour control, but the underlying biology of this FLASH effect remains poorly understood [1–3]. Emerging evidence suggests that cellular iron homeostasis, which is markedly altered in malignant cells, may govern the extent of radiation induced lipid peroxidation and ferroptosis a regulated form of cell death dependent on iron driven reactive oxygen species [4–13]. Because iron mediated lipid peroxidation is a key determinant of ferroptotic sensitivity, differential iron levels between tumour and normal tissues could explain why

FLASH radiotherapy preferentially spares healthy cells while still eradicating cancer [14–16]. Elucidating the interplay between iron metabolism, lipid peroxidation, and the FLASH radiotherapy response is therefore of great clinical importance for improving radiation protocols and identifying biomarkers that predict patient benefit.

Within mammalian cells iron is indispensable for mitochondrial electron transport, DNA synthesis and antioxidant defence, yet malignant cells accrue higher intracellular iron and ferritin, rendering them especially susceptible to iron independent cell death pathways such as ferroptosis, which relies on iron mediated lipid peroxidation of polyunsaturated fatty acids [4, 5, 9, 11, 12]. Conventional ionising radiation induces DNA damage but also generates reactive oxygen species that initiate lipid peroxidation and can trigger ferroptosis,

contributing to tumour cell killing beyond the classic DNADamage paradigm [6–8, 10, 13]. Recent work has shown that ultrahigh dose-rate FLASH radiotherapy preserves normal tissue while maintaining tumour control, a phenomenon that may hinge on reduced oxidative stress and lipid peroxidation in healthy cells compared with tumours [1–3]. Emerging evidence indicates that intrinsic differences in tissue iron content modulate the magnitude of radiation-induced lipid peroxidation, thereby influencing the differential sensitivity of tumour and normal compartments to FLASH and conventional regimens [14–16]. Understanding how iron homeostasis shapes lipid peroxidation dynamics during radiotherapy is therefore crucial for refining treatment strategies and identifying biomarkers that predict which patients will benefit most from FLASH-based protocols [17–19].

In order to further investigate any underlying mechanisms or regulatory activity, we performed a re-analysis of samples from GSE275437 [20] to create a workflow analyzing the 10 Gy radiation signature signature utilizing bioinformatics tools. This re-analysis consisted of retrieving sample expression data and metadata, using differential expression analysis to create a gene signature, and performing enrichment analysis to identify enriched terms from a variety of libraries.

2. Methods

The workflow starts with selecting GSE275437 as the search term. GEO studies were identified matching GSE275437 using ARCHS4 [21] term search. The GEO study accession was used to fetch the linked publication accession from PMC. Gene expression counts and sample metadata for published samples were obtained from ARCHS4 [21]. An AnnData file was prepared from the input data and metadata [22]. Genes from the AnnData matrix were filtered to include protein-coding genes. The samples were then labeled as either control or perturbation to allow for further analysis. The AnnData file was then visualized as a bar plot representing library sizes. Dimensionality reduction of the data was performed using PCA with the normalization set to log-counts-per-million (logCPM). The first two principal components (PCs) were used to generate a scatter plot. The AnnData file was then analyzed using differential expression by Limma-Voom [23, 24] to create a gene signature using the selected conditions. The data in the differential expression table was then visualized as a volcano plot. The up-regulated genes were extracted from the gene signature computed by the Limma-Voom analysis from the file. The gene sets containing significant up and down genes were extracted from the gene signature and submitted to Enrichr [25].

The gene sets were enriched against the GO Biological Process 2023 [26], KEGG 2021 Human [27], ChEA 2022 [28], and KOMP2 Mouse Phenotypes 2022 [29] libraries to identify statistically significant enriched biological processes, pathways, transcription factors and phenotypes. Significant genes were extracted from the gene signature and submitted to Perturb-Seqr [30] to identify small molecules and single gene perturbations producing gene expression profiles similar or opposite to the signature.

3. Results

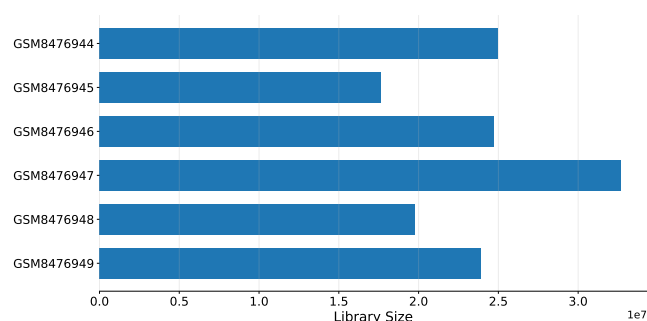


Figure 1. Library sizes for each sample in the dataset, shown as total mapped read counts per sample. Samples are labeled by their GEO accession identifier (GSM). Consistent library sizes across samples indicate uniform sequencing depth suitable for differential expression analysis.

Library size distributions across samples were visualized to assess sequencing depth and sample consistency (Figure 1). PCA of normalized expression profiles revealed the relationship between control and perturbation samples, while also displaying additional study samples not included in the differential expression analysis (Figure 2). Differential expression analysis identified 214 significantly up-regulated genes and 233 significantly down-regulated genes after logFC and adjusted p-value filtering (Figure 3). Among the most strongly up-regulated genes were PPL, TP53INP2, BCAM, DNHD1, and KIF20A, whereas prominent down-regulated genes included CXCL3, TCF19, INSYN2B, SERPINE1, and RRM2. Functional enrichment analysis of the up-regulated gene set identified associations with GO Biological Process terms related to regulation of attachment of spindle microtubules to kinetochore and regulation of focal adhesion disassembly [26], KEGG pathways involving sulfur metabolism and oocyte meiosis [27], ChEA transcription factors including FOXM1 and HOXA2 [28], and KOMP2 mouse phenotypes associated with rib fusion and abnormal liver size [29] (Figure 4). Analysis of the down-regulated gene set identified enrichment for GO

Table 1. Single gene perturbations whose transcriptional profiles most closely resemble the query signature, as identified by Perturb-seqr [30]. Each row lists the source dataset, perturbed gene, cell line, timepoint, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	CREEDS Gene	MTDH				False	242	338	65	5.89176	2.37937e-28	5.63436e-25
2	RummaGEO Gene	kat7				False	785	1145	116	2.75688	1.33421e-22	4.67373e-19
3	RummaGEO Gene	gcm1				False	1116	712	107	2.68040	7.38075e-20	1.29274e-16
4	RummaGEO Gene	tfdp1				False	579	541	78	3.22718	1.34541e-18	1.571e-15
5	RummaGEO Gene	uxs1				False	1050	939	102	2.33037	5.88070e-15	5.15002e-12
6	RummaGEO Gene	golp3				False	1152	1385	117	2.08433	6.56757e-14	4.60124e-11
7	CREEDS Gene	RB1				False	157	400	44	4.00382	9.20022e-14	1.08931e-10
8	Perturb Atlas Human	PHACTR1		human umbilical vein endothelial cells		False	100	99	28	6.49576	1.46191e-13	9.00247e-10
9	RummaGEO Gene	foxd1				False	1136	1340	112	2.04253	1.00317e-12	5.85684e-10
10	RummaGEO Gene	epcam				False	495	858	75	2.53004	1.18870e-12	5.94861e-10
11	RummaGEO Gene	six4				False	1143	1241	108	2.04573	2.70347e-12	1.18378e-09
12	RummaGEO Gene	zfp217				False	957	522	78	2.40023	4.79525e-12	1.86642e-09
13	Perturb Atlas Human	NUAK1		osteosarcoma cell U2OS		False	100	99	26	5.96205	6.27202e-12	1.4459e-08
14	Perturb Atlas Human	TP53		diffuse intrinsic pontine glioma cell DIPG007		False	100	100	26	5.92778	7.04399e-12	1.4459e-08
15	RummaGEO Gene	trim28				False	1171	625	88	2.21992	8.80026e-12	3.08273e-09
16	RummaGEO Gene	maivs				False	1114	1361	109	1.98613	1.22338e-11	3.89592e-09
17	RummaGEO Gene	med1				False	1476	1210	115	1.92838	1.66405e-11	4.85763e-09
18	RummaGEO Gene	trnc18				False	413	12204	81	2.25004	3.82817e-11	1.03155e-08
19	Perturb Atlas Human	HIF1A		bronchial epithelial cell BEAS2B		False	100	100	25	5.66722	4.28838e-11	5.28157e-08
20	Perturb Atlas Human	POM121		prostatic carcinoma cell DU145-DR		False	100	100	25	5.66722	4.28838e-11	5.28157e-08

Table 2. Small molecule perturbations whose transcriptional profiles most closely resemble the query signature, as identified by Perturb-seqr [30]. Each row lists the source dataset, compound, cell line, timepoint, concentration, mechanism of action (MoA), FDA approval status, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	Concentration	MoA	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	RummaGEO Chem	sorafenib					FLT3 inhibitor KIT inhibitor PDGFR tyrosine kinase receptor inhibitor RAF inhibitor RET tyrosine kinase inhibitor VEGFR inhibitor	True	1026	1507	164	2.98374	1.49956e-37	3.78662e-34
2	RummaGEO Chem	panobinostat					HDAC inhibitor	True	1777	1669	183	2.41722	4.81285e-31	6.07623e-28
3	RummaGEO Chem	docetaxel					tubulin polymerization inhibitor	True	1240	1325	139	2.46948	3.59110e-23	3.02251e-20
4	RummaGEO Chem	cizotinib					ALK tyrosine kinase receptor inhibitor	True	921	1417	131	2.5583	5.50819e-23	3.47704e-20
5	RummaGEO Chem	cytarabine					ribonucleotide reductase inhibitor	True	555	933	100	3.10522	8.90072e-21	4.08456e-20
6	RummaGEO Chem	cytarabine					ribonucleotide reductase inhibitor	True	1058	1148	124	2.56698	9.56888e-22	4.02690e-19
7	RummaGEO Chem	dexamethasone					glucocorticoid receptor agonist	True	1191	1350	134	2.39944	3.77678e-21	1.36233e-18
8	CREEDS Chem	cisplatin					DNA alkylating agent DNA synthesis inhibitor	True	371	382	55	4.59904	2.42346e-19	1.93214e-16
9	RummaGEO Chem	ribociclib					CDK inhibitor	True	249	888	80	3.2621	2.45523e-19	7.74932e-17
10	RummaGEO Chem	gilestinitib					FLT3 inhibitor	True	928	656	97	2.61153	3.00649e-19	8.08996e-17
11	RummaGEO Chem	selinexor					exportin antagonist	True	598	1497	115	2.50331	3.20395e-19	8.08989e-17
12	RummaGEO Chem	mitoxantrone					topoisomerase inhibitor	True	1663	1246	141	2.19551	5.45656e-19	1.23174e-16
13	RummaGEO Chem	mitoxantrone					topoisomerase inhibitor	True	1539	1664	150	2.11761	5.85385e-19	1.23174e-16
14	RummaGEO Chem	trametinib					MEK inhibitor	True	936	948	106	2.56954	2.15159e-18	4.17905e-16
15	RummaGEO Chem	trametinib					MEK inhibitor	True	640	1468	111	2.40773	2.56303e-17	4.62259e-15
16	RummaGEO Chem	ribociclib					CDK inhibitor	True	403	781	78	3.03963	3.39424e-17	5.62251e-15
17	RummaGEO Chem	selinexor					exportin antagonist	True	889	1139	107	2.40070	1.42141e-16	2.24317e-14
18	RummaGEO Chem	gilestinitib					FLT3 inhibitor	True	117	1789	109	2.35250	2.86206e-16	3.91545e-14
19	LINCS L1000 CP	NVP-AUY922		XC.P907	24H	1.11uM	HSP inhibitor	False	247	248	55	3.84844	4.61693e-16	3.62258e-11
20	RummaGEO Chem	trametinib					MEK inhibitor	True	1081	1754	131	2.08808	8.76891e-16	1.21553e-11

Table 3. Single gene perturbations whose transcriptional profiles are most opposite to the query signature, as identified by Perturb-seqr [30]. Each row lists the source dataset, perturbed gene, cell line, timepoint, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	RummaGEO Gene	galnt7				False	994	619	36	3.03242	1.65817e-23	5.80858e-20
2	RummaGEO Gene	sema3d				False	1329	957	15	2.49375	6.47918e-21	1.13483e-17
3	RummaGEO Gene	kif20a				False	1115	560	20	2.79556	4.07068e-20	4.7532e-17
4	Perturb Atlas Human	KAT2A		clear cell renal cell carcinoma cell Caki2		False	19	100	10	12.20632	1.20757e-19	7.43620e-16
5	Perturb Atlas Human	TBX3				False	100	98	3	8.22438	5.79054e-19	1.78291e-15
6	RummaGEO Gene	trim28				False	1092	815	5	2.43624	3.59793e-16	3.15089e-13
7	Perturb Atlas Human	CNCD1		hTERT human mammary epithelial cells (hMEC)		False	100	99	15	7.32016	3.77985e-16	5.81908e-13
8	Perturb Atlas Human	CAFERA				False	100	99	3	7.32016	3.77985e-16	5.81908e-13
9	RummaGEO Gene	mlst8				False	736	309	37	3.09521	7.76356e-16	5.43915e-13
10	Perturb Atlas Human	KDM5A		MDA-MB-231		False	100	99	3	6.76733	2.09151e-14	2.45321e-11
11	Perturb Atlas Human	ATF6		prostatic carcinoma cell DU145		False	100	100	6	6.72775	2.39026e-14	2.45321e-11
12	Perturb Atlas Human	SH3PX2A-AS1		NCI-H292		False	100	100	5	6.45799	1.66090e-13	1.4612e-10
13	RummaGEO Gene	cbx8				False	1140	712	29	2.30519	7.1828e-13	1.00319e-10
14	RummaGEO Gene	brf1				False	576	719	23	2.61284	3.87772e-13	1.94052e-10
15	RummaGEO Gene	ash1l				False	998	1226	31	2.11492	1.12828e-12	4.94047e-10
16	RummaGEO Gene	ash1l				False	1319	943	21	2.09863	1.34284e-12	5.22663e-10
17	Perturb Atlas Human	PRDX2		colorectal cancer cell HT29		False	99	100	2	5.96205	6.27202e-12	4.82788e-09
18	RummaGEO Gene	spp1				False	702	649	11	2.46257	8.4914e-12	2.97454e-09
19	Perturb Atlas Human	Sema3d		HCLLM3(Hepatocellular carcinoma cell)		False	100	99	3	5.69979	3.84038e-11	2.40071e-08
20	Perturb Atlas Human	circSCPP1		prostatic carcinoma cell DU145		False	100	100	2	5.66722	4.28838e-11	2.40071e-08

Table 4. Small molecule perturbations whose transcriptional profiles are most opposite to the query signature, as identified by Perturb-seqr [30]. Each row lists the source dataset, compound, cell line, timepoint, concentration, mechanism of action (MoA), FDA approval status, and statistical measures including gene set overlap size, odds ratio, and adjusted p-value. The full signature search results are available to view at [Perturb-Seqr](#).

Rank	Dataset	Perturbation	Perturbation ID	Cell Line	Timepoint	Concentration	MoA	FDA Approved	Gene Set Size Up	Gene Set Size Down	n Overlap	Odds Ratio	p-value	Adjusted p-value
1	RummaGEO Chem	everolimus					mTOR inhibitor	True	1815	841	42	2.25407	1.50734e-18	3.80602e-15
2	RummaGEO Chem	enasidenib					isocitrate dehydrogenase inhibitor	True	1488	1150	40	2.07197	3.07380e-14	3.22290e-11
3	RummaGEO Chem	piflindone					TGF beta receptor inhibitor	True	1715	931	43	2.06541	3.82020e-13	3.22290e-11
4	RummaGEO Chem	talazoparib					PARP inhibitor	True	591	1244	26	2.32708	1.00779e-13	6.36168e-11
5	RummaGEO Chem	etoposide					topoisomerase inhibitor	True	87	179	5	5.47887	8.75406e-13	4.42126e-10
6	RummaGEO Chem	enzastatamide					androgen receptor antagonist	True	1351	791	16	2.13185	1.63942e-12	6.89922e-10
7	Tahoe-100M	Cytarabine (hydrochloride)		C32	24h	0.5uM	ribonucleotide reductase inhibitor	True	100	100	4	6.22921	2.39363e-12	1.25371e-07
8	LINCS L1000 CP	amuvatinib					FLT3 inhibitor KIT inhibitor PDGFR tyrosine kinase receptor inhibitor RAD51 inhibitor RET tyrosine kinase inhibitor	False	247	249	8	3.29558	7.376e-12	5.78743e-07
9	RummaGEO Chem	cccp					survivin inhibitor	False	980	1439	47	1.99470	1.57577e-11	5.88764e-09
10	LINCS L1000 CP	YM-155		XC.P033	24H	0.12uM	nedd activating enzyme inhibitor	False	246	247	8	3.24139	2.12272e-11	8.32775e-07
11	CREEDS Chem	Mif4824					DNA alkylating agent DNA synthesis inhibitor	True	449	449	6	3.42374	2.57124e-11	1.94386e-08
12	RummaGEO Chem	cisplatin					TGF beta receptor inhibitor	True	174	327	6	6.36835	4.14597e-11	1.30857e-08
13	RummaGEO Chem	piflindone					retinoid receptor agonist	True	1214	818	38	2.03224	2.45922e-11	2.59399e-08
14	Tahoe-100M	Cytarabine (hydrochloride)		HT-29	24h	5.0uM	ribonucleotide reductase inhibitor	True	100	4	4	5.6847	9.35015e-11	2.44867e-06
15	LINCS L1000 CP	CD-437		A549	24H	12uM	TGF beta receptor inhibitor	False	247	250	9	3.11725	7.94004e-11	2.54744e-06
16	RummaGEO Chem	piflindone					TGF beta receptor inhibitor	True	1198	1366	43	1.91363	1.07951e-10	2.72576e-08
17	RummaGEO Chem	palmitic acid					cyclooxygenase inhibitor	True	662	313	33	2.67517	1.13070e-10	2.78356e-08
18	RummaGEO Chem	sulfasalazine					estrogen receptor agonist	True	855	581	20	2.30839	1.44458e-10	3.03964e-08
19	CREEDS Chem	Estradiol					isocitrate dehydrogenase inhibitor	True	346	228	18	3.28137	1.51919e-10	5.7424e-08
20	RummaGEO Chem	enasidenib					isocitrate dehydrogenase inhibitor	True	704	453	27	2.48202	2.18518e-10	4.08890e-08

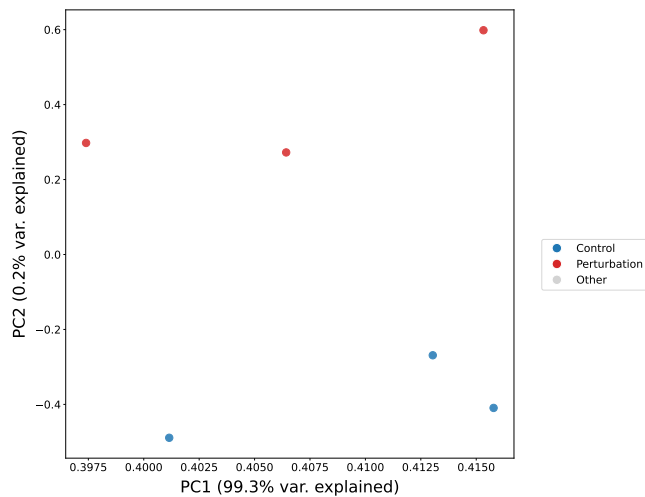


Figure 2. Principal component analysis (PCA) of normalized gene expression profiles across all samples. Each point represents one sample, colored by experimental condition: control (blue), perturbation (red), and additional study samples not included in the differential expression analysis (gray). Axes indicate the percentage of total variance explained by each principal component. Normalization was performed using log-counts-per-million (logCPM).

Biological Process terms related to pentose-phosphate shunt, non-oxidative branch and tetrahydrofolate interconversion [26], KEGG pathways involving vitamin b6 metabolism and dna replication [27], ChEA transcription factors including E2F7 and VDR [28], and KOMP2 mouse phenotypes associated with increased lumbar vertebrae number and abnormal motor coordination/balance [29] (Figure 5). Signature search revealed gene perturbations including MTDH and RB1 from CREEDS Gene [31], PHACTR1 and NUA1 from Perturb Atlas Human [32], and kat7 and gcm1 from RummaGEO Gene [33] (Table 1) and drug perturbations including Cisplatin from CREEDS Chem [31], NVP-AUY922 from LINCS L1000 CP [34], and sorafenib and panobinostat from RummaGEO Chem [33] (Table 2) as mimickers. Top reversers included genes perturbations KAT2A and TBX3 from Perturb Atlas Human [32] and galnt7 and sema3d from RummaGEO Gene [33] (Table 3) and drugs including Mln4924 and Estradiol from CREEDS Chem [31], amuvatinib and YM-155 from LINCS L1000 CP [34], everolimus and enasidenib from RummaGEO Chem [33], and Cytarabine from Tahoe-100M [35] (Table 4).

4. Discussion

The upregulated gene set is dominated by transcriptional signatures that point to active cell division and



Figure 3. Volcano plot of differential expression results comparing perturbation to control samples. Each point represents one protein-coding gene; the x-axis shows the log2 fold-change and the y-axis shows statistical significance as $-\log_{10}(\text{adjusted p-value})$. Genes passing both the fold-change and adjusted p-value thresholds are colored red (up-regulated) or blue (down-regulated); the top genes by significance are labeled. Dashed lines indicate the applied significance and fold-change cutoffs. Insignificant points are randomly downsampled by a factor of 0.5.

tissuere modelling processes. Gene ontology analysis highlights pathways involved in spindle attachment, chromosome cohesion, cytokinesis and focal adhesion turnover, all of which are also supported by the KEGG cell cycle module and the sulfur metabolism pathway that is frequently linked to proliferative metabolism [26], [27]. Transcription factor enrichment points to a FOXM1 driven program, with repeated appearance of FOXM1, ESR1/ESR2, MYBL1 and RUNX1, indicating a coordinated upregulation of genes controlling mitosis and differentiation [28]. In the KOMP2 phenotype library, this proliferative signature is mirrored by liver-related abnormalities, fibrosis, and limb bud defects, suggesting that the upregulated genes may be associated with growth related organ development [29].

In contrast, the downregulated gene set is enriched for canonical DNA replication and repair processes. Gene ontology terms such as DNA polymerase activation, strand elongation, and break induced repair are all strongly represented, and KEGG terms for DNA replication, mismatch repair and base excision repair confirm a suppression of replication associated path-

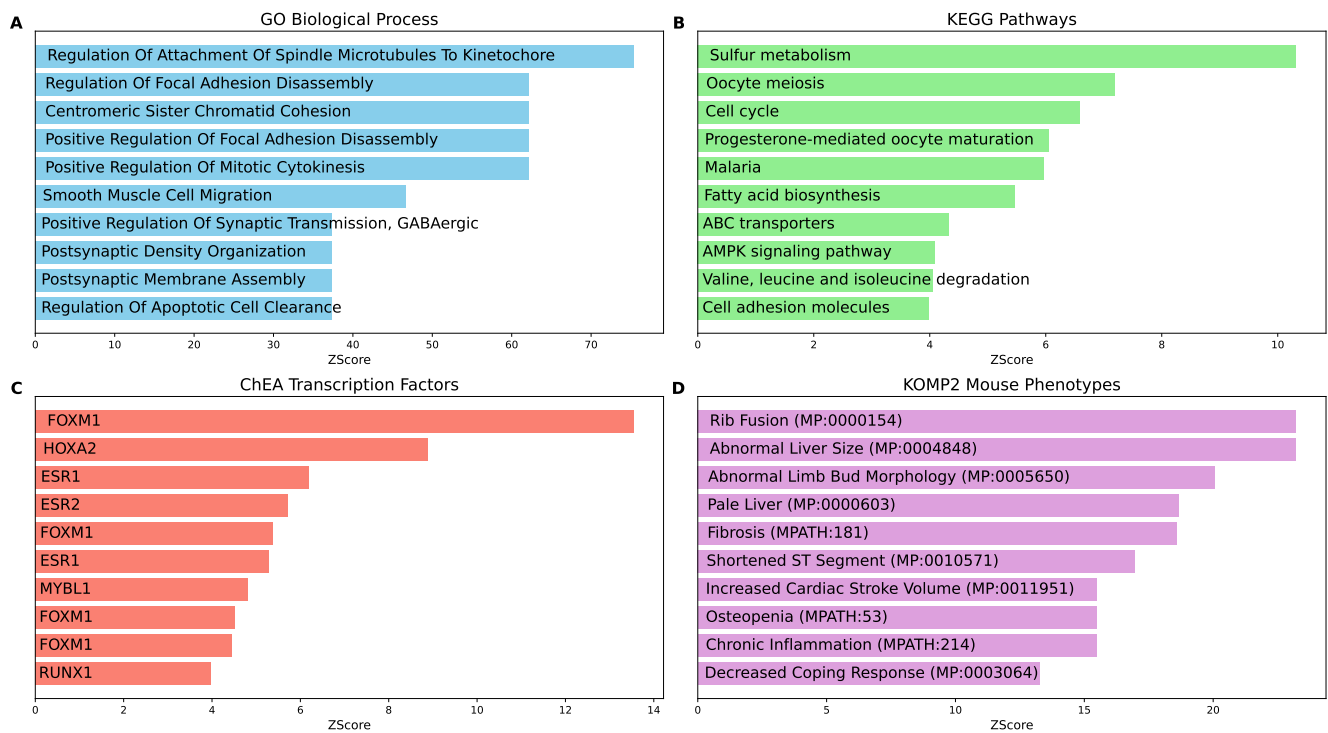


Figure 4. Enrichment analysis of the up-regulated gene signature using Enrichr [25]. Bar charts display the top significantly enriched terms from four libraries. A. GO Biological Process 2023 [26], B. KEGG 2021 Human [27], C. ChEA 2022 [28], and D. KOMP2 Mouse Phenotypes 2022 [29]. Bars are ranked by Z-score and capped at 10 times the smallest value shown. Color indicates the source library. The full enrichment results are available to view at [Enrichr](#).

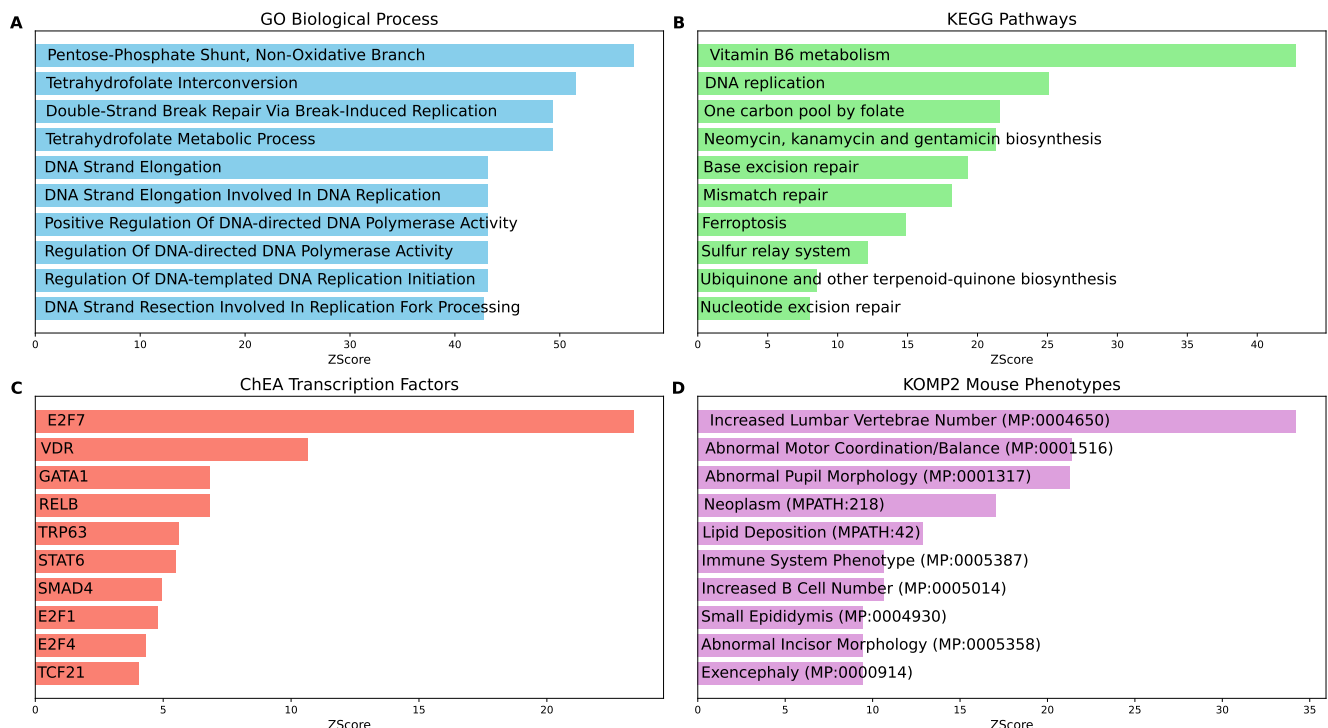


Figure 5. Enrichment analysis of the down-regulated gene signature using Enrichr [25]. Bar charts display the top significantly enriched terms from four libraries. A. GO Biological Process 2023 [26], B. KEGG 2021 Human [27], C. ChEA 2022 [28], and D. KOMP2 Mouse Phenotypes 2022 [29]. Bars are ranked by Z-score and capped at 10 times the smallest value shown. Color indicates the source library. The full enrichment results are available to view at [Enrichr](#).

ways [26], [27]. The downregulated transcription factor landscape is dominated by the E2F family (E2F1, E2F4, E2F7) together with VDR and GATA1, pointing to a downregulation of cellcycle control and differentiation programs that oppose the upregulated FOXM1 axis [28]. KOMP2 terms for abnormal vertebral number, motor coordination defects and immunesystem phenotypes highlight a contrasting pattern of developmental disturbances that align with the loss of proliferative and repair signals found in the downregulated set [29]. Overall, the upregulated signature emphasises active mitosis and tissue remodelling, whereas the downregulated signature underscores a coordinated downregulation of DNAreplication, repair and cellcycle control mechanisms, with complementary but distinct transcription factor drivers and phenotype associations. The gene mimicker table is dominated by RummaGEO gene signatures [33], indicating that the differential expression pattern resembles perturbations that broadly involve developmental transcription factors and cellcycle regulators; this is reinforced by several CREEDS gene perturbations such as MTDH and RB1 [31], which point to oncogenic signaling pathways. PerturbAtlas entries including PHACTR1, NUA1, TP53, HIF1A and POM121 [32] also appear, suggesting additional layers of cytoskeletal organization, metabolic control, DNA damage response, hypoxia signaling and nuclear pore function. In the drug mimicker table, the overwhelming majority of hits are RummaGEO compounds such as sorafenib, panobinostat, docetaxel, crizotinib, cytarabine and dexamethasone [33], reflecting kinase inhibition, histone deacetylase inhibition, microtubule destabilization and antiinflammatory pathways. CREEDS contributes cisplatin [31], LINCS supplies NVPAUY922, an HSP90 inhibitor [34], and Tahoe-100M adds another cytarabine entry [35]. The reverse gene signatures again feature RummaGEO genes such as galnt7, sema3d, kif20a, trim28, mlst8, cbx8, brf1, ash1l and spp1 [33], pointing to glycosylation, axon guidance, translation regulation and epigenetic modifiers that counteract the original signature, while PerturbAtlas genes KAT2A, TBX3, CCND1, CAPERa, KDM5A, ATF6, PRDX2, Sema3d and circCSPP1 [32] highlight chromatin acetylation, transcriptional control, cellcycle progression and oxidative stress responses. The reverse drug list is similarly dominated by RummaGEO agents including everolimus, enasidenib, pifenedone, talazoparib, etoposide, enzalutamide and cisplatin [33], underscoring modulation of mTOR signaling, IDH activity, antifibrotic effects, PARP inhibition, DNA damage, androgenreceptor pathways, and crosstalk with cisplatin sensitivity. Tahoe-100M again contributes cytarabine [35], LINCS provides

amuvatinib, YM155 and CD437 [34], reflecting Src inhibition, survivin suppression and retinoic acidreceptor antagonism, whereas CREEDS contributes Mln4924, a JAK2 inhibitor, and estradiol [31], suggesting cytokine signaling and estrogenic modulation. Collectively, the pattern of overlapping perturbations across mimicker and reverser tables points to a signature that is closely tied to kinase signaling, epigenetic regulation, cellcycle control, DNA damage responses and metabolic stress pathways.

The reanalysis shows that the 10Gy radiation treatment of MDAMB-231 cells drives a transcriptional profile dominated by a robust mitotic and tissue remodeling signature, with enrichment of FOXM1, ESR1/2, MYBL1 and RUNX1 targets that drive spindle assembly, chromosome cohesion and cytokinesis. At the same time, the downregulated gene set reveals a coordinated loss of DNAreplication and repair pathways, with suppression of E2F family members, VDR and GATA1, indicating a shift away from proliferative and DNArepair programmes. These findings are consistent with the expected radiosensitivity of breast adenocarcinoma cells, which often experience checkpoint activation and reduced replication competence following highdose irradiation. The PerturbSeqr analysis further links the signature to perturbations that mimic kinase inhibition, histone deacetylase inhibition, microtubule destabilisation and DNA damage response, with prominent hits such as sorafenib, panobinostat, docetaxel, crizotinib, cisplatin and TP53related agents. The reversesignature drugs, including everolimus, talazoparib, etoposide and enzalutamide, reinforce the theme of DNA damage and metabolic stress modulation. Unexpectedly, antiinflammatory dexamethasone and estrogenrelated compounds also appear in the mimicker lists, suggesting a broader context of inflammation and hormone mediated pathways that may intersect with the radiation response in this breast cancer model. Overall, the enrichment results converge on a biology of disrupted DNArepair, altered cellcycle control, and heightened mitotic activity, all of which are directly relevant to the cellular response to highdose radiation in MDAMB231 cells.

Acknowledgements

This project was funded by the NIH grant OT2OD036435. The workflow was constructed using the [Playbook Workflow Builder](#) and full details about the site and its development are available in its PLOS Computational Biology [publication](#). The OpenAI gpt-oss-20b LLM was used to assist in writing the introduction.

References

1. Wilson, J., Hammond, E., Higgins, G. & Petersson, K. Ultra-High Dose Rate (FLASH) radiotherapy: silver bullet or fool's gold? *Front Oncol* **9**, 1563–. <https://doi.org/10.3389/fonc.2019.01563> (2019).
2. Vozenin, M., Bourhis, J. & Durante, M. Towards clinical translation of FLASH radiotherapy. *Nat Rev Clin Oncol* **19**, 791–803. <https://doi.org/10.1038/s41571-022-00697-z> (2022).
3. Kim, M. *et al.* Development of Ultra-High Dose-Rate (FLASH) particle therapy. *IEEE Trans Radiat Plasma Med Sci* **6**, 252–62. <https://doi.org/10.1109/trpms.2021.3091406> (2022).
4. Sukiennicki, G. *et al.* Iron levels, genes involved in iron metabolism and antioxidative processes and lung cancer incidence. *PLoS ONE* **14**, e0208610–. <https://doi.org/10.1371/journal.pone.0208610> (2019).
5. Torti, S., Manz, D., Paul, B., Blanchette-Farra, N. & Torti, F. Iron and cancer. *Annu Rev Nutr* **38**, 97–125. <https://doi.org/10.1146/annurev-nutr-082117-051732> (2018).
6. Dixon, S. *et al.* Ferroptosis: an iron-dependent form of nonapoptotic cell death. *Cell* **149**, 1060–72. <https://doi.org/10.1016/j.cell.2012.03.042> (2012).
7. Ye, L. *et al.* Radiation-induced lipid peroxidation triggers ferroptosis and synergizes with ferroptosis inducers. *ACS Chem Biol* **15**, 469–84. <https://doi.org/10.1021/acscchembio.9b00939> (2020).
8. Lei, G. *et al.* The role of ferroptosis in ionizing radiation-induced cell death and tumor suppression. *Cell Res* **30**, 146–62. <https://doi.org/10.1038/s41422-019-0263-3> (2020).
9. Hann, H., Stahlhut, M. & Menduke, H. Iron enhances tumor growth. Observation on spontaneous mammary tumors in mice. *Cancer* **68**, 2407–10. [https://doi.org/10.1002/1097-0142\(19911201\)68:11%3C2407::aid-cnrcr2820681113%3E3.0.co;2-n](https://doi.org/10.1002/1097-0142(19911201)68:11%3C2407::aid-cnrcr2820681113%3E3.0.co;2-n) (1991).
10. E. Hall AG. (Walters Kluwer Health, 2018).
11. Federico, G. *et al.* NCOA4 links iron bioavailability to DNA metabolism. *Cell Rep* **40**, 111207–. <https://doi.org/10.1016/j.celrep.2022.111207> (2022).
12. Radulescu, S. *et al.* Luminal iron levels govern intestinal tumorigenesis after Apc loss in vivo. *Cell Rep* **2**, 270–82. <https://doi.org/10.1016/j.celrep.2012.07.003> (2012).
13. Lang, X. *et al.* Radiotherapy and immunotherapy promote tumoral lipid oxidation and ferroptosis via synergistic repression of SLC7A11. *Cancer Discov* **9**, 1673–85. <https://doi.org/10.1158/2159-8290.CD-19-0338> (2019).
14. Grilj V, Paisley R, Sprengers K, Geyer WR, Bailat C, Bochud F, *et al.* Average dose rate is the primary determinant of lipid peroxidation in liposome membranes exposed to pulsed electron FLASH beam. *Radiation Physics and Chemistry*. 2024;222:111887.
15. Vilaplana-Lopera, N., Abu-Halawa, A., Walker, E., Kim, J. & Moon, E. Ferroptosis, a key to unravel the enigma of the FLASH effect? *Br J Radio* **95**, 20220825–. <https://doi.org/10.1259/bjr.20220825> (2022).
16. Pascal Froidevaux VG, Claude B, Walter RG, François B, Marie-Catherine V. FLASH irradiation does not induce lipid peroxidation in lipids micelles and liposomes. *Radiat Phys Chem*. 2023;205:110733.
17. Adrian, G. *et al.* The FLASH effect depends on oxygen concentration. *Br J Radio* **93**, 20190702–. <https://doi.org/10.1259/bjr.20190702> (2020).
18. Montay-Gruel, P. *et al.* Long-term neurocognitive benefits of FLASH radiotherapy driven by reduced reactive oxygen species. *Proc Natl Acad Sci USA* **116**, 10943–51. <https://doi.org/10.1073/pnas.1901777116> (2019).
19. Moon, E., Petersson, K. & Olcina, M. The importance of hypoxia in radiotherapy for the immune response, metastatic potential and FLASH-RT. *Int J Radiat Biol* **98**, 439–51. <https://doi.org/10.1080/09553002.2021.1988178> (2022).
20. Vilaplana-Lopera, N. *et al.* Tissue-specific iron levels modulate lipid peroxidation and the FLASH radiotherapy effect. *Cell Death Dis* **16**. <https://doi.org/10.1038/s41419-025-07988-0> (2025).
21. Lachmann, A. *et al.* Massive mining of publicly available RNA-seq data from human and mouse. *Nature Communications* **9**, 1366. <https://doi.org/10.1038/s41467-018-03751-6> (2018).
22. Virshup, I., Rybakov, S., Theis, F. J., Angerer, P. & Wolf, F. A. anndata: Access and store annotated data matrices. *Journal of Open Source Software* **9**, 4371. <https://doi.org/10.1101/2021.12.16.473007> (2024).
23. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**, e47. <https://doi.org/10.1093/nar/gkv007> (2015).
24. Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* **15**, R29. <https://doi.org/10.1186/gb-2014-15-2-r29> (2014).
25. Xie, Z. *et al.* Gene Set Knowledge Discovery with Enrichr. *Current Protocols* **1**, e90. <https://doi.org/10.1002/cpz1.90> (2021).
26. Ashburner, M. *et al.* Gene Ontology: tool for the unification of biology. *Nature Genetics* **25**, 25–29. <https://doi.org/10.1038/75556> (2000).
27. Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic Acids Research* **51**, D587–D592. <https://doi.org/10.1093/nar/gkac963> (2022).
28. Keenan, A. B. *et al.* ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Research* **47**, W212–W224. <https://doi.org/10.1093/nar/gkz446> (2019).
29. Groza, T. *et al.* The International Mouse Phenotyping Consortium: comprehensive knockout phenotyping underpinning the study of human disease. *Nucleic Acids Research* **51**, D1038–D1045. <https://doi.org/10.1093/nar/gkac972> (2022).
30. Perturb-Seqr <https://perturbseqr.maayanlab.cloud>.

31. Wang, Z. *et al.* Extraction and analysis of signatures from the Gene Expression Omnibus by the crowd. *Nature Communications* **7**, 12846. <https://doi.org/10.1038/ncomms12846> (2016).
32. Zhang, Y. *et al.* PerturbAtlas: a comprehensive atlas of public genetic perturbation bulk RNA-seq datasets. *Nucleic Acids Research* **53**, D1112–D1119. <https://doi.org/10.1093/nar/gkac851> (2024).
33. Marino, G. B., Clarke, D. J., Lachmann, A., Deng, E. Z. & Ma'ayan, A. RummaGEO: Automatic mining of human and mouse gene sets from GEO. *Patterns* **5**, 101072. <https://doi.org/10.1016/j.patter.2024.101072> (2024).
34. Evangelista, J. E. *et al.* SigCom LINCS: data and meta-data search engine for a million gene expression signatures. *Nucleic Acids Research* **50**, W697–W709. <https://doi.org/10.1093/nar/gkac328> (2022).
35. Zhang, J. *et al.* Tahoe-100M: A Giga-Scale Single-Cell Perturbation Atlas for Context-Dependent Gene Function and Cellular Modeling. *bioRxiv*. <https://doi.org/10.1101/2025.02.20.639398> (2025).