Instructions for using code associated with “A Framework for Benchmarking Functional Genomics Signatures of Cancer Cell Radiosensitivity”

Authors: John O’Connor, Ian Overton, and Stephen McMahon

## Summary

The purpose of the paper was to implement gene signatures of radiosensitivity and test them in two datasets, namely the NCI60 and the CCLE.

The 3 scripts (which should be run in order for each dataset) provided here were used for:

* Cleaning/matching radiosensitivity and expression data
* Generating and testing random signatures
* Testing published signatures and producing comparison plots

## Data input

### Radiosensitivity data

**NCI60 –** Radiosensitivity data (SF2) for this dataset comes from the clonogenic assay and was taken from Table 1 and Table 4 of “Eschrich et al. International Journal of Radiation Oncology Biology Physics 2009; 75:497–505”. This is included in the repository.

**CCLE –** Radiosensitivity data (MID) comes from a high throughput assay and is labelled as “AUC” in the “Supplemental Data 1” from “Yard et al. A genetic basis for the variation in the vulnerability of cancer to DNA damage. Nat. Commun. 2016; 7:11428”. This is not included in the repository and will have to be downloaded from the publication and renamed to: “Supplementary\_Data\_1\_Yard\_et\_al.csv” and put in the raw data folder.

### mRNA microarray data

**NCI60 –** This data is from GSE32474 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE32474>). It is downloaded and unzipped within the R code for cleaning the NCI60 data.

**CCLE –** This data is in the “CCLE\_Expression\_2012-09-29.res” file and needs to be downloaded from <https://depmap.org/portal/download/>. It can be found under the “All Downloads” tab on the left of the screen. This needs to be added to the “raw\_data” folder.