

## ***NCI Division of Cancer Biology***

### **Information Related to Element 1A in NIH Data Management and Sharing (DMS) Plans**

#### **IN THIS DOCUMENT**

- [Example of a tabular format to describe all data generated under the proposed project for Element 1A in DMS Plans](#)
- [NIH Genomic Data Sharing \(GDS\) policy thresholds for large-scale data](#)

#### **Example of a tabular format to describe all data generated under the proposed project for Element 1A in DMS Plans**

<b>Type of Data</b>	<b>Species</b>	<b>Sample Type</b>	<b>Platform/Source</b>	<b># of Samples</b>
scRNA sequencing data	Human	Primary tumor	Illumina	10 tumors with matched normal = 20 samples total
Whole RNA sequencing data	Human	Primary tumor	Illumina	15 breast primary tumors with 1 matched metastasis sample= 30 samples
Targeted RNA sequencing data from CRISPR/ORF screen	Human	Cell line	Illumina	20 samples from 1 patient-derived cell line
Bio-TAP-XL proteomics	Mouse	Cell line	Thermo Velos Elite Orbitrap Mass Spectrometer	10 samples from one cell line
Cell biological assays: qRT-PCR, Western Blots, cell viability assays, ELISA	Mouse	Cell line		5 mouse cell lines
HiPlex RNAscope	Mouse	Primary tumor	ACD Bio; ~12 RNA probes	30 mouse tumors
Immunofluorescence, IHC	Mouse	Primary tumor, cell lines	20-25 antibodies	20 mouse tumors, 5 mouse cell lines
Flow Cytometry	Mouse	Primary tumor	BD FACSCanto II	20 mouse tumor and 20 adjacent non-tumor tissue

## NIH Genomic Data Sharing (GDS) policy thresholds for large-scale data

Type of Data	Data Examples	# of Human Specimens (including Human Cell Lines)	# of Model Organisms, Non-Human Cell Lines, Infectious Organisms, etc.
SNP array data from >500k single nucleotide polymorphisms (SNPS)	GWAS Data	1,000	500
DNA sequence data from < 100 genes or regions of interest	Targeted Sequencing	1,000	500
DNA sequence data from $\geq$ 100 genes or regions of interest	Targeted Sequencing, Whole Exome Sequencing, Whole Genome Sequencing	100	50
Genome-wide RNA sequencing data ***	Transcriptomic Data	100	50
Genome-wide DNA methylation data	Bisulfite Sequencing Data	100	50
Genome-wide chromatin immunoprecipitation sequencing (ChIP-seq) data	Transcription Factor ChIP-seq, Histone Modification ChIP-seq	100	50
Metagenome (or microbiome) sequencing data	16S rRNA Sequencing, Shotgun Metagenomics, Whole-Genome Microbial Sequencing	100	50
Metatranscriptome sequencing data	Microbial/Microbiome Transcriptomics	100	50
SNP array data from >500k single nucleotide polymorphisms (SNPS)	GWAS Data	1,000	500

\*\*\*

*If any single cell sequencing data is being generated, the GDS policy applies if over 50 individual cells will be profiled.*