NCI-60 Growth Inhibition Data

DOWNLOAD NCI CELL LINE DATA

BACKGROUND

The NCI cell line screen

- The _overview_ of the NCI cell line screen describes the screening methods and data processing, and describes the _process for investigators to submit compounds to the screen._

The public NSC compounds

- We provide a list of the public NSC numbers.
- Downloadable chemical data are available for public NSC compounds.
- For compounds with inventory, investigators can request samples from the NCI/DTP Open Chemicals Repository.
- Most **NSC numbers** represent single, defined small molecules (as either a free base or as a simple salt). Some NSC numbers have been assigned to more complex biological agents, but they have also been assigned to mixtures, extracts, crude fractions, etc. In the past DTP also assigned NSC numbers to the contents of plated sets from outside suppliers under an agreement whereby DTP never had access to any detailed information regarding the content of the plates. NSC numbers are not intended to identify unique chemical structures in the NSC series, though most of them do.

The NCI cell lines

- Information about the NCI cell lines is in the DTP/DCTD Tumor Repository catalog.
- The process for requesting cell lines is described in the catalog.
- Please note these links for more information on the SNB-19, U251, NCI/ADR-RES, and MDA-MB-435 cell lines.
  - MDA-MB-435
  - U251
  - SNB-19
  - NCI/ADR-RES

FILES FOR DOWNLOAD April 2024 Release

Previous Releases.

	DATA FILE SIZE	DOWNLOAD SIZE	LINK
CONCENTRATION/RESPONSE DATA	2.28 Gb	319 Mb	DOSERESP.zip
GI50 DATA	386 Mb	36 Mb	GI50.zip
TGI DATA	381 Mb	32 Mb	TGI.zip
LC50 DATA	377 Mb	28 Mb	LC50.zip
IC50 DATA	386 Mb	34 Mb	IC50.zip
ONECONC (PRESCREEN) DATA	404 Mb	44 Mb	ONECONC.zip

GENERAL COMMENTS REGARDING THE DOWNLOADABLE FILES

- Previous data releases reported aggregate values across experiments, grouping the data by NSC number and the log of the highest concentration, rounded to one decimal point, in the concentration/response dilution series. Now, we report data for individual experiments identified by an **EXPID**, and all values are reported to 4 decimal places.
- All cell lines for an individual **EXPID** are grown and assayed contemporaneously.
- The format of the **EXPID** is YYMMLLSS, where YY is the last 2 digits of the year (00 - 21 for 2000 to 2021, MM is the month number (01 for January - 12 for December), LL is a pair of letters for internal process tracking and SS is a 2-digit numeric sequence.
- There are 60 cell lines in the current NCI60 cell line screen. There are 11 other cell lines that were part of the NCI60 screen in the past. These 71 cell lines comprise most of the public data. This data release also includes other cell lines which have been assayed at least once using the same protocols as the NCI60 cell line screen.
- Experimental QC checks were performed at the lab-level and during data-processing at the time that the experiments were run. Additional quality control or consistency checks have not been performed.
- Endpoint values (GI50, TGI, LC50) have accompanying concentration/response data; however, not all concentration/response data have accompanying endpoint values.
- Most NSC numbers represent small molecules, and the reported concentrations use a "M" (molar) **CONCENTRATION_UNIT**. For more complex biological agents concentrations may be reported as µg/ml (micrograms per milliliter) with a **CONCENTRATION_UNIT** "u". Mixtures, extracts, crude fractions, etc. in the assay may use units of µg/ml or volume-based measurements designated by **CONCENTRATION_UNIT** "V". (There is no further definition regarding what a volume-based concentration means.)

- At the level of individual experiments almost all endpoint values will have a count of 1 and a standard deviation of 0. In a few cases, though, multiple replicate determinations were run within a single experiment.
- Most of the concentration/response data are from a series of 5 dilutions at log intervals (10-fold dilution); however, a few experiments were run with 10 dilutions at half-log intervals for some cell lines and NSC compounds.
- **PTC** is the Percent of Treated cell growth as a fraction of Control cell growth. The IC50 endpoint values are interpolated from these data.
- **GIPRCNT** is the percent of treated cell growth as a fraction of control cell growth corrected for the count of cells at the time of drug addition in the assay. 100 is control growth, 0 is complete inhibition of growth (cytostasis), and -100 is complete cell kill. The GI50, TGI and LC50 values are determined by simple interpolation of GIPRCNT values above and below 50, 0, and -50 respectively.
- Where GI50, TGI or LC50 values would be outside the concentration range of the dilution series the highest or lowest concentration in the series is reported.
- For historical reasons, PTC values were not stored in our database. They have been recalculated for this data release. There are a few cases where we report no PTC value but do report a GIPRCNT value. The data processing is able to work with certain occurrences of null values within the series of concentration/response data.
- The ONECONC prescreen has changed over the years. Currently all 60 of the NCI60 cell lines are evaluated, but in the past the assay was run against only a small number of cell lines.

FILE COLUMN HEADERS

CONCENTRATION/RESPONSE DATA

1. **RELEASE_DATE** The date of this data release.
2. **EXPID** Please see the General Comments above.
3. **PREFIX** The identifier of the sequence from which an NSC number was assigned. All public data are in the S series.
4. **NSC** The numeric identifier in the S series.
5. **CONCENTRATION_UNIT** Please see the General Comments above.
6. **LOG_HI_CONCENTRATION** The log10 of the highest concentration of the concentration/response data.
7. **CONCENTRATION** The log10 of the concentration in the dilution series.
8. **PANEL_NUMBER** Internal identifier. The combinations of panel_number and cell_number are unique cell line identifiers.
9. **CELL_NUMBER** Internal identifier. The combinations of panel_number and cell_number are unique cell line identifiers.
10. **PANEL_NAME** The name of the NCI cell line panel (cancer type).
11. **CELL_NAME** The name of the NCI cell line.
12. **PANEL_CODE** An abbreviation for the panel_name.
13. **COUNT_GIPRCNT** Count of GIPRCNT values.
14. **AVERAGE_GIPRCNT** Average of GIPRCNT values.
15. **STDDEV_GIPRCNT** Standard deviation of GIPRCNT values.
16. **COUNT_PTC** Count of PTC values.
17. **AVERAGE_PTC** Average of PTC values.
18. **STDDEV_PTC** Standard deviation of PTC values.

GI50, TGI, LC50, IC50 VALUES

1. **RELEASE_DATE** The date of this data release.
2. **EXPID** Please see the General Comments above.
3. **PREFIX** The identifier of the sequence from which an NSC number was assigned. All public data are in the S series.
4. **NSC** The numeric identifier in the S series.
5. **CONCENTRATION_UNIT** Please see the General Comments above.
6. **LOG_HI_CONCENTRATION** The log10 of the highest concentration of the concentration/response data.
7. **PANEL_NUMBER** Internal identifier. The combinations of panel_number and cell_number are unique cell line identifiers.
8. **CELL_NUMBER** Internal identifier. The combinations of panel_number and cell_number are unique cell line identifiers.
9. **PANEL_NAME** The name of the NCI cell line panel (cancer type).
10. **CELL_NAME** The name of the NCI cell line.
11. **PANEL_CODE** An abbreviation for the panel_name.
12. **COUNT** Count of interpolated values.
13. **AVERAGE** Average of interpolated values.
14. **STDDEV** Standard deviation of interpolated values.

ONECONC (PRESCREEN) DATA

1. **RELEASE_DATE** The date of this data release.
2. **EXPID** Please see the General Comments above.
3. **PREFIX** The identifier of the sequence from which an NSC number was assigned. All public data are in the S series.
4. **NSC** The numeric identifier in the S series.
5. **CONCENTRATION_UNIT** Please see the General Comments above.
6. **CONCENTRATION** The concentration in the pre-screen assay.
7. **PANEL_NUMBER** Internal identifier. The combinations of panel_number and cell_number are unique cell line identifiers.
8. **CELL_NUMBER** Internal identifier. The combinations of panel_number and cell_number are unique cell line identifiers.
9. **PANEL_NAME** The name of the NCI cell line panel (cancer type).
10. **CELL_NAME** The name of the NCI cell line.
11. **PANEL_CODE** An abbreviation for the panel_name.
12. **COUNT_GIPRCNT** Count of GIPRCNT values.
13. **AVERAGE_GIPRCNT** Average of GIPRCNT values.
14. **STDDEV_GIPRCNT** Standard deviation of GIPRCNT values.

Cancelled Experiments

The following files contain data that were cancelled by the laboratory due to quality control issues.

	CSV FILE LINK
CONCENTRATION/RESPONSE DATA	DOSERESP_Cancelled.csv
GI50 DATA	GI50_Cancelled.csv
TGI DATA	TGI_Cancelled.csv
LC50 DATA	LC50_Cancelled.csv
IC50 DATA	IC50_Cancelled.csv
ONECONC (PRESCREEN) DATA	ONECONC_Cancelled.csv

The cancelled data were removed from the December 2022 Release files listed above.