

# TCGA DDR Data Resources

## “Genomic and Molecular Landscape of DNA Damage Repair Deficiency Across The Cancer Genome Atlas”

Excel file (**TCGA DDR Data Resources.xlsx**) with the various sheets (detailed below).

TSV files (**TCGA DDR Data Resources.gz**) are also available for the items in blue. The order of the genes, pathways, samples and footprint scores in these TSV files are as given in [Genes.tsv](#), [Pathways.tsv](#), [Samples.tsv](#) and [Scores.tsv](#)

### **DDR genes and pathways** [PathwayMembership.tsv](#)

Annotation of the 276 DDR genes and membership in the inclusive and core DDR pathways.

### **DDR gene alterations** **ONCOPRINT**

ONCOPRINT style data table listing somatic truncating mutations and missense mutations that are predicted to be functionally deleterious (MUT), deep copy number deletions defined by GISTIC (DEL), and epigenetic silencing events (SIL) across all samples (n=9,125) and all genes (n=276).

### **DDR gene mutation** [GeneMutations.tsv](#)

Binary data table listing gene mutations, i.e. somatic truncating mutations and missense mutations that are predicted to be functionally deleterious (MUT), across all samples (n=9,125) and all genes (n=276). A ‘1’ indicates a mutation; a ‘0’ indicates no mutation; an empty cell indicates a missing value.

### **DDR deep deletions** [GeneDeletions.tsv](#)

Binary data table listing deletions, i.e. deep copy number deletions defined by GISTIC (DEL), across all samples (n=9,125) and all genes (n=276). A ‘1’ indicates a deletion; a ‘0’ indicates no deletion; an empty cell indicates a missing value.

### **DDR epigenetic silencing** [GeneSilencing.tsv](#)

Binary data table listing epigenetic silencing events (SIL) across all samples (n=9,125) and all genes (n=276). A ‘1’ indicates a silencing event; a ‘0’ indicates no silencing event; an empty cell indicates a missing value.

### **DDR gene alterations** [GeneAlterations.tsv](#)

Binary data table listing gene alterations, i.e. a MUT, DEL or SIL, across all samples (n=9,125) and all genes (n=276). A ‘1’ indicates at least one of the three alterations is observed (MUT, DEL or SIL); a ‘0’ indicates none of the three alterations types (MUT, DEL or SIL) is observed; an empty cell indicates that none of the three alterations types is observed and at least one of them has a missing value.

### **DDR footprint summary**

Description of the 43 DDR footprint scores.

### **DDR footprints** [DDRscores.tsv](#)

Data table listing the scores of the 43 DDR footprints and the RRPB-based DDR score across all samples (n=9,125).

## **DDR pathway alterations ONCOPR**

ONCOPRINT style data table listing somatic truncating mutations and missense mutations that are predicted to be functionally deleterious (MUT), deep copy number deletions defined by GISTIC (DEL), and epigenetic silencing events (SIL) across all samples (n=9,125) summarized at the DDR pathway level. Specifically, for each of 10 inclusive pathways and 9 core pathways, this table lists 'MUT' if at least one of the genes in the pathway has a somatic truncating mutation or missense mutation that is predicted to be functionally deleterious; 'DEL' if at least one of the genes in the pathway has a deep copy number deletion as defined by GISTIC; 'SIL' if at least one of the genes in the pathway has an epigenetic silencing event.

### **DDR pathway mutations [PathwayMutations.tsv](#)**

Binary data table listing gene mutations, i.e. somatic truncating mutations and missense mutations that are predicted to be functionally deleterious (MUT), across all samples (n=9,125) and all pathways (n=19). A '1' indicates a mutation in at least one gene in the pathway; a '0' indicates no mutations in the genes in the pathway; an empty cell indicates missing values for all the genes in the pathway.

### **DDR pathway deletions [PathwayDeletions.tsv](#)**

Binary data table listing deletions, i.e. deep copy number deletions defined by GISTIC (DEL), across all samples (n=9,125) and all pathways (n=19). A '1' indicates a deletion in at least one gene in the pathway; a '0' indicates no deletions in the genes in the pathway; an empty cell indicates missing values for all the genes in the pathway.

### **DDR pathway silencing [PathwaySilencing.tsv](#)**

Binary data table listing epigenetic silencing events (SIL) across all samples (n=9,125) and all pathways (n=19). A '1' indicates a silencing event in at least one gene in the pathway; a '0' indicates no silencing events in the genes in the pathway; an empty cell indicates missing values for all the genes in the pathway.

### **DDR pathway alterations [PathwayAlterations.tsv](#)**

Binary data table listing gene alterations, i.e. a MUT, DEL or SIL, across all samples (n=9,125) and all pathways (n=19). A '1' indicates at least one of the three alterations is observed (MUT, DEL or SIL) in at least one of the genes in the pathway; a '0' indicates none of the three alterations types (MUT, DEL or SIL) is observed; an empty cell indicates missing values for all the genes in the pathway for all three data types.

## **ME CO analysis core pathways**

Results of the DDR pathway co-occurrence and mutual exclusivity tests for the core pathways.

## **ME CO analysis inclusive pathw**

Results of the DDR pathway co-occurrence and mutual exclusivity tests for the inclusive pathways.

## **DDR Survival Univariate**

Results of survival analysis for the univariate cox proportional hazards models for DDR footprints.

## **DDR Survival Multivariate**

Results of survival analysis for the multivariate cox proportional hazards models for DDR footprints.

## **DDR gene fusions**

Summary of fusion events affecting DDR genes, filtered by read alignment at the break point.

## **TP53 predictor**

Machine learning model coefficients for TP53 status classification.