Data in Brief

Genome wide expression after different doses of irradiation of a three-dimensional (3D) model of oral mucosal

Maria P. Lambros [a,*], Michael K. DeSalvo [b], Hari Chandana Mulamalla [a], Jonathan Moreno [a], Lavanya Kondapalli [a]

[a] College of Pharmacy, Western University of Health Sciences, Pomona, CA 91766, USA
[b] Gran, LLC, Fullerton, CA 92835, USA

ARTICLE INFO

ABSTRACT

We evaluated a three-dimensional (3D) human oral cell culture that consisted of two types of cells, oral keratinocytes and fibroblasts as a model of oral mucositis which is a debilitating adverse effect of chemotherapy and radiation treatment. The 3D cell culture model was irradiated with 12 or 2 Gy, and total RNA was collected 6 h after irradiation to compare global gene expression profiles via microarray analysis. Here we provide detailed methods and analysis on these microarray data, which have been deposited in Gene Expression Omnibus (GEO): GSE62395.

Specifications	
Organism/cell line/tissue	Human primary oral keratinocytes and fibroblasts grown in 3D culture
Sex	Male
Sequencer or array type	Phalanx Human OneArray v4.3 (GPL19298)
Data format	Raw and processed
Experimental factors	Control (non-irradiated) compared to: (1) irradiated with 12 Gy; (2) irradiated with 2 Gy
Experimental features	3D oral tissues were exposed to gamma irradiation at 12 Gy or 2 Gy. After irradiation, the tissues were incubated for 6 h at 37 °C with 5% CO_2. Subsequently, some of the tissues were used for the extraction of total RNA, and others were placed in 10% formalin for histopathological studies.
Consent	N/A
Sample source location	Pomona, CA

1. Direct link to deposited data

The deposited data can be found at: http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE62395.

2. Experimental design, materials and methods

2.1. Tissue culture and irradiation

Tissue culture methods and irradiation procedure have been described in [2]. The three dimensional tissue culture model (3D EpiOral) and media were purchased from MatTek Corporation (Ashland, MA). The 3D EpiOral tissue is a co-culture organotypic model and consisted of human fibroblasts on the bottom and human oral keratinocytes grown on top of the fibroblasts. In order to induce differentiation and stratification of the keratinocytes of the 3D cultures, the cultures were elevated to the air–liquid interface. Then the tissues were irradiated (dose of irradiation was 0, 2, or 12 Gy) at the City of Hope (Duarte, CA) facility. There were at least three 3D tissues per irradiation dose. Subsequently the tissues were incubated for 6 h at 37 °C with 5% CO_2 after irradiation. Then one of the tissues was placed in 10% formalin for histopathological studies and the others were used for the extraction of total RNA.

2.2. RNA isolation and microarray hybridization

The RNeasy Plus Mini Kit (Qiagen, Germantown, MD) was used to extract total RNA of at least two identically treated tissues. The total RNA of the identically treated tissues were combined and used for analysis. The integrity and quality of RNA was determined by evaluating the A260/280 absorbance ratio using an Agilent 2100 Bioanalyzer (Agilent Technologies, Palo Alto, CA). Only RNA samples with absorbance ratio, A260/280 > 2.0, and RIN > 9.5 were used. The RNA was converted to double-stranded cDNA and amplified using in vitro transcription
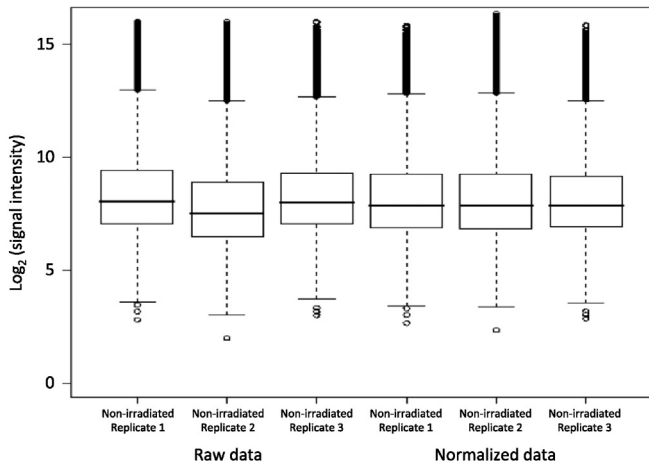
---

* Corresponding author.

Fig. 1. Representative box plot of raw and normalized data from three technical replicate hybridizations of the non-irradiated control sample. The box plot indicates median-centered raw data distributions, which were further refined during normalization (this was true for all samples). Taken as a whole, we see high repeatability of technical replicate hybridizations.

with T7 polymerase; the transcription reaction included aminoallyl UTP (aa-dUTP) and the subsequent product was conjugated to Cy5 NHS ester (GE Healthcare Life Sciences, Pittsburg, PA). The Human Whole Genome OneArray v4.3 (Phalanx Biotech, San Diego, CA) was used to perform the DNA microarray analysis. For this purpose, a quantity of 0.025 mg/mL fragmented Cy5-labeled cDNA was hybridized overnight at 42 °C using the HybBag mixing system with $1 \times$ OneArray Hybridization Buffer (Phalanx Biotech) and 0.01 mg/mL sheared salmon sperm DNA (Promega, Madison, WI). After the hybridization procedure, the microarrays were washed according to the OneArray protocol (Phalanx Biotech). The raw Cy5 intensity signals produced by each of the

Table 1
Representative Pearson correlation table for raw (R) and normalized (N) data from three technical replicate hybridizations of the non-irradiated (NI) control sample. High correlation values were found for all replicate hybridizations for each sample, which indicates high repeatability among replicate hybridizations and minimal data transformation during normalization.

	R_NI_1	R_NI_2	R_NI_3	N_NI_1	N_NI_2	N_NI_3
R_NI_1	1	0.953	0.955	1	0.953	0.955
R_NI_2	0.953	1	0.985	0.953	1	0.985
R_NI_3	0.955	0.985	1	0.955	0.985	1
N_NI_1	1	0.953	0.955	1	0.953	0.955
N_NI_2	0.953	1	0.985	0.953	1	0.985
N_NI_3	0.955	0.985	1	0.955	0.985	1

microarrays were captured using a Molecular Devices Axon™ 4100A scanner and measured using GenePix Pro™ software and stored in GPR format.

2.3. Microarray data pre-processing and statistical analysis

The data from all the arrays of each experimental set were analyzed using Rosetta Resolver System® (Rosetta Biosoftware, USA). Testing was done in triplicate by combining technical replicates while the statistical analyses were done using the proprietary modeling techniques of Rosetta Resolver [1]. The signal intensities were normalized using 75-percentile median centering. The error-weighted approach (which is specifically geared towards combining replicated hybridizations to improve measurement precision and accuracy) was used to calculate the average expression values.

In order to decide whether intensity data was significantly above background, P-values were calculated to test the null hypothesis that expression is absent (called "P-value detected"). A P-value detected <0.05 indicated that the transcript specific to a given probe was truly present or expressed. Lastly, we also
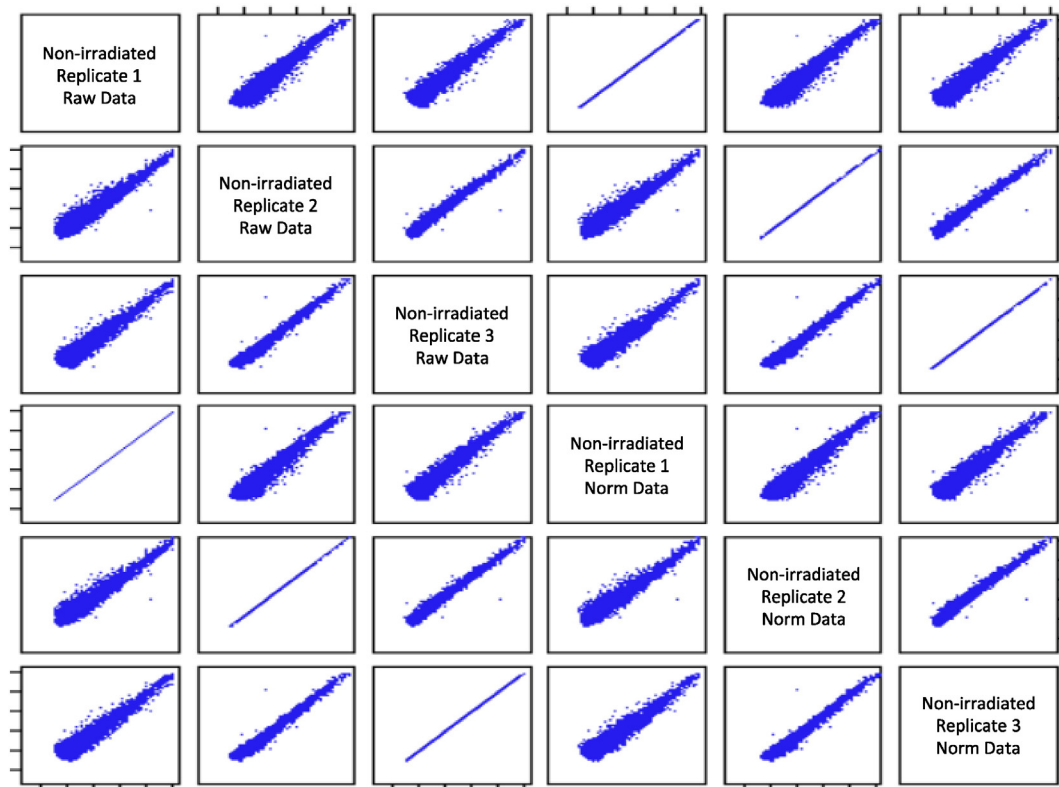

Fig. 2. Representative scatter plot of raw and normalized data from three technical replicate hybridizations of the non-irradiated control sample. The scatter plot indicates tight correlation between raw and normalized data replicates, and this was true for all samples. These results, along with the box plot results, point to high repeatability of technical replicate hybridizations.

Table 2

Selected categories from an enrichment analysis using DAVID Bioinformatics. The complete list of results, including the genes within each enriched term, can be found in Table S1. Irradiated samples were compared to non-irradiated control samples. GO BP = Biological Process. Count = the number of differentially expressed genes annotated with a given enriched term. Adj P-val = Benjamini-adjusted P-values.

Enriched Term	Category	Count	Adj P-val
Up-regulated genes from irradiated samples			
ubl conjugation	SP PIR keyword	37	2.30E − 04
Methylation	SP PIR keyword	21	3.78E − 04
Proto-oncogene	SP PIR keyword	20	5.66E − 04
GO:0051726—regulation of cell cycle	GO BP	26	0.013
GO:0031497—chromatin assembly	GO BP	12	0.016
GO:0010941—regulation of cell death	GO BP	44	0.037
GO:0042127—regulation of cell proliferation	GO BP	43	0.037
GO:0070482—response to oxygen levels	GO BP	14	0.049
Down-regulated genes from irradiated samples			
GO:0006396—RNA processing	GO BP	32	3.17E−05
tRNA processing	SP PIR keyword	8	0.017
Cell cycle	SP PIR keyword	20	0.028
Cell division	SP PIR keyword	14	0.038

calculated P-values for determining whether genes were differentially expressed. Rather than focusing solely on fold changes, Rosetta Resolver uses error-model-based hypothesis tests, which account for both fold change and expression level.

2.4. Microarray data quality control

Three technical replicate hybridizations were averaged while ensuring that the technical replicates were of high repeatability. Using the R function *boxplot*, raw and normalized log_2 data from each sample were plotted but we did not include control and flagged probes. Fig. 1 shows a representative box plot. We did not find any hybridization with intensities different from their technical replicates. This analysis not only helps finding hybridization with aberrant intensity distributions but also helps to confirm that the normalization of each replicate microarray has an appropriately centered distribution.

For each sample the scattered plots of raw and normalized log_2 data were compared using the R function *pairs*. We included only data with a P-value of <0.01. Fig. 2 shows a representative scatter plot. Pearson correlation tables were used to define the linear relationship of the two variables, raw and normalized data, in the scatter plots. For each technical repeat, we calculated correlation values for raw and normalized log_2 intensities and we included in the calculation only probes with P-value

of <0.01. Table 1 is an example of a representative correlation table. Scatter plots showed that technical replicates had high repeatability and all the correlation values were >0.953.

Our article, [2], emphasizes the differentially expressed genes of tissues that received different amounts of radiation (i.e. 0, 2, or 12 Gy), while herein we show the profiling of all the transcriptomic data resulting from these treatments. Enrichment analyses using DAVID Bioinformatics [3] were performed as QC metrics. Genes that were upregulated or downregulated were separately analyzed. We only used genes with |fold change| > 1.5 and P-value < 0.05. We input gene symbols into DAVID Bioinformatics, while throughout we used the default settings. The threshold for significance was a Benjamini-adjusted P-value < 0.05.

Based on our hypothesis, irradiated samples (compared to non-irradiated control samples) would display patterns of gene expression consistent with the physiological effects of irradiation. Table 2 includes selected enriched categories from our QC enrichment analysis (Table S1 contains all enrichment analysis results). We found that our hypothesis agrees with the enriched categories. That is, up-regulated genes in irradiated samples were strongly enriched for functional categories involved in cell death, cell proliferation, oncogenesis, and ubiquitin conjugation. Also, down-regulated genes in irradiated samples were strongly enriched for functional categories involved in RNA processing and cell division. Overall, these results confirmed the quality of the microarray data and facilitated further interpretation of the data presented in [2].

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.gdata.2015.12.013.

Conflict of interest

All the authors in this paper declare no conflict of interest.

Acknowledgments

References

[1] L. Weng, H. Dai, Y. Zhan, Y. He, S.B. Stepaniants, D.E. Bassett, Rosetta error model for gene expression analysis. Bioinformatics 22 (9) (2006) 1111–1121.
[2] M.P. Lambros, C. Parsa, H.C. Mulamalla, R. Orlando, B. Lau, Y. Huang, D. Pon, M. Chow, Identifying cell and molecular stress after radiation in a three-dimensional (3-D) model of oral mucositis. Biochem. Biophys. Res. Commun. 405 (1) (2011) 102–106.
[3] D.W. Huang, B.T. Sherman, R.A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat. Protoc. 4 (1) (2009) 44–57.