Metadata Specifications:
Primary Cells

Importance 1: Required, 2: Required if available, 3: Optional
Common Fields Fields that are common across all LINCS metadata standards
Custom Fields Fields that are unique to a single LINCS metadata standard or common across only a subset of them

Common Fields

LINCS Field Name	Related to	Description	Comments	Importance
PC_LINCS_ID	Canonical	Unique LINCS internal identifier	LINCS internal ID; This is a batch independent ID; Canonical primary cell ID	1
PC_Name	Canonical	The name for the primary cells as chosen by LINCS	Should be descriptive and correspond to existing cell names as much as possible; Batch independent name	1
PC_Alternative_Name	Canonical	Other relevant names	Synonymous or alternative names, but only significant difference	2
PC_Alternative_ID	Canonical	Other relevant IDs for cells	CLO or other synonymous IDs	2
PC_Center_Canonical_ID	Canonical	LINCS DSGC-specific canonical ID. This will be assigned by a given LINCS DSGC according to its primary cell registration scheme.	-	1
PC_Relevant_Citations	Canonical	List of references (with PMIDs) of relevance to cell isolation, etc.	-	2
PC_Center_Name	Batch	LINCS center using the primary cells	-	1
PC_Center_Batch_ID	Batch	LINCS center-specific cell ID; Batch specific ID	-	1
PC_Provider_Name	Batch	Name of vendor or lab (provider) that supplied the primary cell	Vendor(s) or provider	1
PC_Provider_Catalog_ID	Batch	ID or catalogue number or name assigned to the primary cell by the vendor or provider	Primary cell provider's IDs	1
PC_Provider_Batch_ID	Batch	Vendor/Provider Batch ID number; Batch or lot number assigned to the primary cell by the vendor or provider	-	1

Custom Fields

PC_Organism	Canonical	Organism of origin; A controlled vocabulary describing the organism from which the primary cell was derived (e.g. Homo sapiens, Mus musculus, etc.)	-	1
PC_Organ	Canonical	Organ of origin; A controlled term describing the organ from which cell line is derived; (e.g. lung, mammary gland etc.)	-	1
PC_Tissue	Canonical	Organ or tissue of origin; A controlled vocabulary describing the organ or tissue from which the primary cell was derived (e.g. lung, mammary gland etc.)	Some histology information might be provided in this field.	1
PC_Cell_Type	Canonical	A controlled vocabulary describing the cell type from which a primary cell was derived; e.g. epithelial like, fibroblast-like, lymphoblast like, hematopoetic, mesenchymal, neural, etc. This provides information about cell morphology. Also sometimes referred to as cell morphology.	controlled terminology from CL	1
PC_Cell_Type_Detail	Canonical	Additional description of cell type (histology) that is not available in CL, but may be known from other sources like ATCC	Terms from other sources like ATCC; Will develop over time	2
PC_Donor_Sex	Canonical	Describes sex of the organism from which the cell was obtained;	Controlled terms to describe genders and also chromosomal abnormalities	2
PC_Gonosome_Code	Canonical	List of the sex chomosomes (gonosome) of the sample e.g. XX, XY, XXY	-	3
PC_Donor_Age	Canonical	The age of the donor	Numeric number; Donor age in years	2
PC_Donor_Ethnicity	Canonical	For human cells, the ethnicity of the donor	-	2
PC_Donor_Health_Status	Canonical	Controlled vocabulary describing the health status of the donor	Needs to be defined in more detail; Need level of detail required	2
PC_Disease	Canonical	If the primary cell came from a particular diseased tissue, the disease should be noted in terms of a controlled vocabulary (e.g. breast cancer, colon cancer, not diseased, etc.)	The disease hierarchy is captured in the ontology; i.e. DOID	1
PC_Disease_Detail	Canonical	Additional description of a disease related to the primary cell that may not be available in the disease ontology above	Need to develop what exactly should go here and the corresponding terms (e.g. tumor stage, cell from metastatic site, preceding treatments, etc)	2
PC_Disease_Site_Onset	Canonical	Site of disease onset in primary cell donor	Primary cell / cell line of origin Information	2
PC_Disease_Age_Onset	Canonical	Age of disease onset in primary cell donor (in years)	Primary cell / cell line of origin Information	2
PC_Donor_Age_Death	Canonical	Age of death of primary cell donor (in years)	Primary cell / cell line of origin Information	2
PC_Donor_Disease_Duration	Canonical	Disease duration in primary cell donor; Age of Sample Acquisition - Age of Onset. (in years)	Primary cell / cell line of origin Information	2
PC_Known_Mutations	Canonical	Known mutation in primary cell captured explicitly; e.g. if reference is not available	Needs some ontology to describe gene / protein and mutation	2
PC_Mutation_Citations	Canonical	Known mutation in primary cell from a reference; Needs to include the reference source and the reference to the specific cell	reference to cell line inherent mutations	2
PC_Molecular_Features	Canonical	Relevant molecular and morphological features of the Primary Cell. (e.g. ER Status, Luminal Cells)	-	3
PC_Genetic_Modification	Canonical	Stable transfection, viral transduction or any other genetic modifications (de novo mutations, translocations) that were acquired. If yes, the modifications (e.g. expressing GFP-tagged protein) should be described and appropriate references provided.	MIACA is minimal information that may be a guidance	1
PC_Cell_Markers	Canonical	A controlled vocabulary describing the markers used to isolate / identify the cell type	Controlled terms of markers; At this point no reference	2
PC_Growth_Properties	Canonical	A controlled vocabulary describing the growth properties of the primary cell (e.g. adherent, suspension)	-	1
PC_Recommended_Culture_Conditions	Canonical	A description of the standard tissue culture conditions (media, supplements, culture dish treatment) used to maintain the primary cell. Description of culture dish treatment conditions would include information about coating of culture dish with fibronectin, collagen, etc. prior to cell plating. If special culture vessels are required to grow the cells, these should also be mentioned and details provided.	Recommended standard culturing conditions go here; Not a required field; The actual culture conditions are captured as experimental conditions; see EXP_PC:2	2
PC_Related_Projects	Canonical	Other projects in which the primary cells have been studied / used; A controlled vocabulary describing other large scale projects in which the cells have been used (e.g. ENCODE, TCGA, ICBP, Epigenomics, etc.)	Need some defined project code	3
PC_Verification_Reference_Profile	Canonical	expected STR (reference) profile of the cell based on provider information, if available	From cell provider / reference	2
PC_Production_Details	Canonical	This field specifies the procedure(s) by which the cells were derived from the parent/precursor cell, including genetic transformations and phenotypic selections. Citations / source information for constructs and citations for procedures should be included here when appropriate. If this cell line is derived from another registered cell line, this field also should specify the Center LINCS batch ID of the specific batch from which it was derived.	-	2
PC_Precursor_Cell_Name	Canonical	This field specifies the name of the parent cell from which the primary cell was derived. It is left blank if this cell line was not known to be derived from another. The particular batch of that parent line that was used to generate the new line should be specified in the Production_Details (canonical) field.	-	1
PC_Precursor_Cell_LINCS_ID	Canonical	This field specifies the global LINCS ID of the parent cell from which the primary cell was derived. It is left blank if this cell line was not known to be derived from another. The particular batch of the parent line that was used to generate the new line should be specified in the Production Details (canonical) field.	-	1
PC_Precursor_Cell_Center_Batch_ID	Canonical	If this primary cell is derived from another registered cell, this field should specify the Center LINCS Batch ID of the specific batch from which it was derived.	-	2
PC_Center_Specific_Code	Batch	LINCS center-specific coded information that can include in its format information regarding the parent / protocol used / date	-	3
PC_Quality_Verification	Batch	Information pertaining to experimental verification of the primary cell identity; Batch-specific ID; STR profile	Acceptable protocols for verification will be determined by LINCS participants and a controlled vocabulary will be developed. Comment: We should at least make an effort to ensure cells within LINCS are the same either by STR / SNP profiling or by actually exchanging vials previously matched to repository	2
PC_Culture_Conditions	Batch	A description of the culture conditions that were used and are suitable for this type of cell	-	2
PC_Passage_Number	Batch	The number of times, if any the primary cells have been re-plated and allowed to grow back to confluency or to some maximum density if using suspension cultures.	-	2
PC_Transient_Modification	Batch	Transient transfection or viral transduction	Need to capture transfection agent	1
PC_Source_Information	Batch	This is a free-text field that provides detailed source information for this particular batch, which may include information on from whom and when the provider obtained the cells and for what purpose the cells were obtained by the end user.	-	2
PC_Date_Received	Batch	This field specifies when this batch was obtained from the provider. Because YEAR-MO-DY is not always known, this field may only contain partial date information (e.g. YEAR-MO, YEAR only).	-	2