nls_logsig

This program fits parameters and compute area under curve (AUC) for drug killing curves with non-linear least square regression. The input "dose" should be log transformed. The input "Kill" represents the fraction of cells killed. We used a three parameters (Ainf, Hill, EC50) log sigmoid model:
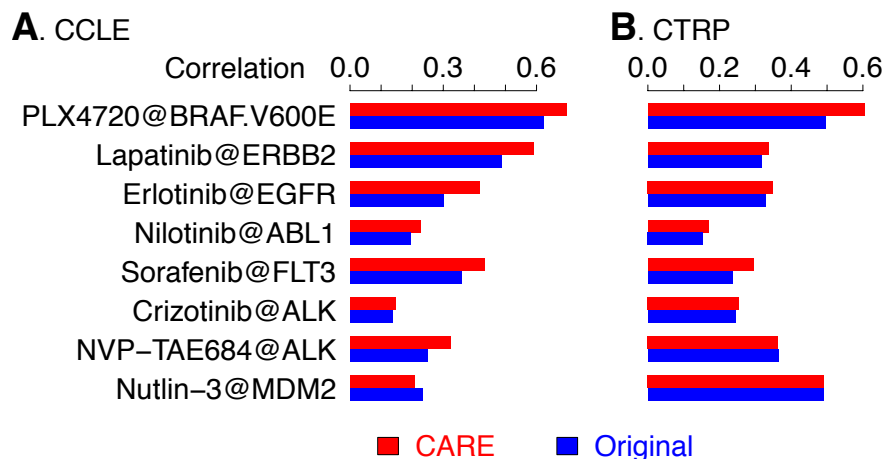
$$Kill = Ainf * (1 - \frac{1}{1 + e^{Hill(dose - EC50)}})$$

Ainf: Maximum killing fraction when dose is infinitely high.

Hill: Shape parameter of log sigmoid curve.

EC50: dose of compound at 50% maximum efficacy.

Many pharmacological screen projects already fitted their drug response curves and computed AUC. However, our program outperformed their results (Figure 1). As an example, for targeted therapy, the correlation between drug inhibition and drug target gene molecular status (e.g., expression or mutation) should be significantly high. Compared on commonly screened compounds in both CCLE[1] and CTRP[2] projects, our AUC values have overall better correlation values than the original results released in each project.

Figure 1: drug target correlation comparison.

The Pearson correlation values between drug inhibition effects and drug target molecular features are shown for both CARE fitted results and the original results.

Usage: nls_logsig -i drug_curve -o output

Example: get into folder "data", then type:

nls_logsig -i drug.curve.sample -o drug.curve.sample.output

drug_curve:

Input drug curves. This file follows a four columns format. The first column is cell line name. The second column is compound name. The third column contains all doses used in screen, separated by ",". The fourth column contains all compound inhibition values between the range of 0 (no inhibition) and 1 (full inhibition), separated by ",". As an example, please see "data/drug.curve.sample".

output:

This output file contains all parameters (Ainf, EC50, Hill) of drug response curve fitted. "MSE" (mean square error) reflects the error of curve fitting. "AUC" (area under curve) reflects the overall drug inhibition effect. "Fit" indicates whether nonlinear least square fitting converge. Certain noisy input may make the nonlinear least square procedure fail, and our program will use grid searching to find best parameters within pre-defined range. As an example of output, please see "data/drug.curve.sample.output".

Parameter range:

All parameters are fitted with in the range of [lower, upper].

Input flag	Parameter	Default
-a	Ainf	0,1
-h	Hill	0,10
-e	EC50	-10,10

Other options:

Flag	Parameter	Default	Comment
-l	log transform dose 1: yes. 0: no	1	
-s	step size in grid search	0.01	If the non-linear least square procedure failed to converge, grid search will start within parameter bound.
-t	Numerical tolerance	1e-08	Used in testing convergence.
-n	max iteration	1000	Maximum number of iterations in non-linear least square fitting.

Reference

1.   Barretina, J.*, et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603-607 (2012).

2.   Seashore-Ludlow, B.*, et al.* Harnessing Connectivity in a Large-Scale Small-Molecule Sensitivity Dataset. *Cancer discovery* **5**, 1210-1223 (2015).