**M**ethod for **E**valuating **D**eath **U**sing a **S**imulation-assisted **A**pproach **(MEDUSA).**

MEDUSA is an analysis strategy that scores changes in the drug-induced death rate. MEDUSA uses a simple model of population dynamics in the presence and absence of DNA damage to simulate all possible combinations of growth rates and drug-induced death rates. The results of this comprehensive simulation can be used to infer changes in the drug-induced death rate, from a combination of the relative population size and the relative growth rate in the absence of drug (Honeywell et al. 2023, bioRxiv).

**Data collection**

* + Calculation of the DNA damage-induced death rate requires:
    1. **Experimental measurement of the population size in the context of DNA damage over time:** Population size can be measured by counting live cells that have been treated with the desired concentration of DNA-damaging drug, or a vehicle control. Measurements of live cells should be collected with sufficient frequency to capture the biphasic nature of DNA damaging drugs, i.e. an initial phase of slow growth, followed by a second phase of cell death.
    2. **Calculation of the fold-change for untreated/T0 and treated/untreated populations from a chemo-genetic screen:** Analysis of the drug-induced death rate requires the collection of untreated, treated, and T0 populations from a chemo-genetic screen. From these populations, the fold-change of each sgRNA should be determined for 2 different comparisons: log2(untreated/T0), and log2(treated/untreated). Fold-change can be calculated using DESeq2.

**MEDUSA structure**

* + The MEDUSA function requires 7 inputs, ordered as shown below:

MEDUSA(untreated\_gr, drug\_gr, drug\_dr, onset, endpoint, L2FC\_dataset, sgRNA\_num)

**untreated\_gr** = fold-increase in the untreated population after 1 day

**drug\_gr** = fold-change in the treated population at death-onset, including plating day

**drug\_dr** = fold-change in the treated population at assay endpoint, including plating day

**onset** = time of death-onset, in days

**endpoint** = time of assay endpoint, in days

**L2FC\_dataset** = table containing 4 columns: ‘ID’, ’Gene’, ’UTvT0’, and ‘TRvUT’

**sgRNA\_num** = maximum number of sgRNAs belonging to a single gene

**Data preparation**

* + The calculated fold-changes for the untreated/T0 comparison (UTvT0) and the treated/untreated comparison (TRvUT) should be stored in two separate columns of a table. This table should also contain the gene and sgRNA identifiers for the calculation of sgRNA-level and gene-level growth and death rates.
  + Non-targeting sgRNAs should be assigned to artificial non-targeting “genes”. These non-targeting genes should contain an equal number of sgRNAs as a typical gene within the library. Non-targeting sgRNAs should begin with the prefix “Nont” so that they can be identified during the empiric p-value calculation.

A picture containing text, screenshot, menu, font

Description automatically generated

(Example data format for ‘L2FC\_dataset’)

**Running MEDUSA**

* To calculate the drug-induced death rate, call the function MEDUSA:

[simtable guideLevelRates GeneLevelRates] =

MEDUSA(untreated\_gr, drug\_gr, drug\_dr, onset, endpoint, L2FC\_dataset, sgRNA\_num)

* The supplied example data (README-ex1.mat) contains L2FC values for 1000 genes from a chemo-genetic screens. For this screen:

**untreated\_gr** = 2 🡪 untreated cell number increases two-fold after one day

**drug\_gr** = 1.5 🡪 cell number increases 1.5-fold prior to death onset

**drug\_dr** = 1 🡪 cells have returned to their baseline level at assay endpoint

**onset** = 3 🡪 death onset occurs after 2 days (+1 day for plating)

**endpoint** = 5 🡪 total assay length is 4 days (+1 day for plating)

**sgRNA\_num** = 6 🡪 sgRNA library used contains 6 guides per gene

* To run example data:

load README-ex1.mat

[simtable guideLevelRates GeneLevelRates] = MEDUSA(2, 1.5, 1, 3, 5, L2FC\_dataset, 6)

* This should yield three outputs:

**simtable** – a table of the simulated relative growth and death rates along with their associated L2FC

**guideLevelRates** – a table of sgRNA-level growth and death rates

**GeneLevelRates** – a table of gene-level growth rates, death rates, and FDR-corrected p-values

A screenshot of a table

Description automatically generated with low confidence

(Example data output for ‘simtable’)

A picture containing text, screenshot, font, number

Description automatically generated

(Example data output for ‘guideLevelRates’)

A screenshot of a table

Description automatically generated with low confidence

(Example data output for ‘GeneLevelRates’)