Metadata Specifications:

Nucleic Acid Reagents

Natural or engineered assay perturbagens built of a polynucleic structure, often used in the interference of specific sequences of DNA/RNA in cells, tissues, or animal models, including siRNA, shRNA, and guide RNAs for TALEN and CRISPR/CAS9.

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

	LINCS Field Name	Related to	Description	Comments	Importance
Common Fields	NA_LINCS_ID	Canonical	Unique LINCS internal identifier	LINCS internal ID; This is a batch independent ID	1
	NA_Name	Canonical	The primary name of the siRNA or shRNA or any other genetic perturbagen as chosen by LINCS	Should be descriptive and correspond to existing siRNA or shRNA names as much as possible; Batch independent name	1
	NA_Alternative_Name	Canonical	List of synonymous reagent names or other alternative names	-	2
	NA_Alternative_ID	Canonical	List of synonymous reagent IDs, or other alternative IDs	-	2
	NA_Center_Canonical_ID	Canonical	LINCS DSGC-specific canonical ID; will be assigned by a given LINCS DSGC according to its reagent registration scheme	-	1
	NA_Relevant_Citations	Canonical	Appropriate literature reference(s) (with PMIDs) for reagent's derivation, production, and/or validation	-	2
	NA_Center_Name	Batch	LINCS center using the genetic perturbagen	-	1
	NA_Center_Batch_ID	Batch	LINCS DSGC-specific batch ID; will be assigned by a given LINCS DSGC according to its reagent registration scheme	-	1
	NA_Provider_Name	Batch	Vendor or laboratory that supplied the reagent	-	1
	NA_Provider_Catalog_ID	Batch	ID or catalogue number assigned to the reagent by the vendor or provider	-	1
	NA_Provider_Batch_ID	Batch	Batch or lot number assigned to the reagent by the vendor or provider	-	1
	NA_Comments	Batch	DSGC Comments regarding reagent	-	3
Custom Fields	NA_Probe_ID	Canonical	ID of the siRNA, shRNA, sgRNA, etc. as listed in NCBI Probe database	-	2
	NA_Type	Canonical	A controlled vocabulary specifying whether the reagent is Coding or Non-coding	-	1
	NA_Subtype	Canonical	A controlled vocabulary specifying the nucleic acid subtype. Coding subtypes: cDNA, Other. Non-coding subtypes: siRNA, sgRNA, shRNA, miRNA, Other	-	1
	NA_Delivery_Format	Canonical	A controlled vocabulary specifying the format in which the perturbation was introduced. Can be DNA or RNA	-	1
	NA_Mode	Canonical	A controlled vocabulary specifying the mode of action of the reagent	Can be Disruption / Element Introduction / Editing	1
	NA_Target_Locus	Canonical	The gene name of the target. Can be a gene name/ miRNA gene name or coordinates if no gene name is available	-	2
	NA_Target_Locus_Species	Canonical	A controlled vocabulary specifying the species of the target locus	-	1
	NA_Target_Gene_ID	Canonical	The NCBI Entrez Gene ID for the gene targeted by the siRNA, shRNA or sgRNA	-	2
	NA_Transcript_ID	Canonical	The NCBI Reference Sequence IDs of the transcripts that are being targeted by the reagent	-	1
	NA_Sense_Sequence	Canonical	The nucleotide sequence of the sense (passenger) strand of the siRNA, the processed shRNA or the sgRNA. The ORF sequence of the cDNA which eventually gets transcribed and translated	If the sequence is not provided by the vendor, specify the Probe ID from the NCBI Probe DB	2
	NA_Validation_Information	Canonical	Information about experimental verification of siRNA/shRNA activity; a reference (PubMed or other suitable reference) should be provided	Information about the cell line/cell type and organism used for validation, as well as the % reduction in protein expression and mRNA observed in the validation experiments; whether the target monitored in validation studies was the endogenous mRNA or a transfected mRNA	2
	NA_Driver_System	Batch	A free text field that describes the expression/driver system that was used. Examples include information about promoter, whether the expression was constitutive or inducible (and level/method of induction), or other dependencies that could effect output when the nucleic acid-based reagent is designed to be transcribed (and possible translated as protein)	-	2