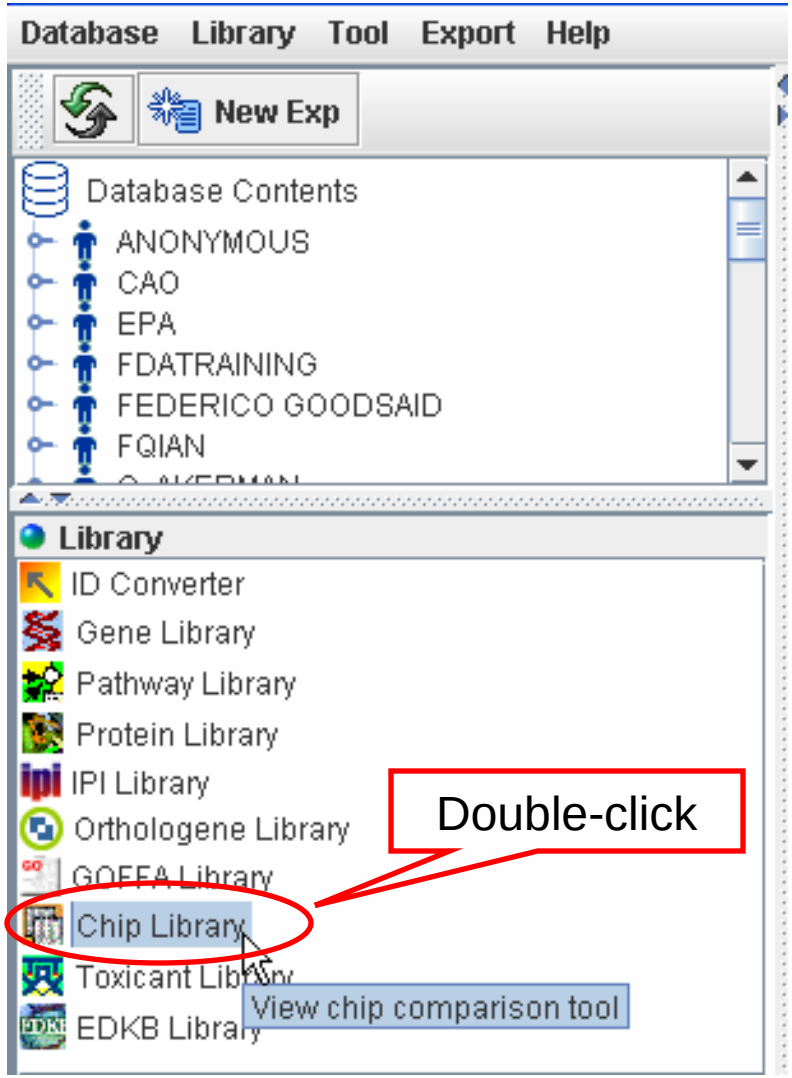
Array Track

Tutorial 9A: Data Import

Data Import

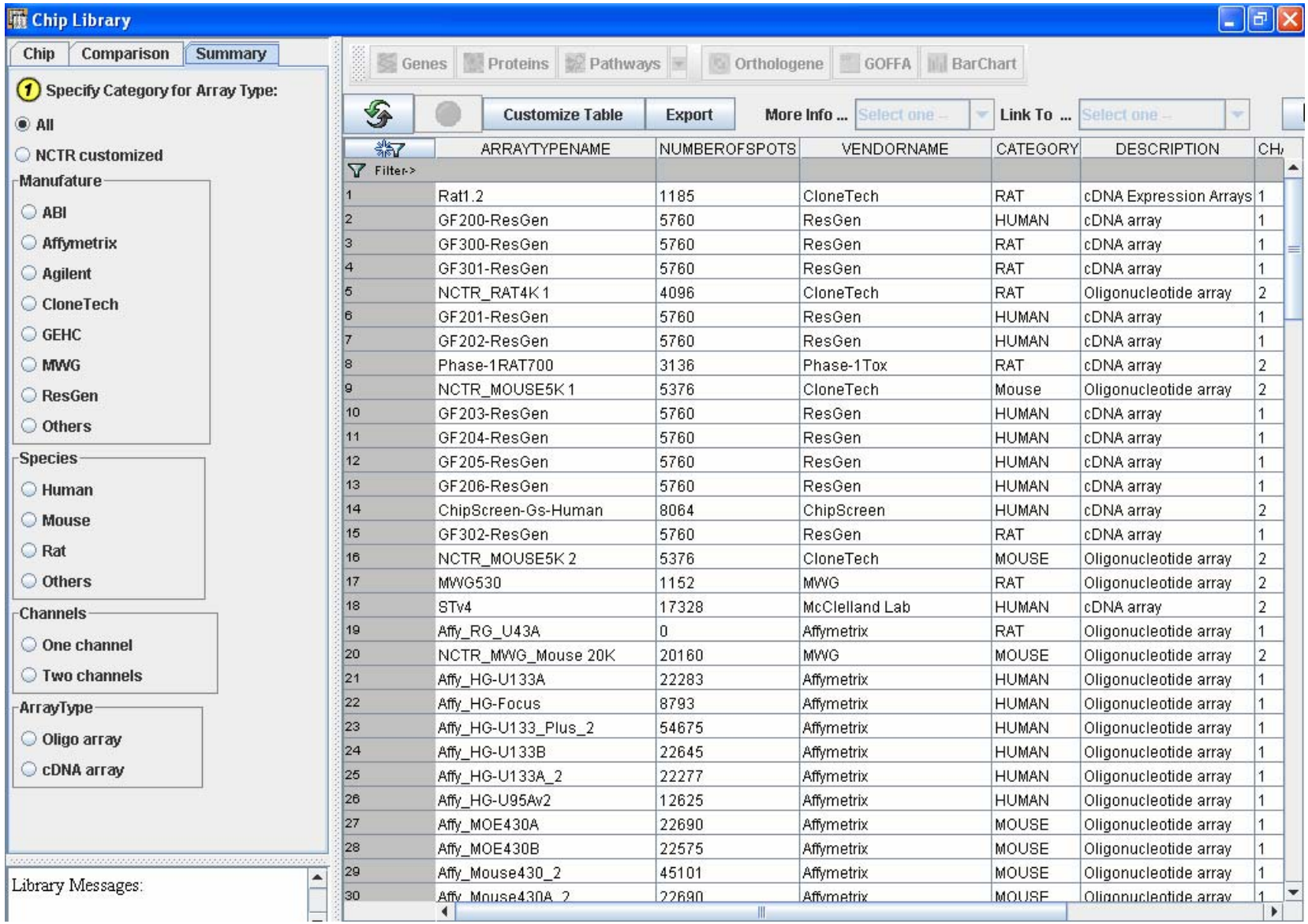
Before you import any data, you need to create array type unless the array type is already exist in ArrayTrack (you can check all the existing array type from Chip Library). You also need to create hybridization file. The following slides will lead you through all the steps.

Check Existing Array Type



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

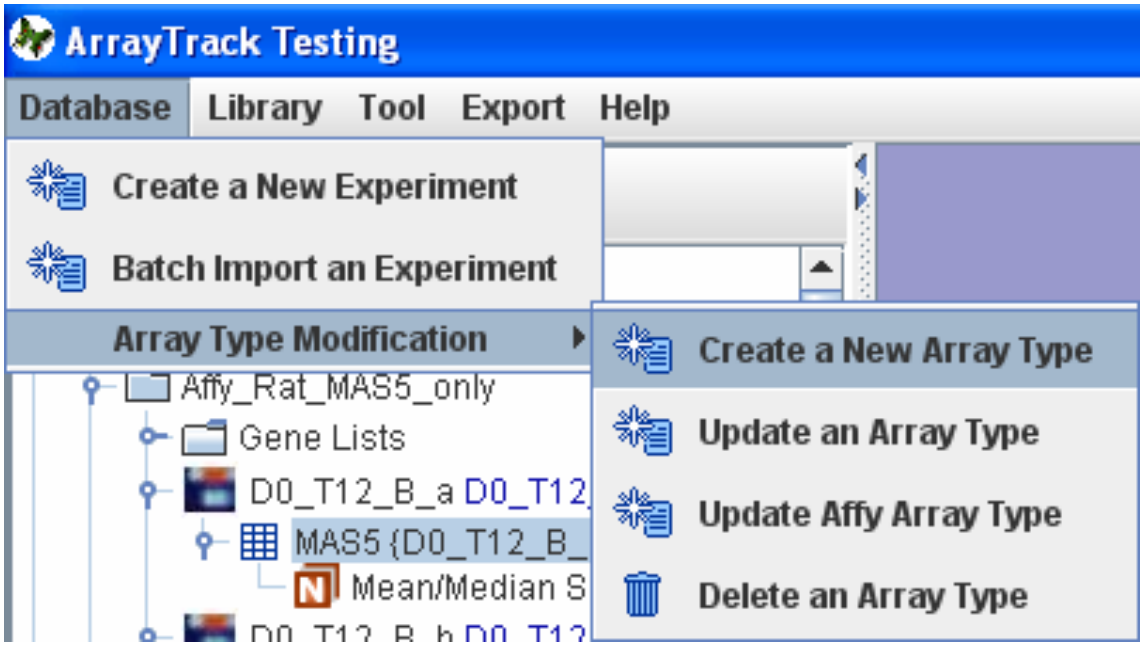
Check Existing Array Type



Create Array Type

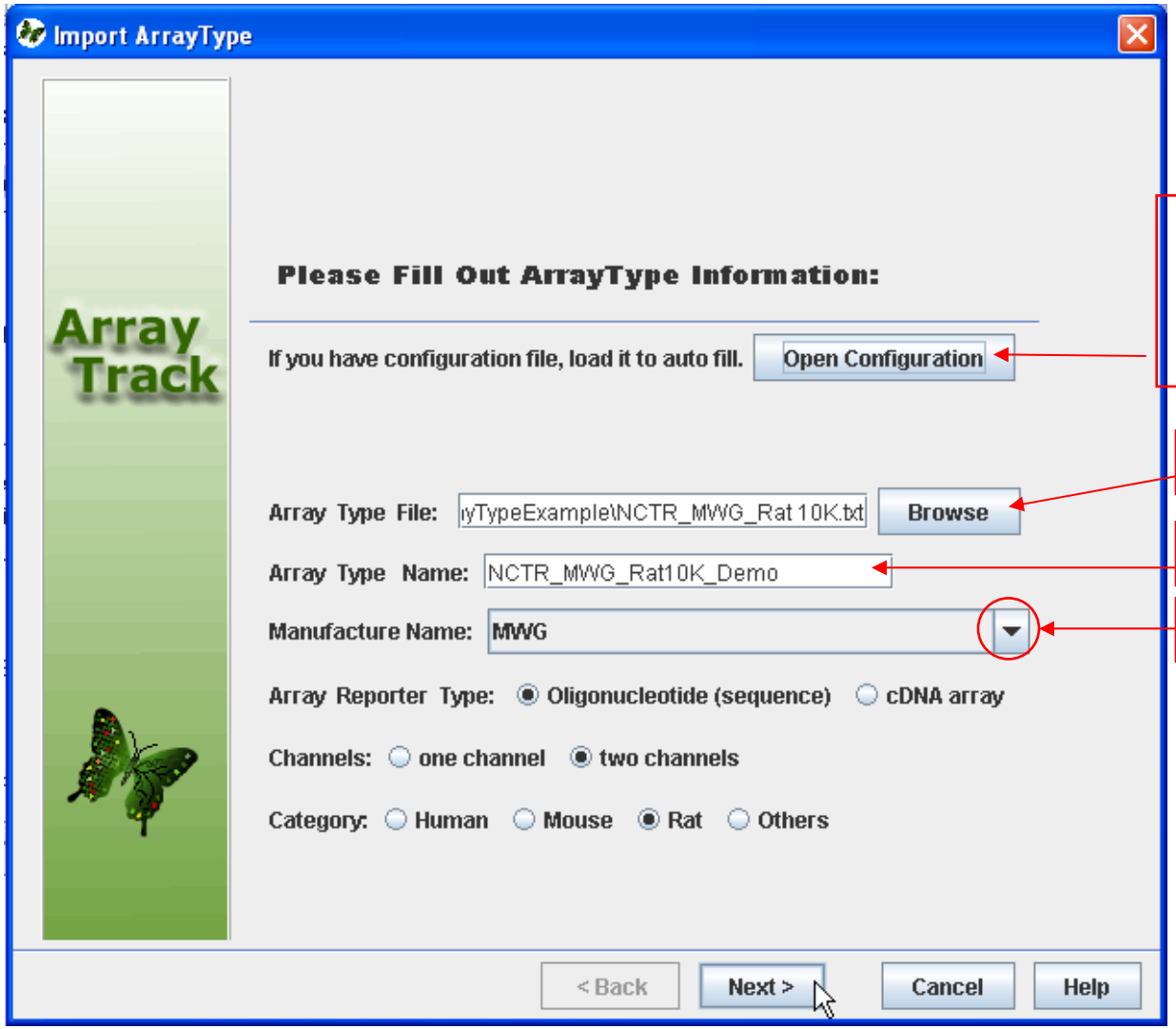
If the array type for your data (e.g. Affy_HG-U133_Plus_2) already exists in ArrayTrack, you can skip this step and go to slide #11(create experiment).

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.



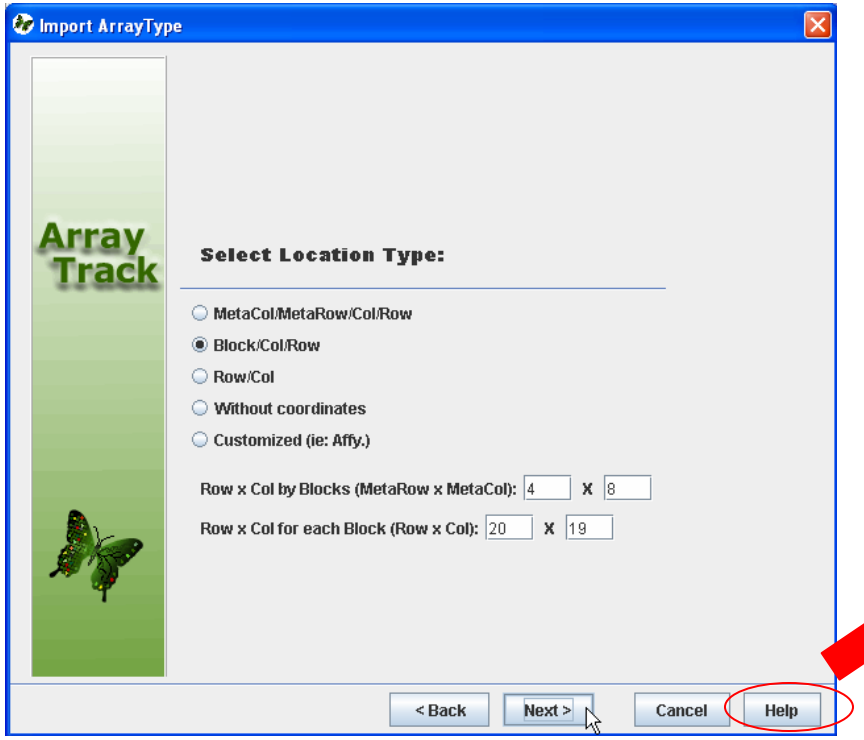
Ignore this button If you don't have configuration file. You will be asked to save the configuration file later, see slide #26

Locate the array type file

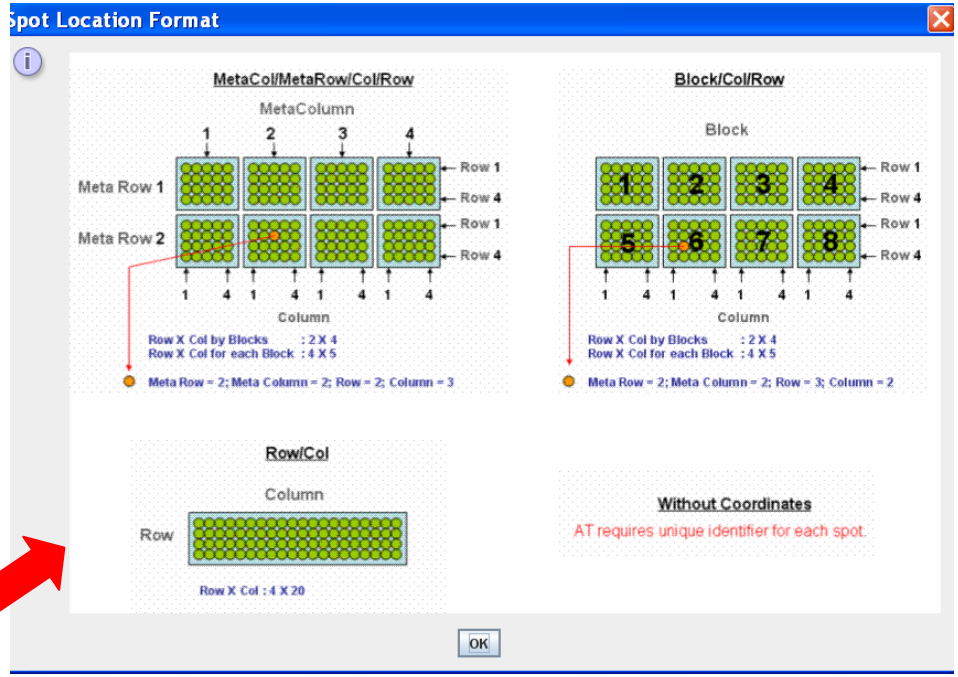
Type in the array type name

Choose the manufacture name

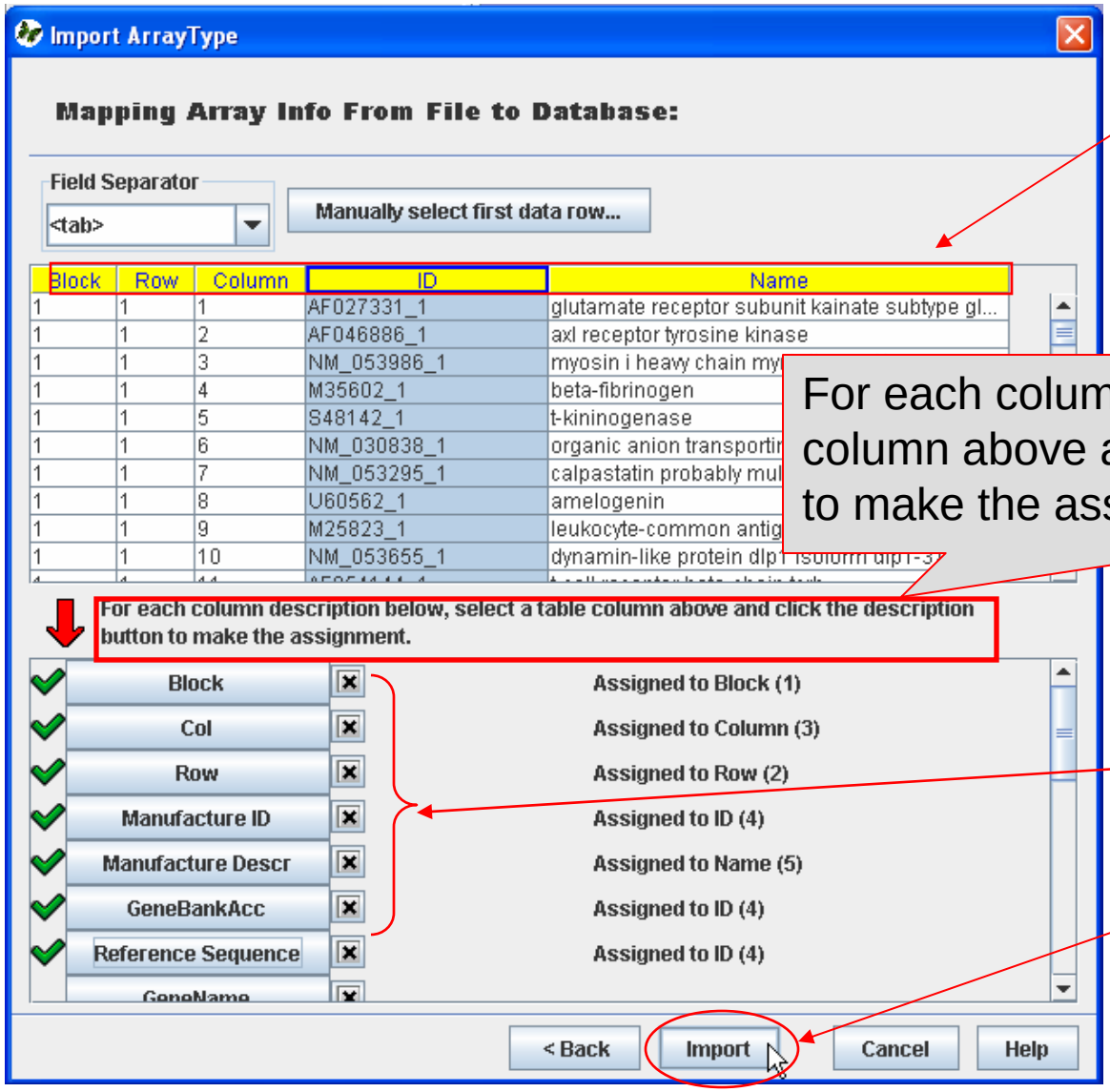
Create Array Type - continued



Click Help button to view the location definition



Create Array Type - continued



Column header

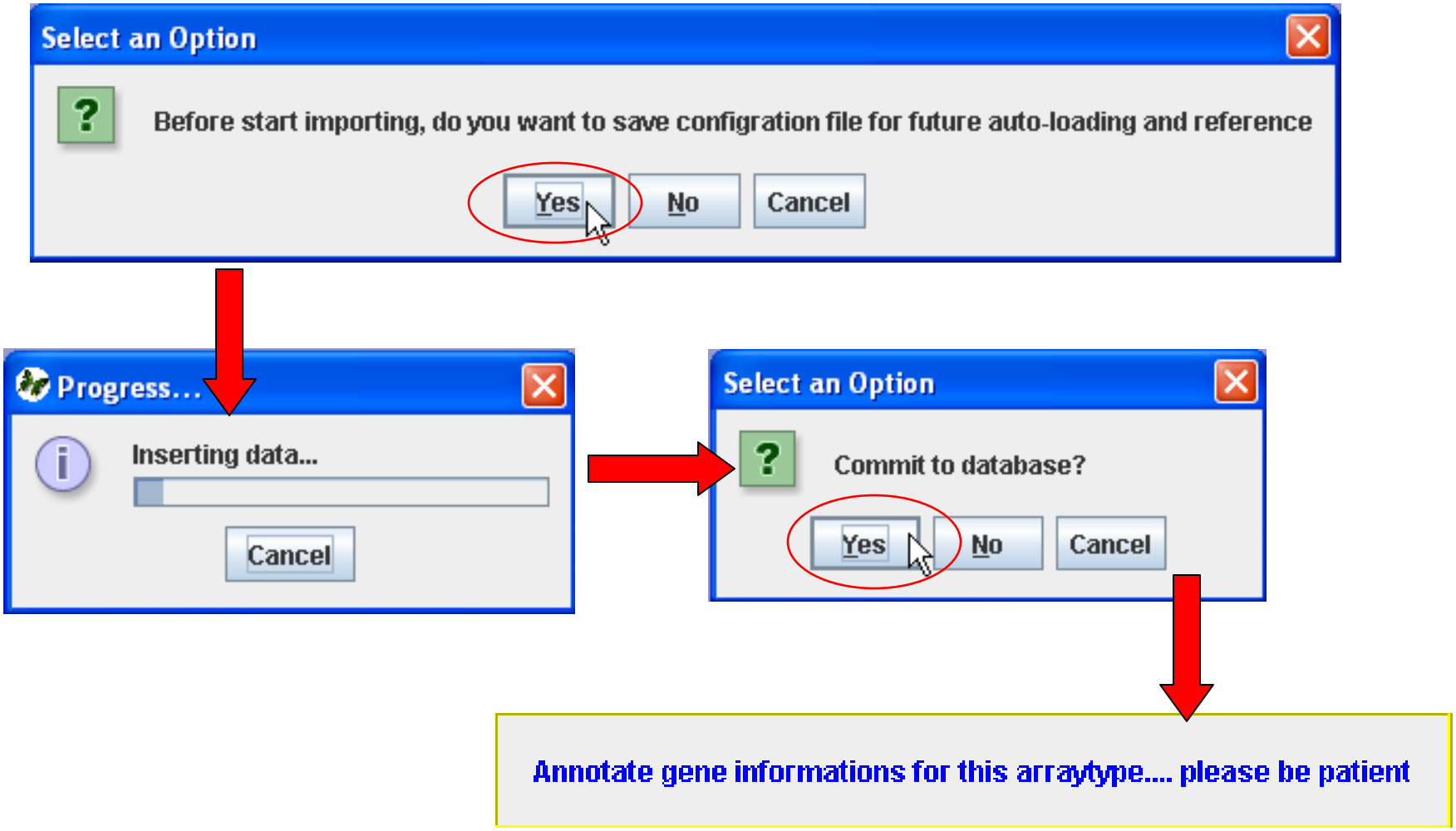
For each column description, select a table column above and click the description button to make the assignment

Description buttons

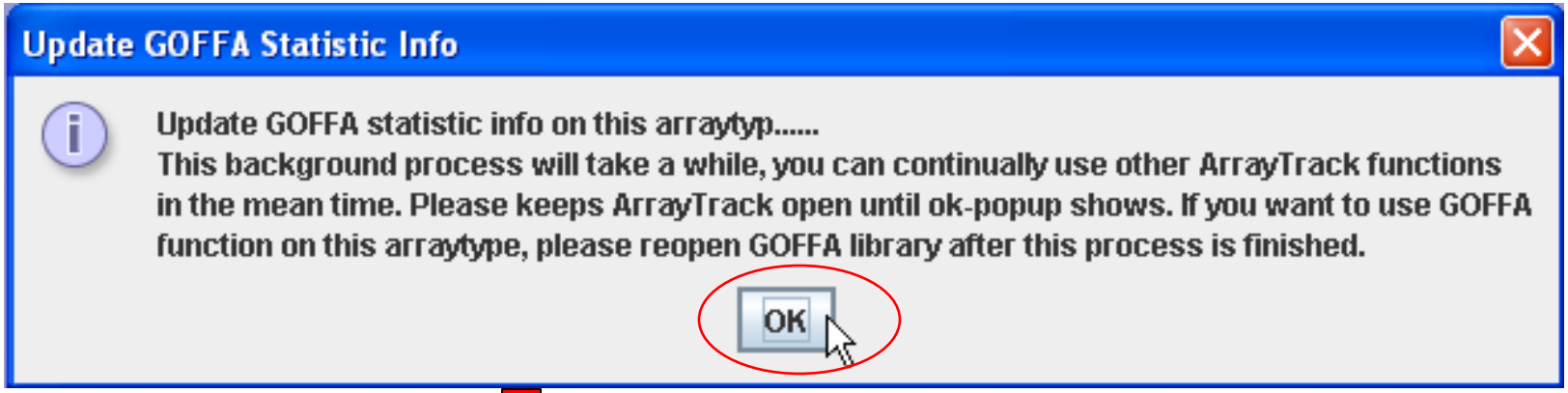
Click Import button

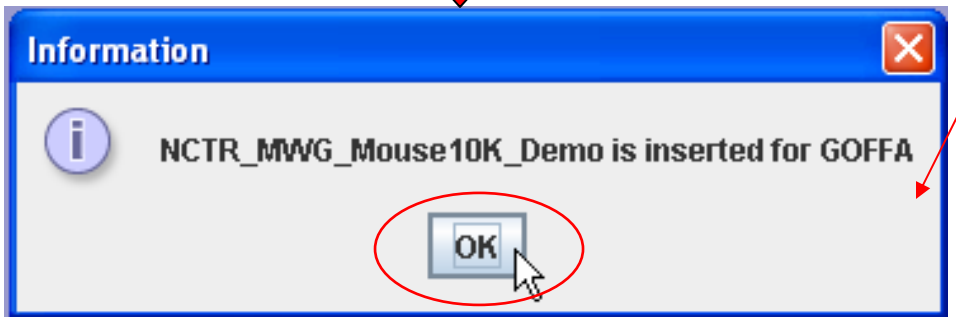
Create Array Type - continued

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



Create Array Type - continued



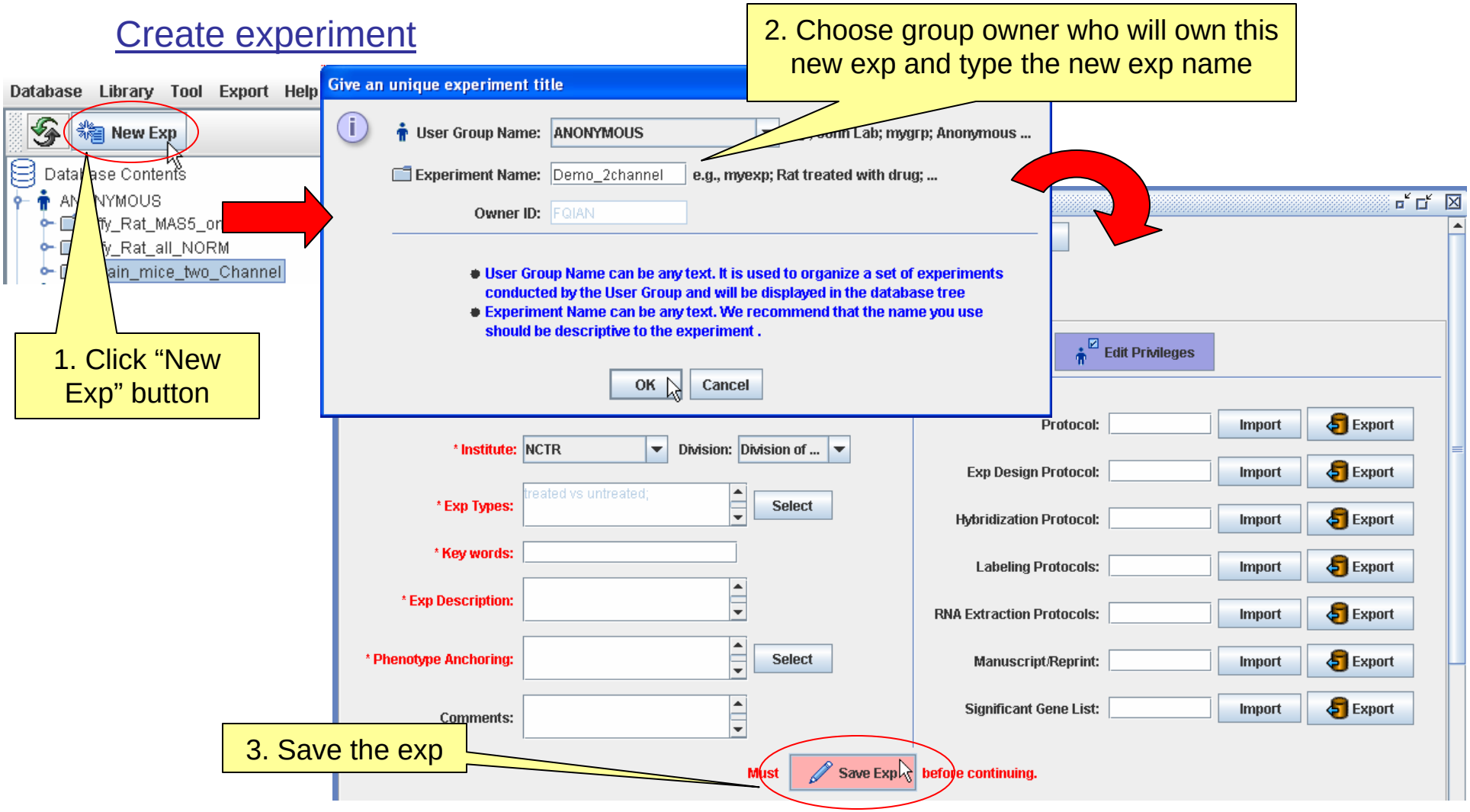


The array type is created successfully.

Create experiment

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

Create experiment



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

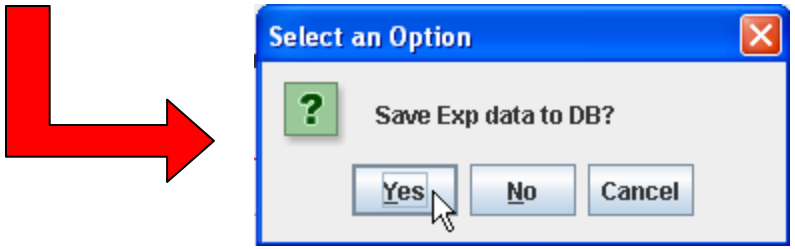
Labeling Protocols: | Import | Export

RNA Extraction Protocols: | Import | Export

Manuscript/Reprint: | Import | Export

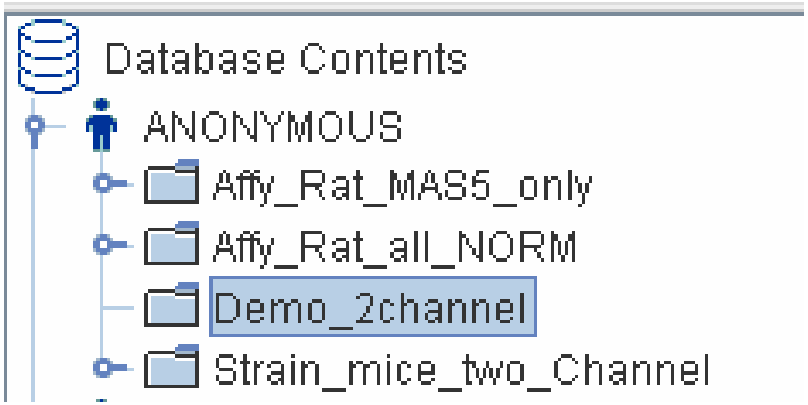
Significant Gene List: | Import | Export

Must Save Exp before continuing.



Create experiment

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

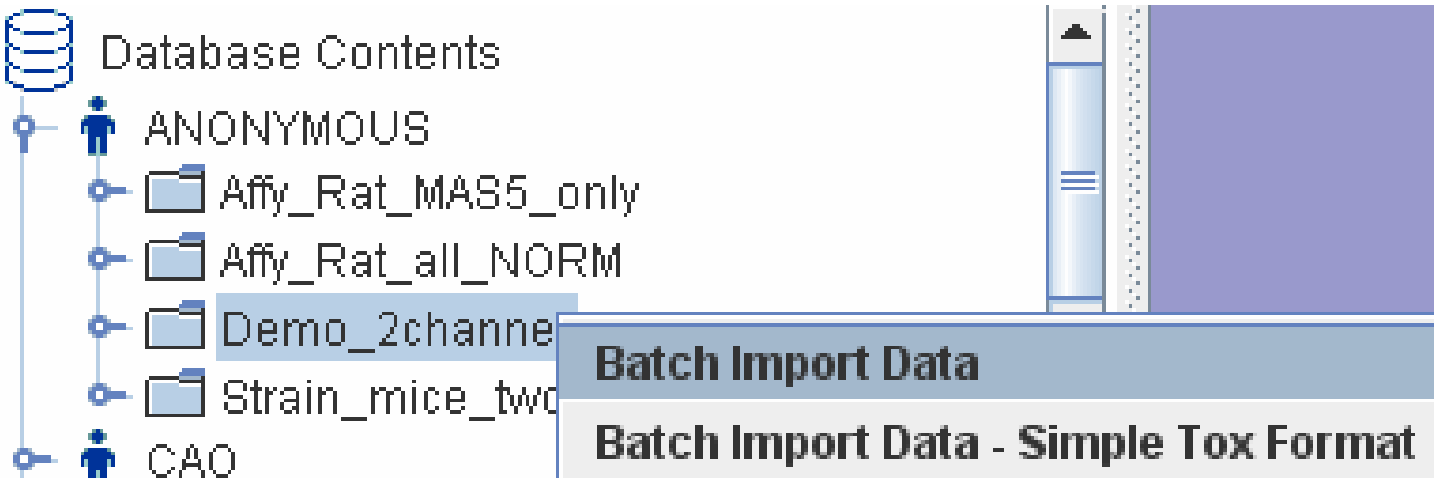


Batch Import

There are two formats for batch import:

- Regular batch import
- Simple Tox format – for animal toxicology study data submission, with combined and extracted main fields from SEND (Standard for Exchange of Nonclinical Data)

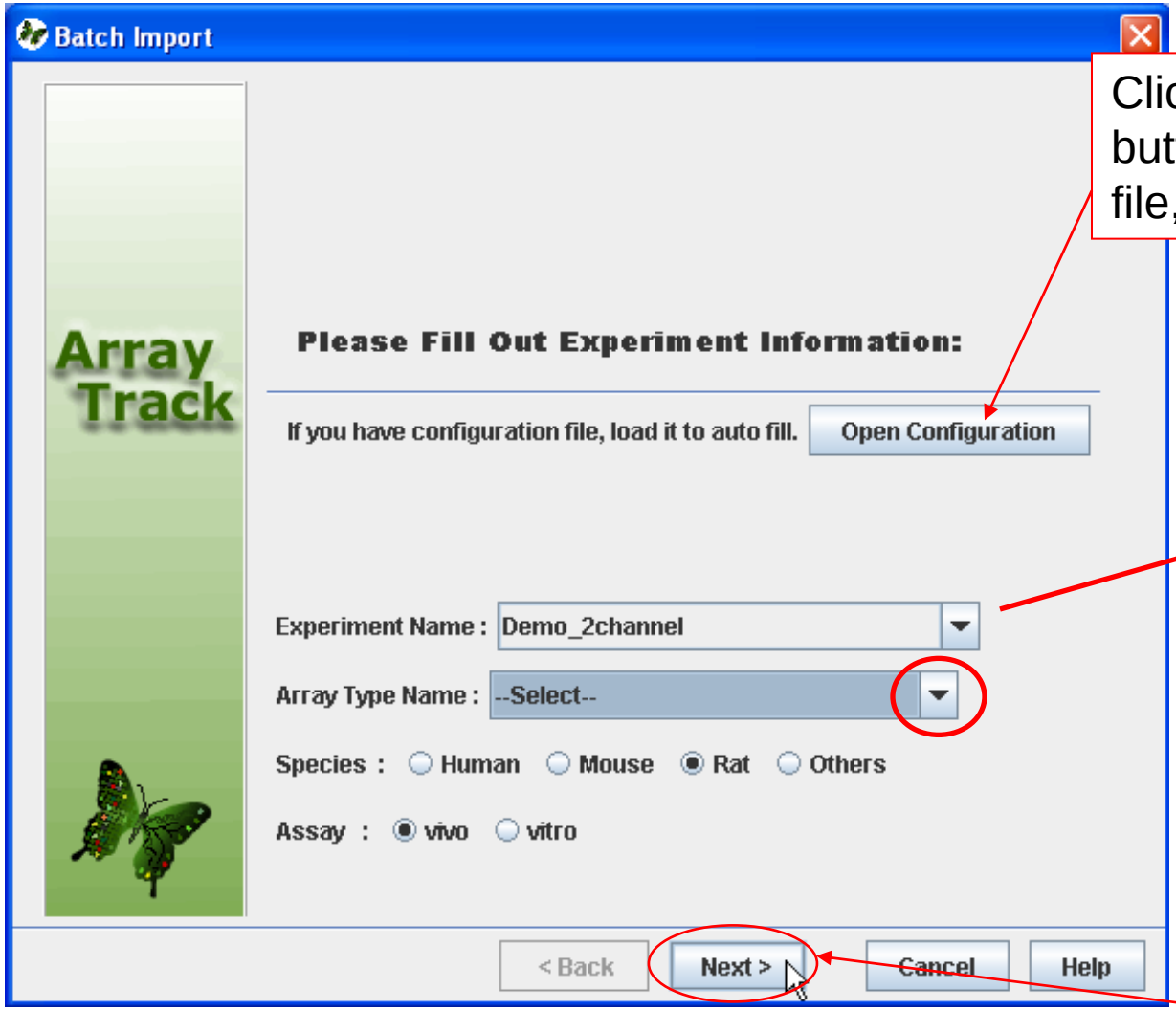
Batch Import - regular



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

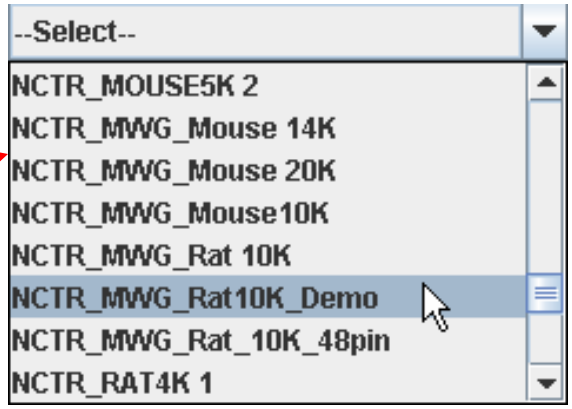
Batch Import Wizard

This is an example of importing 2-channel data.



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

Choose the array type



Click Next button

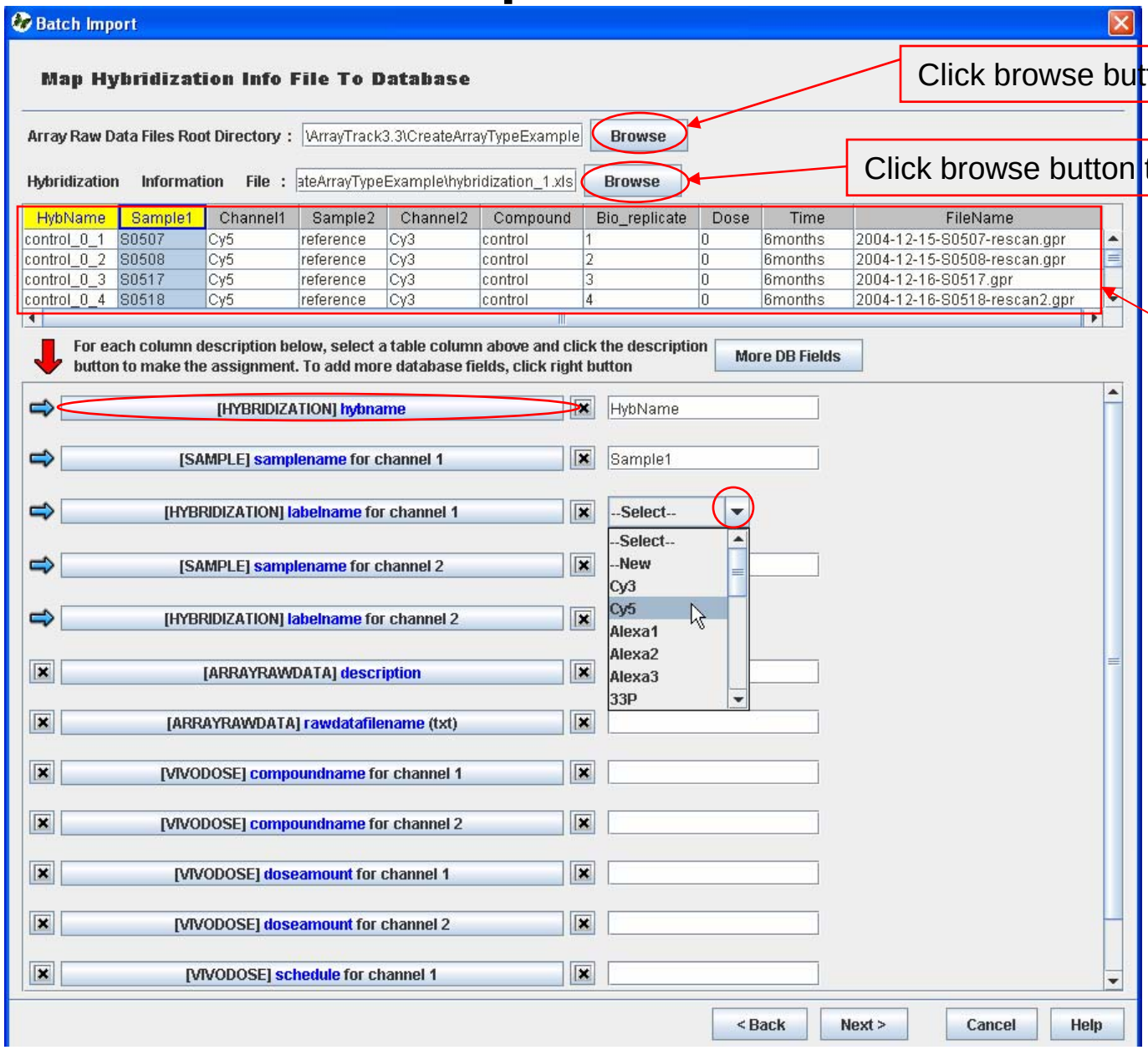
Batch Import

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

hybridization_1

	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

Batch Import Wizard – continued

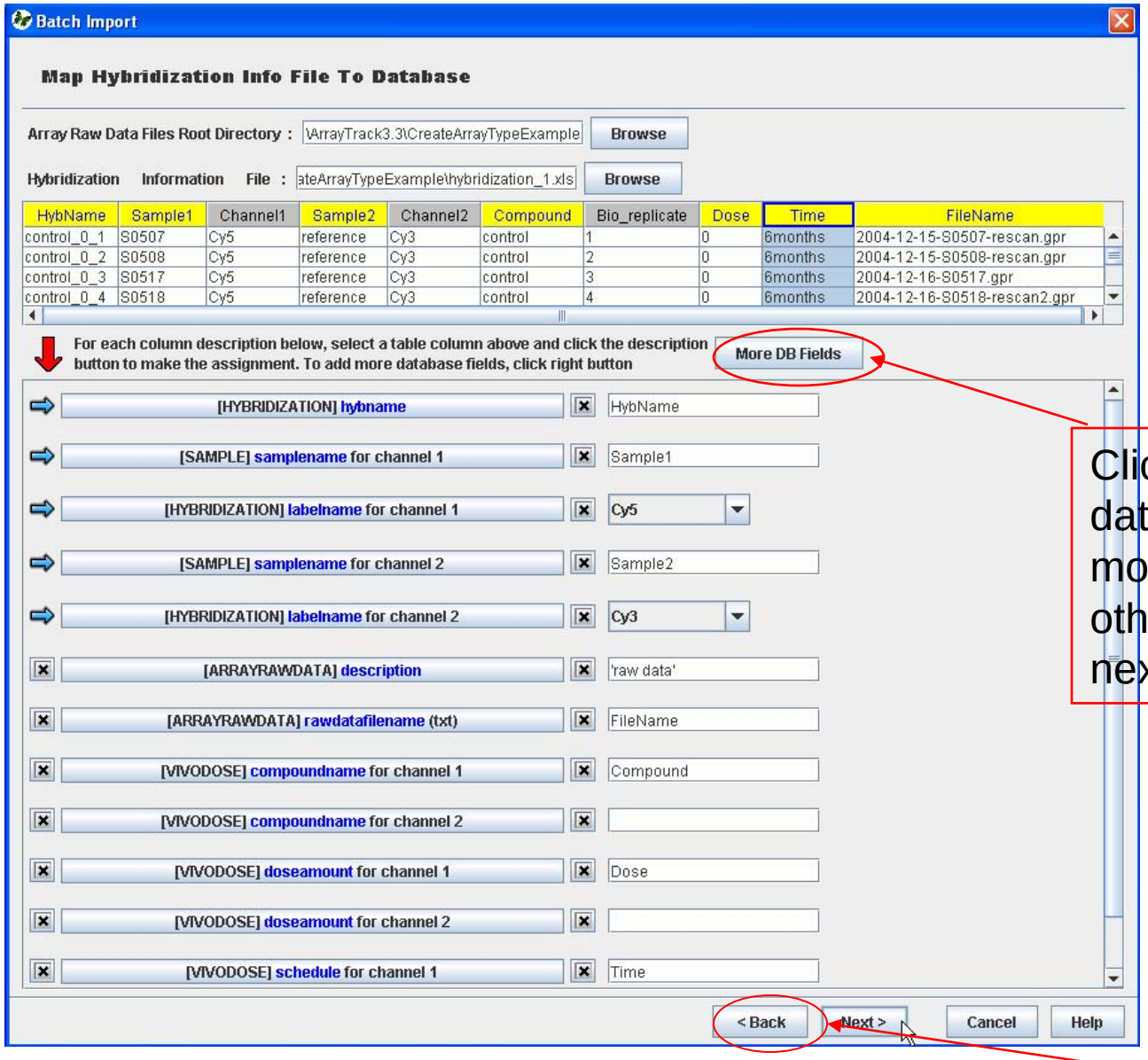


Click browse button to locate the data files

Click browse button to locate the hybridization files

Hybridization file

Batch Import Wizard - – continued



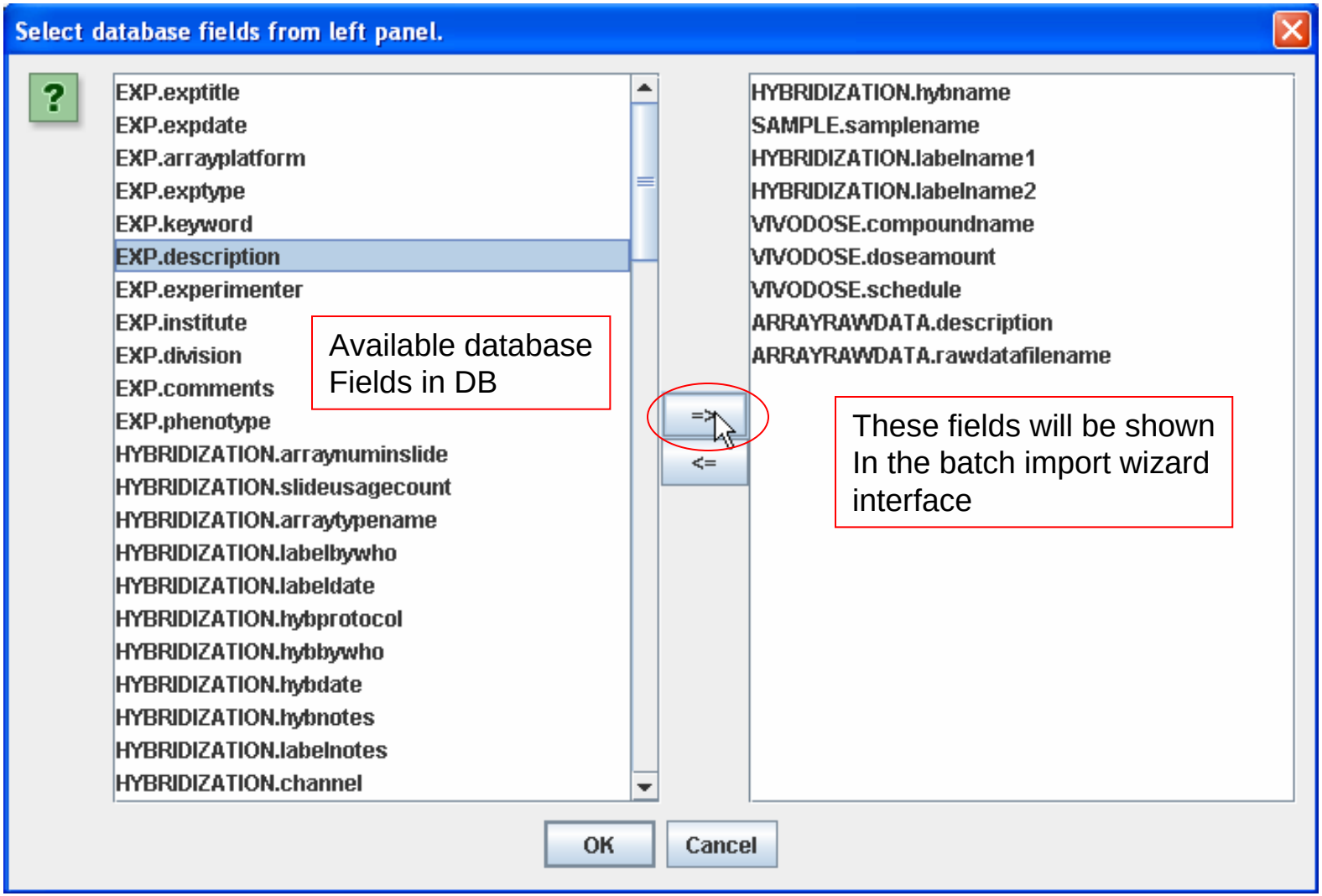
Assign the other columns to the corresponding fields

Click this button to get more database fields if you have more information to put in, otherwise skip this step. See next slide.

Click Next button

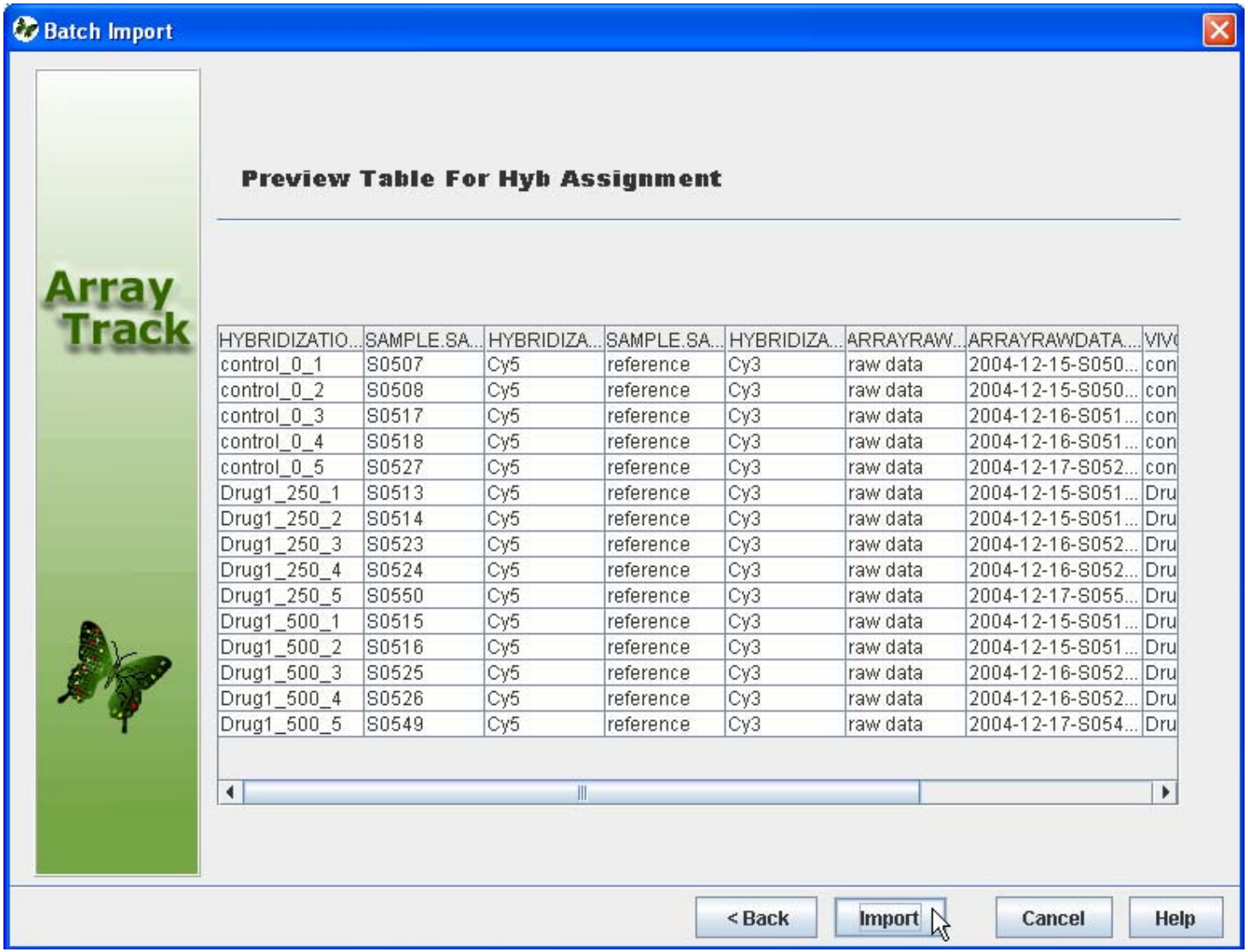
Batch Import Wizard - continued

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.



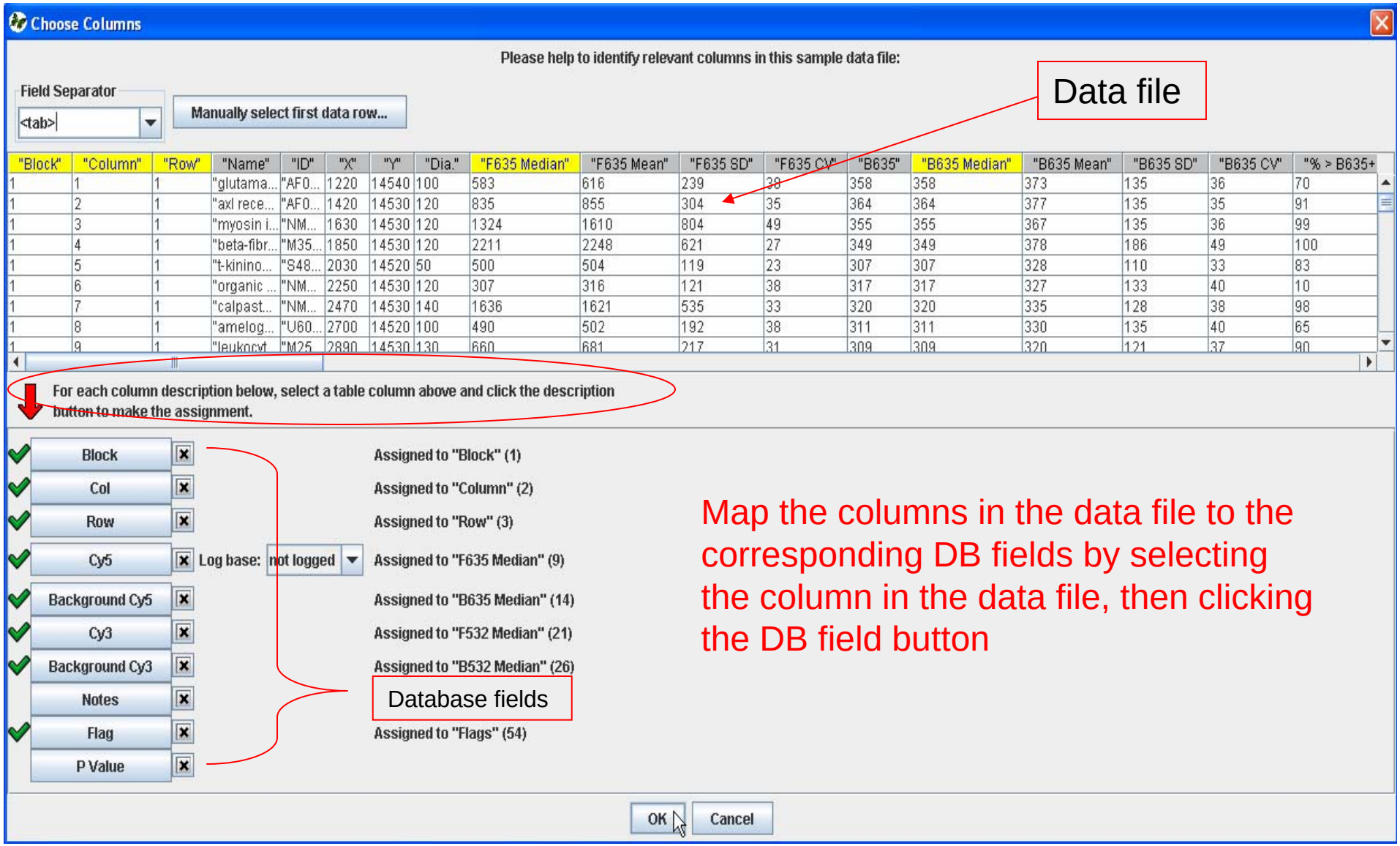
Batch Import Wizard - continued

Preview the table for hybridization assignment before import



HYBRIDIZATIO...	SAMPLE.SA...	HYBRIDIZA...	SAMPLE.SA...	HYBRIDIZA...	ARRAYRAW...	ARRAYRAWDATA....	VIV
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

Batch Import Wizard - continued

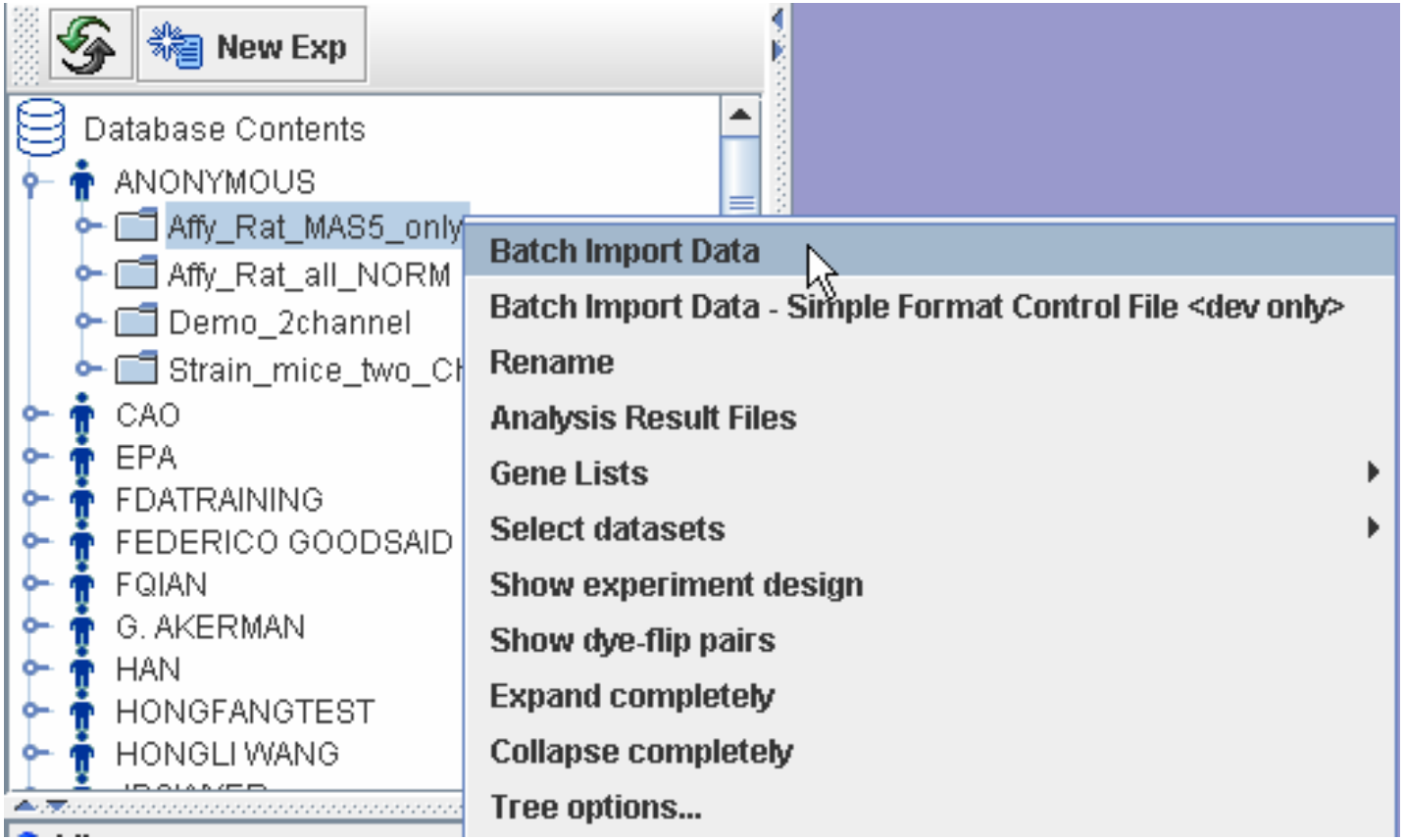


Map the columns in the data file to the corresponding DB fields by selecting the column in the data file, then clicking the DB field button

Batch Import Wizard - continued

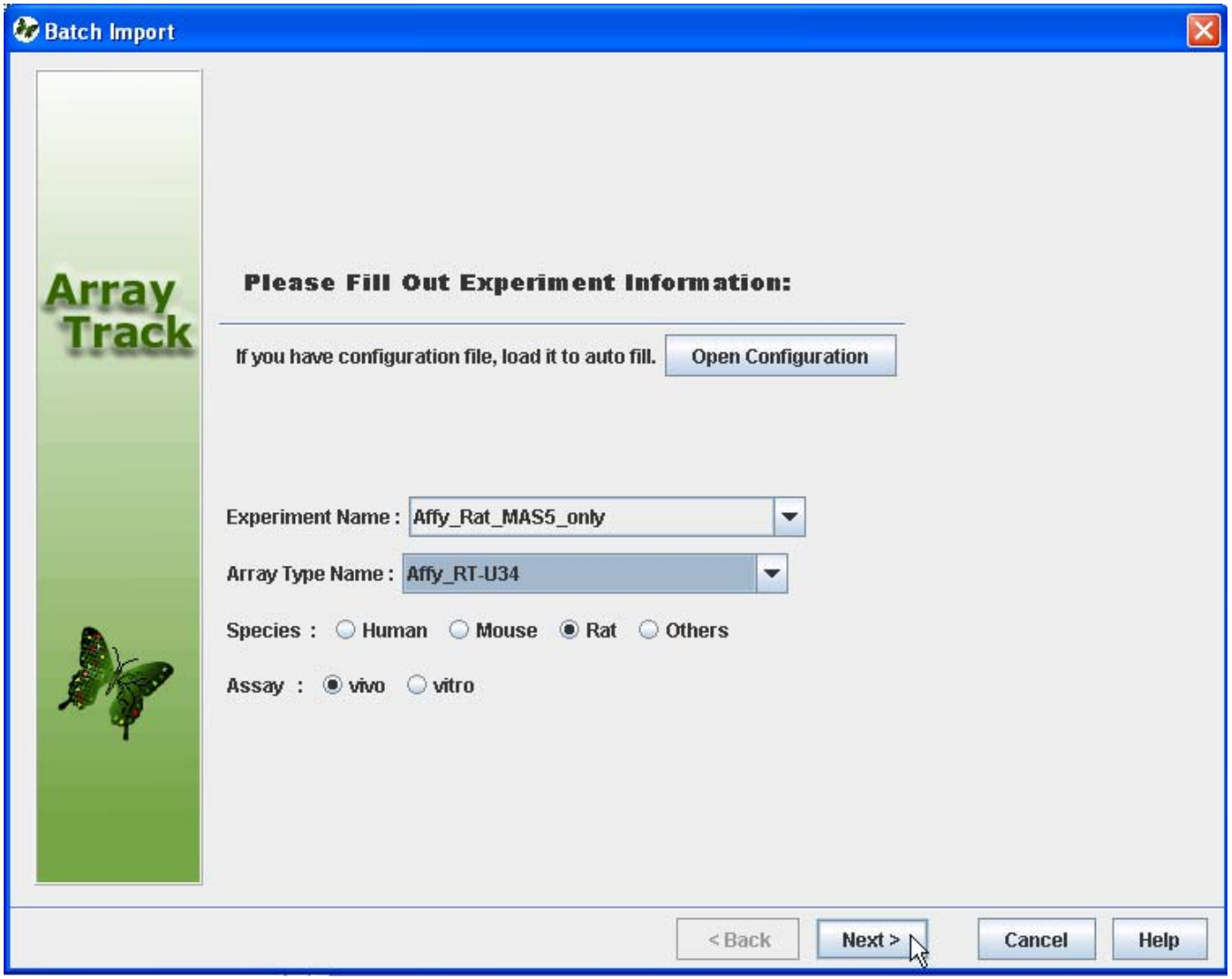
Data import is finished!

Batch Import Wizard
Import CEL File Data

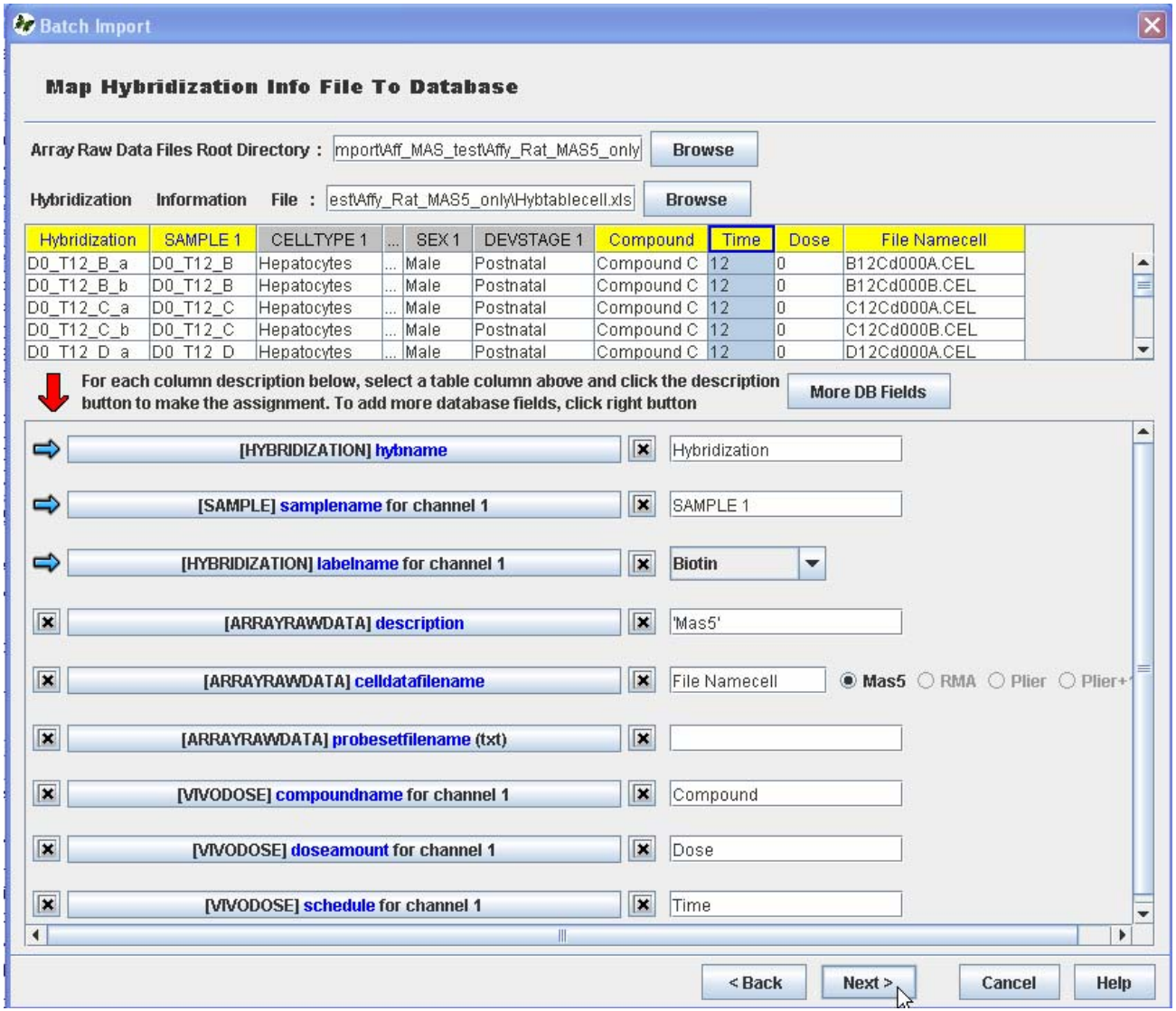


Right-click the experiment and choose “Batch Import Data”.

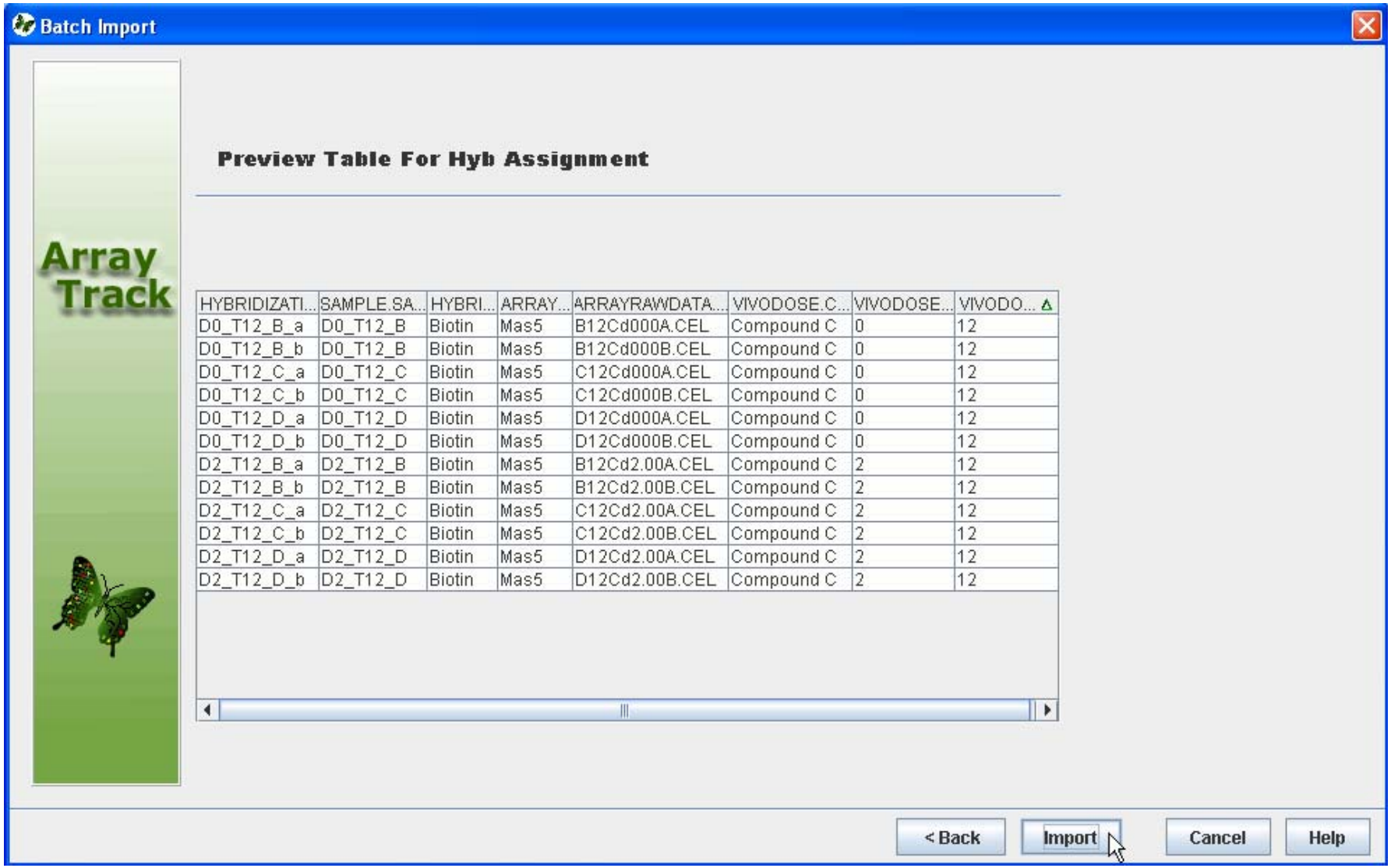
Batch Import Wizard
Import CEL File Data



Batch Import Wizard
Import CEL File Data



Batch Import Wizard
Import CEL File Data

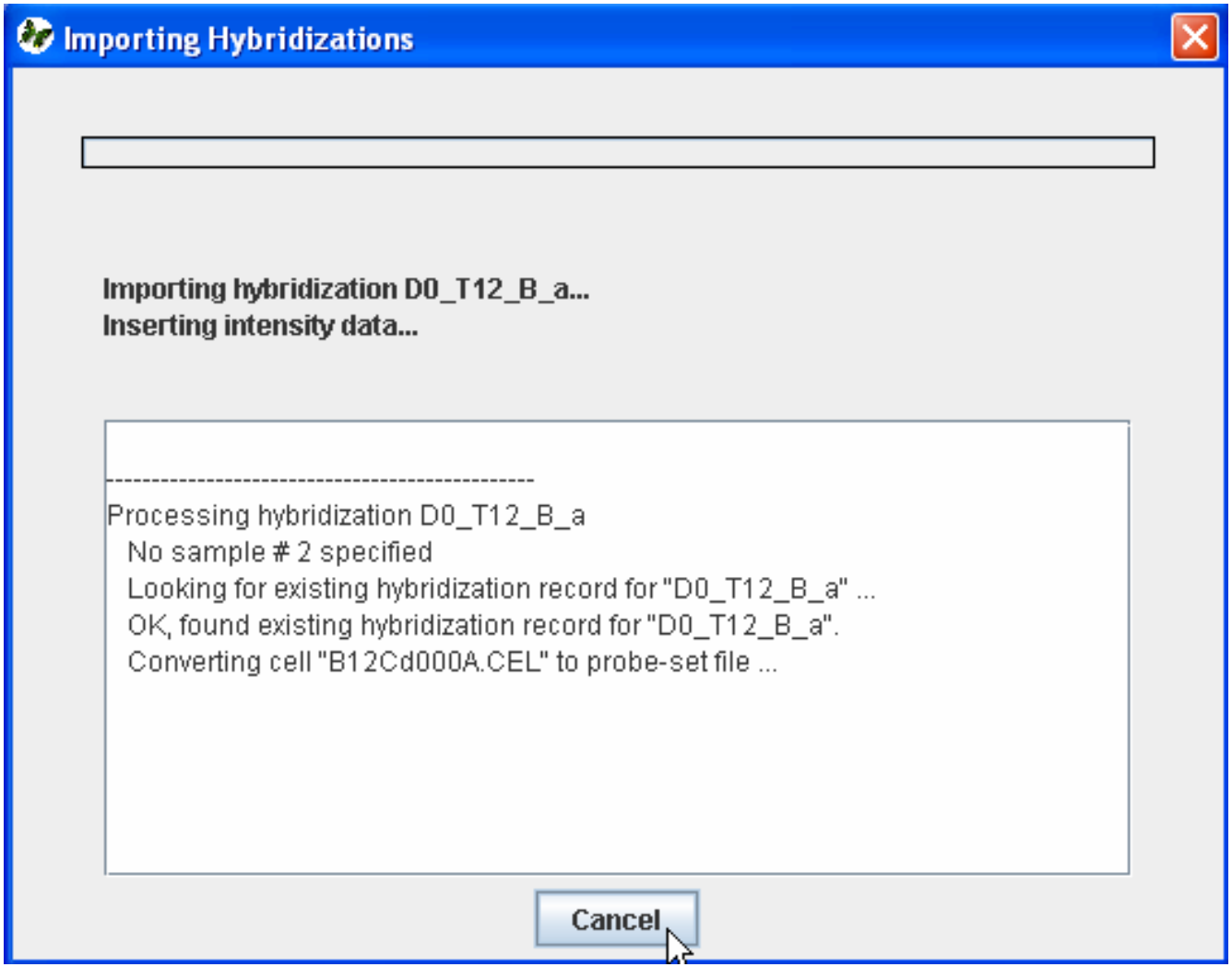


HYBRIDIZATI...	SAMPLE.SA...	HYBRI...	ARRAY...	ARRAYRAWDATA...	VIVODOSE.C...	VIVODOSE...	VIVODO...
D0_T12_B_a	D0_T12_B	Biotin	Mas5	B12Cd000A.CEL	Compound C	0	12
D0_T12_B_b	D0_T12_B	Biotin	Mas5	B12Cd000B.CEL	Compound C	0	12
D0_T12_C_a	D0_T12_C	Biotin	Mas5	C12Cd000A.CEL	Compound C	0	12
D0_T12_C_b	D0_T12_C	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

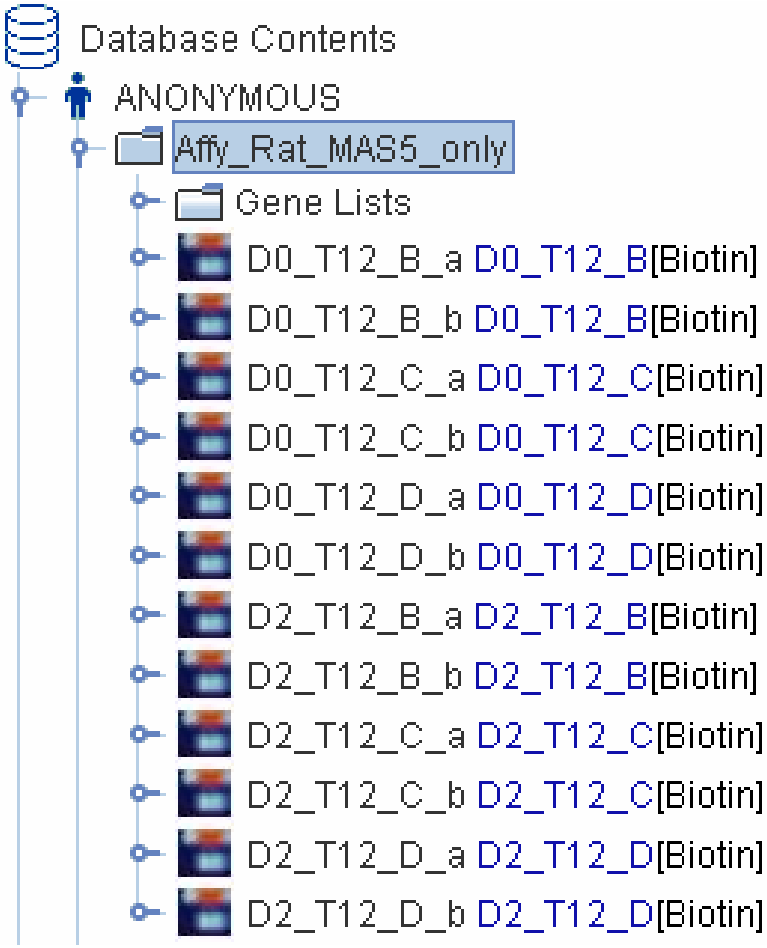
Batch Import Wizard
Import CEL File Data

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

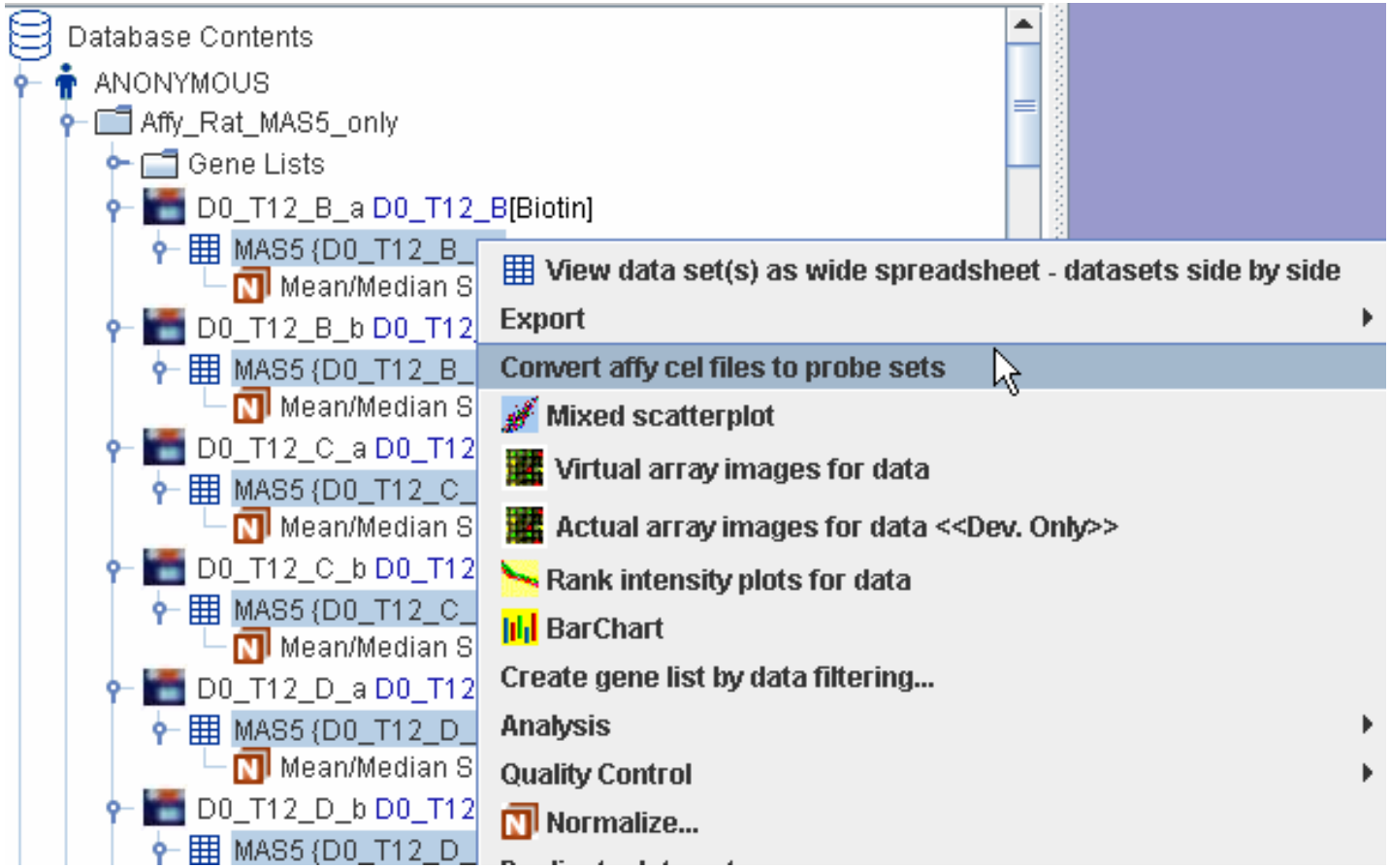


Batch Import Wizard
Import CEL File Data



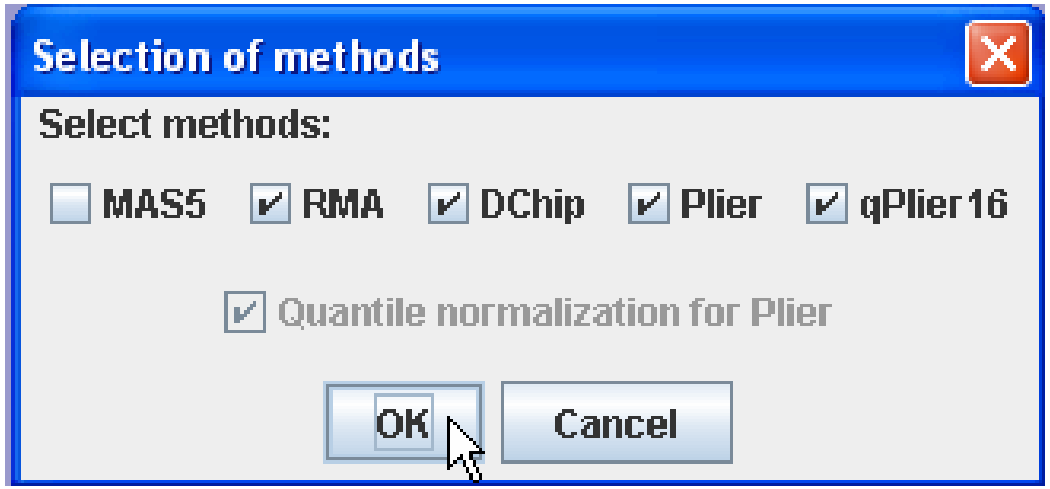
Import is finished.

Convert CEL file to probe set file

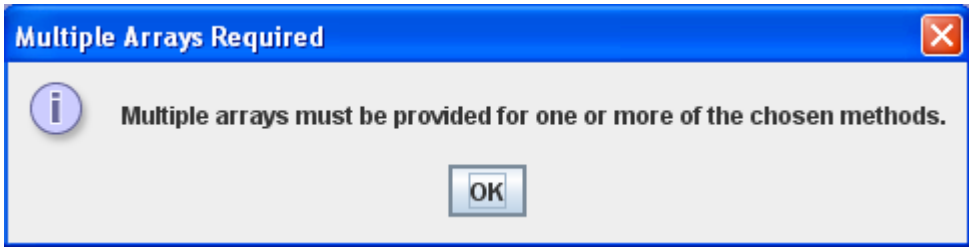


Convert CEL file to probe set file

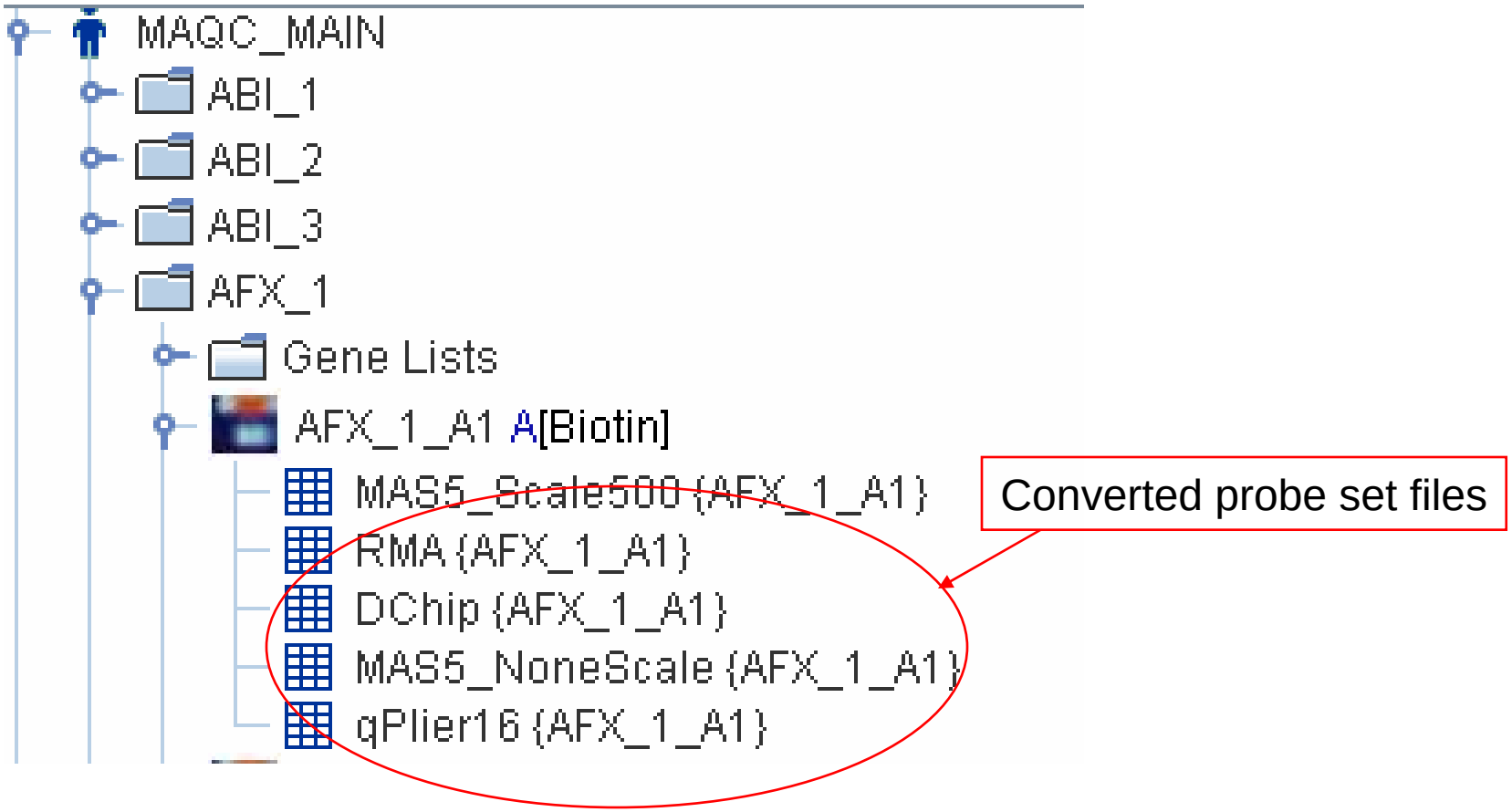
Select the data, right-click, choose “Convert affy cel files to probe sets. ArrayTrack can convert cel to RMA, DChip, qPlier and qPlier16



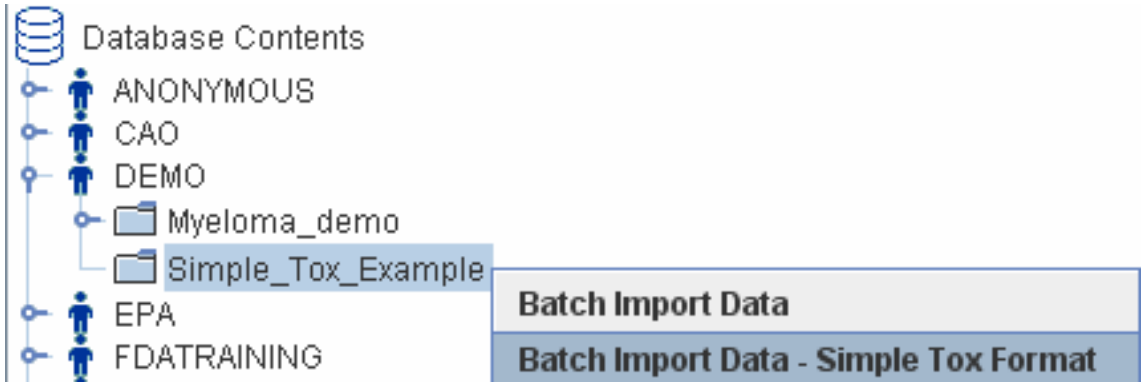
Keep in mind that you have to choose multiple data to do the converting of all the methods except MAS5. For example, you can not select just one data to do RMA converting, otherwise you will get the following message:



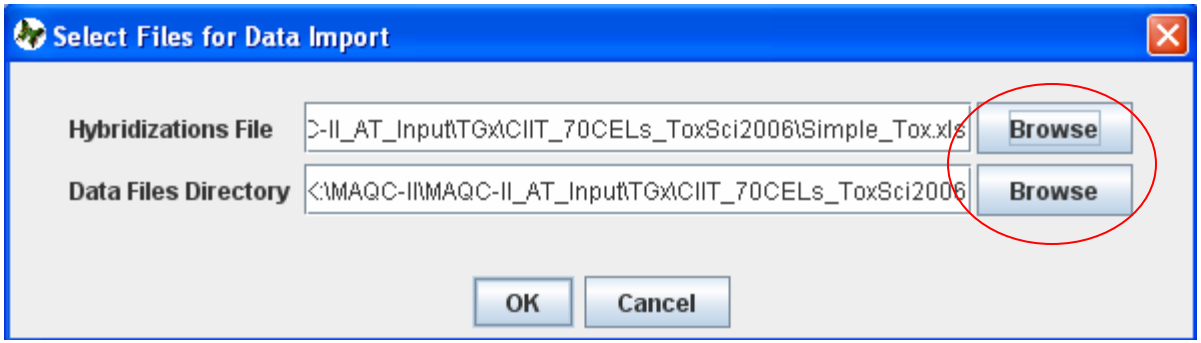
Convert CEL file to probe set file



Batch Import – Simple Tox Format



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



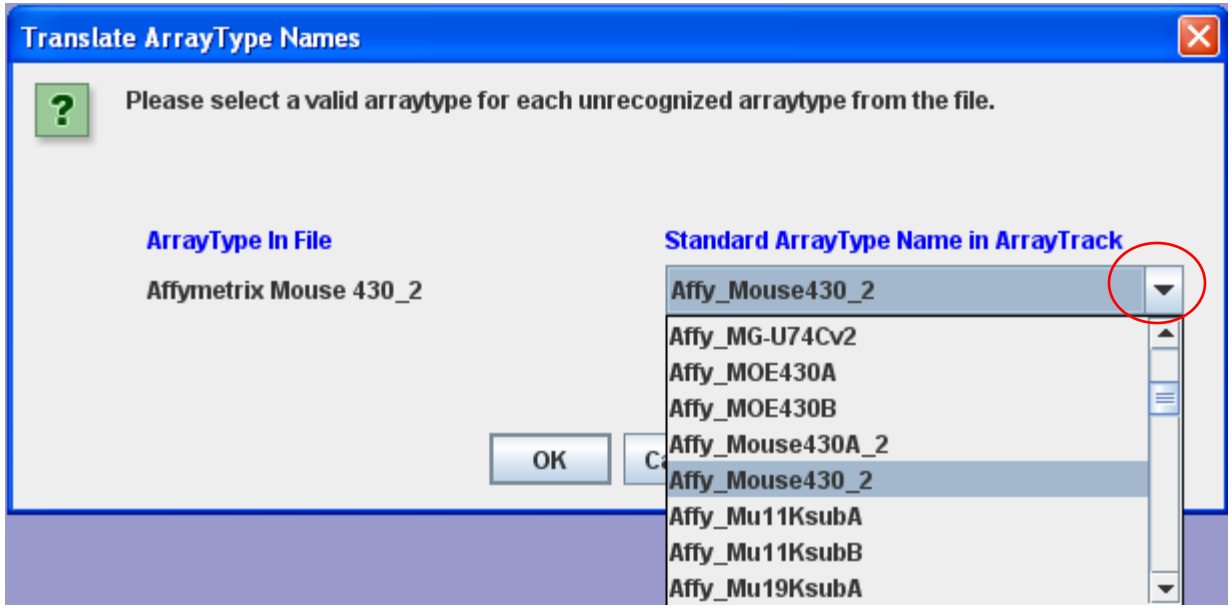
Click Browse button to locate the hybridization file and data files directory.
The next slide shows an example of hybridization table file.

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.

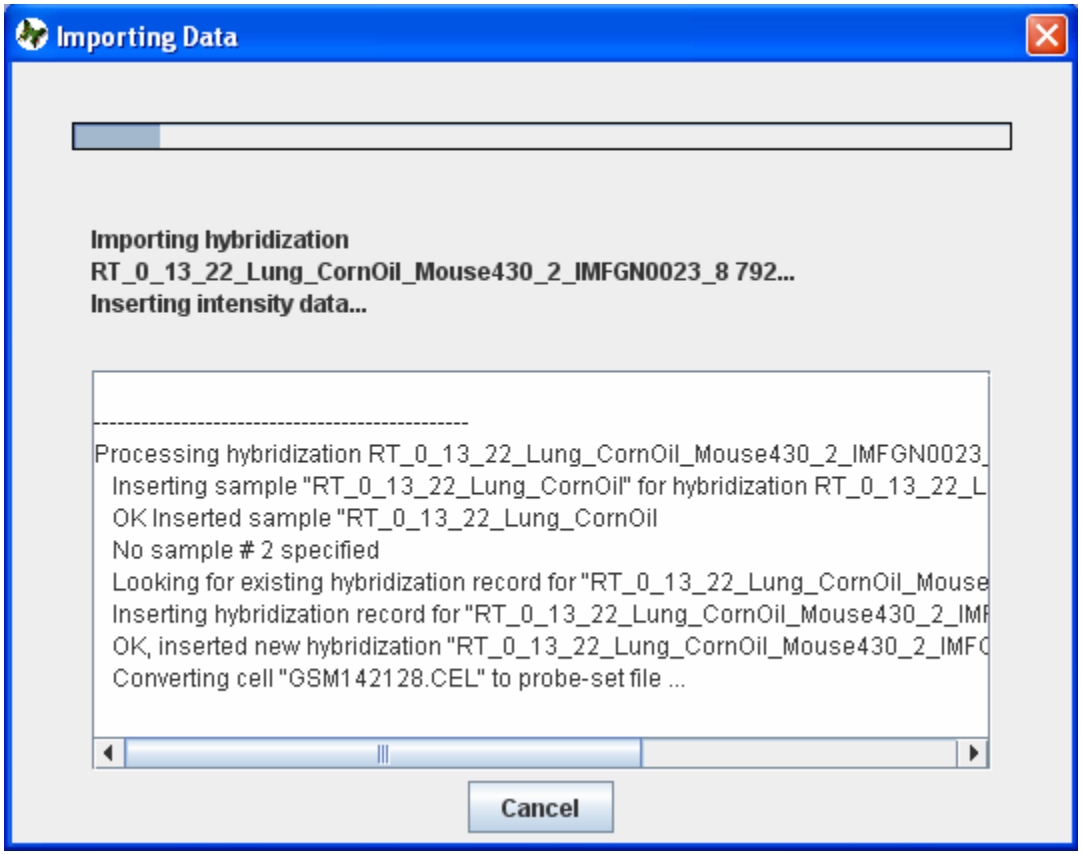
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	Array_ID	Animal_ID	Institutio	StudyTitle	StudyTy	Class	ClassCo	ClassRe	Compou	CAS	Control	Treatme	Dose	DoseUn	HybName	Sample
2	792	1-22	Hamner	Hamner Mice Lung	Repeat Dose				Corn Oil		Y	C	0	mg/kg	RT_0_13_22_Lu	RT_0_13_
3	793	1-24	Hamner	Hamner Mice Lung	Repeat Dose				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-Naphth	2243-62-1	N	LT	2000	ppm	RT_2000_13_2_	RT_2000_

Batch Import – Simple Tox Format

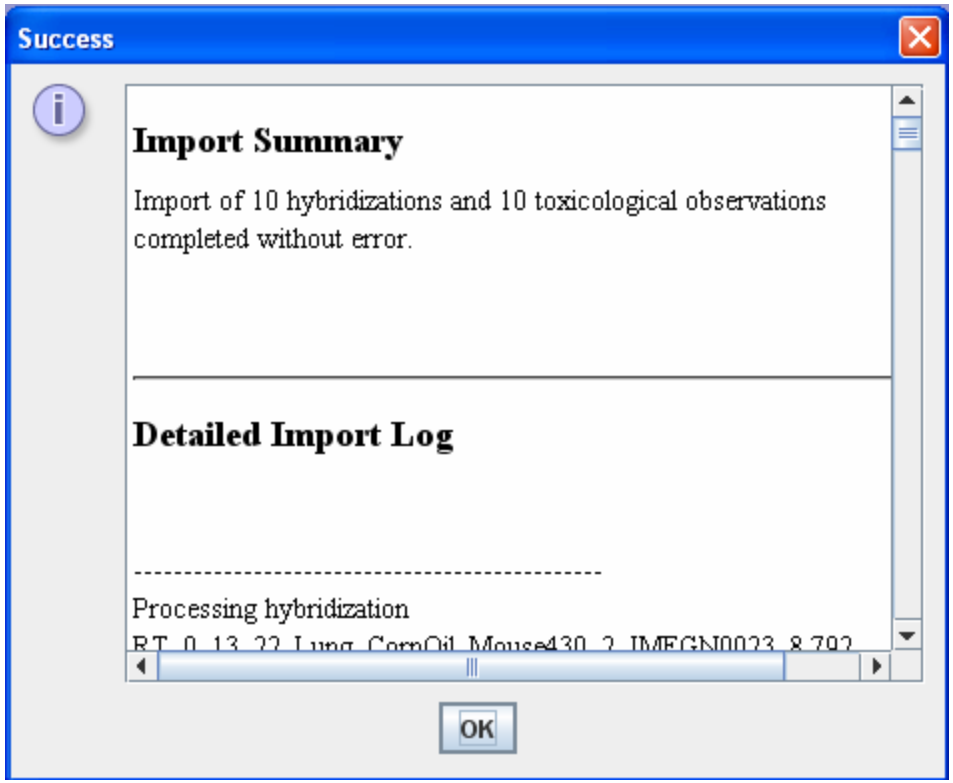


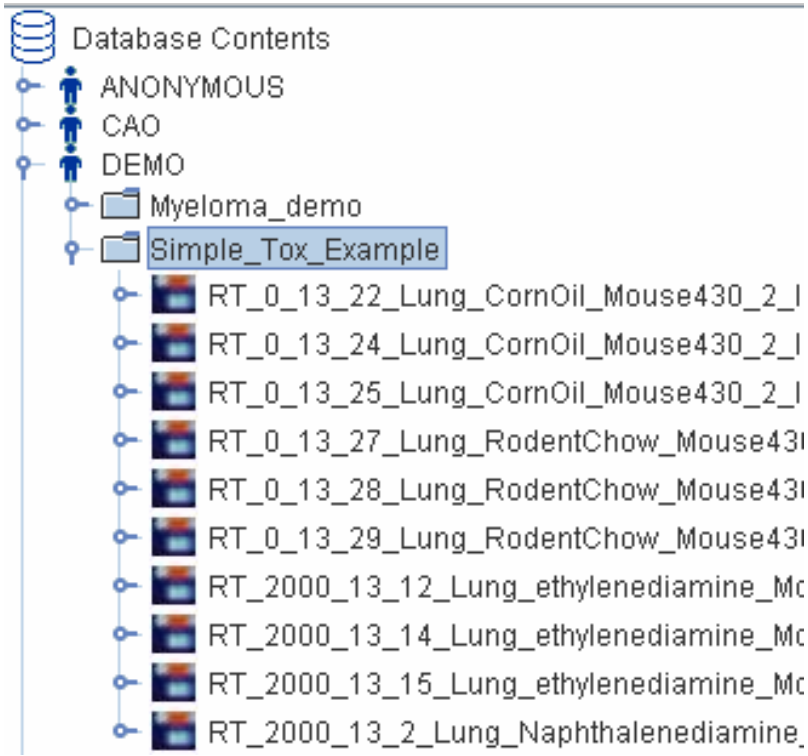
Select the right array type.

Batch Import – Simple Tox Format



Batch Import – Simple Tox Format





Batch import is successfully finished.