The effect of radiotherapy on the transcriptional profile of human cell lines

Stuart Lee[1], Lipi Shukla[1], Ramin Shayan[1], Matthew Ritchie[1,1]

Abstract

placeholder

Specifications

Organism/cell line/tissue	Human, Normal Dermal Fibroblasts, Normal Epidermal Keratinocytes, Pericytes, Microvascular Blood Endothelial Cells, Lymphatic Endothelial Cells and Adipose Derived Stem Cell
Sex	N/A
Data format	Unmapped fastq and summarised counts.
Experimental factors	RNA was obtained from each cell line under normal conditions and after being treated with a single dose of radiotherapy. The adipose stem cells and lymphatic endothelial cells were treated with additional fractionated doses of radiotherapy.
Experimental features	Effect of radiotherapy at dosages of 0Gy versus (2Gy x 5) versus 10Gy.
Consent	N/A
Sample source location	Melbourne, Australia

Direct link to deposited data

Introduction

Experimental Design, Materials and Methods

Sample information

The effect radiotherapy has on gene expression was measured in 6 different human embryonic cell-lines consisting of normal dermal fibroblasts (NDF), normal epidermal keratinocytes (NEK), pericytes (PC), microvascular blood endothelial cells (MEC), lymphatic endothelial cells (LEC) and adipose derived stem cells (ASC). Each cell-line was subjected to a single dose of radiotherapy (10Gy) or no radiotherapy (0Gy, control). For the ASC and LEC cell-lines, there was also a fractionated dose (5 x 2Gy over a 48 hour period). Two replicates of each cell-line/treatment combination were available.

Email addresses: `lee.s@wehi.edu.au` (Stuart Lee), `placeholder` (Lipi Shukla), `placeholder` (Ramin Shayan), `mritchie@wehi.edu.au` (Matthew Ritchie)

RNA-seq sample preparation and sequencing

[1]

Standardized numbers of MEC were plated in cell culture flasks and irradiated once 80-90% confluence was achieved. RNA extraction was undertaken at 4 hours using the QIAGEN® RNEasy Plus Universal Kit as per manufacturers instructions. Samples were then tested for purity and quality control using the Nanodrop Spectrophotometer (Thermo Fischer Scientific) and stored at -80 °C, until further processing. Extraction of RNA at a 4-hour time point post radiotherapy-injury was utilized. Each sample underwent RNA sequencing (100 base pair single end) in the Illumina HiSeq machine at the Australian Genome Research Facility (AGRF) in Melbourne.

Quality-control and data preprocessing

Sequences were mapped to the *hg19* reference genome using the *Rsubread* program with default settings and gene-level counts were obtained by the *featureCounts* procedure using default settings. Transcripts were annotated using the *org.Hs.eg.db* package. Analysis of the resulting counts matrix was performed using the *edgeR* and *limma* R/Bioconductor packages. First, counts per million (cpm) were computed for each gene to remove differences caused by different library sizes. One sample, PERI1_10Gy, had its libraries split across three sequencing lanes, thus its cpm values were merged together. Genes were retained for further analysis if they had a baseline expression level of 0.5 cpm in at least 2 samples. Counts were normalised using the trimmed mean of M-values (TMM) method.

During exploratory data analysis we performed multidimensional scaling (MDS) on the expression matrix with cpm values. The first and second dimension explain 39 per cent and 27 of the variation in the data, respectively (figure 1). There is a clear separation between the different cell-lines and the replicate number from the MDS plots. We model the heteroscedasticty in gene counts using the *voomWithQualityWeights* procedure using the cell-line specific radiotherapy treatment as a main effect (figure 1). This results in down-weighting genes with systematically higher variation. Furthermore, given the clear difference between replicate samples observed in the MDS plot, the *voomWithQualityWeights* procedure also down-weights samples that are more variable compared to others.

Differential expression analysis

After modelling the heteroscedasticty in the counts matrix, we used empirical Bayes moderation to compute more precise estimates of gene-wise variablilty and remove the dependence between the variance and mean expression level (figure 2. Differential expression between all contrasts were assessed using moderated t-statistics using a false-discovery rate (FDR) of 5 per cent. Using this criteria, there were 6720 differentially expressed genes in at least one contrast (figure 2).

Results

Due to the large numbers of contrasts we focus on two cellines and their response to radiotherapy treatment: MEC and LEC. For microvascular blood endothelial cells, there were 2 genes differentially upregulated between the full dosage (10Gy) radiotherapy and control treatments.

While for lymphatic endothelial cells there were 73 differentialy expressed genes between full dosage (10Gy) radiotherapy and contorl treatments. Of these 7 genes were downregulated and 66 were upregulated. There were 4 genes differentially expressed between LEC cells under fractionated dosage of radiotherapy and control treatment. Of these 2 were downregulated and 2 were upregulated.

Between LEC and MEC cells under no treatment there were 3115 differentially expressed genes, 1423 downregulated and 1692 , respectively. Between celllines under full radiotherapy treatment, there were 2379 differentially expressed genes, 973 downregulated and 1406 upregulated, respectively.

[1] note: this is only for MEC need to check with Lipi the same protocol was used for each cell type.

Discussion

List of Figures

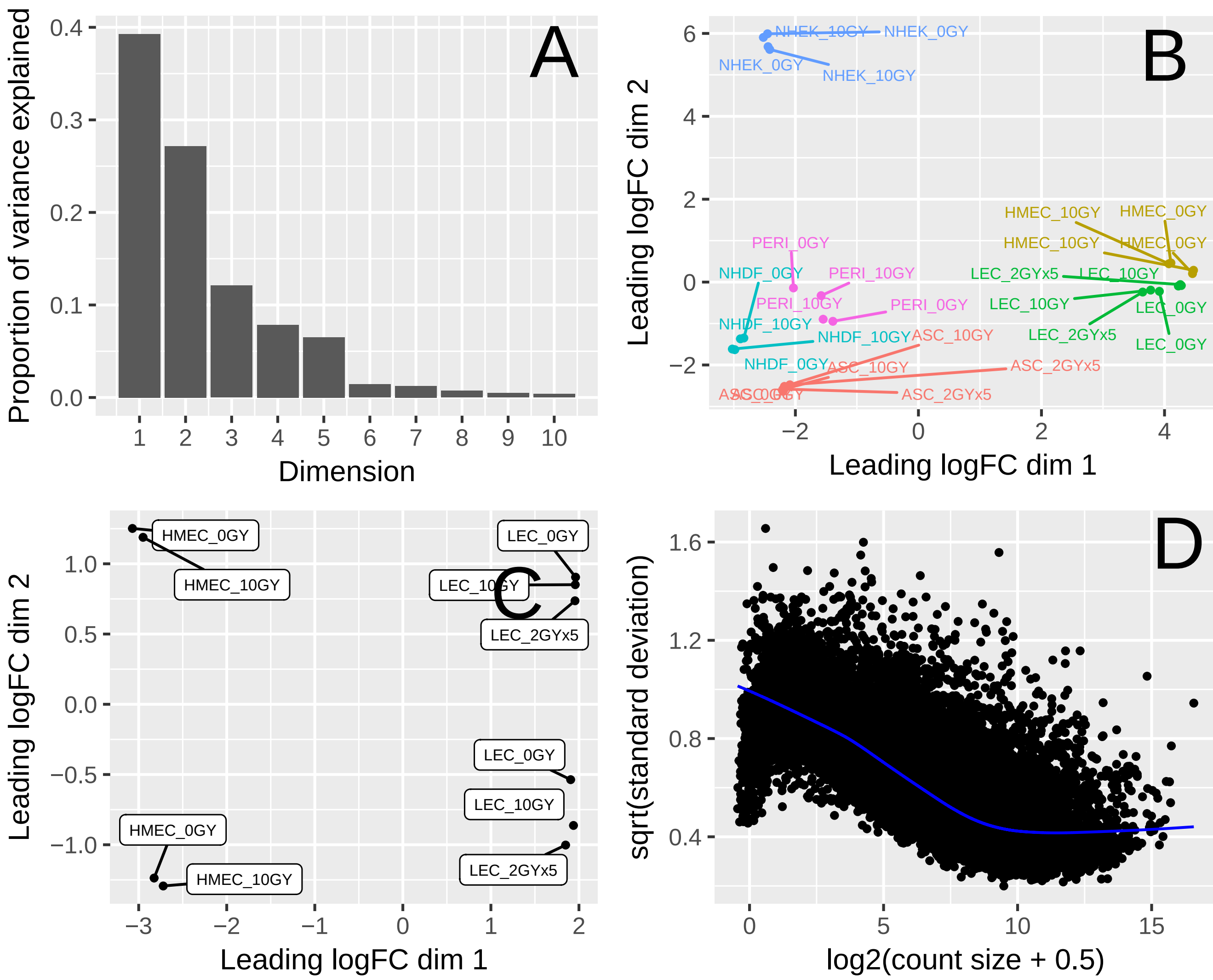

Figure 1: **A** The variance explained by each dimension from the eigenvalues of the multidimensional scaling. **B** MDS plot coloured by cell-line. **C** MDS plot of the LEC and MEC celllines. There is clear differential expression between treatment and celltype, but also confounding caused by replicate number. **D** Mean-variance trend in expression estimated by voom.

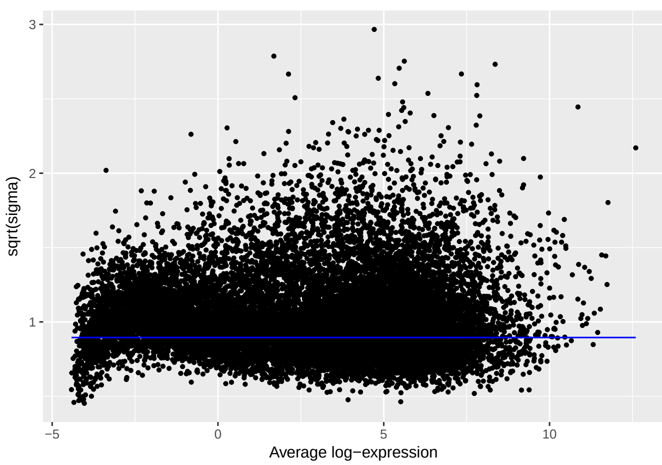

Figure 2: **A** The corrected mean-variance plot. **B** Common DE genes counts between and within contrasts.