

# Molecular Target Data

Data from the [DTP Molecular Target program](#).

Email questions concerning DTP's molecular targets program to: [Molecular Target Team \[moltarget@mail.nih.gov\]](mailto:Molecular Target Team [moltarget@mail.nih.gov])

## Primary molecular target data (excluding microarray data)

includes protein, mRNA, miRNA, DNA methylation, mutations, SNPs, enzyme activity, metabolites

<a href="#">WEB_DATA_ALL_MT.ZIP</a>	A 3.4 Mb zip file - The uncompressed file is approximately 82.7 Mb.
-------------------------------------	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), GENE, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, ENTITY\_MEASURED, GeneID, UNITS, METHOD, VALUE, TEXT

## Protein Data Only

A subset of WEB\_DATA\_ALL\_MT containing just the protein data

<a href="#">WEB_DATA_PROTEIN.ZIP</a>	A 219 Kb zip file - The uncompressed file is approximately 3.2 Mb.
--------------------------------------	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), GENE, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, ENTITY\_MEASURED, GeneID, UNITS, METHOD, VALUE, TEXT

## DNA Data Only

A subset of WEB\_DATA\_ALL\_MT containing just the DNA data

<a href="#">WEB_DATA_DNA.ZIP</a>	A 26 kb zip file - The uncompressed file is approximately 580 Kb.
----------------------------------	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), GENE, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, ENTITY\_MEASURED, GeneID, UNITS, METHOD, VALUE, TEXT

## DNA Methylation Data From Sequenom

A subset of WEB\_DATA\_ALL\_MT containing just the DNA methylation data, Proc Natl Acad Sci USA 2008, Mar 25; 105(12): 4844-9.

<a href="#">WEB_DATA_SEQUENOM_METHYLATION.ZIP</a>	A 5.2 Mb zip file - The uncompressed file is approximately 43.8 Mb.
---	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), GENE, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, ENTITY\_MEASURED, GeneID, VALUE

## microRNA Data From the Israel Lab

A subset of WEB\_DATA\_ALL\_MT containing just the microRNA data from the Israel lab. Cancer Res. 2007, Mar 15; 67(6): 2456-68.

<a href="#">WEB_DATA_ISRAEL_MIR.ZIP</a>	A 98 Kb zip file - The uncompressed file is approximately 1.1 Mb.
---	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, ENTITY\_MEASURED, UNITS, METHOD, VALUE, TEXT

## microRNA Data From the Weinstein and Croce Labs

A subset of WEB\_DATA\_ALL\_MT containing just the microRNA data from the Weinstein and Croce labs. Mol Cancer Ther 2007, May; 6(5): 1483-91.

<a href="#">WEB_DATA_CROCE-WEINSTEIN_MIR.ZIP</a>	A 601 Kb zip file - The uncompressed file is approximately 6.3 Mb.
--	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, ENTITY\_MEASURED, UNITS, METHOD, VALUE, TEXT

## Metabolomic Data From Metabolon- data averaged from triplicate experiments

A subset of WEB\_DATA\_ALL\_MT containing just the metabolomic data from Metabolon.

<a href="#">WEB_DATA_METABOLON.ZIP</a>	A 171 Kb zip file - The uncompressed file is approximately 1.2 Mb.
--	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, VALUE, STD DEV

## Metabolomic Data From Metabolon - individual data from each of the triplicate experiments

<a href="#">WEB_DATA_METABOLON_ALL.ZIP</a>	A 404 Kb zip file - The uncompressed file is approximately 3.2 Mb.
--	---

When uncompressed the file is comma delimited in the following format:

**File Format:** SAMPLENAME, TITLE, PANELNBR, CELLNBR, pname, cellname, VALUE

## THE FOLLOWING DATASETS ARE DERIVED FROM LARGE-SCALE EXPERIMENTS

**Estimated chromosomal band copy number**, extracted from spectral karyotyping Cancer Res 63, 8634-47 (2003).

Data is provided as an Excel file listing copy number of each chromosomal band for each cell line. [Download excel file](#)

**Affymetrix 125K SNP array data** from the Sellers' lab Nature 436, 117-122 (2005).

Data is provided as 3 datasets: Copy number, allele calls and identifiers.

### The Copy Number Data:

<a href="#">COPYNUM.ZIP</a>	A 51.6 Mb zip file - The uncompressed file is approximately 307 Mb.
-----------------------------	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MARKER, CellID, COPYNBR, PANELNBR, CELLNBR, pname, cellname

### The Allele Call Data:

<a href="#">ALLELECALL.ZIP</a>	A 29.6 Mb zip file - The uncompressed file is approximately 290 Mb.
--------------------------------	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MARKER, CellID, AlleleCall, PANELNBR, CELLNBR, pname, cellname

### The Identifiers Data:

<a href="#">IDENTIFIERS.ZIP</a>	A 3.6 Mb zip file - The uncompressed file is approximately 11 Mb
---------------------------------	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MARKER, RefSNPID, Chromosome, Chromosome Location, Allele\_A, Allele\_B, FlankingSeqA, FlankingSeqB

### Array CGH DNA copy number data

(relative to normal female DNA) from the Weinstein (NCI) and Gray (UCSF) labs Mol Cancer Ther. 2006 Apr;5(4):853-67.)

<a href="#">CGH_COPYNUM.ZIP</a>	A 105 Kb zip file - The uncompressed file is approximately 1.0 Mb
---------------------------------	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), GENE, CHROMOSOMAL LOCATION, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, cellname, pname, VALUE

### Microarray Data- Affymetrix U133 array data from Chiron

Data was processed with the Affymetrix MAS5 algorithm, with a scaling factor of 100.

<a href="#">WEB_DATA_CHIRON.ZIP</a>	A 21.9 Mb zip file - The uncompressed file is approximately 305 Mb
-------------------------------------	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), ACC, GENE, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, CHIP, FEATURE\_ID, UniGene, GeneID, PRESENT\_CALL, VALUE Gene assignments are based on Unigene Build #U225 (August 2010)

### Microarray data- Affymetrix U133 array data from Gene Logic, Inc.

Data was processed with the Affymetrix MAS5 algorithm, with a scaling factor of 100.

<a href="#">WEB_DATA_GENELOGIC_U133.ZIP</a>	A 26.0 Mb zip file - The uncompressed file is approximately 314 Mb
---	--

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), ACC, Gene, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, CHIP, FEATURE\_ID, UniGene, GeneID, VALUE Gene assignments are based on Unigene Build #U225 (August 2010)

### Microarray data - Affymetrix U95 array data from Gene Logic, Inc.

<a href="#">WEB_DATA_GENELOGIC_U95.ZIP</a>	A 34.1 Mb zip file - The uncompressed file is approximately 401 Mb.
--	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), ACC, Gene, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, CHIP, FEATURE\_ID, UniGene, GeneID, VALUE Gene assignments are based on Unigene Build #U225 (August 2010)

### Microarray data - Affymetrix U95A data from Novartis, averaged data (from triplicate arrays)

<a href="#">WEB_DATA_NOVARTIS.ZIP</a>	A 9.4 Mb zip file - The uncompressed file is approximately 84 Mb.
---------------------------------------	---

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), ACC, Gene, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, CHIP, FEATURE\_ID, UniGene, GeneID, VALUE Gene assignments are based on Unigene Build #U214 (June 2008)

### Microarray data - Affymetrix U95A data from Novartis, data from individual arrays

<a href="#">WEB_DATA_NOVARTIS_ALL.ZIP</a> -	A 26.8 Mb zip file - The uncompressed file is approximately 145 Mb.
---	---

When uncompressed the file is comma delimited in the following format:

**File Format:** Probe Set Name, ID (composite of the moltid derived from this measurement, and a letter to distinguish individual arrays), GENE, cellname, pname, PANELNBR, CELLNBR, Signal, Detection, P value Gene assignments are based on Unigene Build #U225 (August 2010)

### Microarray data - cDNA array data from the Weinstein (NCI) and Brown & Botstein (Stanford) groups

Nat Genet. 2000 Mar;24(3):236-44.

Nat Genet. 2000 Mar;24(3):227-35.

Units are log<sub>2</sub> of signal from the test cell/reference pool

[WEB\\_DATA\\_STANFORD.ZIP](#)

A 5.5 Mb zip file - The uncompressed file is approximately 60.1 Mb.

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), ACC, Gene, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, CHIP, FEATURE\_ID, UniGene, GeneID, VALUE Gene assignments are based on Unigene Build #U225 (August 2010)

#### **Microarray data - Affymetrix HUM6000 array data from Millenium Pharmaceuticals**

[Details of this experiment.](#)

[WEB\\_DATA\\_MILLENIUM.ZIP](#)

A 4.2 Mb zip file - The uncompressed file is approximately 47.2 Mb.

When uncompressed the file is comma delimited in the following format:

**File Format:** MOLTID (NCI pattern #), ACC, Gene, TITLE, MOLTNBR (NCI exp. id #), PANELNBR, CELLNBR, pname, cellname, CHIP, FEATURE\_ID, UniGene, GeneID, VALUE Gene assignments are based on Unigene Build #U225 (August 2010)