

# Tutorial 9A: Data Import

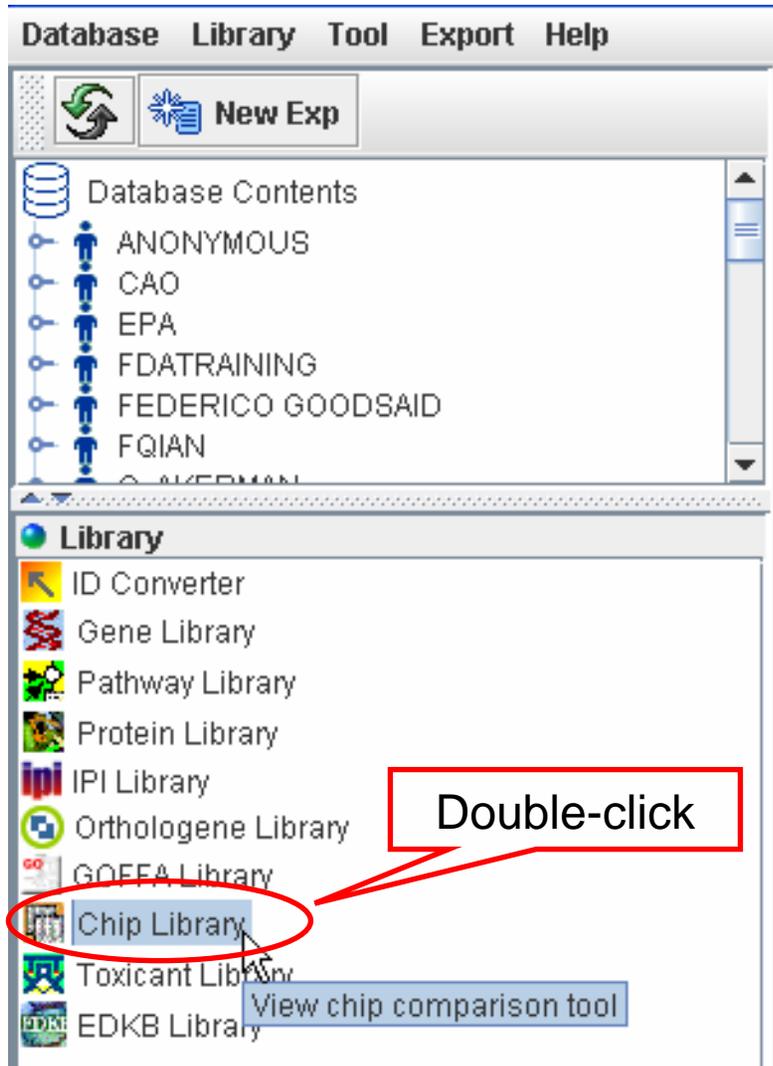


# Data Import

Before you import any data, you need to create array type unless the array type is already exist in ArrayTrack (you can check all the existing array type from Chip Library). For Affymetrix array type, you can take it as it is. But for other array type (e.g. Agilent) you need to check the coordinates and gene ID to make sure they are exactly match with your data file.

You also need to create hybridization file which will be used for batch import. The following slides will lead you through all the steps.

# Check Existing Array Type



Chip Library stores all the array types that are available in ArrayTrack.

# Check Existing Array Type

The screenshot shows the 'Chip Library' application window. The 'Summary' tab is active, and the 'Specify Category for Array Type' section is expanded. The 'All' radio button is selected. The 'Manufacture' section includes options for ABI, Affymetrix, Agilent, CloneTech, GEHC, MWG, ResGen, and Others. The 'Species' section includes options for Human, Mouse, Rat, and Others. The 'Channels' section includes options for One channel and Two channels. The 'ArrayType' section includes options for Oligo array and cDNA array.

The main table displays the following data:

	ARRAYTYPENAME	NUMBEROFSPOTS	VENDORNAME	CATEGORY	DESCRIPTION	CH
1	Rat1_2	1185	CloneTech	RAT	cDNA Expression Arrays	1
2	GF200-ResGen	5760	ResGen	HUMAN	cDNA array	1
3	GF300-ResGen	5760	ResGen	RAT	cDNA array	1
4	GF301-ResGen	5760	ResGen	RAT	cDNA array	1
5	NCTR_RAT4K 1	4096	CloneTech	RAT	Oligonucleotide array	2
6	GF201-ResGen	5760	ResGen	HUMAN	cDNA array	1
7	GF202-ResGen	5760	ResGen	HUMAN	cDNA array	1
8	Phase-1RAT700	3136	Phase-1Tox	RAT	cDNA array	2
9	NCTR_MOUSE5K 1	5376	CloneTech	Mouse	Oligonucleotide array	2
10	GF203-ResGen	5760	ResGen	HUMAN	cDNA array	1
11	GF204-ResGen	5760	ResGen	HUMAN	cDNA array	1
12	GF205-ResGen	5760	ResGen	HUMAN	cDNA array	1
13	GF206-ResGen	5760	ResGen	HUMAN	cDNA array	1
14	ChipScreen-Gs-Human	8064	ChipScreen	HUMAN	cDNA array	2
15	GF302-ResGen	5760	ResGen	RAT	cDNA array	1
16	NCTR_MOUSE5K 2	5376	CloneTech	MOUSE	Oligonucleotide array	2
17	MWG530	1152	MWG	RAT	Oligonucleotide array	2
18	STv4	17328	McClelland Lab	HUMAN	cDNA array	2
19	Affy_RG_U43A	0	Affymetrix	RAT	Oligonucleotide array	1
20	NCTR_MWG_Mouse 20K	20160	MWG	MOUSE	Oligonucleotide array	2
21	Affy_HG-U133A	22283	Affymetrix	HUMAN	Oligonucleotide array	1
22	Affy_HG-Focus	8793	Affymetrix	HUMAN	Oligonucleotide array	1
23	Affy_HG-U133_Plus_2	54675	Affymetrix	HUMAN	Oligonucleotide array	1
24	Affy_HG-U133B	22645	Affymetrix	HUMAN	Oligonucleotide array	1
25	Affy_HG-U133A_2	22277	Affymetrix	HUMAN	Oligonucleotide array	1
26	Affy_HG-U95Av2	12625	Affymetrix	HUMAN	Oligonucleotide array	1
27	Affy_MOE430A	22690	Affymetrix	MOUSE	Oligonucleotide array	1
28	Affy_MOE430B	22575	Affymetrix	MOUSE	Oligonucleotide array	1
29	Affy_Mouse430_2	45101	Affymetrix	MOUSE	Oligonucleotide array	1
30	Affy_Mouse430A_2	22690	Affymetrix	MOUSE	Oligonucleotide array	1

# Check Existing Array Type

- If the array type for your data (e.g. Affy\_HG U133\_Plus\_2) does not exist in ArrayTrack, you need to go to slide #34 (create array type). After you created array type, then come back to continue the next slides (create experiment).
- If the array type for your data does exist in ArrayTrack, you can continue to the next slide (create experiment).

# Create experiment

Before importing data, you need to create an experiment to hold the data.

## Create experiment

1. Click "New Exp" button

2. Choose group owner who will own this new exp and type the new exp name

3. Save the exp

Must Save Exp before continuing.

Give an unique experiment title

User Group Name: ANONYMOUS

Experiment Name: Demo\_2channel e.g., myexp; Rat treated with drug; ...

Owner ID: FQIAN

- User Group Name can be any text. It is used to organize a set of experiments conducted by the User Group and will be displayed in the database tree
- Experiment Name can be any text. We recommend that the name you use should be descriptive to the experiment .

OK Cancel

\* Institute: NCTR Division: Division of ...

\* Exp Types: treated vs untreated; Select

\* Key words:

\* Exp Description:

\* Phenotype Anchoring: Select

Comments:

Protocol: Import Export

Exp Design Protocol: Import Export

Hybridization Protocol: Import Export

Labeling Protocols: Import Export

RNA Extraction Protocols: Import Export

Manuscript/Reprint: Import Export

Significant Gene List: Import Export

# Create experiment

Input Form

Clean Form Print Form Help

\* is required

1 Experiment Design

\* Experiment ID: Demo\_2channel Delete Exp Edit Privileges

\* Experimenter: [ ] Date(mm/dd/yyyy): [ ] / [ ] / [ ]

\* InSTITUTE: --Select--

\* Exp Types: [ ] Select

\* Key words: [ ]

\* Exp Description: [ ]

\* Phenotype Anchoring: [ ] Select

Comments: [ ]

Protocol: [ ] Import Export

Exp Design Protocol: [ ] Import Export

Hybridization Protocol: [ ] Import Export

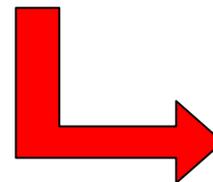
Labeling Protocols: [ ] Import Export

RNA Extraction Protocols: [ ] Import Export

Manuscript/Reprint: [ ] Import Export

Significant Gene List: [ ] Import Export

Must Save Exp before continuing.



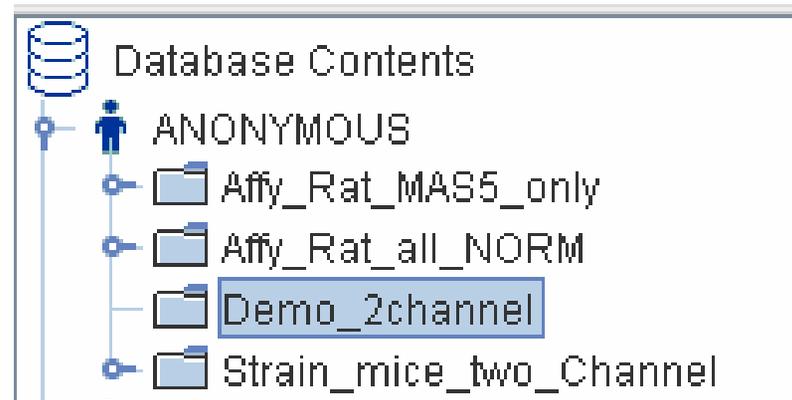
Select an Option

? Save Exp data to DB?

Yes No Cancel

# Create experiment

After clicking the save exp button, you will see the new experiment displaying under the designated owner.

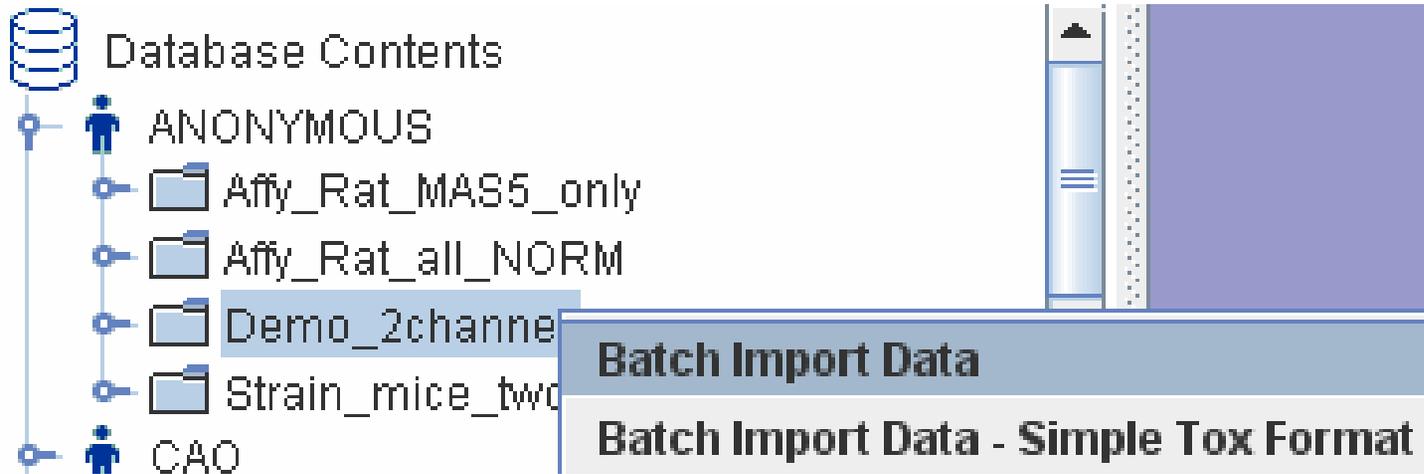


# Batch Import

There are two formats for batch import:

- Regular batch import
- SimpleTox format – for animal toxicology study data submission, with combined and extracted main fields from SEND (Standard for Exchange of Nonclinical Data) \*See tutorial 13 – SimpleTox Format.

# Batch Import - regular



Right-click the experiment, choose “Batch Import Data”

# Batch Import Wizard

This is an example of importing 2-channel data.

**Batch Import**

**Please Fill Out Experiment Information:**

If you have configuration file, load it to auto fill.

Experiment Name :

Array Type Name :

Species :  Human  Mouse  Rat  Others

Assay :  vivo  vitro

Inserts or Updates?

Insert new datas, with no updates of existing data

Only update existing hybridizations and samples, ignoring intensity datas and images

Click Open Configuration button if you have configure file, if not just skip it.

Choose the array type

--Select--

- NCTR\_MOUSE5K 2
- NCTR\_MWG\_Mouse 14K
- NCTR\_MWG\_Mouse 20K
- NCTR\_MWG\_Mouse10K
- NCTR\_MWG\_Rat 10K
- NCTR\_MWG\_Rat10K\_Demo
- NCTR\_MWG\_Rat\_10K\_48pin
- NCTR\_RAT4K 1

Click Next button

# Batch Import

Use Excel to create a hybridization file like following. You will need this file for batch import.

	A *	B *	C	D	E	F	G	H	I	J *
1	HybName	Sample1	Channel1	Sample2	Channel2	Compound	Bio_replicate	Dose	Time	FileName
2	control_0_1	S0507	Cy5	reference	Cy3	control	1	0	6months	2004-12-15-S0507-rescan.gpr
3	control_0_2	S0508	Cy5	reference	Cy3	control	2	0	6months	2004-12-15-S0508-rescan.gpr
4	control_0_3	S0517	Cy5	reference	Cy3	control	3	0	6months	2004-12-16-S0517.gpr
5	control_0_4	S0518	Cy5	reference	Cy3	control	4	0	6months	2004-12-16-S0518-rescan2.gpr
6	control_0_5	S0527	Cy5	reference	Cy3	control	5	0	6months	2004-12-17-S0527-rescan.gpr
7	Drug1_250_1	S0513	Cy5	reference	Cy3	Drug1	1	250	6months	2004-12-15-S0513.gpr
8	Drug1_250_2	S0514	Cy5	reference	Cy3	Drug1	2	250	6months	2004-12-15-S0514.gpr
9	Drug1_250_3	S0523	Cy5	reference	Cy3	Drug1	3	250	6months	2004-12-16-S0523.gpr
10	Drug1_250_4	S0524	Cy5	reference	Cy3	Drug1	4	250	6months	2004-12-16-S0524.gpr
11	Drug1_250_5	S0550	Cy5	reference	Cy3	Drug1	5	250	6months	2004-12-17-S0550.gpr
12	Drug1_500_1	S0515	Cy5	reference	Cy3	Drug1	1	500	6months	2004-12-15-S0515.gpr
13	Drug1_500_2	S0516	Cy5	reference	Cy3	Drug1	2	500	6months	2004-12-15-S0516.gpr
14	Drug1_500_3	S0525	Cy5	reference	Cy3	Drug1	3	500	6months	2004-12-16-S0525.gpr
15	Drug1_500_4	S0526	Cy5	reference	Cy3	Drug1	4	500	6months	2004-12-16-S0526.gpr
16	Drug1_500_5	S0549	Cy5	reference	Cy3	Drug1	5	500	6months	2004-12-17-S0549.gpr

\* is required column. The others are optional columns.

# Batch Import Wizard – continued

**Map Hybridization Info File To Database**

Array Raw Data Files Root Directory :  **Browse**

Hybridization Information File :  **Browse**

HybName	Sample1	Channel1	Sample2	Channel2	Compound	Bio_replicate	Dose	Time	FileName
control_0_1	S0507	Cy5	reference	Cy3	control	1	0	6months	2004-12-15-S0507-rescan.gpr
control_0_2	S0508	Cy5	reference	Cy3	control	2	0	6months	2004-12-15-S0508-rescan.gpr
control_0_3	S0517	Cy5	reference	Cy3	control	3	0	6months	2004-12-16-S0517.gpr
control_0_4	S0518	Cy5	reference	Cy3	control	4	0	6months	2004-12-16-S0518-rescan2.gpr

For each column description below, select a table column above and click the description button to make the assignment. To add more database fields, click right button **More DB Fields**

- [HYBRIDIZATION] **hybname**
- [SAMPLE] **samplename** for channel 1
- [HYBRIDIZATION] **labelname** for channel 1
- [SAMPLE] **samplename** for channel 2
- [HYBRIDIZATION] **labelname** for channel 2
- [ARRAYRAWDATA] **description**
- [ARRAYRAWDATA] **rawdatafilename** (txt)
- [MVODOSE] **compoundname** for channel 1
- [MVODOSE] **compoundname** for channel 2
- [MVODOSE] **doseamount** for channel 1
- [MVODOSE] **doseamount** for channel 2
- [MVODOSE] **schedule** for channel 1

Dropdown menu options: --Select--, --New, Cy3, Cy5, Alexa1, Alexa2, Alexa3, 33P

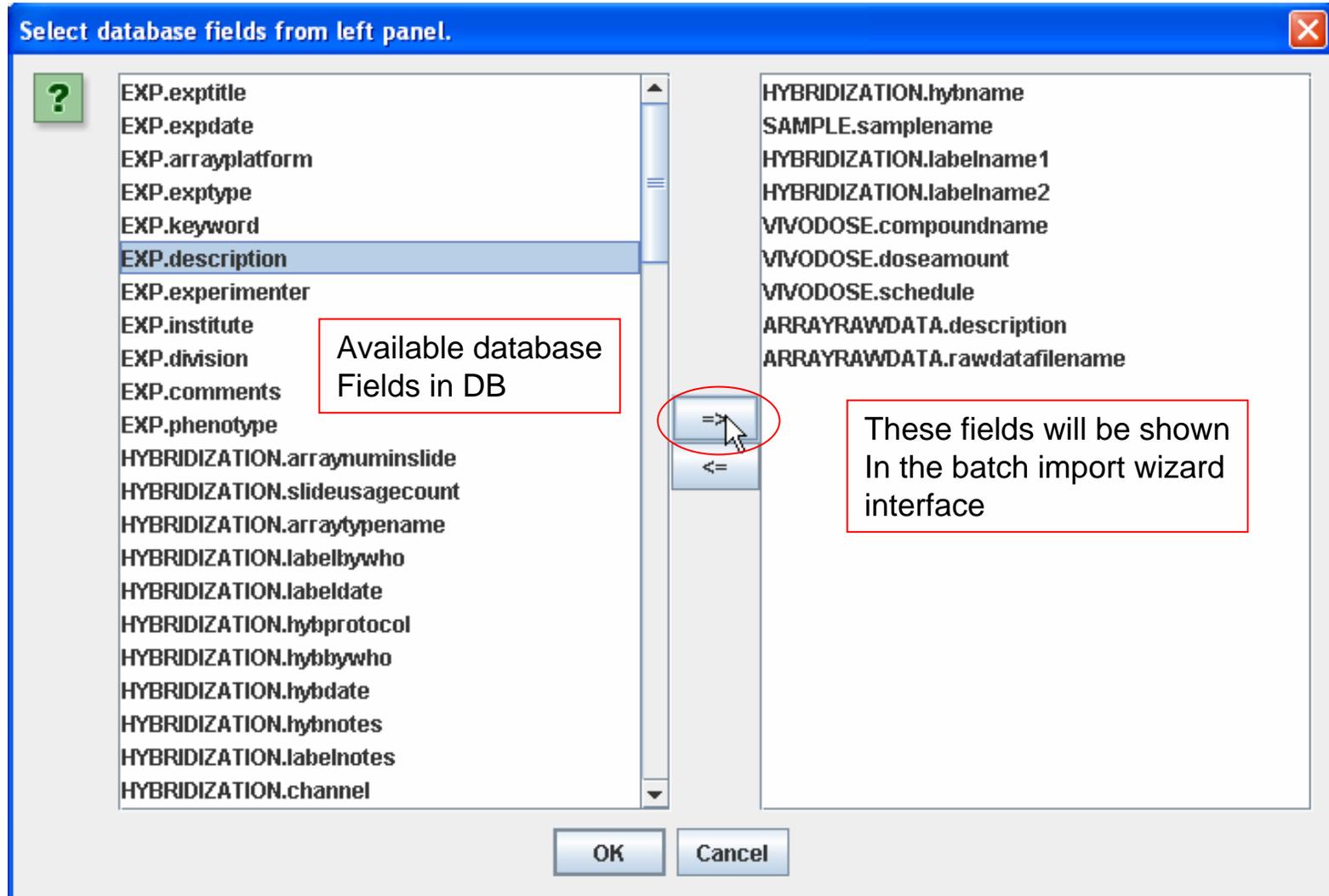
**Hybridization file**

< Back   Next >   Cancel   Help



# Batch Import Wizard - continued

Select the database fields in the left panel, click => to add in the right panel and this added field will be shown in the batch import interface.



# Batch Import Wizard - continued

Preview the table for hybridization assignment before import

**Batch Import**

**Preview Table For Hyb Assignment**

**Array Track**

HYBRIDIZATIO...	SAMPLE.SA...	HYBRIDIZA...	SAMPLE.SA...	HYBRIDIZA...	ARRAYRAW...	ARRAYRAWDATA...	VIVO
control_0_1	S0507	Cy5	reference	Cy3	raw data	2004-12-15-S050...	con
control_0_2	S0508	Cy5	reference	Cy3	raw data	2004-12-15-S050...	con
control_0_3	S0517	Cy5	reference	Cy3	raw data	2004-12-16-S051...	con
control_0_4	S0518	Cy5	reference	Cy3	raw data	2004-12-16-S051...	con
control_0_5	S0527	Cy5	reference	Cy3	raw data	2004-12-17-S052...	con
Drug1_250_1	S0513	Cy5	reference	Cy3	raw data	2004-12-15-S051...	Dru
Drug1_250_2	S0514	Cy5	reference	Cy3	raw data	2004-12-15-S051...	Dru
Drug1_250_3	S0523	Cy5	reference	Cy3	raw data	2004-12-16-S052...	Dru
Drug1_250_4	S0524	Cy5	reference	Cy3	raw data	2004-12-16-S052...	Dru
Drug1_250_5	S0550	Cy5	reference	Cy3	raw data	2004-12-17-S055...	Dru
Drug1_500_1	S0515	Cy5	reference	Cy3	raw data	2004-12-15-S051...	Dru
Drug1_500_2	S0516	Cy5	reference	Cy3	raw data	2004-12-15-S051...	Dru
Drug1_500_3	S0525	Cy5	reference	Cy3	raw data	2004-12-16-S052...	Dru
Drug1_500_4	S0526	Cy5	reference	Cy3	raw data	2004-12-16-S052...	Dru
Drug1_500_5	S0549	Cy5	reference	Cy3	raw data	2004-12-17-S054...	Dru

< Back   **Import**   Cancel   Help

# Batch Import Wizard - continued

Choose Columns

Please help to identify relevant columns in this sample data file:

Field Separator: <tab> Manually select first data row...

**Data file**

"Block"	"Column"	"Row"	"Name"	"ID"	"X"	"Y"	"Dia."	"F635 Median"	"F635 Mean"	"F635 SD"	"F635 CV"	"B635"	"B635 Median"	"B635 Mean"	"B635 SD"	"B635 CV"	"% > B635"
1	1	1	"glutama...	"AF0...	1220	14540	100	583	616	239	38	358	358	373	135	36	70
1	2	1	"axl rece...	"AF0...	1420	14530	120	835	855	304	35	364	364	377	135	35	91
1	3	1	"myosin i...	"NM...	1630	14530	120	1324	1610	804	49	355	355	367	135	36	99
1	4	1	"beta-fibr...	"M35...	1850	14530	120	2211	2248	621	27	349	349	378	186	49	100
1	5	1	"t-kinino...	"S48...	2030	14520	50	500	504	119	23	307	307	328	110	33	83
1	6	1	"organic ...	"NM...	2250	14530	120	307	316	121	38	317	317	327	133	40	10
1	7	1	"calpast...	"NM...	2470	14530	140	1636	1621	535	33	320	320	335	128	38	98
1	8	1	"amelog...	"U60...	2700	14520	100	490	502	192	38	311	311	330	135	40	65
1	9	1	"leukocyt	"M25	2890	14530	130	660	681	217	31	309	309	320	121	37	90

For each column description below, select a table column above and click the description button to make the assignment.

- Block Assigned to "Block" (1)
- Col Assigned to "Column" (2)
- Row Assigned to "Row" (3)
- Cy5 Log base: not logged Assigned to "F635 Median" (9)
- Background Cy5 Assigned to "B635 Median" (14)
- Cy3 Assigned to "F532 Median" (21)
- Background Cy3 Assigned to "B532 Median" (26)
- Notes
- Flag Assigned to "Flags" (54)
- P Value

**Database fields**

Map the columns in the data file to the corresponding DB fields by selecting the column in the data file, then clicking the DB field button

OK Cancel

# Batch Import Wizard - continued

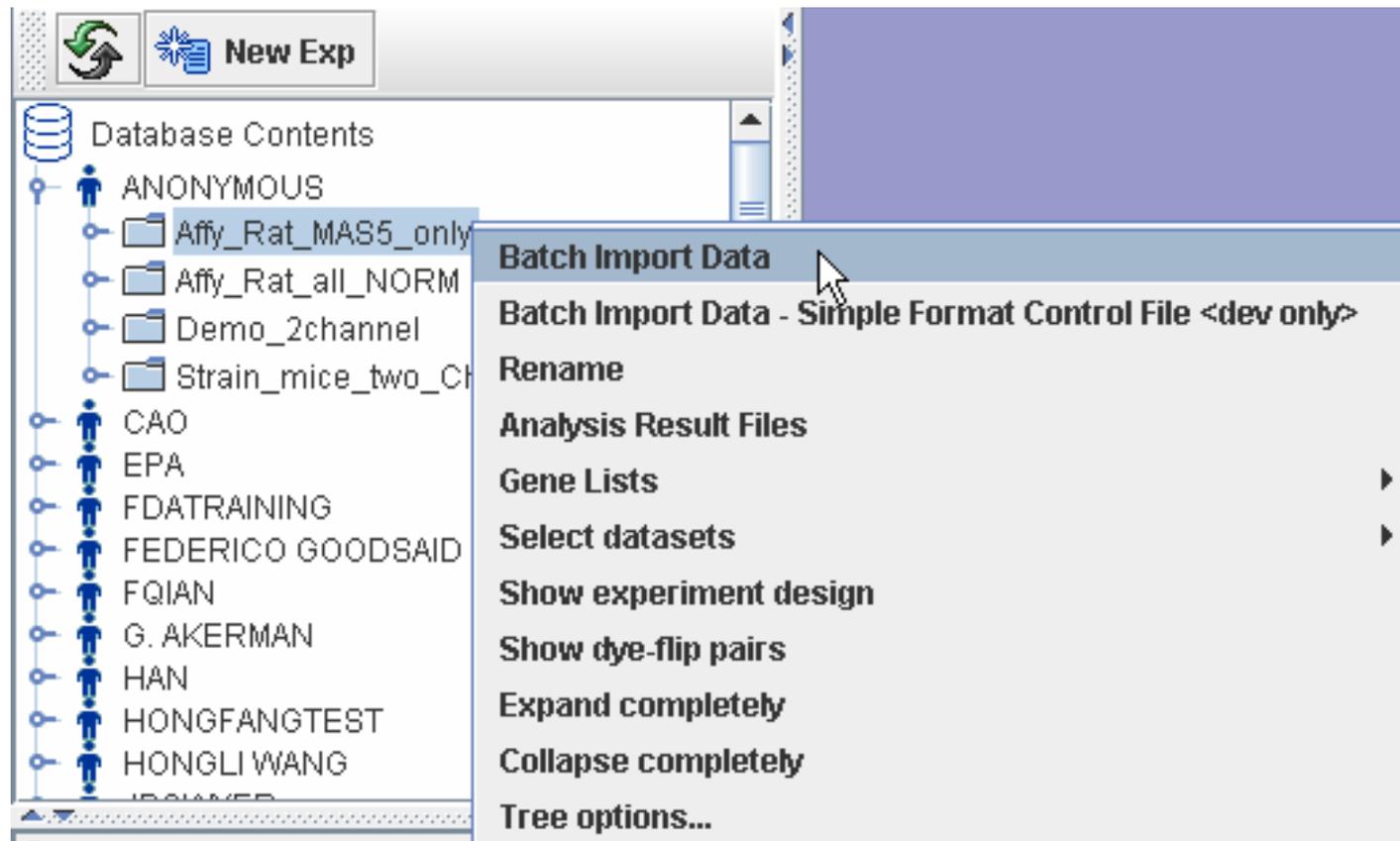
Database Contents

- ANONYMOUS
  - Affy\_Rat\_MAS5\_only
  - Affy\_Rat\_all\_NORM
  - Demo\_2channel
    - Drug1\_250\_1 S0513[Cy5] | reference[Cy3]
      - raw {Drug1\_250\_1} [file: 2004-12-15-S0513.gpr]
    - Drug1\_250\_2 S0514[Cy5] | reference[Cy3]
      - raw {Drug1\_250\_2} [file: 2004-12-15-S0514.gpr]
    - Drug1\_250\_3 S0523[Cy5] | reference[Cy3]
      - raw {Drug1\_250\_3} [file: 2004-12-16-S0523.gpr]
    - Drug1\_250\_4 S0524[Cy5] | reference[Cy3]
      - raw {Drug1\_250\_4} [file: 2004-12-16-S0524.gpr]
    - Drug1\_250\_5 S0550[Cy5] | reference[Cy3]
      - raw {Drug1\_250\_5} [file: 2004-12-17-S0550.gpr]
    - Drug1\_500\_1 S0515[Cy5] | reference[Cy3]
      - raw {Drug1\_500\_1} [file: 2004-12-15-S0515.gpr]
    - Drug1\_500\_2 S0516[Cy5] | reference[Cy3]
      - raw {Drug1\_500\_2} [file: 2004-12-15-S0516.gpr]
    - Drug1\_500\_3 S0525[Cy5] | reference[Cy3]
      - raw {Drug1\_500\_3} [file: 2004-12-16-S0525.gpr]
    - Drug1\_500\_4 S0526[Cy5] | reference[Cy3]
      - raw {Drug1\_500\_4} [file: 2004-12-16-S0526.gpr]
    - Drug1\_500\_5 S0549[Cy5] | reference[Cy3]
      - raw {Drug1\_500\_5} [file: 2004-12-17-S0549.gpr]
    - control\_0\_1 S0507[Cy5] | reference[Cy3]
      - raw {control\_0\_1} [file: 2004-12-15-S0507-rescan.gpr]
    - control\_0\_2 S0508[Cy5] | reference[Cy3]
      - raw {control\_0\_2} [file: 2004-12-15-S0508-rescan.gpr]
    - control\_0\_3 S0517[Cy5] | reference[Cy3]
      - raw {control\_0\_3} [file: 2004-12-16-S0517.gpr]
    - control\_0\_4 S0518[Cy5] | reference[Cy3]
      - raw {control\_0\_4} [file: 2004-12-16-S0518-rescan2.gpr]
    - control\_0\_5 S0527[Cy5] | reference[Cy3]
      - raw {control\_0\_5} [file: 2004-12-17-S0527-rescan.gpr]

**Data import is finished!**

# Batch Import Wizard

## Import CEL File Data



Right-click the experiment and choose “Batch Import Data”.

# Batch Import Wizard

## Import CEL File Data

**Batch Import**

**Array Track**

**Please Fill Out Experiment Information:**

If you have configuration file, load it to auto fill.

Experiment Name :

Array Type Name :

Species :  Human  Mouse  Rat  Others

Assay :  vivo  vitro

# Batch Import Wizard

## Import CEL File Data

**Batch Import**

**Map Hybridization Info File To Database**

Array Raw Data Files Root Directory :

Hybridization Information File :

Hybridization	SAMPLE 1	CELLTYPE 1	SEX 1	DEVSTAGE 1	Compound	Time	Dose	File Name	cell
D0_T12_B_a	D0_T12_B	Hepatocytes	Male	Postnatal	Compound C	12	0	B12Cd000A.CEL	
D0_T12_B_b	D0_T12_B	Hepatocytes	Male	Postnatal	Compound C	12	0	B12Cd000B.CEL	
D0_T12_C_a	D0_T12_C	Hepatocytes	Male	Postnatal	Compound C	12	0	C12Cd000A.CEL	
D0_T12_C_b	D0_T12_C	Hepatocytes	Male	Postnatal	Compound C	12	0	C12Cd000B.CEL	
D0_T12_D_a	D0_T12_D	Hepatocytes	Male	Postnatal	Compound C	12	0	D12Cd000A.CEL	

 For each column description below, select a table column above and click the description button to make the assignment. To add more database fields, click right button

	<input type="text" value="[HYBRIDIZATION] hyname"/>	<input type="button" value="X"/>	<input type="text" value="Hybridization"/>
	<input type="text" value="[SAMPLE] samplename for channel 1"/>	<input type="button" value="X"/>	<input type="text" value="SAMPLE 1"/>
	<input type="text" value="[HYBRIDIZATION] labelname for channel 1"/>	<input type="button" value="X"/>	<input type="text" value="Biotin"/>
<input type="checkbox"/>	<input type="text" value="[ARRAYRAWDATA] description"/>	<input type="button" value="X"/>	<input type="text" value="'Mas5'"/>
<input type="checkbox"/>	<input type="text" value="[ARRAYRAWDATA] celldatafilename"/>	<input type="button" value="X"/>	<input type="text" value="File Namecell"/> <input checked="" type="radio"/> Mas5 <input type="radio"/> RMA <input type="radio"/> Plier <input type="radio"/> Plier+
<input type="checkbox"/>	<input type="text" value="[ARRAYRAWDATA] probesetfilename (txt)"/>	<input type="button" value="X"/>	<input type="text"/>
<input type="checkbox"/>	<input type="text" value="[VIVODOSE] compoundname for channel 1"/>	<input type="button" value="X"/>	<input type="text" value="Compound"/>
<input type="checkbox"/>	<input type="text" value="[VIVODOSE] doseamount for channel 1"/>	<input type="button" value="X"/>	<input type="text" value="Dose"/>
<input type="checkbox"/>	<input type="text" value="[VIVODOSE] schedule for channel 1"/>	<input type="button" value="X"/>	<input type="text" value="Time"/>

# Batch Import Wizard

## Import CEL File Data

Batch Import

**Preview Table For Hyb Assignment**

Array Track

HYBRIDIZATI...	SAMPLE_SA...	HYBRI...	ARRAY...	ARRAYRAWDATA...	VIVODOSE.C...	VIVODOSE...	VIVODO... ▲
D0_T12_B_a	D0_T12_B	Biotin	Mas5	B12Cd000A.CEL	Compound C	0	12
D0_T12_B_b	D0_T12_B	Biotin	Mas5	B12Cd000B.CEL	Compound C	0	12
D0_T12_C_a	D0_T12_C	Biotin	Mas5	C12Cd000A.CEL	Compound C	0	12
D0_T12_C_b	D0_T12_C	Biotin	Mas5	C12Cd000B.CEL	Compound C	0	12
D0_T12_D_a	D0_T12_D	Biotin	Mas5	D12Cd000A.CEL	Compound C	0	12
D0_T12_D_b	D0_T12_D	Biotin	Mas5	D12Cd000B.CEL	Compound C	0	12
D2_T12_B_a	D2_T12_B	Biotin	Mas5	B12Cd2.00A.CEL	Compound C	2	12
D2_T12_B_b	D2_T12_B	Biotin	Mas5	B12Cd2.00B.CEL	Compound C	2	12
D2_T12_C_a	D2_T12_C	Biotin	Mas5	C12Cd2.00A.CEL	Compound C	2	12
D2_T12_C_b	D2_T12_C	Biotin	Mas5	C12Cd2.00B.CEL	Compound C	2	12
D2_T12_D_a	D2_T12_D	Biotin	Mas5	D12Cd2.00A.CEL	Compound C	2	12
D2_T12_D_b	D2_T12_D	Biotin	Mas5	D12Cd2.00B.CEL	Compound C	2	12

< Back Import Cancel Help

# Batch Import Wizard

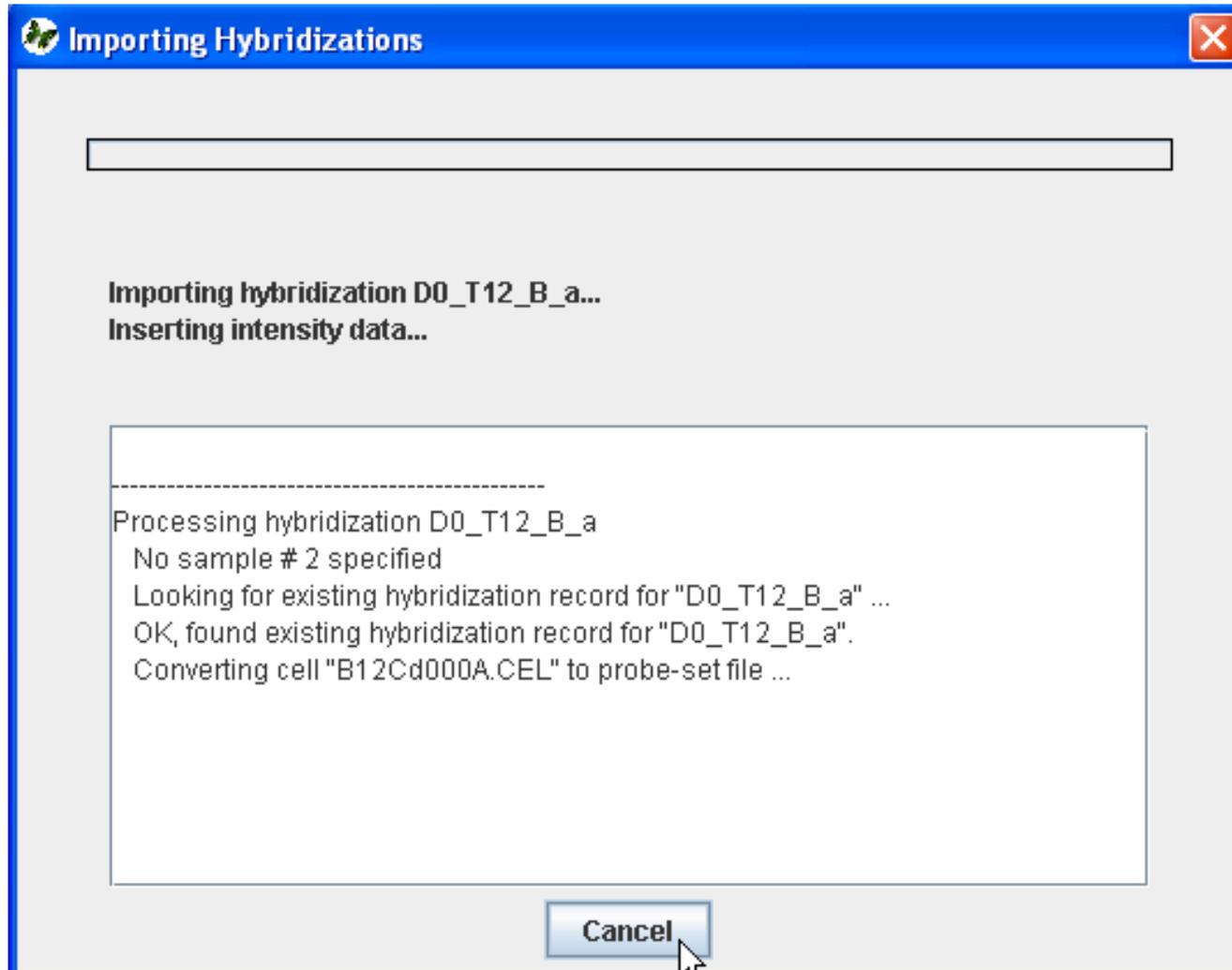
## Import CEL File Data



Click Yes to save the configuration file which stores all the assignments, making it easier for the second time import in case of the first import failure.

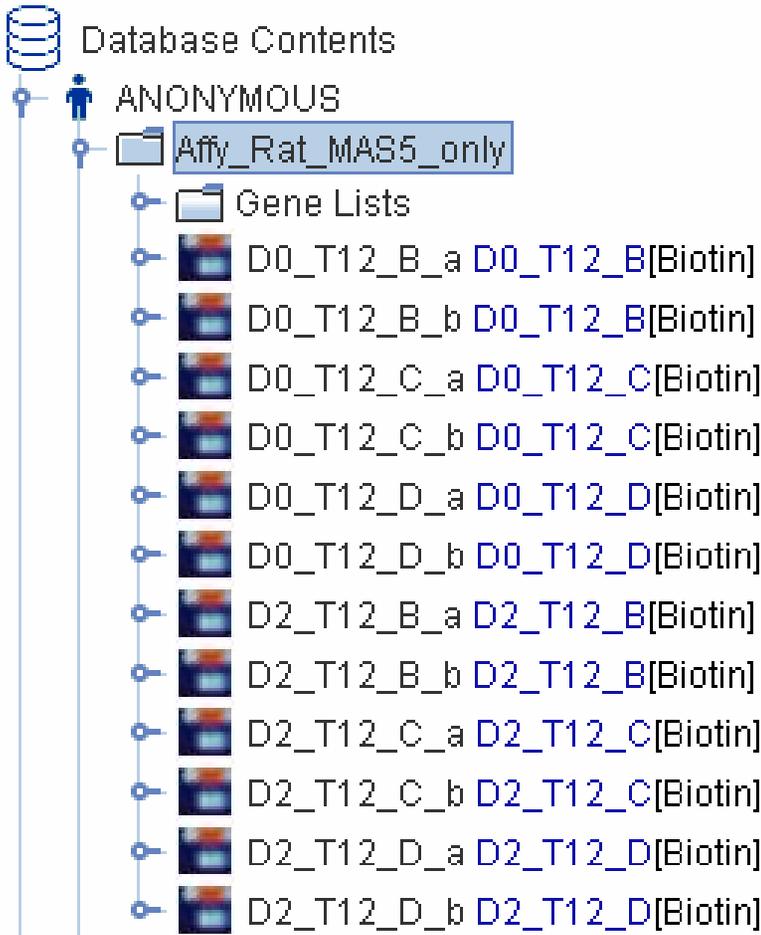
# Batch Import Wizard

## Import CEL File Data



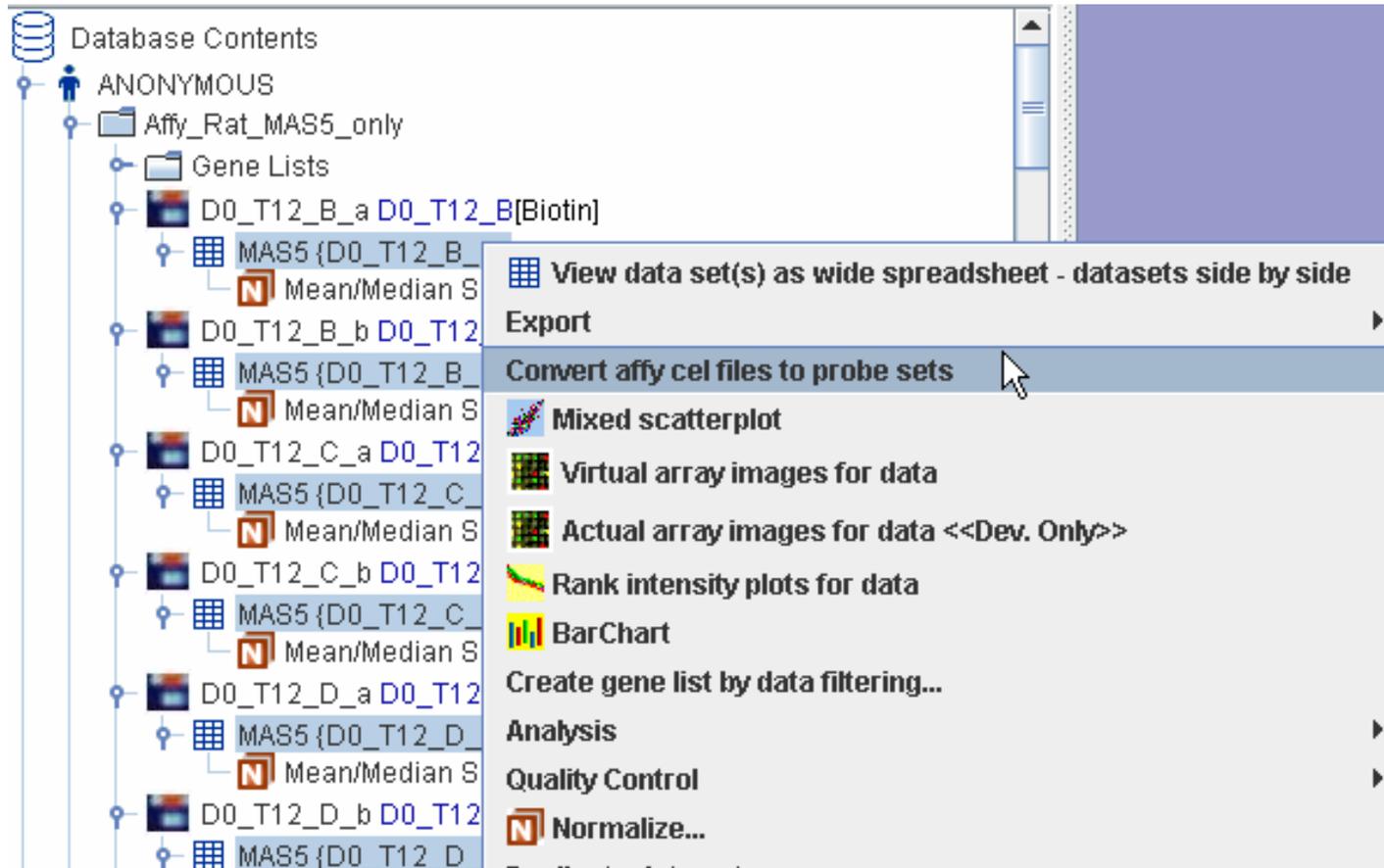
# Batch Import Wizard

## Import CEL File Data



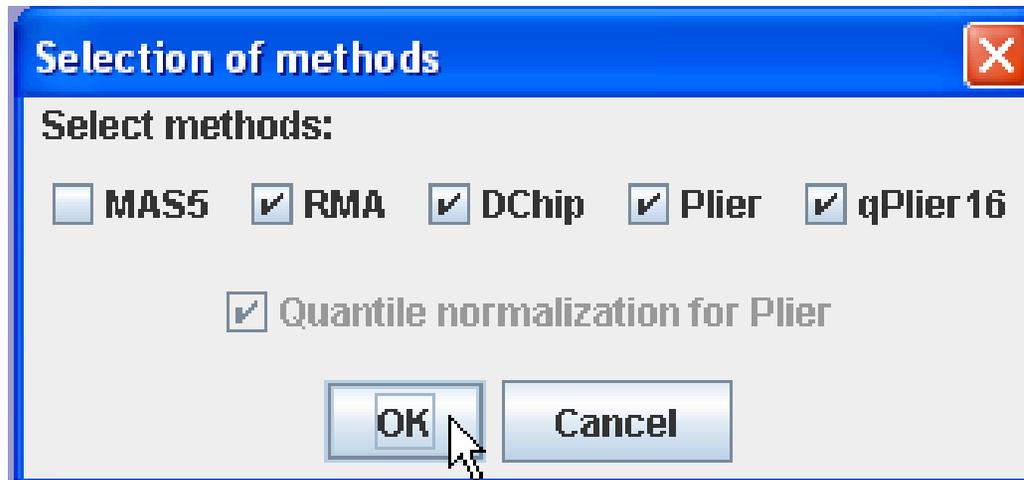
Import is finished.

# Convert CEL file to probe set file

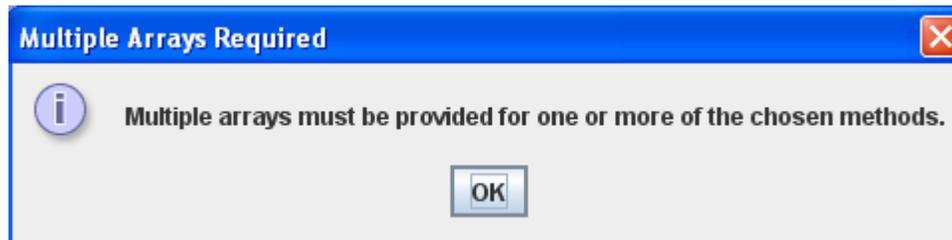


# Convert CEL file to probe set file

Select the data, right-click, choose “Convert affy cel files to probe sets. ArrayTrack can convert cel to RMA, DChip, qPlier and qPlier16



Keep in mind that you have to choose multiple data to do the converting of all the methods except MAS5. For example, you can not select just one data to do RMA converting, otherwise you will get the following message:



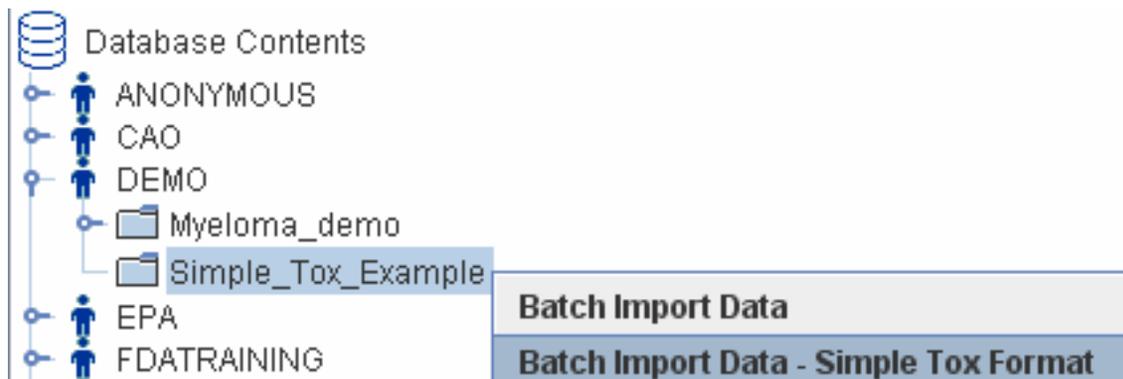
# Convert CEL file to probe set file

MAQC\_MAIN

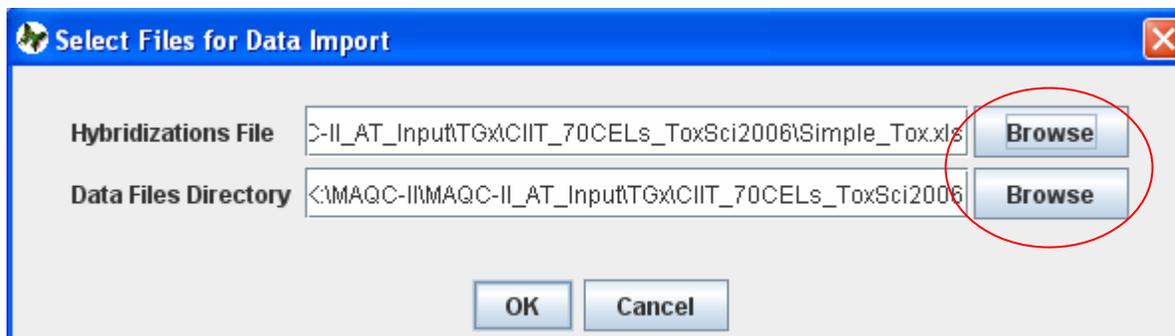
- ABI\_1
- ABI\_2
- ABI\_3
- AFX\_1
  - Gene Lists
  - AFX\_1\_A1 A[Biotin]
    - MAS5\_Scale500 {AFX\_1\_A1}
    - RMA {AFX\_1\_A1}
    - DChip {AFX\_1\_A1}
    - MAS5\_NoneScale {AFX\_1\_A1}
    - qPlier16 {AFX\_1\_A1}

Converted probe set files

# Batch Import – Simple Tox Format



Right-click the experiment, choose “Batch Import Data – Simple Tox Format”



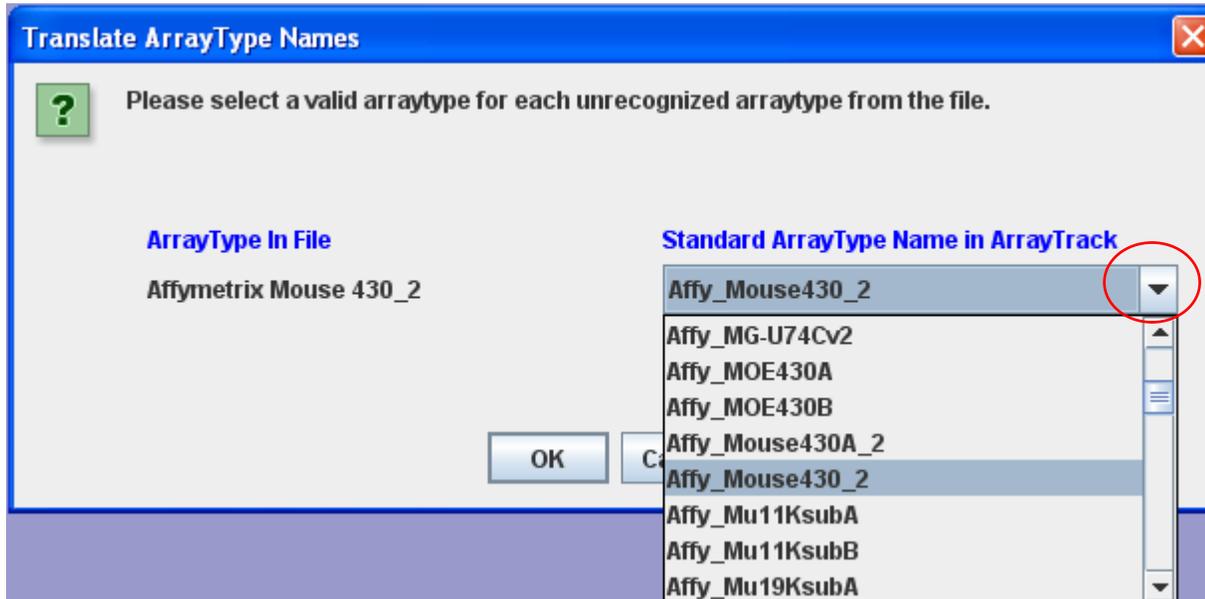
Click Browse button to locate the hybridization file and data files directory. The next slide shows an example of hybridization table file.

# Batch Import – Simple Tox Format

This is an example of hybridization table file (which is attached). User can use this as a template to make his own hybridization file. Just make sure that your column titles are exactly same as this template.

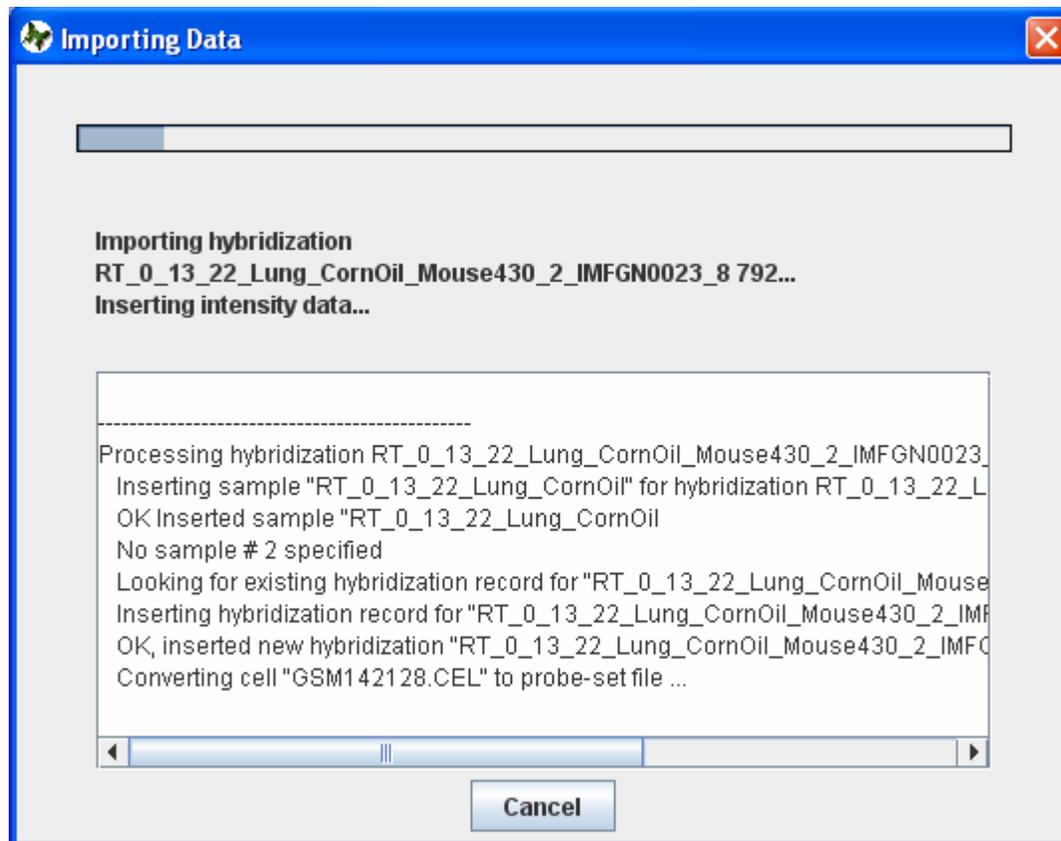
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Array_ID	Animal_ID	Institutio	StudyTitle	StudyTy	Class	ClassCo	ClassRe	Compo	CAS	Control	Treatme	Dose	DoseUn	HybName	Sample
2	792	1-22	Hamner	Hamner Mice Lung	Repeat Dose				Corn Oil		Y	C	0 mg/kg		RT_0_13_22_Lu	RT_0_13
3	793	1-24	Hamner	Hamner Mice Lung	Repeat Dose				Corn Oil		Y	C	0 mg/kg		RT_0_13_24_Lu	RT_0_13
4	794	1-25	Hamner	Hamner Mice Lung	Repeat Dose				Corn Oil		Y	C	0 mg/kg		RT_0_13_25_Lu	RT_0_13
5	795	1-27	Hamner	Hamner Mice Lung	Repeat Dose				Rodent Chow		Y	C	0 ppm		RT_0_13_27_Lu	RT_0_13
6	796	1-28	Hamner	Hamner Mice Lung	Repeat Dose				Rodent Chow		Y	C	0 ppm		RT_0_13_28_Lu	RT_0_13
7	797	1-29	Hamner	Hamner Mice Lung	Repeat Dose				Rodent Chow		Y	C	0 ppm		RT_0_13_29_Lu	RT_0_13
8	798	1-12	Hamner	Hamner Mice Lung	Repeat Do	Non-lungtu	NLT	NTP_No16	N-(1-napht	1465-25-4	N	NLT	2000 ppm		RT_2000_13_12	RT_2000
9	799	1-14	Hamner	Hamner Mice Lung	Repeat Do	Non-lungtu	NLT	NTP_No16	N-(1-napht	1465-25-4	N	NLT	2000 ppm		RT_2000_13_14	RT_2000
10	800	1-15	Hamner	Hamner Mice Lung	Repeat Do	Non-lungtu	NLT	NTP_No16	N-(1-napht	1465-25-4	N	NLT	2000 ppm		RT_2000_13_15	RT_2000
11	801	1-2	Hamner	Hamner Mice Lung	Repeat Do	lung tumor	LT	NTP_No14	1,5-Napht	2243-62-1	N	LT	2000 ppm		RT_2000_13_2	RT_2000

# Batch Import – Simple Tox Format

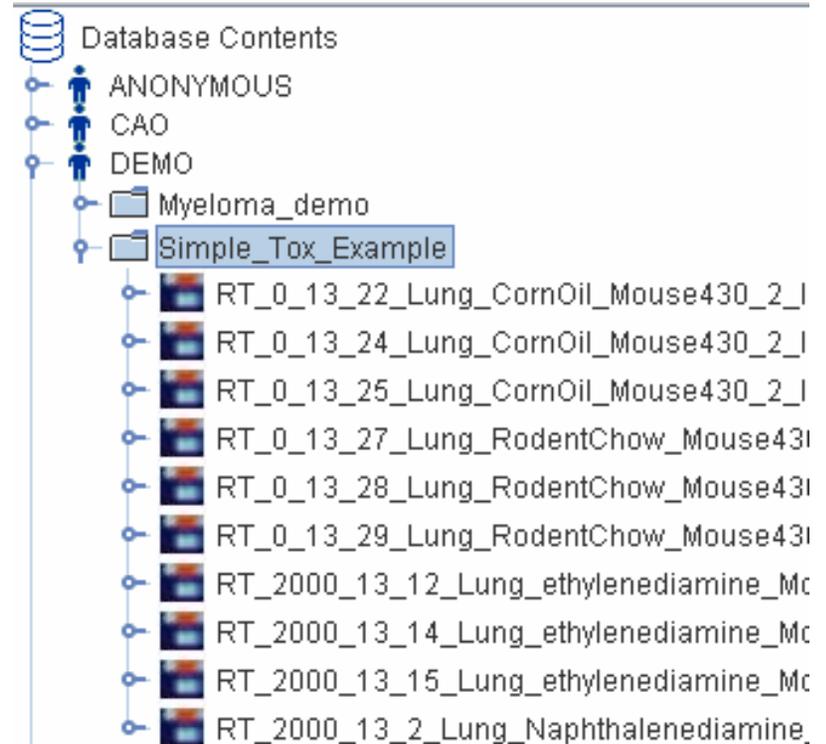
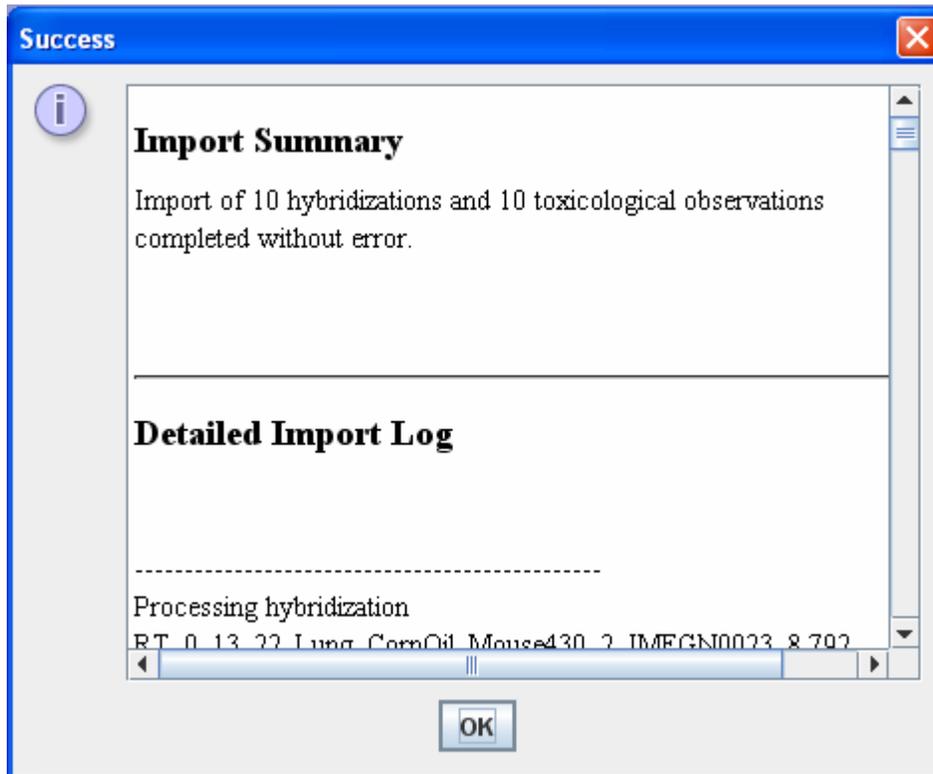


Select the right array type.

# Batch Import – Simple Tox Format



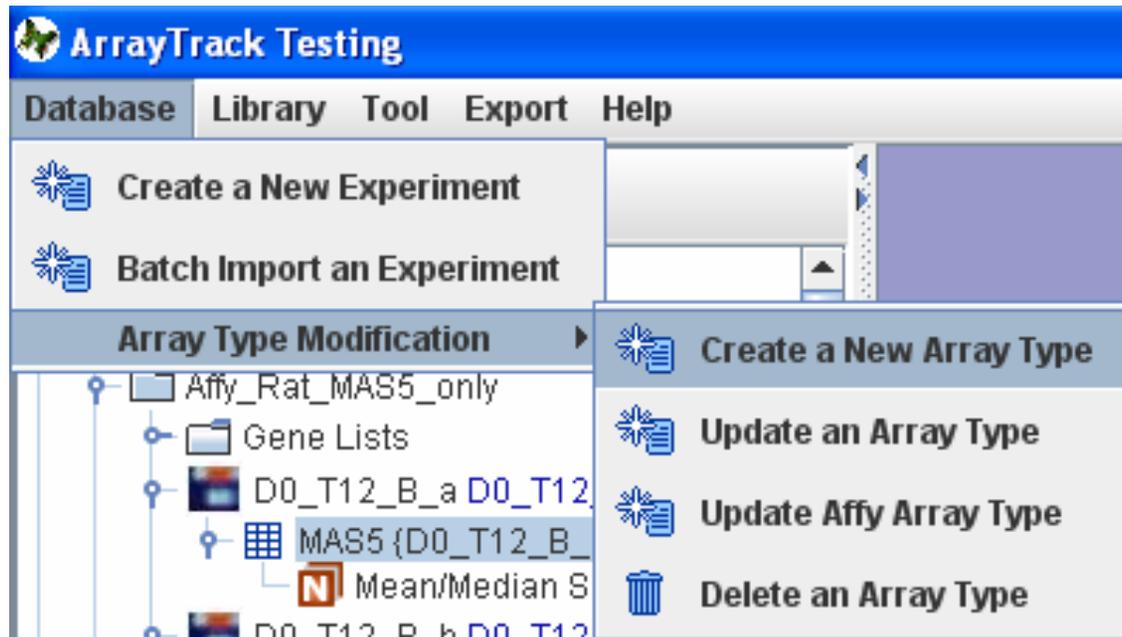
# Batch Import – Simple Tox Format



Batch import is successfully finished.

# Create Array Type

Click “Database” menu, choose “Array Type Modification” ->”Create a new Array Type”.



# Create Array Type - continued

The following is the example of creating array type for 2 channel data.

**Import ArrayType**

**Please Fill Out ArrayType Information:**

If you have configuration file, load it to auto fill.

Array Type File:

Array Type Name:

Manufacture Name:

Array Reporter Type:  Oligonucleotide (sequence)  cDNA array

Channels:  one channel  two channels

Category:  Human  Mouse  Rat  Others

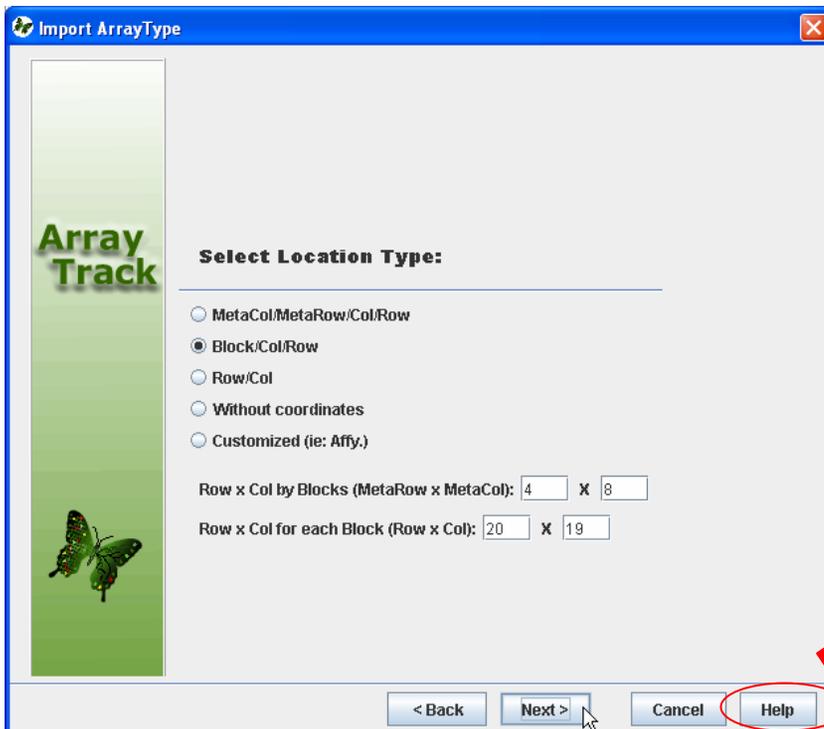
Ignore this button If you don't have configuration file. You will be asked to save the configuration file later, see slide #28

Locate the array type file

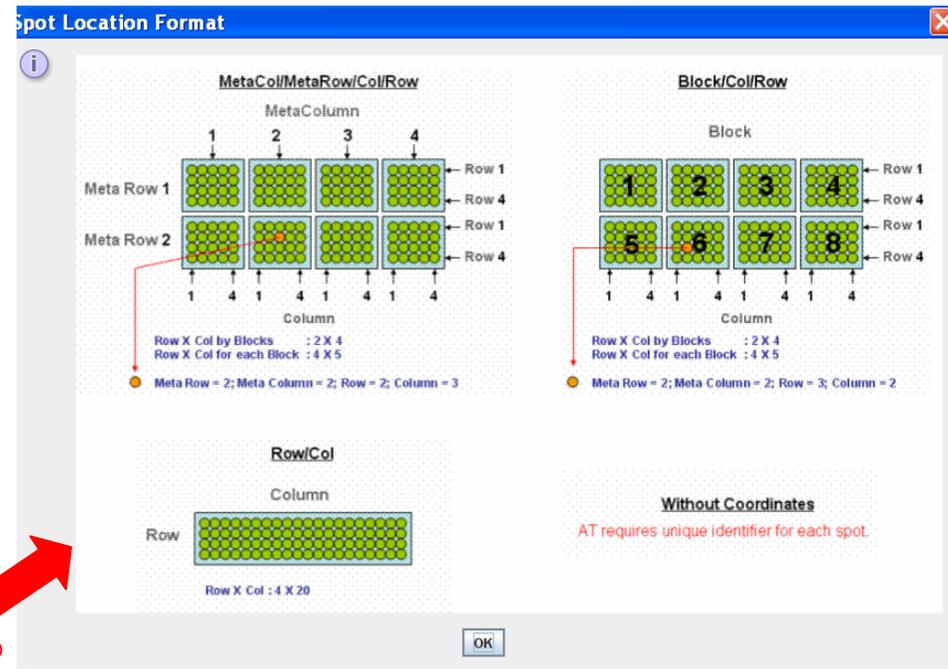
Type in the array type name

Choose the manufacture name

# Create Array Type - continued



Click Help button to view the location definition



# Create Array Type - continued

**Import ArrayType**

**Mapping Array Info From File to Database:**

Field Separator:

Block	Row	Column	ID	Name
1	1	1	AF027331_1	glutamate receptor subunit kainate subtype gl...
1	1	2	AF046886_1	axl receptor tyrosine kinase
1	1	3	NM_053986_1	myosin i heavy chain my...
1	1	4	M35602_1	beta-fibrinogen
1	1	5	S48142_1	t-kininogenase
1	1	6	NM_030838_1	organic anion transportin
1	1	7	NM_053295_1	calpastatin probably mul...
1	1	8	U60562_1	amelogenin
1	1	9	M25823_1	leukocyte-common antig...
1	1	10	NM_053655_1	dynamin-like protein dlp1 isoform dlp1-3

For each column description below, select a table column above and click the description button to make the assignment.

<input checked="" type="checkbox"/>	Block	<input type="button" value="x"/>	Assigned to Block (1)
<input checked="" type="checkbox"/>	Col	<input type="button" value="x"/>	Assigned to Column (3)
<input checked="" type="checkbox"/>	Row	<input type="button" value="x"/>	Assigned to Row (2)
<input checked="" type="checkbox"/>	Manufacture ID	<input type="button" value="x"/>	Assigned to ID (4)
<input checked="" type="checkbox"/>	Manufacture Descr	<input type="button" value="x"/>	Assigned to Name (5)
<input checked="" type="checkbox"/>	GeneBankAcc	<input type="button" value="x"/>	Assigned to ID (4)
<input checked="" type="checkbox"/>	Reference Sequence	<input type="button" value="x"/>	Assigned to ID (4)
<input type="checkbox"/>	GeneName	<input type="button" value="x"/>	

Column header

For each column description, select a table column above and click the description button to make the assignment

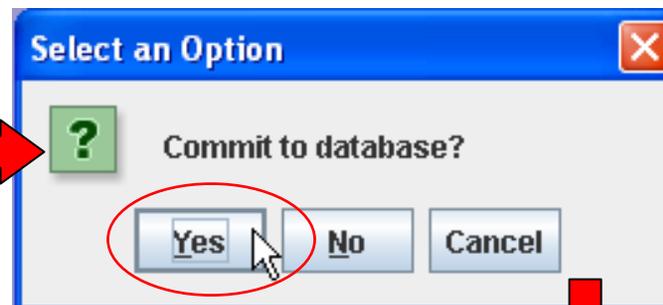
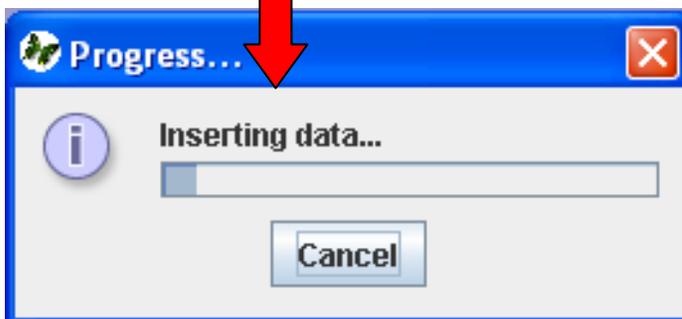
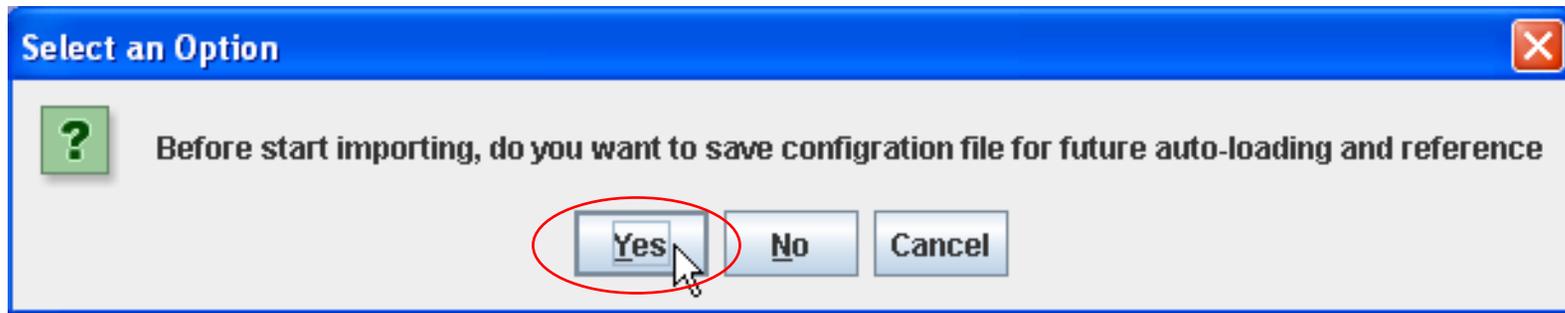
For each column description below, select a table column above and click the description button to make the assignment.

Description buttons

Click Import button

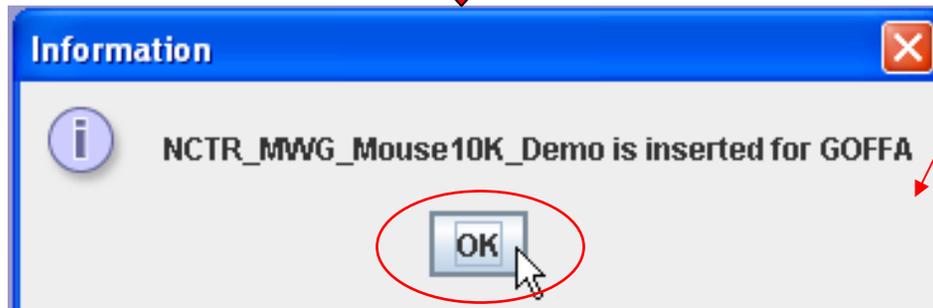
# Create Array Type - continued

The user can save configuration file for future auto-loading and reference. Click Yes button to save to the local drive.



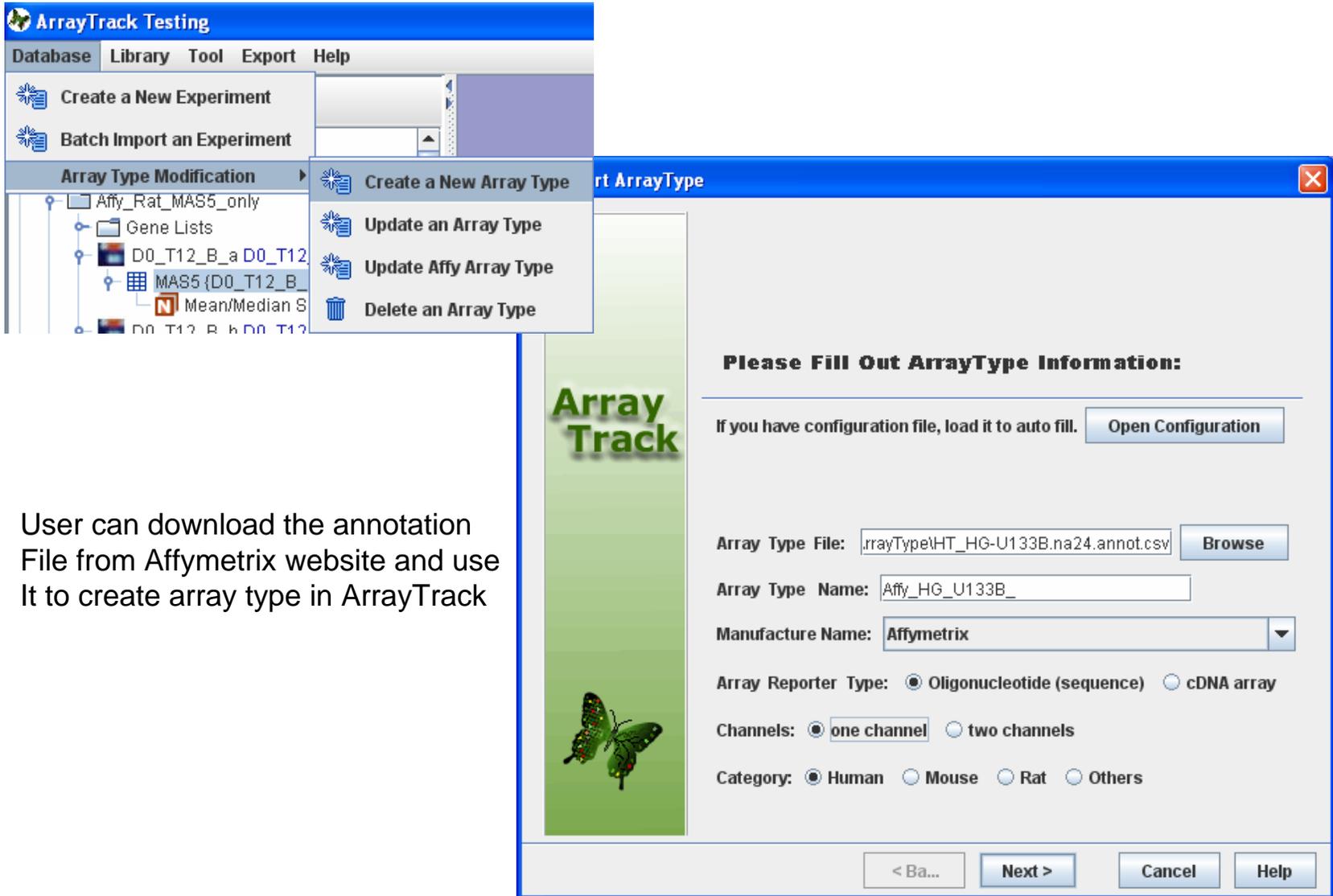
Annotate gene informations for this arraytype... please be patient

# Create Array Type - continued



The array type is created successfully.

# Create Array Type - Affymetrix



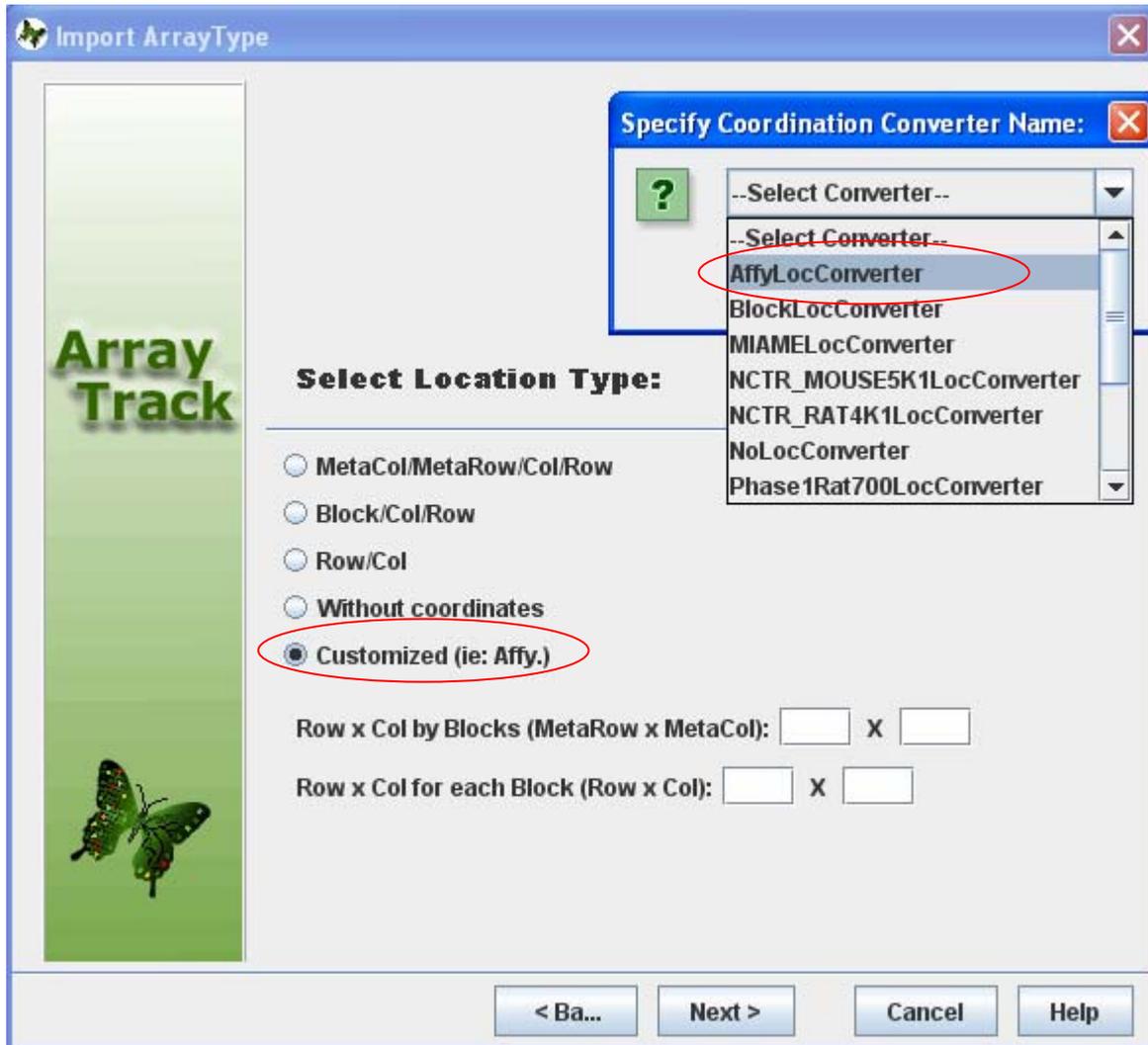
The screenshot shows the 'ArrayTrack Testing' application interface. The 'Array Type Modification' menu is open, displaying options: 'Create a New Array Type', 'Update an Array Type', 'Update Affy Array Type', and 'Delete an Array Type'. The 'Create a New Array Type' dialog box is active, titled 'Create ArrayType'. It contains the following fields and options:

- Please Fill Out ArrayType Information:**
- If you have configuration file, load it to auto fill.
- Array Type File:
- Array Type Name:
- Manufacture Name:
- Array Reporter Type:  Oligonucleotide (sequence)  cDNA array
- Channels:  one channel  two channels
- Category:  Human  Mouse  Rat  Others

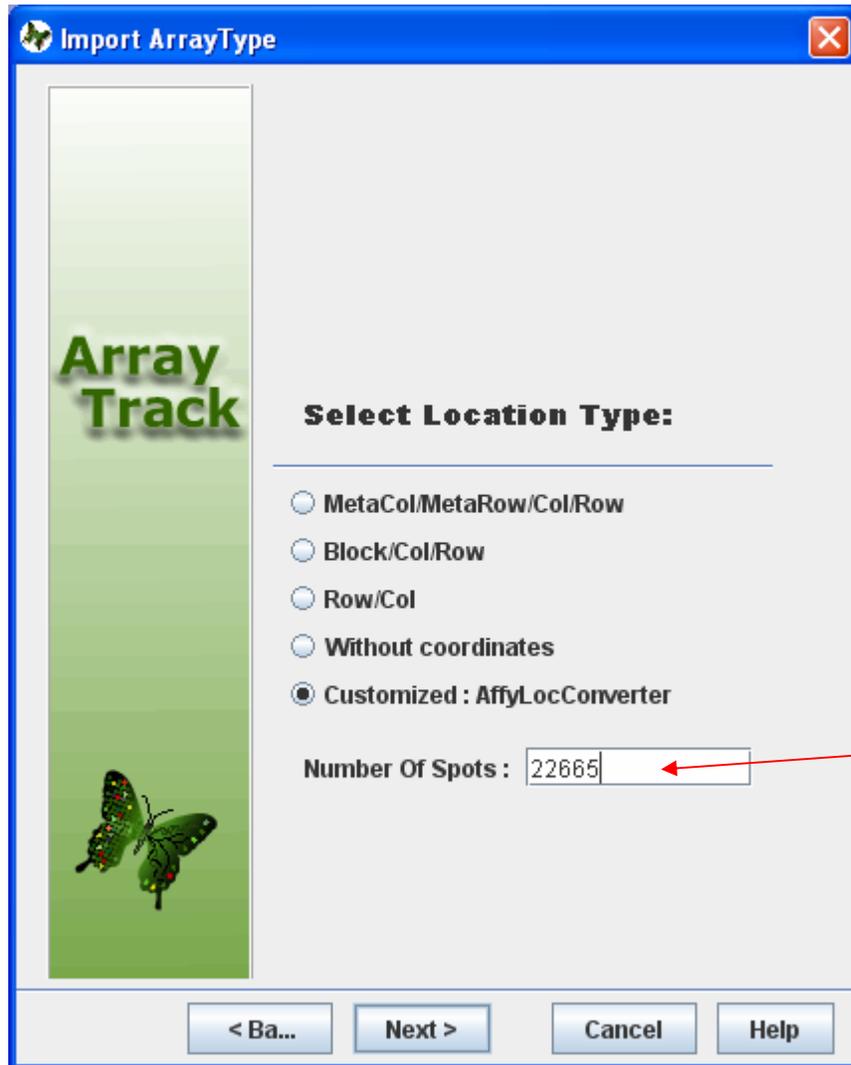
At the bottom of the dialog are buttons for '< Ba...', 'Next >', 'Cancel', and 'Help'.

User can download the annotation File from Affymetrix website and use it to create array type in ArrayTrack

# Create Array Type - Affymetrix



# Create Array Type - Affymetrix



Specify the number of spots

# Create Array Type - Affymetrix

Map the columns to database fields then click "Import" button.  
See the next slide for the suggested mapping fields.

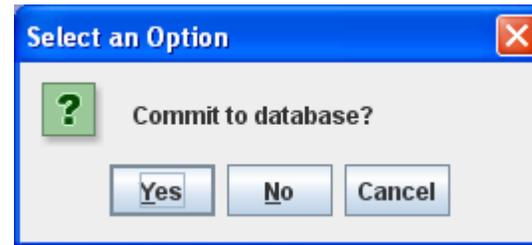
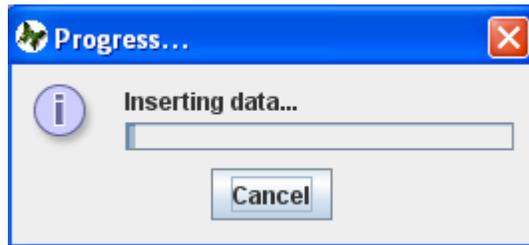
The screenshot shows the 'Import ArrayType' dialog box with the following components:

- Field Separator:** A dropdown menu set to '<tab>' and a button labeled 'Manually select first data row...'.
- Table:** A table with 9 columns: Probe Set ID, GeneChip Array, Sequence Type, Species Scientific Name, Annotation Date, Sequence Source, Transcript ID(Array Design), Target Description, and Representative Public. It contains 6 rows of data.
- Instructions:** A red arrow points to the text: 'For each column description below, select a table column above and click the description button to make the assignment.'
- Mapping List:** A list of database fields with checkboxes and assignment counts:
  - Manufacture ID [X] Assigned to Probe Set ID (1)
  - Manufacture Descr [X] Assigned to Target Description (8)
  - GeneBankAcc [X] Assigned to Representative Public ID (9)
  - Reference Sequence [X]
  - GeneName [X] Assigned to Gene Symbol (15)
  - LocusID [X]
  - UnigenelD [X] Assigned to UniGene ID (11)
  - SwissProtAcc [X] Assigned to SwissProt (20)
- Buttons:** '< Ba...', 'Import', 'Cancel', and 'Help'.

We suggest mapping the columns of Affy annotation file to the ArrayTrack fields according the following table.

Affymetrix annotation CSV file	ArrayTrack chip fields
Probe Set ID	GEN_ID_MFR
Strand	<b>Strand</b>
Sart	<b>BP_Start</b>
End	<b>BP_End</b>
Target Description	GEN_DESCR_MFR
Representative Public ID	GENEBANKACC
Archival UniGene Cluster	
UniGene ID	UNIGENEID
Genome Version	
Alignments	
Gene Title	DESCRIPTION
Gene Symbol	GENENAME
Chromosomal Location	<b>CHROMLOCATION</b>
Unigene Cluster Type	
Ensembl	<b>Ensembl</b>
Entrez Gene	LOCUSID
SwissProt	SWISSPROT_ACC_NUMBER
EC	
OMIM	OMIM
RefSeq Protein ID	<b>PROTEIN_REFSEQ</b>
RefSeq Transcript ID	REFSEQ
Gene Ontology Biological Process	BIOLOGICAL_PROCESS
Gene Ontology Cellular Component	CELULAR_COMPONENT
Gene Ontology Molecular Function	MOLECULAR_FUNCTION
Pathway	<b>PATHWAYS</b>

# Create Array Type - Affymetrix



Annotate gene informations for this arraytype.... please be patient

