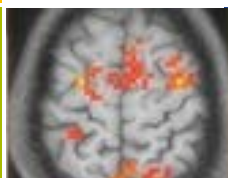




NDAR Webinar | NDAR Data Submission – Genomics Data

May 2010

Questions? Contact NDARhelp@mail.nih.gov



Agenda

- NDAR Overview
- NDAR Data Submission Preparedness
- NDAR Genomics Data Submission
- Questions and Answers

NDAR Overview

NDAR Overview

- The NDAR research portal is a secure bioinformatics platform for scientific collaboration and data sharing for Autism Spectrum Disorder (ASD) investigators

"Open access to data from many people and many studies is paramount in ASD research because of the tremendous range of symptom type and severity among those affected."

—Thomas R. Insel, M.D.,
Director, National Institute of Mental Health (NIMH)

NDAR Overview

- The NDAR research portal is a secure bioinformatics platform for scientific collaboration and data sharing for Autism Spectrum Disorder (ASD) investigators
 - Facilitates data sharing and scientific collaboration
 - Provides bioinformatics solutions to address community-wide needs
 - Enables the effective communication of detailed research data, tools, information, and results

NDAR Overview

■ NIH Funding Organizations

- National Institute of Mental Health (NIMH)
- Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD)
- National Institute of Neurological Disorders and Stroke (NINDS)
- National Institute of Environmental Health Sciences (NIEHS)

■ NIH Contributing Organizations

- Center for Information Technology (CIT)

■ NIH Advisory Groups

- NDAR Governance
- Autism Coordinating Committee (ACC)
- NDAR Implementation Team (NIT)

NDAR Overview

■ Federation Partners

- NIH NIMH Genetics Repository
- NIH NIMH Transcriptional Atlas of Human Brain Development
- NIH Pediatric MRI Data Repository
- Autism Genetic Resource Exchange (AGRE)
- Interactive Autism Network (IAN)

■ Autism Informatics Consortium

- Autism Genetic Resource Exchange (AGRE)
- Interactive Autism Network (IAN)
- Simons Foundation Autism Research Initiative (SFARI)

NDAR Team

- Mike Huerta, PhD – Director
- Anne Sperling, PhD – Policy Analyst
- Dan Hall – Manager
- Matt McAuliffe, PhD – Technical Manager, Imaging Lead
- B. Lynn Young, PhD – Genomics Lead
- Gretchen Navidi – Principal Analyst, Clinical Assessments
- Kristin Mead – ARRA Outreach Coordinator
- Eric Stanton – ARRA Web Programming

NDAR Data Submission Preparedness

Data Submission Preparedness

- Request an Account on the NDAR Portal
<http://ndar.nih.gov/ndarpublicweb/policies.go#sop1>
- Request Data Submission Privileges
<http://ndar.nih.gov/ndarpublicweb/policies.go#sop2>
- Create an NDAR Collection
<http://ndarportal.nih.gov>
- Generate GUIDs
<http://ndar.nih.gov/ndarpublicweb/policies.go#sop8>

NDAR Genomics Data Submission

NDAR Genomics Workflow

1. Data Preparation

- NG_Templates (<http://ndar.nih.gov>; click on “Standards” and “Genomics”)

2. Data Validation

- Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Data Validation Tool”

3. Data Submission

- Via the NDAR Collection you created in the NDAR Portal (<https://ndarportal.nih.gov/>); click on “Submit Data” in the Navigation section

4. Query Data (final quality check)

- Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Query Tool”

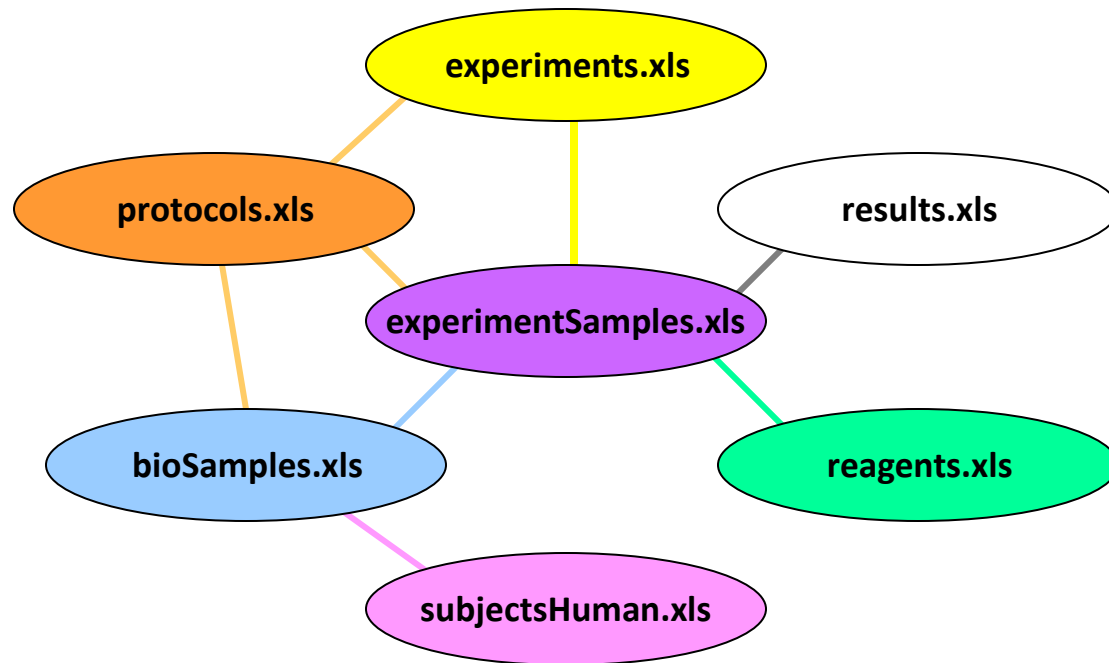
NDAR Genomics Workflow

Step 1: Data Preparation

Genomics Data Preparation

- NDAR Genomics requires detailed descriptions of experiments to enhance data sharing by query
 - Enough information such that, in theory, the experiment can be repeated
 - Required attributes, optional attributes
 - Reuse of templates and of descriptions of protocols, reagents, experiments, and biological samples

Genomics Data Submission Templates



■ NG_Templates Characteristics

- Identifiers for subjects, protocols, reagents, experiments, and biological samples
- Color coding
- Comments
- Drop down menus for filling in selected items
- Gray shading for optional columns



NDAR National Database
for Autism Research
Serving the autism research community

[Home](#) [About NDAR](#) [Data Sharing](#) **[Standards](#)** [Tools](#) [Policy & Procedures](#) [Training](#) [FAQs](#)

Back to: [Home](#) > Standards

Standards

[Subject GUIDs](#) | [Clinical Assessments](#) | [Imaging](#) | [Genomics](#)

NDAR has established a foundation of initial standards and conventions for data-sharing within the ASD research community. As a research portal, NDAR provides a platform for investigators to define, refine, and standardize data definition for autism research.

NDAR has established the following special interest groups for community involvement:

- The **ASD Ontology Group** works toward establishing an ASD Ontology using the NDAR data dictionary and tools.
- The **ASD Data Dictionary Group** works through the requirements and procedures needed to support a community-wide ASD data dictionary.
- The **ASD Data Managers Group** works through requirements and procedures for data submission to NDAR.

Please contact us at ndar@mail.nih.gov to join a special interest group. Meeting announcements will be sent to your e-mail inbox.

Subject GUIDs

- [GUID User Manual](#) [PDF 1.6 MB] *Last updated: Oct 20, 2009*
- [Request free GUID custom client software](#)
- [Brochure for Parents](#) [PDF 108 KB] *Last updated: Dec 4, 2009*

Data Definition

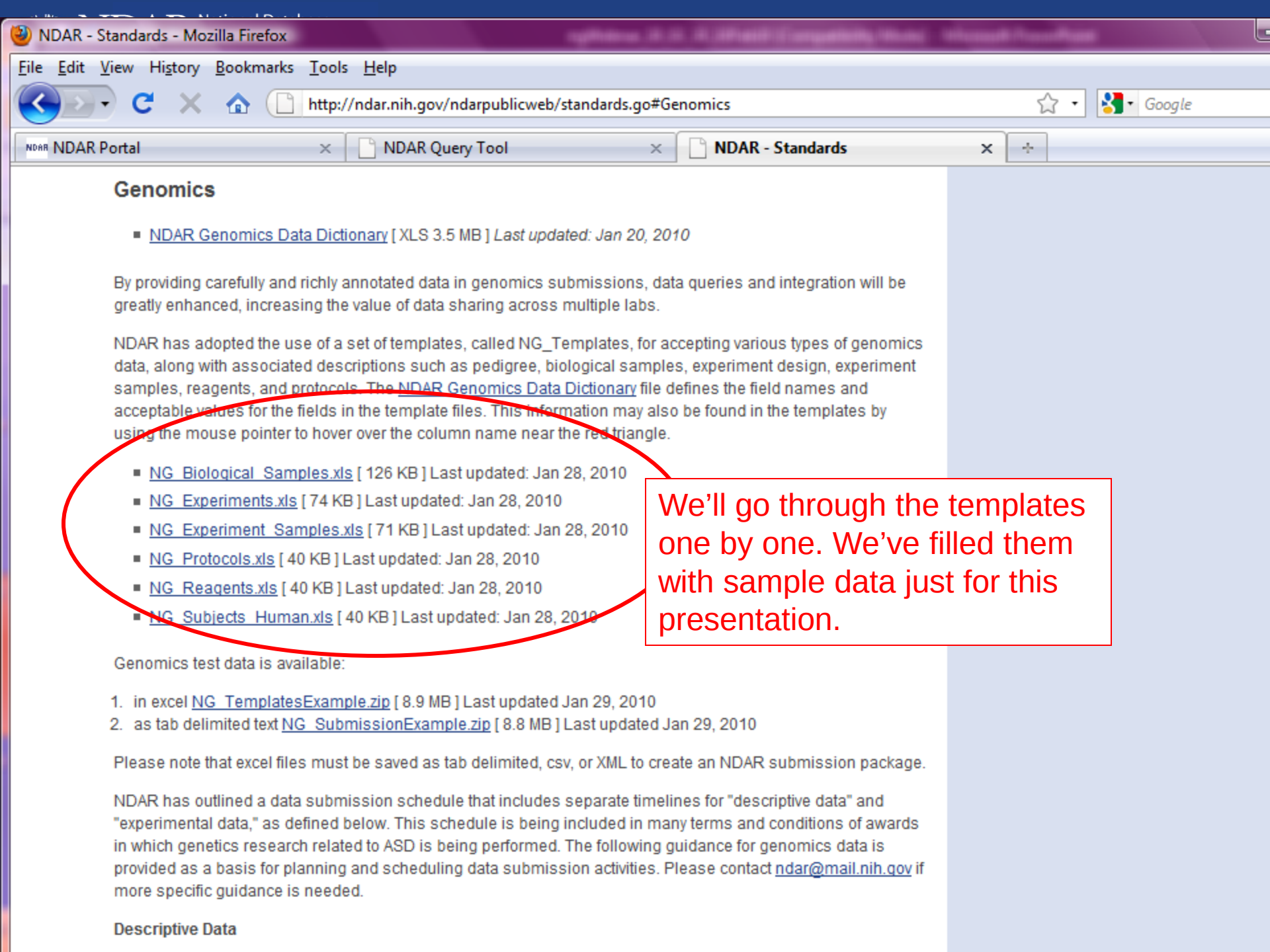
- [NDAR Data Structure Definition Template](#)
[XLS 80 KB]
Last updated: Apr 29, 2010
- [NDAR Genomics Templates](#)

Software

- [Medical Image Processing, Analysis, and Visualization \(MIPAV\) Application](#)
- [Request free GUID custom client software](#)

Manuals

- [GUID User Manual](#)
[PDF 1.6 MB]
Last updated: Oct 20, 2009



Genomics

- [NDAR Genomics Data Dictionary](#) [XLS 3.5 MB] *Last updated: Jan 20, 2010*

By providing carefully and richly annotated data in genomics submissions, data queries and integration will be greatly enhanced, increasing the value of data sharing across multiple labs.

NDAR has adopted the use of a set of templates, called NG_Templates, for accepting various types of genomics data, along with associated descriptions such as pedigree, biological samples, experiment design, experiment samples, reagents, and protocols. The [NDAR Genomics Data Dictionary](#) file defines the field names and acceptable values for the fields in the template files. This information may also be found in the templates by using the mouse pointer to hover over the column name near the red triangle.

- [NG Biological Samples.xls](#) [126 KB] *Last updated: Jan 28, 2010*
- [NG Experiments.xls](#) [74 KB] *Last updated: Jan 28, 2010*
- [NG Experiment Samples.xls](#) [71 KB] *Last updated: Jan 28, 2010*
- [NG Protocols.xls](#) [40 KB] *Last updated: Jan 28, 2010*
- [NG Reagents.xls](#) [40 KB] *Last updated: Jan 28, 2010*
- [NG Subjects Human.xls](#) [40 KB] *Last updated: Jan 28, 2010*

We'll go through the templates one by one. We've filled them with sample data just for this presentation.

Genomics test data is available:

1. in excel [NG_TemplatesExample.zip](#) [8.9 MB] *Last updated Jan 29, 2010*
2. as tab delimited text [NG_SubmissionExample.zip](#) [8.8 MB] *Last updated Jan 29, 2010*

Please note that excel files must be saved as tab delimited, csv, or XML to create an NDAR submission package.

NDAR has outlined a data submission schedule that includes separate timelines for "descriptive data" and "experimental data," as defined below. This schedule is being included in many terms and conditions of awards in which genetics research related to ASD is being performed. The following guidance for genomics data is provided as a basis for planning and scheduling data submission activities. Please contact ndar@mail.nih.gov if more specific guidance is needed.

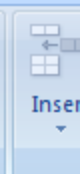
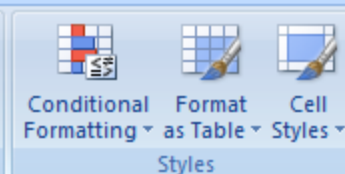
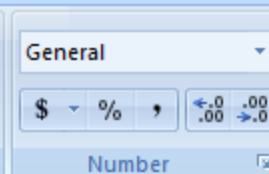
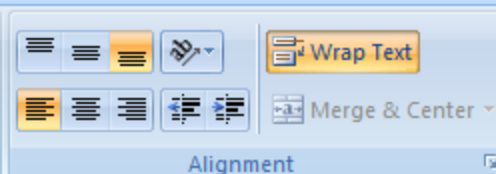
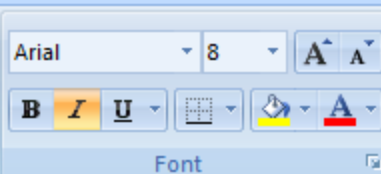
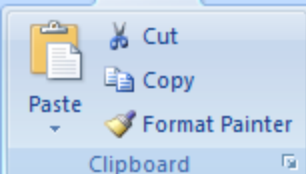
Descriptive Data

NG_Experiments [Compatibility Mode] - Microsoft Excel						
Home Insert Page Layout Formulas Data Review View						
Clipboard Font Alignment Number Styles						
A12						
A B C D E F						
1	Experiments	(Template courtesy of import.org)				
2	Experiment User-Defined ID*	Experiment NDAR Accession	Experiment Name*	Experiment Type*	Measurement Technique*	Analysis Technique*
3	experiment_user_def_id	experiment_ndar_id	experiment_name	experiment_type	measurement_technique	analysis_technique
4	GSE6754		Mapping autism risk loci using genetic linkage and chromosomal rearrangements	Genotyping	Array	Linkage analysis
5						
6						
7						
8						
9						
10						
11						
12						





Home Insert Page Layout Formulas Data Review View



A1 fx subject

	A	B	C	D	E	F	
	Subjects	(Template courtesy of import.org)					
1	<i>subject</i>						
2	Source Subject NDAR Global Unique Identifier (GUID)*	Family User-Defined ID*	Subject User-Defined ID*	Father's Pedigree User-Defined ID*	Mother's Pedigree User-Defined ID*	Analysis Technique*	Species
	<i>subject_guid</i>	<i>family_user_def_id</i>	<i>subject_user_def_id</i>	<i>father_pedigree_id</i>	<i>mother_pedigree_id</i>	<i>analysis_technique</i>	<i>species</i>
3	NDARDEMOWX878LVN	9059	9059-1			Linkage analysis	Homo s
4	NDARDEMODA301EZZ	9059	9059-2			Linkage analysis	Homo s
	NDARDEM0EZ736GR9	9059	9059-4	9059-1	9059-2	Linkage analysis	Homo s



Home Insert Page Layout Formulas Data Review View

Cut Copy Paste Format Painter Clipboard

Arial 10 A A

B I U Font

Alignment

Wrap Text Merge & Center

General Number

Conditional Formatting Format as Table Cell Styles

A3 SBSU1

	A	B	C	D	E	F	G
	Biological Samples	(Template courtesy of import.org)					
1	<i>biosample</i>						
	Biological Sample User-Defined ID*	Biological Sample NDAR Accession	Biological Sample Type*	Biological Sample Name	Biological Sample Description*	Biological Sample sub-type	Source Biological Sample Use
2	<i>biosample_user_def_id</i>	<i>biosample_ndar_id</i>	<i>sample_type</i>	<i>biosample_name</i>	<i>biosample_description</i>	<i>sample_subtype</i>	<i>source_biosample_user_def_id</i>
3	SBSU1		Body_Fluid		blood or lymphoblast		
4	SBSU2		Body_Fluid		blood or lymphoblast		
5	SBSU3		Body_Fluid		blood or lymphoblast		
6	SampleFromNDAR1		DNA		genomic DNA		SBSU1
7	SampleFromNDAR2		DNA		genomic DNA		SBSU2
8	SampleFromNDAR3		DNA		genomic DNA		SBSU3
9							
10							
11							
12							
13							
14							
15							
16							
17							
18							
19							
20							
21							
22							



Home Insert Page Layout Formulas Data Review View

Cut Copy Paste Format Painter Clipboard

Arial 8 Bold Italic Underline Font

Wrap Text Merge & Center Alignment

General Number

Conditional Formatting Format as Table Cell Styles Styles

A9						
fx						
	A	B	C	D	E	F
	Reagents-Array	(Template courtesy of import.org)				
1	<i>reagent</i>					
	Array Reagent User-Defined ID*	Array Reagent NDAR Accession	Array Platform Name*	Array Platform Description	Array Platform Annotation File Name	Manufacturer*
2	<i>reagent_user_def_id</i>	<i>reagent_ndar_id</i>	<i>reagent_name</i>	<i>reagent_description</i>	<i>annotation_file_name</i>	<i>manufacturer</i>
	GPL2641		Affymetrix GeneChip Mapping 10K 2.0 Array (Xba142_SNP)	As with the previous-generation GeneChip® Mapping 10K Array, the GeneChip Mapping 10K 2.0 Array enables researchers to genotype more than 10,000 SNPs in a single experiment with one primer. The GeneChip Mapping 10K 2.0 Array contains equivalent content and performance as the original GeneChip Mapping 10K Array, but advances in photolithographic technology enable it to be produced in a smaller, more cost-efficient format. Now, researchers can genotype 10,000 SNPs in a single experiment at a cost comparable to microsatellite research.	Mapping10K_Xba142.na24.annot.csv.zip	Affymetrix
3						
4						
5						
6						
7						
8						
9						

A22  f_x

	A	B	C	D	E	F
	Experiment Samples	1 (Template courtesy of import.org)				
1	<i>expsample</i>					
	Experiment Sample User-Defined ID*	Experiment Sample NDAR Accession	Experiment User-Defined ID*	Experiment Sample Name	Result File or Folder Name*	Type of Result
2	<i>expsample_user_def_id</i>	<i>expsample_ndar_id</i>	<i>experiment_user_def_id_ref</i>	<i>expsample_name</i>	<i>resultset_name</i>	<i>resultset_type</i>
3	GSM167064		GSE6754		GSM167064.CEL.gz;GSM167064.CHP.gz	pre-publication
4	GSM168493		GSE6754		GSM168493.CEL.gz;GSM168493.CHP.gz	pre-publication
5	GSM166354		GSE6754		GSM166354.CEL.gz;GSM166354.CHP.gz	pre-publication
6						
7						
8						
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						

Click on the triangle or hover over the short name for instructions

Protocol NDAR session	Protocol File Name*	Protocol Name*	Protocol Type*	Protocol Type (Required)
protocol_ndar_id	file_name	protocol_name	protocol_type	pr
	Extract1.pdf	Extract	Biomaterial_Transformation	Ex
	Label1.pdf	Label	Biomaterial_Transformation	La ex m in co Ru fo
	Hyb1.pdf	Hybridize	Assay	Hybridization: See Affymetrix GeneChip Mapping 10K 2.0 Assay. 1. Prepare the Hybridization Mix (for multiple samples make 5% excess). 2. Transfer each of the labeled sample from the Eppendorf tube. Aliquot 190 uL of the Hybridization Cocktail Master Mix into the 70 uL of labeled DNA. 3. Heat the 260 uL of hybridization mix and labeled DNA at 95°C in a heat block for 10 minutes. Cool down on crushed ice for 10 seconds. 5. Spin briefly at 2,000 rpm for 1 minute in a microfuge. 6. Place the tubes at 48°C for 2 minutes. 7. Inject 80 uL denatured hybridization mix. 8. Hybridize at 48°C for 16 to 18 hours at 60 rpm.
	Scan1.pdf	Scan	Assay	Scan: See Affymetrix GeneChip Mapping 10K 2.0 Assay Manual. Fluidics Station 450/2500 automate the washing and staining of GeneChip® Mapping probe arrays. Arrays were scanned on a GeneChip® Scanner 3000 using GeneChip Operating Software (GCOS)
	Growth1.pdf	Growth	Biomaterial_Transformation	Lymphoblastoid cell line was established under the following conditions
	Design1.pdf	Design	Data_Transformation	This is a large linkage study undertaken by The Autism Genome Project (AGP) Consortium to identify candidate genes underlying the etiology of autism. 168 Multiplex families with affected individuals



Name	Date modified	Type	Size
NG_Biological_Samples	1/12/2010 4:03 PM	Text Document	2 KB
NG_Experiment_Samples	1/13/2010 11:34 AM	Text Document	1 KB
NG_Experiments	1/13/2010 11:34 AM	Text Document	2 KB
NG_Protocols	1/28/2010 9:33 AM	Text Document	3 KB
NG_Reagents	1/28/2010 9:33 AM	Text Document	
NG_Subjects_Human	1/12/2010 4:03 PM	Text Document	
GSM166354.CEL.gz	1/11/2010 3:18 PM	GZ File	
GSM166354.CHP.gz	1/11/2010 3:18 PM	GZ File	
GSM167064.CEL.gz	1/11/2010 3:18 PM	GZ File	
GSM167064.CHP.gz	1/11/2010 3:18 PM	GZ File	
GSM168493.CEL.gz	1/11/2010 3:18 PM	GZ File	
GSM168493.CHP.gz	1/11/2010 3:18 PM	GZ File	
Mapping10K_Xba142.n...	6/5/2008 4:07 PM	Compressed (zipp...	
Design1	10/1/2008 3:15 PM	Adobe Acrobat D...	
Extract1	5/13/2008 4:30 PM	Adobe Acrobat D...	3 KB
Growth1	5/13/2008 4:30 PM	Adobe Acrobat D...	3 KB
Hyb1	5/13/2008 4:30 PM	Adobe Acrobat D...	3 KB
Label1	5/13/2008 4:30 PM	Adobe Acrobat D...	3 KB
Scan1	5/13/2008 4:30 PM	Adobe Acrobat D...	3 KB

Save your completed work in “Tab Delimited Text” format. Store it in the same folder with the data files, protocol description files, and feature mapping files you will be submitting.

Data Definition for Additional Data Types

- For every data submission to NDAR, a data structure must be defined and published to the NDAR Data Dictionary.
- The NG_Templates define the data structures for Genomics data.
- For other data types, such as Clinical Assessments data, NDAR provides these useful tools for data definition:
 - http://ndar.nih.gov/ndarpublicweb/Documents/NDAR_Data_Structure_Mapping_Instructions.ppt
 - http://ndar.nih.gov/ndarpublicweb/Documents/NDAR_Data_Structure_Definition_Template.xls

NDAR Genomics Example Files

- The archive of the NG_Templates with sample annotations is at
http://ndar.nih.gov/ndarpublicweb/Documents/NG_TemplatesExample.zip
- The archive of the NG_Templates saved as tab-delimited text, packaged for validation is at
http://ndar.nih.gov/ndarpublicweb/Documents/NG_SubmissionExample.zip

NDAR Genomics Workflow

-  **Data Preparation**
 - NG_Templates (<http://ndar.nih.gov>; click on “Standards” and “Genomics”)
- 2. Data Validation**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Data Validation Tool”
- 3. Data Submission**
 - Via the NDAR Collection you created in the NDAR Portal
- 4. Query Data (final quality check)**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Query Tool”

NDAR Genomics Workflow

Step 2: Data Validation



NDAR National Database
for Autism Research
Serving the greater community in autism research

NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Login

Request An Account

Change Password

Home

Welcome to NDAR!

The National Database for Autism Research (NDAR) is a resource housed at the National Institutes of Health (NIH) that will allow the formation of a world-wide "virtual team" of autism researchers. NDAR will provide seamless integration of data, research tools and institutions across the U.S. and beyond. Currently, data gathered in autism-related research is parochial, separated by format, location and method of analysis. As a result, to perform meta-analyses or cross relationships must first deal with retrieval, leading to delays in the common platform for data collection, data entry and analysis.

Warning Notice

This is a U.S. Government computer system. Unauthorized access or use of this computer system may subject violators to criminal, civil, and administrative penalties.

All information on this computer system may be intercepted, recorded, read, copied, and disclosed by and to authorized personnel for official purposes, including criminal investigations. Such information includes sensitive data encrypted to comply with confidentiality and privacy requirements. Access or use of this computer system by any person, whether authorized or unauthorized, constitutes consent to these terms. There is no right of privacy in this system.

For this presentation, we used the
NDAR Demonstration Environment.
Be sure to submit **real research data**
to your **NDAR Portal account** at
<https://ndarportal.nih.gov>

Login

DEMONSTRATION ENVIRONMENT

User Name

Password

Login

Trouble accessing NDAR? Click [here](#).

personnel. Unauthorized access or use of this computer system

[About NDAR](#) | [Contact Us](#) | [Government Warning Notice](#) | [Privacy Policy](#)



NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young |

Home

NDAR Portal Home



Welcome to the National Database for Autism Research (NDAR) Portal

NDAR Query Tool

The NDAR Query Tool provides a way for the research community to query into the shared data environment for data discovery, processing, and analysis.

NDAR Data Validation Tool

The NDAR Data Validation Tool is an application that validates data to be submitted to NDAR. The tool also creates the submission package (2 files) that can be submitted. Only the submission package created by this tool can be submitted to NDAR.

Navigation

- [Create New Study](#)
- [Submit Data](#)

Version 3.0
February 8, 2010

[About NDAR](#) | [Contact Us](#) | [Government Warning Notice](#) | [Privacy Policy](#)



NDAR Portal - Mozilla Firefox

File Edit View History Bookmarks Tools Help

NDAR

ni

h.gov

https://ndardemo.nih.gov/gridsphere/gridsphere?cid=login

☆

Google

NDAR NDAR Portal

National Institutes of Health

NDAR

National Database for Autism Research

Serving the greater community in autism research

NOTICE

This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young

NDAR

Permissions

Collection

Study

Home

NDAR Portal Home

Welcome to the National Database for Autism Research

NDAR Query Tool

The NDAR Query Tool provides a way for data discovery, processing, and analysis.

NDAR Data Validation Tool

The NDAR Data Validation Tool is an application package (2 files) that can be submitted.

Navigation

Create New Study

Submit Data

Version 3.0

February 8, 2010

About NDAR

Contact Us

Government Warning Notice

Privacy Policy

National Institutes of Health

National Center for Human Genome Research

Opening datavalidation.jnlp

You have chosen to open

datavalidation.jnlp

which is a: JNLP File

from: https://ndardemo.nih.gov

What should Firefox do with this file?

☒ Open with

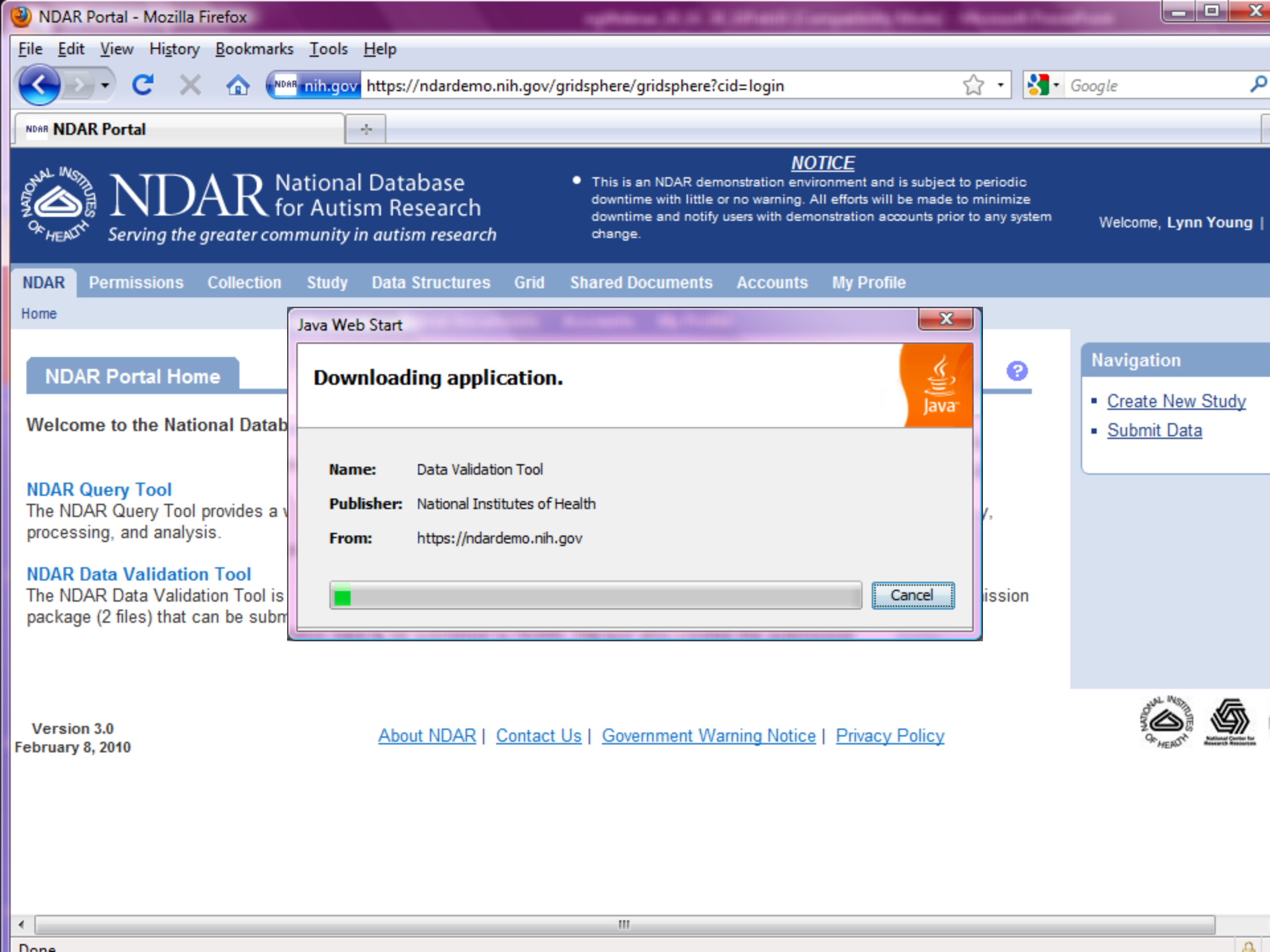
Java(TM) Web Start Launcher (default)

☐ Save File

☐ Do this automatically for files like this from now on.

OK

Cancel



NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young |

Home

NDAR Portal Home

Welcome to the National Database for Autism Research

NDAR Query Tool

The NDAR Query Tool provides a web-based interface for data processing, and analysis.

NDAR Data Validation Tool

The NDAR Data Validation Tool is a software package (2 files) that can be submitted for data validation.

Java Web Start

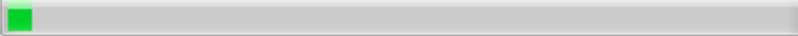
Downloading application.



Name: Data Validation Tool

Publisher: National Institutes of Health

From: https://ndardemo.nih.gov

 Cancel

Navigation

- [Create New Study](#)
- [Submit Data](#)

NDAR Data Validation [v3.1.1.328]



Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

Load Files

Files Loaded

File Name

File Type

Structure Names

Validation Results

Result

Summary

Exclude Selected Files

Include Selected Files

Mapping Details

Validate Files

Result Details

Detail Type

Description

Create Submission Package

Export Result Details

Help

Exit

NDAR Data Validation [v3.1.1.328]

Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

Load Files

Files Loaded

File Name

Exclude Selected Files

Include Selected Files

Result Details

Detail Type	Description

Select directory to validate

Look In:

ngSubmissionExample

ngSubmissionExampleErr

File Name:

Files of Type:

OK

Cancel

Create Submission Package

Export Result Details

Help

Exit

NDAR Data Validation [v3.1.1.328]



Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

C:\ndar\ngSubmissionExample\

Load Files

Files Loaded

File Name

File Type

Structure Names

Validation Results

Result

Summary

Exclude Selected Files

Include Selected Files

Mapping Details

Validate Files

Result Details

Detail Type

Description

Create Submission Package

Export Result Details

Help

Exit

NDAR Data Validation [v3.1.1.328]



Utilities
 Data Source
 Please specify the directory containing the files you would like to validate for submission to NDAR:

Files Loaded

File Name	File Type	Structure Names
/		
GSM166354.CEL.gz	UNKNOWN	Multiple Structures
Scan1.pdf	UNKNOWN	
GSM166354.CHP.gz	UNKNOWN	
Growth1.pdf	UNKNOWN	
Label1.pdf	UNKNOWN	
NG_Experiment_Samples.txt	TAB FILE	expsample01
Hyb1.pdf	UNKNOWN	

Validation Results

Result	Summary

Result Details

Detail Type	Description

<http://ndar.nih.gov>

36







NDAR Data Validation [v3.1.1.328]





Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Files Loaded

File Name	File Type	Structure Names
/		Multiple Structures
GSM166354.CEL.gz	UNKNOWN	
Scan1.pdf	UNKNOWN	
GSM166354.CHP.gz	UNKNOWN	
Growth1.pdf	UNKNOWN	
Label1.pdf	UNKNOWN	
NG_Experiment_Samples.txt	TAB FILE	expsample01
Hyb1.pdf	UNKNOWN	

Validation Results

Result	Summary

Result Details

Detail Type	Description

Validate Files 13%





Validating Files

13%

NDAR Data Validation [v3.1.1.328]

Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

C:\ndar\ngSubmissionExample\

Load Files

Files Loaded

File Name	File Type	Structure Names
/		Multiple Structures
GSM166354.CEL.gz	ASSOCIATED	
Scan1.pdf	ASSOCIATED	
GSM166354.CHP.gz	ASSOCIATED	
Growth1.pdf	ASSOCIATED	
Label1.pdf	ASSOCIATED	
NG_Experiment_Samples.txt	TAB FILE	
Hyb1.pdf	ASSOCIATED	

Validation Results

Result	Summary
Valid	[13 files]
Valid	
Valid	
Valid	
Valid	
Valid	

Exclude Selected Files

Include Selected Files

Mapping Details

Validation results



All meta-data files valid

OK

Result Details

Detail Type	Description

Create Submission Package

Export Result Details

Help

Exit

NDAR Data Validation [v3.1.1.328]

Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

C:\ndar\ngSubmissionExample\

Load Files

Files Loaded

File Name

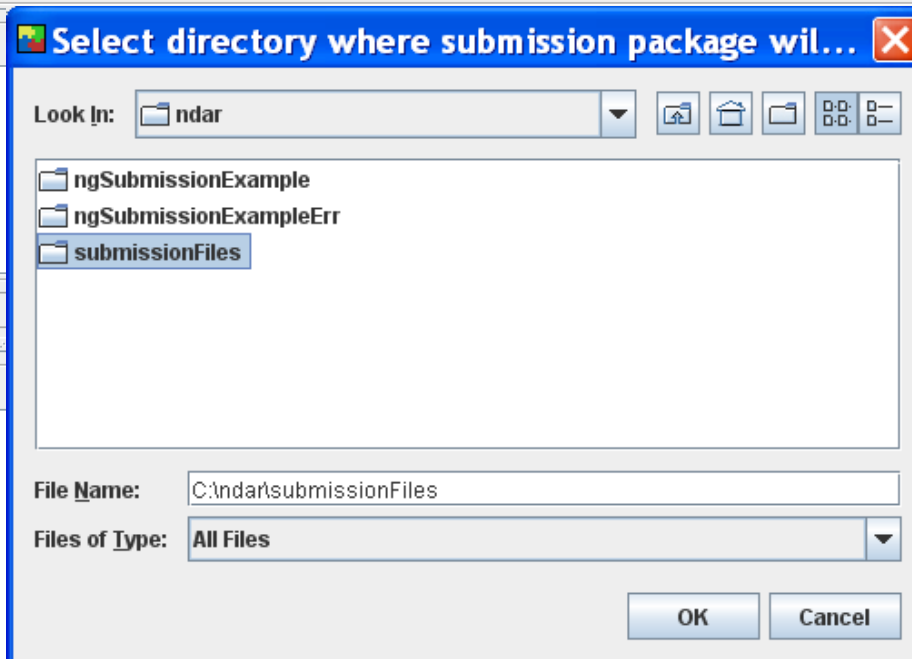
/
 GSM166354.CEL.gz
 Scan1.pdf
 GSM166354.CHP.gz
 Growth1.pdf
 Label1.pdf
 NG_Experiment_Samples.txt
 Hyb1.pdf

Exclude Selected Files

Include Selected Files

Result Details

Detail Type	Description



Create Submission Package

Export Result Details

Help

Exit

NDAR Data Validation [v3.1.1.328]

Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

C:\nda\ngSubmissionExample\

Load Files

Files Loaded

File Name	File Type	Structure Names
GSM166354.CEL.gz	ASSOCIATED	Multiple Structures
Scan1.pdf	ASSOCIATED	
GSM166354.CHP.gz	ASSOCIATED	
Growth1.pdf		
Label1.pdf		
NG_Experiment_Samples.txt		
Hyb1.pdf		

Exclude Selected Files

Include Selected Files

Ma

Result Details

Detail Type	Description

Validation Results

Result	Summary
Valid	[13 files]
Valid	
Valid	
Valid	

Description



Submission package saved to:

C:\nda\submissionFiles\NDARSubmissionPackage-1265743755040.zip

Submission ticket saved to:

C:\nda\submissionFiles\NDARSubmissionPackage-1265743755040.xml

Open directory

OK

Create Submission Package

Export Result Details

Help

Exit

NDAR Data Validation [v3.1.1.328]

Utilities

Data Source

Please specify the directory containing the files you would like to validate for submission to NDAR:

Browse

C:\ndar\ngSubmissionExampleErr\

Load Files

Files Loaded

File Name	File Type	Structure Names
/		Multiple Structures
GSM168493.CHP.gz	ASSOCIATED	
GSM166354.CEL.gz	ASSOCIATED	
NG_Protocols.txt	TAB FILE	protocol01
NG_Biological_Samples.txt	TAB FILE	biosample01
GSM166354.CHP.gz	ASSOCIATED	
NG_Experiments.txt	TAB FILE	experiment01
Label1.pdf	ASSOCIATED	

Validation Results

Result	Summary
Failed	[1 error]
Valid	
Valid	
Valid	
Valid	
Valid	
Valid	

Exclude Selected Files

Include Selected Files

Mapping Details

Validate Files

Result Details

Detail Type	Description
ZIP	Mapping10K_Xba142.na24.annot.csv.zip
GZ	GSM167064.CEL.gz
GZ	GSM167064.CHP.gz
GZ	GSM168493.CEL.gz
GZ	GSM168493.CHP.gz
GZ	GSM166354.CEL.gz
GZ	GSM166354.CHP.gz
PDF	Extract1.pdf
PDF	Label1.pdf
PDF	Hyb1.pdf
PDF	Scan1.pdf
PDF	Growth1.pdf
PDF	Design1.pdf
Error	Rules: NDARDEMO99999999 does not exist, or is invalid. Required type: NDAR*

If a dataset fails validation, as in this example, specific errors are reported. Correct these errors and re-run the Validation Tool until all data passes validation.



Create Submission Package

Export Result Details

Help

Exit

NDAR Genomics Workflow

-  **Data Preparation**
 - NG_Templates (<http://ndar.nih.gov>; click on “Standards” and “Genomics”)
-  **Data Validation**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Data Validation Tool”
- 3. Data Submission**
 - Via the NDAR Collection you created in the NDAR Portal
- 4. Query Data (final quality check)**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Query Tool”

NDAR Genomics Workflow

Step 3: Data Submission

NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young |

NDAR | [Permissions](#) | [Collection](#) | [Study](#) | [Data Structures](#) | [Grid](#) | [Shared Documents](#) | [Accounts](#) | [My Profile](#)

Home

NDAR Portal Home

Welcome to the National Database for Autism Research (NDAR) Portal

[NDAR Query Tool](#)

The NDAR Query Tool provides a way for the research community to query into the shared data environment for data discovery, processing, and analysis.

[NDAR Data Validation Tool](#)

The NDAR Data Validation Tool is an application that validates data to be submitted to NDAR. The tool also creates the submission package (2 files) that can be submitted. Only the submission package created by this tool can be submitted to NDAR.

Navigation

- [Create New Study](#)
- [Submit Data](#)

Version 3.0
February 8, 2010

[About NDAR](#) | [Contact Us](#) | [Government Warning Notice](#) | [Privacy Policy](#)



NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young

[NDAR](#) [Permissions](#) [Collection](#) [Study](#) [Data Structures](#) [Grid](#) [Shared Documents](#) [Accounts](#) [My Profile](#)

My Collections

 Please select a collection to submit data

My Collections 

ID	Collection Title	Owner	Institution	Data	Permission
NDARCOL0000093	Genomics			G	Administrator

Navigation

- [NDAR Home](#)
- [Submit Data](#)
- [Create New Study](#)



Version 3.0
February 9, 2010

[About NDAR](#) | [Contact Us](#) | [Government Warning Notice](#) | [Privacy Policy](#)



NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young |

My Collections

Collection Overview

Collection ID	NDARCOL0000093	Owner	
Collection Title	Genomics	Owner's Email	ndardemo@mail.nih.gov
Collection Status	General Collection		

Actions

- Collection Details
- View My Collections
- Edit Collection

Data Submission

Submit Data to NDAR

NDAR is a collaborative environment with privacy rules that pertain to the collection and display of imaging data. Before accessing and using NDAR, please ensure that you familiarize yourself with our privacy rules, available through the NDAR Rules of Behavior document and supporting documentation.

Collection of this information is authorized under 42 U.S.C. 241, 242, 248, 281(a)(b)(1)(P) and 44 U.S.C. 3101. The primary use of this information is to facilitate medical research around autism and autism treatment. This information may be disclosed to researchers for research purposes, and to system administrators for evaluation and data normalization.

Rules governing submission of this information are based on the data sharing rules defined in the Notice of Grant Award (NOGA). If you do not have a grant defining data sharing requirements, data submission is voluntary. Data entered into NDAR will be used solely for scientific and research purposes and is designed to further the understanding of autism and autism treatments.

Modification of NDAR information may be addressed by contacting your NDAR system administrator at ndarhelp@nih.gov

Navigation

- NDAR Home
- Submit Data
- Create New Study

Collection Overview

Collection ID	NDARCOL0000093	Owner	
Collection Title	Genomics	Owner's Email	ndardemo@mail.nih.gov
Collection Status	General Collection		

- Collection Details
- View My Collections
- Edit Collection

Data Submission

Submit Data to NDAR

NDAR is a collaborative environment with privacy rules that pertain to the collection and display of imaging data. Before accessing and using NDAR, please ensure that you familiarize yourself with our privacy rules, available through the NDAR Rules of Behavior document and supporting documentation.

Collection of this information is authorized under 42 U.S.C. 241, 242, 248, 281(a)(b)(1)(P) and 44 U.S.C. 3101. The primary use of this information is to facilitate medical research around autism and autism treatment. This information may be disclosed to researchers for research purposes, and to system administrators for evaluation and data normalization.

Rules governing submission of this information are based on the data sharing rules defined in the Notice of Grant Award (NOGA). If you do not have a grant defining data sharing requirements, data submission is voluntary. Data entered into NDAR will be used solely for scientific and research purposes and is designed to further the understanding of autism and autism treatments. Modification of NDAR information may be addressed by contacting your NDAR system administrator at ndarhelp@nih.gov. Significant system update information may be posted on the NDAR site as required.

Accept

Navigation

- NDAR Home
- Submit Data
- Create New Study

NDAR NDAR Portal

Rules governing submission of this information are based on the data sharing rules defined in the Notice of Grant Award (NOGA). If you do not have a grant defining data sharing requirements, data submission is voluntary. Data entered into NDAR will be used solely for scientific and research purposes and is designed to further the understanding of autism and autism treatments. Modification of NDAR information may be addressed by contacting your NDAR system administrator at ndarhelp@nih.gov. Significant system update information may be posted on the NDAR site as required.

Accept

Data Submission Details

Please enter the following information. Fields marked with an asterisk (*) are required.

Please upload the Submission Ticket file*

Browse...

Please upload the Submission Zip file*

Browse...

Title

Abstract

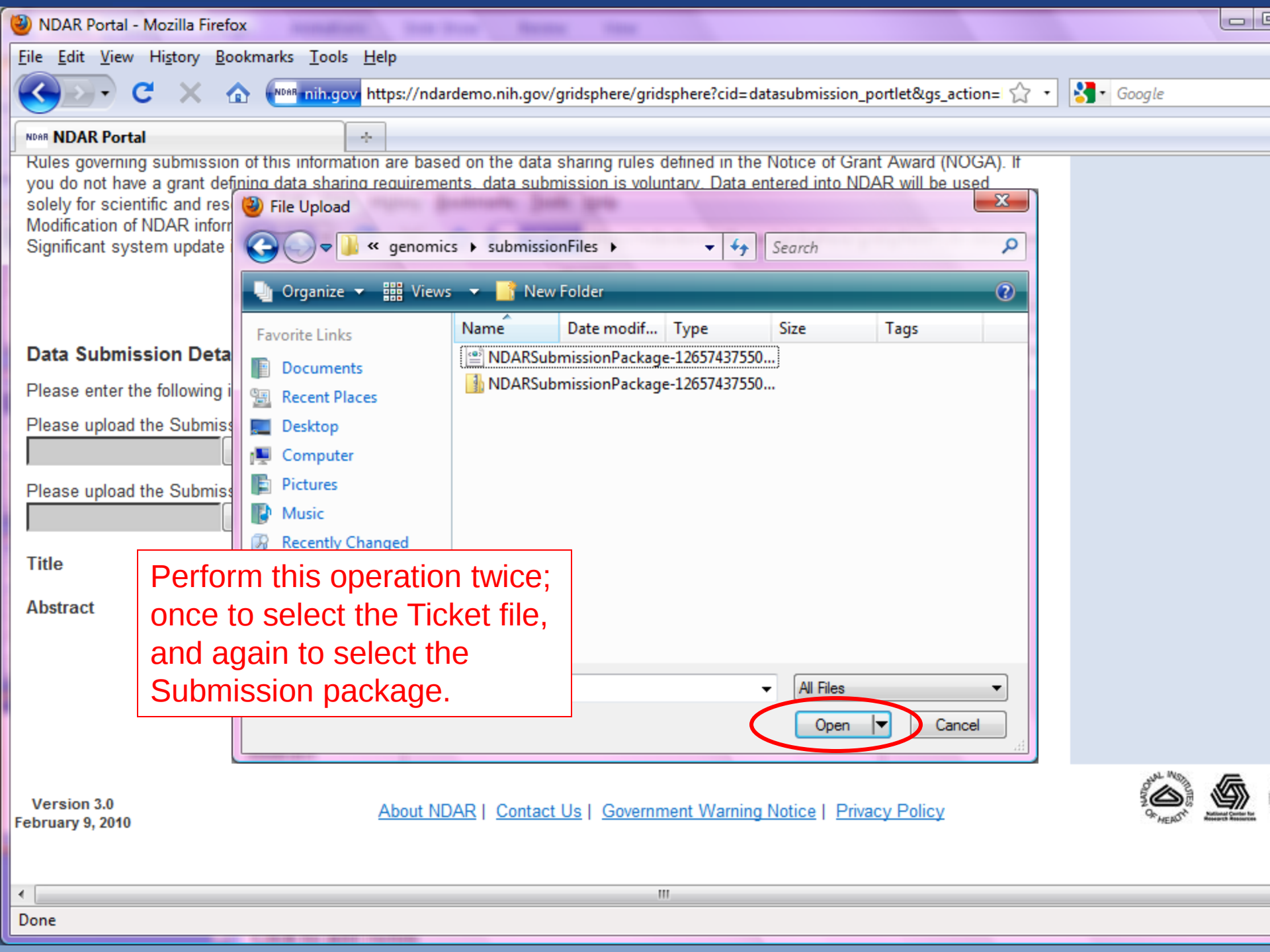
[Cancel](#)

Submit

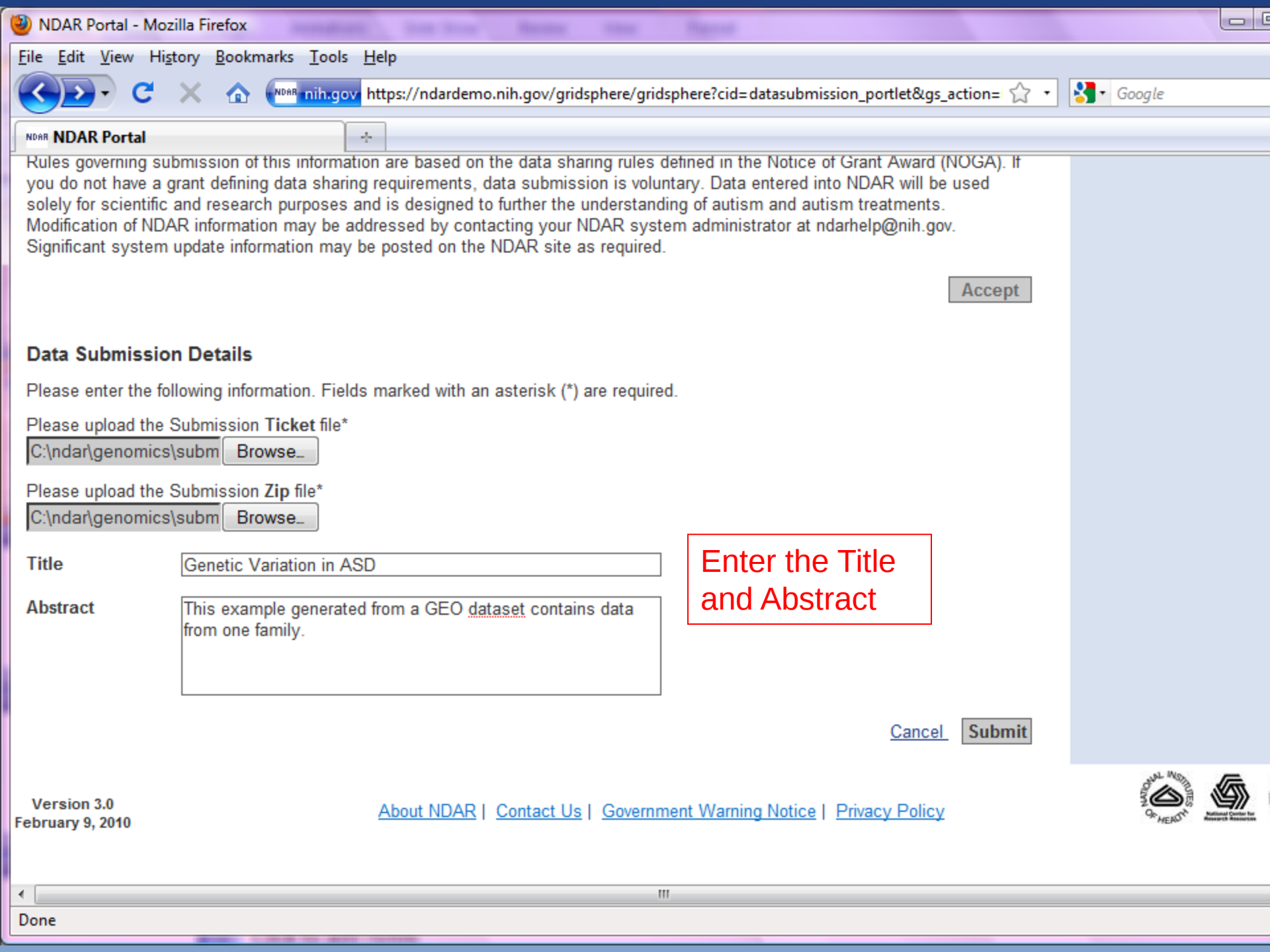
Version 3.0
February 9, 2010

[About NDAR](#) | [Contact Us](#) | [Government Warning Notice](#) | [Privacy Policy](#)





Perform this operation twice;
once to select the Ticket file,
and again to select the
Submission package.



NDAR NDAR Portal

Rules governing submission of this information are based on the data sharing rules defined in the Notice of Grant Award (NOGA). If you do not have a grant defining data sharing requirements, data submission is voluntary. Data entered into NDAR will be used solely for scientific and research purposes and is designed to further the understanding of autism and autism treatments. Modification of NDAR information may be addressed by contacting your NDAR system administrator at ndarhelp@nih.gov. Significant system update information may be posted on the NDAR site as required.

[Accept](#)**Data Submission Details**

Please enter the following information. Fields marked with an asterisk (*) are required.

Please upload the Submission Ticket file*

 [Browse...](#)

Please upload the Submission Zip file*

 [Browse...](#)

Title

Abstract

Enter the Title
and Abstract

[Cancel](#) [Submit](#)

Loading...

Rules governing submission of this information are based on the data sharing rules defined in the Notice of Grant Award (NOGA). If you do not have a grant defining data sharing requirements, data submission is voluntary. Data entered into NDAR will be used solely for scientific and research purposes and is designed to further the understanding of autism and autism treatments. Modification of NDAR information may be addressed by contacting your NDAR system administrator at ndarhelp@nih.gov. Significant system update information may be posted on the NDAR site as required.

Accept

Data Submission Details

Please enter the following information. Fields marked with an asterisk (*) are required.

Please upload the Submission Ticket file*

C:\ndar\genomics\subm

Please upload the Submission Zip file*

C:\ndar\genomics\subm

Title

Abstract

[Cancel](#)

NDAR Genomics Workflow

-  **Data Preparation**
 - NG_Templates (<http://ndar.nih.gov>; click on “Standards” and “Genomics”)
-  **Data Validation**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Data Validation Tool”
-  **Data Submission**
 - Via the NDAR Collection you created in the NDAR Portal
- 4. Query Data (final quality check)**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Query Tool”

NDAR Genomics Workflow

Step 4: Query Data

NOTICE

- This is an NDAR demonstration environment and is subject to periodic downtime with little or no warning. All efforts will be made to minimize downtime and notify users with demonstration accounts prior to any system change.

Welcome, Lynn Young |

Home

NDAR Portal Home



Welcome to the National Database for Autism Research (NDAR) Portal

NDAR Query Tool

The NDAR Query Tool provides a way for the research community to query into the shared data environment for data discovery, processing, and analysis.

NDAR Data Validation Tool

The NDAR Data Validation Tool is an application that validates data to be submitted to NDAR. The tool also creates the submission package (2 files) that can be submitted. Only the submission package created by this tool can be submitted to NDAR.

Navigation

- [Create New Study](#)
- [Submit Data](#)

Version 3.0
February 8, 2010

[About NDAR](#) | [Contact Us](#) | [Government Warning Notice](#) | [Privacy Policy](#)



NDAR Query Tool v2.1

Welcome to the NDAR Query Tool! Listed below are two methods for searching the NDAR data repository. Click on a hyperlink to query NDAR using the method

[Find Collections of Data](#)

This method allows you to search for Collections of data within NDAR by matching the keywords you enter with the Titles, Abstracts and Principal Investigators of Collections. You will be returned a list of Collections meeting the criteria of your search and you can drill down into each Collection to see its descriptive information.

[Query NDAR Data Sources](#)

This method allows you to query NDAR data across NDAR and federated data sources using the Bioinformatics Research Network (BIRN) Mediator. The results of your query will be returned in tabular format and can be exported into CSV, Microsoft Excel or XML formats. Please note that the data sources and records you have access to query will be restricted based upon the permissions of your NDAR account.

1. Select Data Sources

The NDAR data query tool provides a mechanism to access data stored within the NDAR shared data repository. Select the desired data sources, output fields, and inclusion/exclusion criteria to obtain a set of results matching the input parameters.

Note: Query execution may take several minutes. Please do not click the [Submit Query] button more than once.

[Load Query From File](#)

[Query by Collection](#) | [Query by Type](#) | [Query by Study](#) 

Data SourceDescription

☒ Genomics

[Clear All Selections](#)

[Start Over](#)

[New Query](#)

1. Select Data Sources ?

The NDAR data query tool provides a mechanism to access data stored within the NDAR shared data repository. Select the desired data sources, output fields, and inclusion/exclusion criteria to obtain a set of results matching the input parameters.

Note: Query execution may take several minutes. Please do not click the [Submit Query] button more than once.
[Load Query From File](#)

Query by Collection | [Query by Type](#) | [Query by Study](#) ?

Data SourceDescription

- ☐ Genomics
 - ☐ Genomics
 - ☐ 183 - Test Submission 2 - Genomics
 - ☐ Affections Affections
 - ☐ Biological Sample Protocols Biological Sample to Protocol links
 - ☐ Biological Sample Treatment Information Biological Sample Treatment Information
 - ☐ Biological Samples Biological Samples
 - ☐ Experiment Protocols Experi
 - ☐ Experiment Sample Biological
 - ☐ Experiment Sample Protocol
 - ☐ Experiment Sample Reagent
 - ☐ Experiment Samples Experi
 - ☐ Experiments Experiments
 - ☐ Image Image
 - ☐ Protocols Protocols
 - ☒ Reagents Reagents
 - ☐ Subject Protocols Subject to Protocol links
 - ☐ Subject Treatment Information Subject Treatment Information
 - ☐ Subjects Subjects

Select one or more data structures to query and view and click the Next >> button

https://ndardemo.nin.gov/ndarquery/showfieldCriteria.action

Google

NDAR NDAR Portal

NDAR Query Tool

Output Field Type Description ?

☒ Reagents

☒ DATASET_ID number DATASET_ID

☒ reagent_name string Reagent Name

☒ reagent_description string Reagent Description

☒ reagent_user_def_id string Reagent User Defined ID

☒ reagent_ndar_id string NDAR Accessions are used when indicating a Reagent that has already been uploaded into NDAR. Accession assigned to previously uploaded Reagents can be found via the NDAR Query Tool.

☒ annotation_file_name string Reagent Annotation File Name

☒ manufacturer string Manufacturer

☒ catalog_number string Catalog Number

☒ lot_number string Lot Number

☒ contact string Contact

☒ web_link string Web Link

Limit the number of
results with a query

Clear All Sele

Query Criteria ?

▼ ndarx_ng_reagent01.manufacturer ▼

= ▼

Affymetrix

▼ Remove

▼

▼

▼ Remove

Add Row

Clear

--- Please Select a Study --- ▼

Add Query to Study



Save Q

3. Query Results ?

Note: When attempting to add subjects to a study, the system will not allow the addition of any subjects unless all study members have access to all associated data sets. Study members are users with "Read-Only" or "Administrator" access to the study.

One item found.

<< Back Start Over

ndarx_ng_reagent01.DATASET_ID DATASET_ID	ndarx_ng_reagent01.reagent_name Reagent Name	ndarx_ng_reagent01.reagent_description Reagent Description	ndarx_ng_reagent01.reagent_user_defined_id Reagent User Defined ID
183	Affymetrix GeneChip Mapping 10K 2.0 Array (Xba142_SNP)	Generation GeneChip Mapping 10K 2.0 Array enables researchers to genotype more than 10,000 SNPs in a single experiment with one primer. The GeneChip Mapping 10K 2.0 Array contains equivalent content and performance as the original GeneChip Mapping 10K Array, but advances in photolithographic technology enable it to be produced in a smaller, more cost-efficient format. Now, researchers can genotype 10,000 SNPs in a single experiment at a cost comparable to microsatellite research."	GPL2641

Verify that the data is displayed as expected

Export options:  CSV |  XML

<< Back Start Over

Please Select a Study: Add Query to Study

1. Select Data Sources

The NDAR data query tool provides a mechanism to access data stored within the NDAR shared data repository. Select the desired data sources, output fields, and inclusion/exclusion criteria to obtain a set of results matching the input parameters.

Note: Query execution may take several minutes. Please do not click the [Submit Query] button more than once.

[Load Query From File](#)

[Query by Collection](#) | [Query by Type](#) | [Query by Study](#)

Data SourceDescription

☐ Genomics

☐ Genomics

☐ 183 - Test Submission 2 - Genomics

☐ Affections Affections

☐ Biological Sample Proto

☐ Biological Sample Treat

☐ Biological Samples Biolo

☐ Experiment Protocols E

☐ Experiment Sample Biol

☐ Experiment Sample Pro

☐ Experiment Sample Rea

☒ Experiment Samples Experiment Samples

☐ Experiments Experiments

☐ Image Image

☐ Protocols Protocols

☐ Reagents Reagents

☐ Subject Protocols Subject to Protocol links

☐ Subject Treatment Information Subject Treatment Information

☐ Subjects Subjects

We'll query a different dataset to make sure that the links to the data files are working properly

Clear All Selections

Next

Output Field Type Description

Experiment Samples

- ☐ dataset_id number DATASET ID

☐ expsample_name string Experiment Sample Name

☐ expsample_description string Experiment Sample Description

☒ expsample_user_def_id string Experiment Sample User-Defined ID

☐ expsample_ndar_id string NDAR Accessions are used when indicating an Experiment Sample that has already been uploaded into NDAR. Accessions assigned to previously uploaded Experiment Samples can be found via the NDAR Query Tool.

☒ experiment_user_def_id_ref string experiment_user_def_id_ref

☐ experiment_id number NDAR Accession ID for the Experiment to which this Experiment Sample belongs.

☒ resultset_name string File or folder name linked to the experiment sample.

☒ resultset_type string Type of the file or folder linked to the experiment sample. Publication indicates data used in the analysis for a published result. Pre-publication indicates data gathered but not necessarily used in the analysis for the published result.

☒ replicate_group_id string Unique Replicate Group ID to which sample belongs (e.g. duplicates, triplicates).

Select desired data elements/fields

Clear All Selections

Query Criteria

Remove

Remove

Add Row

Clear All

Please Select a Study --- ▼

Add Query to Study

Start Over  Back  Submit  Query

3. Query Results ?

Note: When attempting to add subjects to a study, the system will only allow you to add subjects with associated data sets. Study members are users with "Read-Only" access to all data sets.

3 items found, displaying all items.

ndarx_ng_expsample01.expsample_user_def_id Experiment Sample User-Defined ID	ndarx_ng_expsample01.experiment_user_def_id_ref experiment_user_def_id_ref	ndarx_ng_expsample01.resultset_name File or folder name linked to the experiment sample.
GSM167064	GSE6754	gsm167064.cel.gz
GSM168493	GSE6754	gsm168493.cel.gz
GSM166354	GSE6754	gsm166354.cel.gz

Export options:  CSV |  XML

--- Please Select a Study --- ▼

Add Query to Study

Opening gsm167064.cel.gz

You have chosen to open

 **gsm167064.cel.gz**
which is a: Firefox Document
from: <https://ndardemo.nih.gov>

What should Firefox do with this file?

- ☐ Open with Firefox (default)
- ☒ Save File
- ☐ Do this automatically for files like this from now on.

OK

Cancel

NDAR Genomics Workflow

-  **Data Preparation**
 - NG_Templates (<http://ndar.nih.gov>; click on “Standards” and “Genomics”)
-  **Data Validation**
 - Log into NDAR (<http://ndar.nih.gov>) and click on “NDAR Data Validation”
-  **Data Submission**
 - Via the NDAR Collection you created in the NDAR Portal
-  **Query Data (final quality check)**
 - Log into NDAR Portal (<https://ndarportal.nih.gov/>) and click on “NDAR Query Tool”



Submission
Complete!

NDAR Technical Assistance

- E-mail: ndarhelp@mail.nih.gov
- Web site: <http://ndar.nih.gov>
 - Standard Operating Procedures (Policy & Procedures page)
 - Webinars (Training page)
 - Online Tutorials (Training page, Spring 2010)
- Listservs: <http://ndar.nih.gov/ndarpublicweb/Listserv.go>
 - **NDAR General:** System enhancements, Web updates, training schedules, meetings and special events
 - **NDAR Data:** Timely technical updates, system releases, new data structures, data management-related tools and trainings
- Demonstration Environment: <http://ndardemo.nih.gov>
 - Request an Account
 - Practice setting up data structures and submitting test data