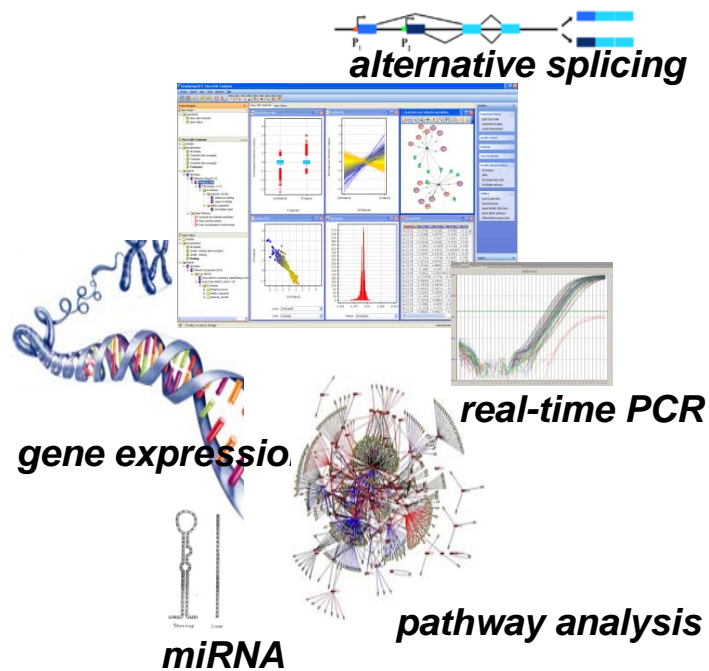


Analysis of Real-time PCR Data in GeneSpring GX and Comparison to Microarray Data



***Pam Tangvoranuntakul, PhD
Agilent Technologies, Inc.***

Agenda for this eSeminar

What annotations and “intensity” values are being used for RT-PCR analysis in GeneSpring?

How do I import and analyze data from non-ABI Platforms?

Analysis of ABI 7900 HT RT-PCR dataset

- Import and normalization
- QC on samples
- Statistical Analysis to identify differential expression

Updating Annotations for RT-PCR experiment

- Build Biological Genome
- Update RT-PCR technology annotations from Biological Genome
- Biological contextualization of RT-PCR results

Comparison of RT-PCR and microarray gene expression experiments

- Translation of RT-PCR data into gene expression experiment
 - Venn Diagram
 - Plot List Associated Values view

RT-PCR Technology

Definition of Technology: A package of information that defines a template of probes, and their associated annotations, AND the file format of the data to be imported

GeneSpring GX 10 supports SDS RQ 2.1, 2.2, and 2.3 and RQ Manager 1.2

“Detector” values are used as Identifier

- If values are gene symbol-ABI identifier, GeneSpring creates Gene Symbol column in technology
- If values are synonym(gene symbol)-ABI identifier, i.e. ATIR(AGTRI)-Hs00241341_m1, GeneSpring creates Synonym and Gene Symbol columns in technology

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	RQ Manager 1.2 RQ Study Results													
2	StudyName CytokineValidation_set1_24jul2008.sdm													
3	Operator													
4														
5	Well	PlateID	Sample	Detector	Task	Ct	delta Rn	delta Ct	Ct Avg	Ct SD	Avg Delta	delta Ct St	Endo Ct A	Endo
6	1	35520021	sensitive_c	Actb-Rn00667869_m1	Endogenou	15.03677	4.572966		15.03677				15.03677	
7	8	35520021	sensitive_c	Cdkn1a-Rn00589996_m1	Target	24.50579	4.189357		24.50579		9.469026		15.03677	
8	14	35520021	sensitive_c	Cxcl1-Rn00578225_m1	Target	27.22445	2.715754		27.22445		12.18769		15.03677	
9	15	35520021	sensitive_c	Casp3-Rn00573260_m1	Target	18.51014	6.041115		18.51014		3.473373		15.03677	
10	16	35520021	sensitive_c	Cxcr4-Rn01483207_m1	Target	25.05786	4.584239		25.05786		10.02109		15.03677	
11	29	35520021	sensitive_c	IL6-Rn99999011_m1	Target	30.64603	1.703982		30.64603		15.60927		15.03677	

RT-PCR Technology- Data File Format

CytokineValidation.sdm-Amplification Data.txt														
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	RQ Manager 1.2 RQ Study Results													
2	StudyName CytokineValidation_set1_24jul2008.sdm													
3	Operator													
4														
5	Well	PlateID	Sample	Detector	Task	Ct	delta Rn	delta Ct	Ct Avg	Ct SD	Avg Delta	delta Ct St	Endo Ct A	Endo
6	1	35520021	sensitive	Actb-Rn00667869_m1	Endogen	15.03677	4.572966		15.03677				15.03677	
7	8	35520021	sensitive	Cdkn1a-Rn00589996_m1	Target	24.50579	4.189357		24.50579	9.469026			15.03677	
8	14	35520021	sensitive	Cxcl1-Rn00578225_m1	Target	27.22445	2.715754		27.22445	12.18769			15.03677	
9	15	35520021	sensitive	Casp3-Rn00573260_m1	Target	18.51014	6.041115		18.51014	3.473373			15.03677	
10	16	35520021	sensitive	Cxcr4-Rn01483207_m1	Target	25.05786	4.584239		25.05786	10.02109			15.03677	
11	29	35520021	sensitive	cll6-Rn99999011_m1	Target	30.64603	1.703982		30.64603	15.60927			15.03677	

“Sample” values indicate samples in GeneSpring

“Task” values used to indicate probe type

“Ct Avg” values used as raw signal

Raw signal: Ct Avg value for each probe

Normalized: Averaged Ct Avg of all endogenous controls – target Ct Avg value; baseline transformed (if applied)

RT-PCR Technology Annotation

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	RQ Manager 1.2 RQ Study Results													
2	StudyName CytokineValidation_set1_24jul2008.sdm													
3	Operator													
4														
5	Well	PlateID	Sample	Detector	Task	Ct	delta Rn	delta Ct	Ct Avg	Ct SD	Avg Delta	delta Ct St	Endo Ct A	Endo
6	1	35520021	sensitive_c	Actb-Rn00667869_m1	Endogenou	15.03677	4.572966		15.03677				15.03677	
7	8	35520021	sensitive_c	Cdkn1a-Rn00589996_m1	Target	24.50579	4.189357		24.50579		9.469026		15.03677	
8	14	35520021	sensitive_c	Cxcl1-Rn00578225_m1	Target	27.22445	2.715754		27.22445		12.18769		15.03677	
9	15	35520021	sensitive_c	Casp3-Rn00573260_m1	Target	18.51014	6.041115		18.51014		3.473373		15.03677	
10	16	35520021	sensitive_c	Cxcr4-Rn01483207_m1	Target	25.05786	4.584239		25.05786		10.02109		15.03677	
11	29	35520021	sensitive_c	U6-Rn99999011_m1	Target	30.64603	1.703982		30.64603		15.60927		15.03677	

Problem: SDS and RQ Manager output files do not contain probe annotations

Solution: GeneSpring automatically parses “Detector” column and creates a column of Gene Symbol (and Synonym if present), which can be used to bring in more annotations from NCBI

Bring in more annotations for your RT-PCR experiment

Entities Attributes

Detector	AssayID	GeneSymbol
Actb-Rn00667869_m1	Rn00667869_m1	Actb
Cdkn1a-Rn00589996_m1	Rn00589996_m1	Cdkn1a
Bmp1-Rn00585674_s1	Rn00585674_s1	Bmp1
Il6-Rn99999011_m1	Rn99999011_m1	Il6
Ptgs2-Rn01483828_m1	Rn01483828_m1	Ptgs2
Cxcr4-Rn01483207_m1	Rn01483207_m1	Cxcr4
Cxcl1-Rn00578225_m1	Rn00578225_m1	Cxcl1

Find: Find Next Find Previous Match Case

Configure Columns

From the menu:
**Annotation > Update
Technology**
**Annotation > from file
or Biological Genome**

Entities Attributes

Detector	AssayID	GeneSym...	description	Entrez Ge...	GeneID	chromoso...	GO_ID_Co...	GO_ID_Fu...	GO_ID_Pr...	assembly	Categr
Actb-Rn0...	Rn00667...	Actb	actin, beta	81822	81822	12	GO:0043...	GO:0019...	GO:0007...	Reference...	Process/
Cdkn1a-R...	Rn00589...	Cdkn1a	cyclin-de...	114851	114851	20	GO:0005...	GO:0032...	GO:0055...	Reference...	Process/
Bmp1-Rn...	Rn00585...	Bmp1	bone mor...	83470	83470	15		GO:0005...	GO:0006...	Reference...	Process/
Il6-Rn999...	Rn99999...	Il6	interleukin 6	24498	24498	4	GO:0009...	GO:0005...	GO:0051...	Reference...	Process/
Ptgs2-Rn...	Rn01483...	Ptgs2	prostagla...	29527	29527	13	GO:0043...	GO:0046...	GO:0055...	Reference...	Process/
Cxcr4-Rn	Rn01483	Cxcr4	chemokin	60628	60628	13	GO:0031	GO:0032	GO:0060	Reference	Process/

Find: Find Next Find Previous Match Case

Configure Columns

Help OK Cancel

What is a Biological Genome in GeneSpring GX 10?

Annotation	Value
Detector	Bmp1-Rn00585674_s1
AssayID	Rn00585674_s1
assembly	Reference assembly (based on RGSC v3.4)...
Category	Process//Function
chromosome	15
dbXrefs	RGD:620739 Ensembl:ENSRNOG0000001...
description	bone morphogenetic protein 1
end_position_on_the_genomic_accession	2033042///50922863///22784830///4...
entityType	Target
Entrez Gene	83470
Evidence	ISS//IEA
Full_name_from_nomenclature_authority	bone morphogenetic protein 1
GeneID	83470
GeneSymbol	Bmp1
genomic_nucleotide_accession.version	AC_000083.1//NW_047454.2//NC_00...
genomic_nucleotide_gi	109649553///62662024///109502978...
GO_Avadis_	GO:0006508 , GO:0005509
GO_ID_Component	
GO_ID_Function	GO:0005509
GO_ID_Process	GO:0006508
GO_term	calcium ion binding//proteolysis
LocusTag	
map_location	15p11
Modification_date	20080903
Nomenclature_status	0

Biological Genome: a technology containing annotations for a particular organism

Annotations are obtained from NCBI

Update annotations of other technologies using information from Biological Genome

Biological Genome

The image shows two overlapping software windows. The background window is titled 'Create Biological Genome' and contains a 'Description' section, a 'Biological Genome' section with a 'Choose Organisms' list (including Homosapiens, Musmusculus, Rattusnorvegicus, Canis familiaris, Gallus gallus, Danio rerio, Drosophila melano, Saccharomyces ce, Micrococcus fulvu, and Canis lupus), and a 'Genomic Data' section with radio buttons for 'Use from local folder' and 'Download from NCBI ftp site'. A 'Confirmation' dialog box is open over the organisms list, asking if the user wants to continue with the online update. The foreground window is titled 'Update Annotations of Technology (Step 1 of 3)' and contains an 'Input Parameters' section with a 'Choose technology' dropdown menu (set to 'ABI.RTPCR.RQ1.2-RtPCR_Lungs') and radio buttons for 'Update from Biological Genome' (selected) and 'Update from file'. A 'Choose file' dropdown menu is also present. A yellow callout box at the bottom of the foreground window contains the text: 'From the menu: Annotation > Update Technology Annotation > from file or Biological Genome'. Another yellow callout box at the bottom of the background window contains the text: 'From the menu: Annotation > Create Biological Genome'.

Step 1: Download or select necessary files from NCBI

- *Download from NCBI ftp site* option will automatically download necessary files and automatically create Biological Genome
- *Use from local folder* option allows user to create biological genome from files that already exist on local computer

Step 2: Update technology using annotations from Biological Genome of the organism

Updating Technology Annotation from Biological Genome

Update Annotations of Technology (Step 3 of 3)

Choose Annotation Columns

Choose technology column to match: GeneSymbol

Choose genome/file column to match: Symbol

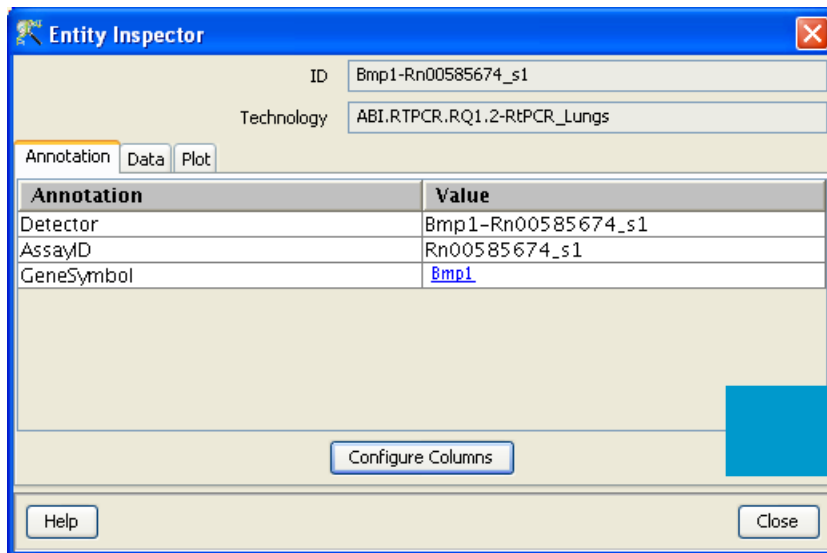
Choose Update Method: Overwrite

<input checked="" type="checkbox"/>	#	Column Name	Data Type	Update Method	Match Column
<input checked="" type="checkbox"/>	0	GeneID	string	Overwrite	Fill missing values
<input checked="" type="checkbox"/>	1	assembly	string	Categorical	None
<input checked="" type="checkbox"/>	2	Category	string	Categorical	None
<input checked="" type="checkbox"/>	3	chromosome	string	Categorical	Chromosome Number
<input checked="" type="checkbox"/>	4	dbXrefs	string	Categorical	None
<input checked="" type="checkbox"/>	5	description	string	Categorical	None
<input checked="" type="checkbox"/>	6	end_position...	string	Categorical	Chromosome End Index
<input checked="" type="checkbox"/>	7	Entrez Gene	string	Categorical	Entrez Gene Id
<input checked="" type="checkbox"/>	8	Evidence	string	Categorical	None
<input checked="" type="checkbox"/>	9	Full_name_fr...	string	Categorical	None
<input checked="" type="checkbox"/>	10	genomic_nuc...	string	Categorical	None
<input checked="" type="checkbox"/>	11	genomic_nuc...	string	Categorical	None
<input checked="" type="checkbox"/>	12	GO_ID	string	Categorical	Gene Ontology accession
<input checked="" type="checkbox"/>	13	GO_ID:Com...	string	Categorical	Cellular Component
<input checked="" type="checkbox"/>	14	GO_ID:Func...	string	Categorical	Molecular Function
<input checked="" type="checkbox"/>	15	GO_ID:Process	string	Categorical	Biological Process

Help << Back Next >> Finish Cancel

Annotation columns in window are columns in Biological Genome available to be added to current technology to be updated

Updating Technology Annotation from Biological Genome

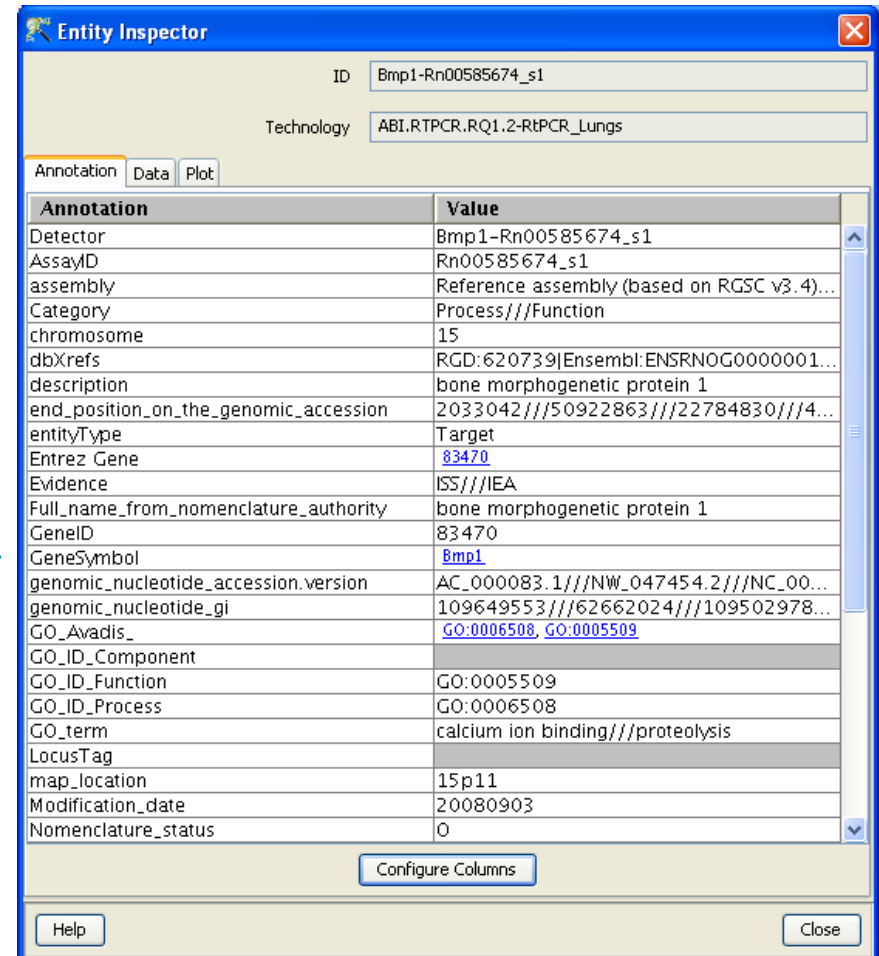


The Entity Inspector window shows the following data:

Annotation	Value
Detector	Bmp1-Rn00585674_s1
AssayID	Rn00585674_s1
GeneSymbol	Bmp1

Buttons: Configure Columns, Help, Close

Before Update



The Entity Inspector window shows the following data:

Annotation	Value
Detector	Bmp1-Rn00585674_s1
AssayID	Rn00585674_s1
assembly	Reference assembly (based on RGSC v3.4)...
Category	Process//Function
chromosome	15
dbXrefs	RGD:620739 Ensembl:ENSRNOG0000001...
description	bone morphogenetic protein 1
end_position_on_the_genomic_accession	2033042///50922863///22784830///4...
entityType	Target
Entrez Gene	83470
Evidence	ISS//IEA
Full_name_from_nomenclature_authority	bone morphogenetic protein 1
GeneID	83470
GeneSymbol	Bmp1
genomic_nucleotide_accession.version	AC_000083.1///NW_047454.2///NC_00...
genomic_nucleotide_gi	109649553///62662024///109502978...
GO_Avadis_	GO:0006508 , GO:0005509
GO_ID_Component	
GO_ID_Function	GO:0005509
GO_ID_Process	GO:0006508
GO_term	calcium ion binding//proteolysis
LocusTag	
map_location	15p11
Modification_date	20080903
Nomenclature_status	0

Buttons: Configure Columns, Help, Close

After Update

Import Data From Other RT-PCR Platforms

	A	B	C	D	E
1	Sample	Detector	Task	Ct Avg	
2	Control	ADCYAP1R1-Hs00153869_m1	Target	26.61352	
3	Control	ADMR-Hs00200681_m1	Target	40	
4	Control	ADORA1-Hs00181231_m1	Target	26.8935	
5	Control	ADORA2A-Hs00169123_m1	Target	29.33256	
6	Control	ADORA2B-Hs00386497_m1	Target	30.41094	
7	Control	ADORA3-Hs00181232_m1	Target	30.95388	
8	Control	ADRA1A-Hs00169124_m1	Target	28.47502	
9	Control	ADRA1B-Hs00171263_m1	Target	28.9379	
10	Control	ADRA1D-Hs00169865_m1	Target	30.26497	
11	Control	ADRA2A-Hs00265081_s1	Target	29.97208	
12	Control	18S	Endogenous	9.516792	

Data needs to be in above format



- Column headers need to be exactly “Sample”, “Detector”, “Task”, and “Ct Avg”
- Order of column headers does not matter
- Values in “Sample”, “Detector”, and “Ct Avg” can be anything, but “Task” needs to be either “Target” or “Endogenous”
- If Detector value format is gene symbol-identifier or synonym(gene symbol)-identifier, GeneSpring creates Synonym and Gene Symbol columns in technology

Experimental Design of Gene Expression Dataset

Background: Acute lung injury (ALI) is a condition that requires mechanical ventilation with subsequent risk for ventilator-associated lung injury (VALI)

Aim: To elucidate the underlying mechanisms of ventilator-associated lung injury (VALI)

Experimental Design: Affymetrix GeneChip Rat230_2 microarrays

	Control	High-tidal Volume Ventilation
VALI-Resistant Rats		
VALI-Sensitive Rats	