

DPCC Data Standard Reference for Plasmid Reagent v1.2

	Project_Identifier	Contributing_Institution	Reagent_Name	Sample_Identifier	Strain_Name	Vector
Input Type	Text Field	Text Field	Text Field	Text Field	Text Field	Text Field
Definition	A unique Project Identifier generated by the DPCC by combining the Center-generated Project Code and a random 4-digit number	The institution code for the group that CREATED THE REAGENT. Must be one of the CEIRS institution codes assigned by NIAID.	Name of the reagent assigned by the investigator	Identifier assigned to the virus used to create the reagent	The WHO strain name of the virus that was used to create the reagent	Name of the vector used to create the plasmid
Format	Project_Code_XXXX Maximum length: 21 characters	Center three-letter code followed by three digits. Maximum length: 6 characters	Text Maximum length: 150 characters	Center-specific Allowed characters include alphanumeric, hyphen, and underscore: a-z, A-Z, 0-9, -, _ Maximum length: 50 characters	Antigenic Type/Host of Origin/Geographical Origin/Strain Number/Year of Isolation (Subtype) rg-Antigenic Type/Host of Origin/Geographical Origin/Strain Number/Year of Isolation [rg details] (Subtype) Maximum length: 150 characters	Text Maximum length: 50 characters
Value List	None	None	None	Text NA	Text U NA	None
Curation	The entry must be a Project Identifier value registered with the DPCC.	The entry must be an Institution Code value registered with the DPCC. Center 3-letter codes are case-sensitive and must be entered in all-caps.	None	It should match with the Sample_Identifier provided in related data submissions.	WHO strain naming convention: Fields must be ordered as follows and separated with the '/' character: 1. The antigenic type (e.g., A, B, C, D) 2. The host of origin (e.g., swine, equine, chicken. For human-origin viruses, no host of origin designation is given.) 3. Geographical origin (e.g., Denver, Taiwan) 4. Strain number (e.g., 15, 7) 5. Year of isolation (e.g., 2009, 1934) 6. For influenza A viruses, the hemagglutinin and neuraminidase antigen description in parentheses (e.g., (H1N1), (H3N2)) If the strain is a reverse genetic virus, prefix Strain_Name with rg-. Provide reverse genetic details in brackets between year of isolation and subtype. Enter NA if the reagent is not related to a virus strain.	None
Examples	SJCProj02_4001	SJC101	pHW2000 VN1203 PB2 E627K	21145-001a	A/Hong Kong/1/1968 (H3N2), A/chicken/Fujian/4/2002 (H3N6), A/chicken/Fujian/4/2002 (HxNx), A/chicken/Fujian/4/2002 (mixed), A/swine/Iowa/233-56/2011 (H3N2v), A/duck/Alberta/35/1976 (H1N1), B/Hong Kong/432/2014, C/Texas/9576/2011, rg-A/Puerto Rico/8/1934 (H1N1), or rg-A/Egypt/N03072/2010 [PR8 internal R 6+2] (H5N1)	pHW2000
Notes				Enter NA if the reagent is not related to a surveillance submission.	Reference: http://www.cdc.gov/flu/about/viruses/types.htm (HxNx) can be used in cases where a partial subtype has been determined (e.g., H5Nx). If there are mixed subtypes contained within a sample use A/chicken/Fujian/4/2002 (mixed) for Strain_Name or list multiple comma-separated Strain_Names: A/mallard/Alaska/2/2007 (H3N8), A/mallard/Alaska/2/2007 (H3N9). For other virus types, enter the common Strain_Name.	
Dependent Fields						
Validation	Project_Identifier should be a valid project identifier.	Must be an Institution Code value registered with the DPCC	Validate field length	Validate field length	Validate field length	Validate field length
Message Code	Error_9_PROJECT_NOT_FOUND	Error_1_INVALID_VALUE	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH
Validation	Project_Identifier exists but user does not have permission to access or edit the project.				Host of origin element of strain name must be alphanumeric, underscore, dash, period, or single quote: a-z, A-Z, 0-9, -, _, ., , '	
Message Code	Error_4_DENIED_USER_ACCESS				Error_137_INVALID_HOST_SPECIES_STRAIN_TEXT	
Validation					Geographical origin element of strain name must be alphanumeric, underscore, dash, period, or single quote: a-z, A-Z, 0-9, -, _, ., , '	
Message Code					Error_138_INVALID_LOCATION_STRAIN_TEXT	
Validation					Strain number element of strain name must be alphanumeric, underscore, dash, or period: a-z, A-Z, 0-9, -, _, ., , '	
Message Code					Error_139_INVALID_STRAIN_NUMBER_STRAIN_TEXT	
Validation					Year of isolation element of strain name must be 4-digit year.	
Message Code					Error_110_INVALID_STRAIN_YEAR	
Validation					If strain is influenza A, subtype must be present as text between parentheses.	
Message Code					Error_112_INVALID_STRAIN_SUBTYPE	
Validation					Brackets cannot be present if strain does not have rg- prefix.	
Message Code					Error_119_INVALID_STRAIN_BRACKETS	

Input Type	Segment	Mutations	BEI Number
	Text Field	Text Field	Text Field
Definition	The influenza virus segment or other element that the plasmid contains	A description of the mutation(s) or non-coding change(s) made in the plasmid reagent	The identification number assigned to the reagent by BEI
Format	Text Maximum length: 30 characters	An HGVS-formatted list describing sequence changes Maximum length: 200 characters	BEI Number Maximum length: 20 characters
Value List	HA HE MP NA NP NS P3 P42 PA PB1 PB2 OTH-	Text NA	Text NA
Curation	The entry must be one and only one member of the Value List.	The entry must use the HGVS reference for annotation guidelines to describe mutations, insertions, or deletions in any gene segment order. http://www.hgvs.org/mutnomen/quickref.html	The entry must be a valid BEI number.
Examples	PB2, or OTH-Luciferase	PB2:p.[E627K;T255K;H155del]	NR-12277
Notes		<p><u>Format for amino acid mutation:</u> protein_name:p.[mutation_information]</p> <p>For a single mutation, list the original amino acid, its position, and the new amino acid in brackets: PB2:p.[E627K]</p> <p>For a multiple mutations within a single protein, separate individual mutations by a semicolon: HA:p.[N220K;Q222L]</p> <p>For a deletion within a single protein, list the amino acid, its position, and 'del': NS:p.[H155del]</p> <p>For a stop codon, use an asterisk: NA:p.[K471_*472insHHHHH]</p> <p><u>Format for non-coding mutation:</u> gene_segment_name:c.[mutation_information]</p> <p>For a single mutation: enter the position, the original nucleotide, >, and the new nucleotide: PB2:c.[273C>T]</p> <p>For multiple mutations within a single gene, separate individual mutations by a semicolon: PB1:c.[273C>T;274A>G;275G>T]</p> <p>For a deletion: enter the position, del, and the nucleotide deleted: NA:c.[73delT]</p> <p>To list both amino acid and non-coding changes, comma-separate the individual mutation information in any order: HA:p.[N220K;Q222L],HA:c.[273C>T;274A>G;275G>T]</p> <p>Enter NA if the gene segment has no mutation.</p>	Enter NA if the reagent was not deposited with BEI.
Dependent Fields			
Validation	Field value should be one of valid values as in list.	Validate field length	Validate field length
Message Code	Error_1_INVALID_VALUE	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH
Validation	Validate field length		
Message Code	Error_75_INVALID_FIELD_LENGTH_OTH		
Validation			
Message Code			
Validation			
Message Code			
Validation			
Message Code			
Validation			
Message Code			
Validation			
Message Code			

	Contact_Name	Contact_Email	Availability	Publication PMID	Comments
Input Type	Text Field	Text Field	Text Field	Text Field	Text Field
Definition	The name of the point-of-contact for reagent requests or the name of the BEI depositor	Email address for the point-of-contact	Whether the reagent is available to share with other investigators on the reagents page of the CEIRS Public Website	The PubMed Unique Identifier (PMID) for the publication in which the reagent was published.	Text describing anything else of interest related to the submission
Format	FirstName LastName FirstName MI. LastName Maximum length: 50 characters	Text@domain.com Maximum length: 50 characters	Text Maximum length: 1 character	PMID Maximum length: 50 characters	Text Maximum length: 2000 characters
Value List	None	None	Y N	Text NA	Text NA
Curation	None	The entry must be a valid email address.	The entry must be one and only one member of the Value List.	The entry must be a valid PMID number: 7 or 8-digit number with no leading zeros. http://www.nlm.nih.gov/bsd/mms/medlineelements.htm#pmid	None
Examples	Jane Smith, or Davey M. Jones	jsmith@email.edu	Y	19465683, or NA	NA
Notes			Y = Yes N = No	If the reagent was referenced in multiple publications, comma-separate the individual PMIDs. Enter NA if the reagent was not referenced in a publication.	If there are no comments, enter NA.
Dependent Fields					
Validation	Validate field length	Validate field length	Field value should be one of valid values as in list.	Validate field length	Validate field length
Message Code	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH	Error_1_INVALID_VALUE	Error_70_INVALID_FIELD_LENGTH	Error_70_INVALID_FIELD_LENGTH
Validation		Must be a valid email address.		Must be a valid PMID.	
Message Code		Error_114_INVALID_EMAIL		Error_96_INVALID_PMIID	
Validation					
Message Code					
Validation					
Message Code					
Validation					
Message Code					
Validation					
Message Code					
Validation					
Message Code					