GEOarchive metadata guidelines table

Field	Content Guidelines
SERIES	
title	Provide a unique title that describes the overall study.
summary	Provide a thorough description of the goals and objectives of this study. The abstract from the associated publication may be suitable. Multiple summary lines can be included.
overall design	Provide a description of the experimental design. Indicate how many Samples are analyzed, if replicates are included, are there control and/or reference Samples, dye-swaps, etc... Multiple lines can be included.
type	Enter keyword(s) that generally describe the type of study. Examples include: time course, dose response, comparative genomic hybridization, ChIP-chip, cell type comparison, disease state analysis, stress response, genetic modification, etc.
contributor	[Optional] List all people associated with this study. Format: 'firstname, middleinitial, lastname' with each contributor on a separate line.
web link	[Optional] Specify a Web link that directs users to supplementary information about the study. Please restrict to Web sites that you know are stable.
pubmed id	[Optional] Specify a valid PubMed identifier (PMID) that references a published article describing this study. Most commonly, this information is not available at the time of submission - it can be added later once the data are published.
variable	[Optional] The format should be "variable type: variable description: list of Sample names" where the variable type can be one of the following: dose, time, tissue, strain, gender, cell line, development stage, age, agent, cell type, infection, isolate, metabolism, shock, stress, temperature, specimen, disease state, protocol, growth protocol, genotype/variation, species, individual, or other. For example: age: 2 months: Sample name 1, Sample name 3 age: 12 months: Sample name 2, Sample name 4 NOTE - this information is optional and does not appear in Series records or downloads, but will be used to assemble corresponding GEO DataSet records.
repeat	[Optional] The format should be "repeat type: list of Sample names" where the repeat type can be one of these three: biological replicate, technical replicate - extract, or technical replicate - labeled-extract. For example: biological replicate: Sample name 1, Sample name 3 biological replicate: Sample name 2, Sample name 4 NOTE - this information is optional and does not appear in Series records or downloads, but will be used to assemble corresponding GEO DataSet

	records.
SAMPLES	
Sample name	A unique name that matches a corresponding header in the matrix file.
title	Provide a unique title that describes this Sample. We suggest that you use the convention [biomaterial]-[condition(s)]-[replicate number], e.g., Muscle_exercised_60min_rep2.
raw data file	Raw data file name, e.g. GPR file. For non-Affymetrix submissions. More than one raw data file column can be included.
CEL file	CEL file name. Affymetrix submissions only.
EXP file	EXP file name. Affymetrix submissions only, if available.
CHP file	CHP file name. Affymetrix submissions only. Use this only if you are planning to submit CHP files instead of a matrix table. If your manuscript discusses data processed by RMA or another algorithm, we recommend providing a matrix of those values rather than CHP files.
source name	Briefly identify the biological material and the experimental variable(s), e.g., vastus lateralis muscle, exercised, 60 min.
organism	Identify the organism(s) from which the biological material was derived.
characteristics	List all available characteristics of the biological source, including factors not necessarily under investigation, e.g., Strain: C57BL/6, Gender: female, Age: 45 days, Tissue: bladder tumor, Tumor stage: Ta. Multiple characteristics columns can be included.
biomaterial provider	[Optional] Specify the name of the company, laboratory or person that provided the biological material.
molecule	Specify the type of molecule that was extracted from the biological material. Include one of the following: total RNA, polyA RNA, cytoplasmic RNA, nuclear RNA, genomic DNA, protein, or other.
label	Specify the compound used to label the extract e.g., biotin, Cy3, Cy5, 33P.
description	Include any additional information not provided in the other fields, or paste in broad descriptions that cannot be easily dissected into the other fields.
platform	Reference the Platform accession number (GPLxxx) if the Platform already exists in GEO. To identify the accession number of an existing Platform in GEO, use the find platform tool (link on the website). Omit this column if a new Platform is included in your GEOarchive submission.
PROTOCOLS	
growth protocol	[Optional] Describe the conditions that were used to grow or maintain organisms or cells prior to extract preparation.
treatment protocol	[Optional] Describe any treatments applied to the biological material prior to extract preparation.

extract protocol	Describe the protocol used to isolate the extract material.
label protocol	Describe the protocol used to label the extract.
hyb protocol	Describe the protocols used for hybridization, blocking and washing, and any post-processing steps such as staining.
scan protocol	Describe the scanning and image acquisition protocols, hardware, and software.
data processing	Provide details of how data in the matrix table were generated and calculated, i.e., normalization method, data selection procedures and parameters, transformation algorithm (e.g., MAS5.0, GCOS, RMA for Affymetrix data), and scaling parameters.
value definition	Provide a short description for the values in the matrix table, for example: - lowess normalized log2 ratio (test/reference) - signal calculated by GCOS1.2 software
PLATFORM	
title	Provide a unique title that describes your Platform. We suggest that you use the convention [institution/lab]-[species]-[number of features]-[version], e.g. FHCRC Mouse 15K v1.0.
distribution	Microarrays are 'commercial', 'non-commercial', or 'custom-commercial' in accordance with how the array was manufactured.
technology	Select the category that best describes the Platform technology: spotted DNA/cDNA, spotted oligonucleotide, in situ oligonucleotide, antibody, tissue, SARST, RT-PCR, MS, or MPSS
organism	Identify the organism(s) from which the features on the Platform were designed or derived.
manufacturer	Provide the name of the company, facility or laboratory where the array was manufactured or produced.
manufacture protocol	Describe the array manufacture protocol. Include as much detail as possible, e.g., clone/primer set identification and preparation, strandedness/length, arrayer hardware/software, spotting protocols.
description	[Optional] Provide any additional descriptive information not captured in another field, e.g., array and/or feature physical dimensions, element grid system.
catalog number	[Optional] Provide the manufacturer catalog number for commercially-available arrays.
web link	[Optional] Specify a Web link that directs users to supplementary information about the array. Please restrict to Web sites that you know are stable.
support	[Optional] Provide the surface type of the array, e.g., glass, nitrocellulose, nylon, silicon, unknown.
coating	[Optional] Provide the coating of the array, e.g., aminosilane, quartz, polysine, unknown.

contributor	[Optional] List all people associated with this array design. Each name in the form 'firstname, middleinitial, lastname'.
pubmed id	[Optional] Specify a valid PubMed identifier (PMID) that references a published article that describes the array.