

Labkey User's Meeting

Richard Green

greener@uw.edu

November 14th 2011

Outline

- Background
- Pepsi challenge
- Labkey in a production system
- Katze lab integration of labkey server:
some real world examples
- Customizing labkey tools
- R/Bioconductor tools
- Future directions

Why We Use LabKey Server

- A development platform to archive and mine data (proper data dissemination)
- To answer biological questions
- We don't always know the questions we will have: We often revisit and repurpose data

Why We Use LabKey Server

- We no longer work in silos



Having a platform where you and your collaborators can exchange data/analysis is necessary

Extend the Platform

- To submit data in the public repositories and integrate information out of the repositories (GEO for microarray and PRIDE for proteomics) * Focused in grants
- Web analytics is necessary
- Meta data will become the next challenge
- Data integration (protein, arrays, NGS, more)
- Data sets will only grow larger and more complex (NGS)

Pepsi Challenge

- SBEAMS and CPAS
- Demo installations in Fall 2005



Canary Grant 2007

- Expanded to operating in a production environment
- Integration with high performance cluster
- Customized proteomics pipelines
- Proteomics: development in analyzing data sets
- Beginning of the enterprise pipeline
- Plan ahead for data size



Katze Lab

- 3 installations of labkey: Katze informatics, Systems Virology, RIG-I adjuvant contract
- LIMS-like system: Tracking experimental design, custom wikis for results, procedures, tutorials
- Recording multiple data types: Proteomics, Genomics, NGS: meta data

Manuscript Tracker

- Uniform way to track all manuscripts generated by Katze Lab
- 1) Tracker queue shows manuscripts by status
 - 2) Discussion board for revisions of manuscripts and discussions

Manuscript Tracker

With this information, proteins and RNA polymerase accessory factors

Defining the Genomic Correlates for Swine Influenza H1N1 Pathogenicity in Mice

Bryan, Janine

Chang, Stewart

Compendium paper

Cilloniz, Cristian

FUNCTIONAL GENOMICS REVEALS AN ESSENTIAL AND SPECIFIC ROLE FOR STAT1 IN PROTECTION OF THE CENTRAL NERVOUS SYSTEM FROM VIRUS INFECTION

Transcriptional signatures associated with Ebola virus pathogenicity

Das, Trina

Diamond, Deborah

Jones, Daniel

Josset, Laurence

Korth, Marcus

Krasnoselsky, Alexei

Law, Lynn

Li, Yu

Navare, Arti

Palermo, Bob

Peng, Xinxia

Pereira, Arema

Messages

NEW


Showing: all messages

Edits on 2011-3-29 draft Marcus Korth 2011-03-30 16:53

Cristian,

Here are my edits. I trimmed some material from the Discussion that I thought was repetitive. Once you accept the Track Changes, I'd say this is good to go.

Marcus






 2011-3-29 MJK FINAL VERSION.docx

[VIEW MESSAGE OR RESPOND >](#)

Final Ebola version Cristian Cillóniz 2011-03-29 14:53






Hello Marcus,

Attached is the latest version of the Ebola manuscript.

 Supplementary Table 1.xlsx  Supplementary Table 2.xlsx  Supplementary Table 3.xlsx  FINAL VERSION.docx  FIGURES Final Version.docx

[VIEW MESSAGE OR RESPOND >](#)


Mouse Ebola manuscript Cristian Cillóniz 2011-03-15 15:04

 Supplementary Table 1.xlsx  Supplementary Table 2.xlsx  Supplementary Table 3.xlsx  Fifth manuscript draft.docx  Figures 3.11.11.docx

[VIEW MESSAGE OR RESPOND >](#)

Edits and comments on 2011-3-8 draft (1 response) Marcus Korth 2011-03-10 14:06

Cristian, Please accept track changes and address comments. I'm looking forward to the revised manuscript. Marcus

 MJK_2011-3-8_Fourth manuscript draft.docx






[VIEW MESSAGE OR RESPOND >](#)

Fourth Ebola manuscript Cristian Cillóniz 2011-03-08 15:25

Hello Michael and Marcus,

Attached is the fourth manuscript draft of the Ebola study.

Thank you

 Supplementary Table 1.xlsx  Supplementary Table 2.xlsx  Supplementary Table 3.xlsx  Fourth manuscript draft.docx  Figures 3.8.11.docx

[VIEW MESSAGE OR RESPOND >](#)

Edits and comments on 2011-2-10 draft Marcus Korth 2011-02-16 10:21

- Screen shot dialogue
- Changes to figures, tables, edits and comments dated and identified by user

Systems Virology LabKey Development

Contract-wide
communication

Wiki, Message
Board

Issue
tracking

Account
management

Secure
login

Resource
management

Project/Core -
specific
communication

Announcement
boards

The screenshot displays the Systems Virology LabKey Development web interface. The top navigation bar includes links for Home, Contract Collaboration, Portal, Messages, Issues, and Wiki. The left sidebar contains a tree view of Project Folders and Projects. The main content area features a File Sharing section with a message about finding 'PAGES', a General Announcement Board with a recent announcement about a meeting agenda, and a Pages section with links to various documents and reports. A search bar is located at the bottom right.

FTP - File
sharing

Meeting
minutes

Protocols

Web
metrics

SV / MMIC Curated Database



PROJECT FOLDERS

- Mathematical Modeling & Informatics Core
- Probe
- Studies
- Curated_database**
- Members Only

PROJECTS

MANAGE PROJECT

MANAGE SITE

Home

Permanent Link

Support

Help

PROJECT FOLDERS

- Mathematical Modeling & Informatics Core
- Probe
- Studies
- Curated_database**
- Members Only

PROJECTS

MANAGE PROJECT

MANAGE SITE

Home

Permanent Link

Support

Help

Assay Dashboard

MMIC_DB_Curated_Data_base

#MMIC_SEARCH #SEARCH

KEYWORDS: Please Enter your key words:

Search by

Stringency

Submit Query

View all MMIC

MMIC_CDB

Assay Dashboard > MMIC_CDB Batches > MMIC_CDB Runs >

MMIC_CDB Results

Mathematical Modeling & Informatics Core Curated Database

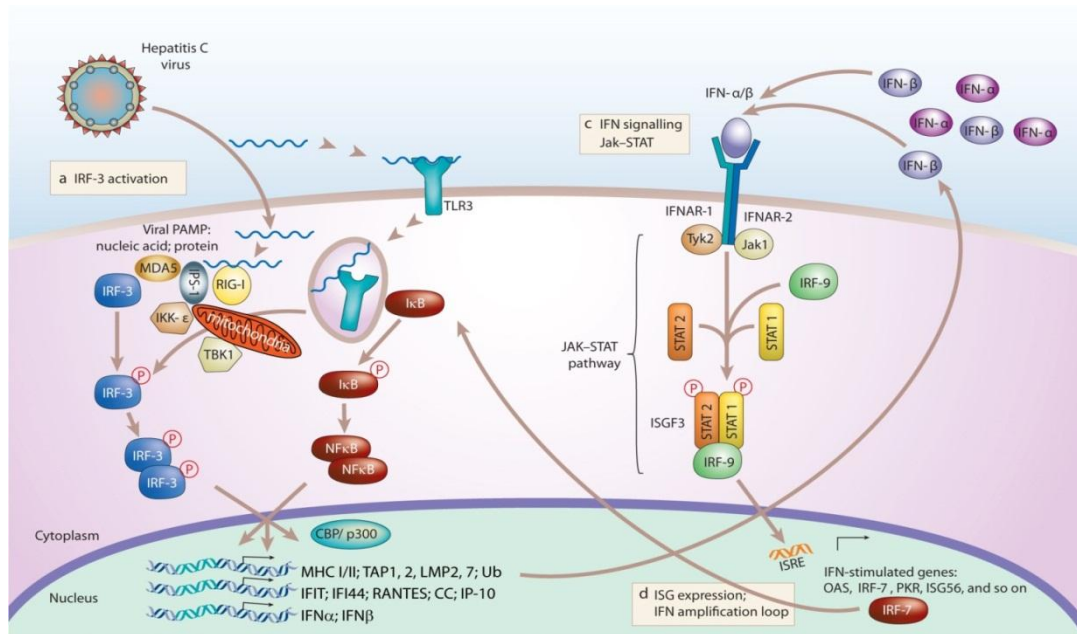
[MANAGE ASSAY DESIGN](#) > [VIEW BATCHES](#) > [VIEW RUNS](#) > [VIEW RESULTS](#) > [VIEW COPY-TO-STUDY HISTORY](#) >

VIEWS ▾ EXPORT ▾ PRINT PAGE SIZE ▾ COPY TO STUDY IMPORT DATA ▾											1 - 56 of 56
This view is filtered: (Run = 6965) Clear all filters											
<input type="checkbox"/>	Probe_ID	ENTREZ_ID	SIGNIFICANCE	PROTOCOL_USED	SYNONYMS	EXPERIMENTAL_METHOD	ORIGINATING_EXPERIMENT(S)	DIRECTION_OF_CHANGE	TISSUE_CONTEXT	CELL_LINE_CONTEXT	SPECIES DATE
<input type="checkbox"/>	A_23_P0001	1021	WGCNA	WGCNA	CXCL10	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010
<input type="checkbox"/>	A_23_P0002	1022	WGCNA	WGCNA	CXCL11	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010
<input type="checkbox"/>	A_23_P0003	1023	WGCNA	WGCNA	CXCL11	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010
<input type="checkbox"/>	A_23_P0004	1024	WGCNA	WGCNA	CXCL2	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010
<input type="checkbox"/>	A_23_P0005	1025	WGCNA	WGCNA	CXCL2	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010
<input type="checkbox"/>	A_23_P0006	1026	WGCNA	WGCNA	DDX58	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010
<input type="checkbox"/>	A_23_P0007	1027	WGCNA	WGCNA	DHX58	microarray	IM001	up regulated	Mouse Lung		Mouse 10/18/2010

- Allows SV MMIC members to query manually curated datasets
- Members upload datasets by text, Excel, or manually
- Stores results into database
- Why are these genes interesting? How were they identified? What statistical test used?

RIG-I Contract

- Find new ways to stimulate, enhance innate antiviral immunity against pathogenic RNA viruses.
- Screen candidate adjuvants for biological activity



Stores assay results:

- Flow data
- Luminex
- Microarray raw data

Customizing LabKey Tools



Custom Assays: UI

- User interfaces
- Content stored into the data base
- Metrics to be generated for future Grants or Contracts

Project Request Form

Please use the form below to request an NGS project. All starred fields are required.

Name is required.
Email is required.

Experiment Name:*

Project Managers
E-mail:*

Origin:*

Name of Grant:*

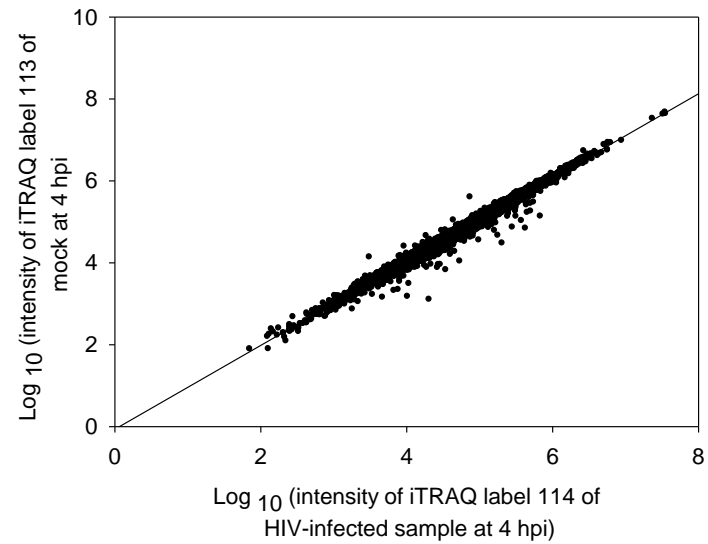
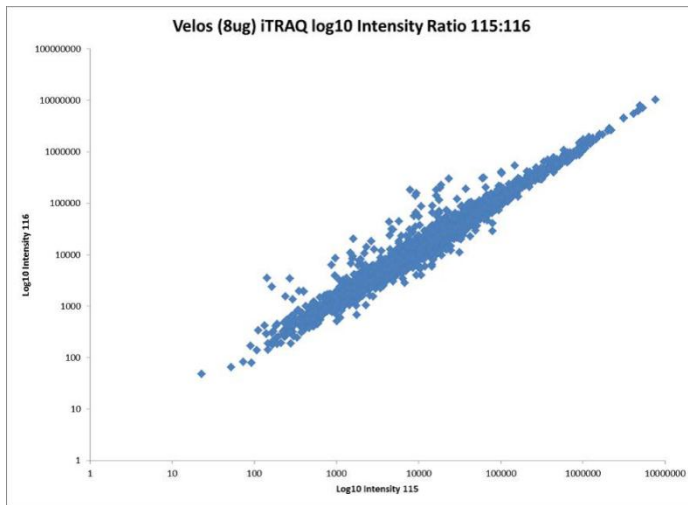
Budget Number:*

External_Collaborators:

Experimental
Overview:

Labkey Project:

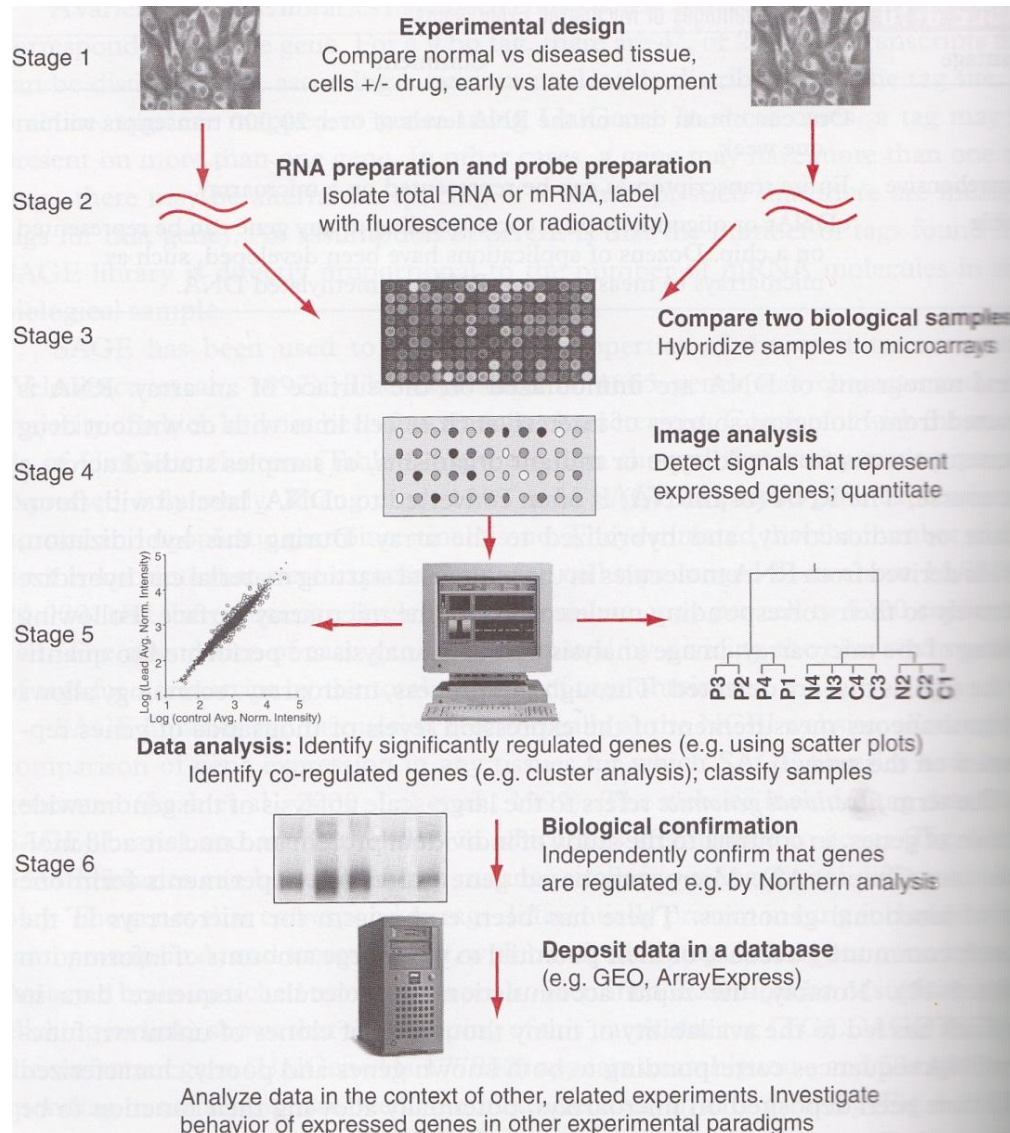
Integration of Libra in 11.3



1:1 correlation plot of all peptide raw intensity values for 2 reporter ions from Libra's quantitation output

Automation and Analysis in R and Bioconductor

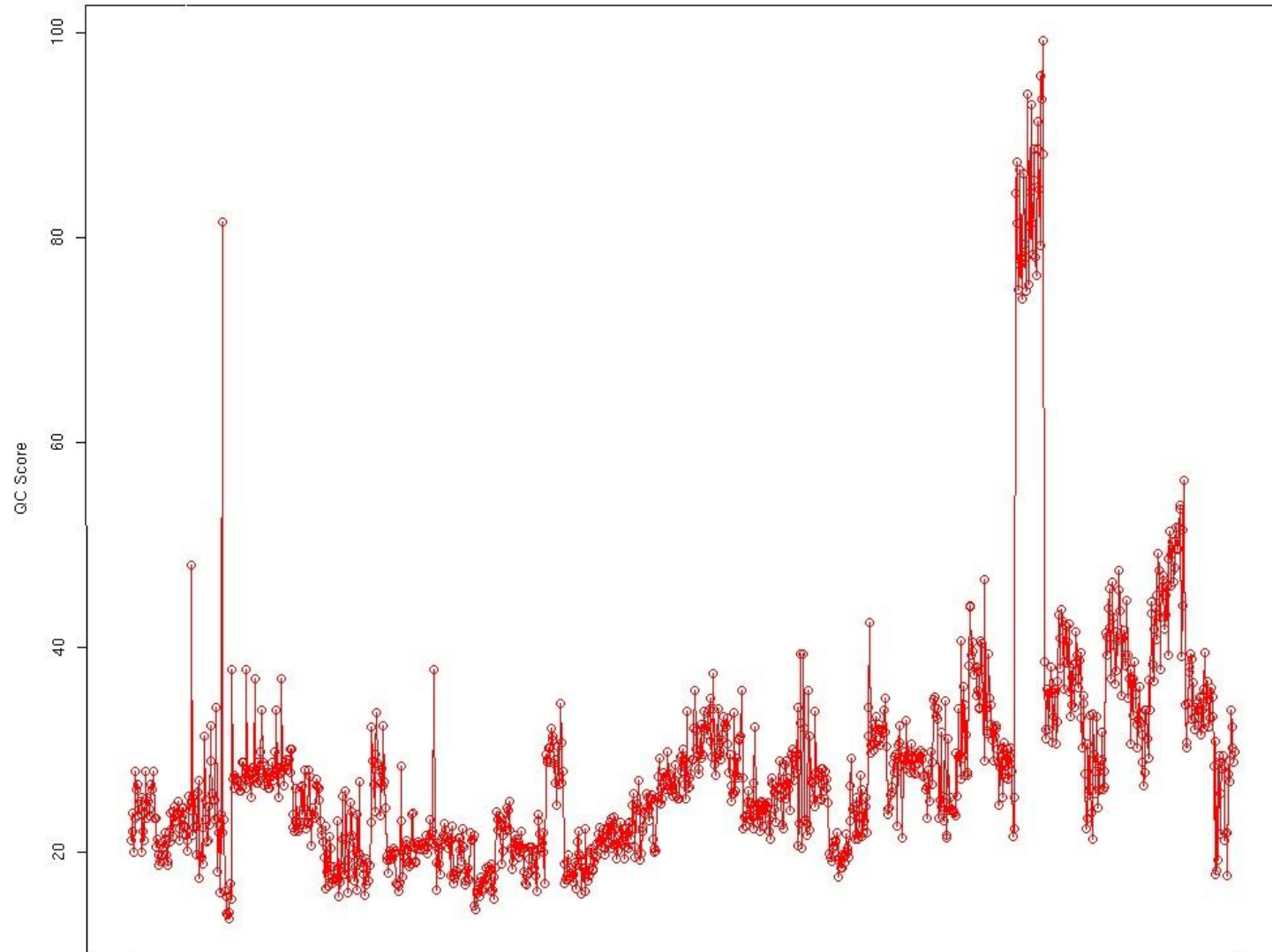
Standard MA Analysis Workflow



Bioinformatics
and Functional
Genomics, p.314
Jonathan Pevsner

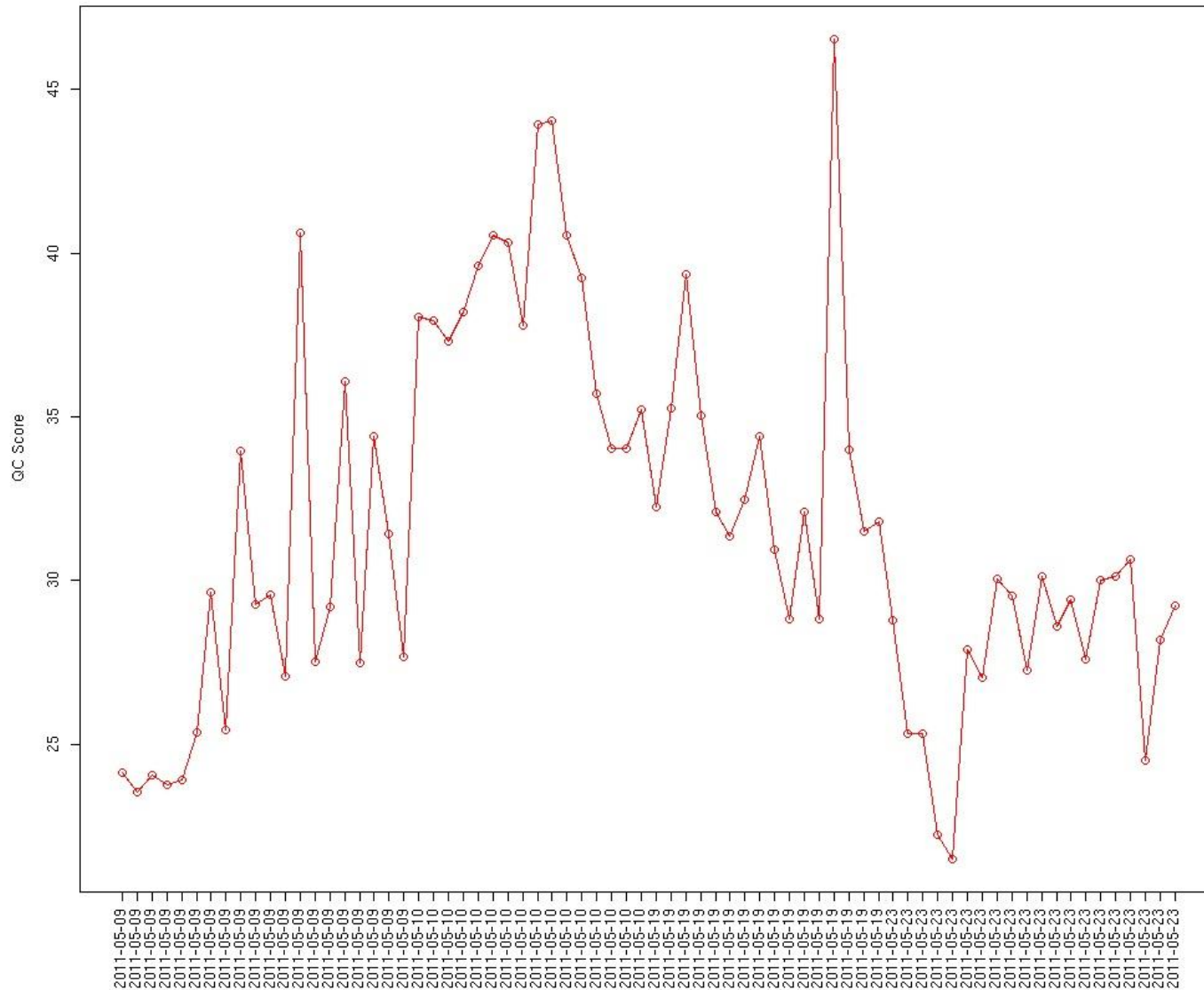
Plotting QC Scores

QC_neg_cntl vs Hyb Date: sort by date



Plotting QC Scores

QC_neg_cntl vs Hyb Date: sort by date



QC Analysis

- Macaque
- siRNA
 - Members Only
 - Shared Information
- Miscellaneous

PROJECTS

MANAGE PROJECT

MANAGE SITE

Home

Permanent Link

Support

Help

Agilent mRNA 1-Color Microarray v10.1 Runs

Single channel custom RNA microarray processing protocol specific for Katze lab processing procedures and built for the Labkey Server v10.1 platform.

[MANAGE ASSAY DESIGN](#) [VIEW BATCHES](#) [VIEW RUNS](#) [VIEW RESULTS](#) [VIEW COPY-TO-STUDY HISTORY](#)

VIEWS [PRINT](#)

default

All fields

for GDC file

QC view

Run View

QC_MA

Manage Views

☒ Apply View Filter

Folder Filter

ce: default

DOES NOT START WITH Mu_Ref_Br)

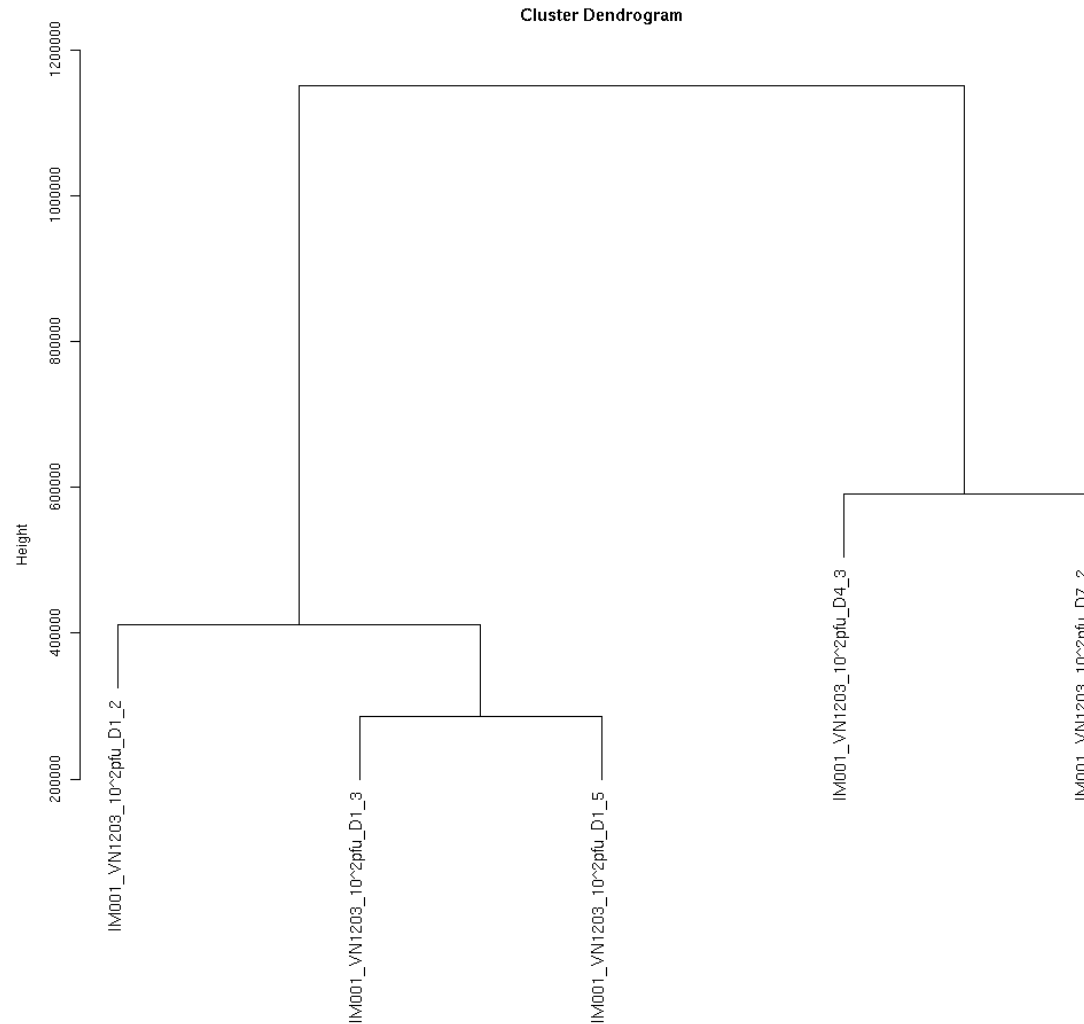
source

[Help](#)

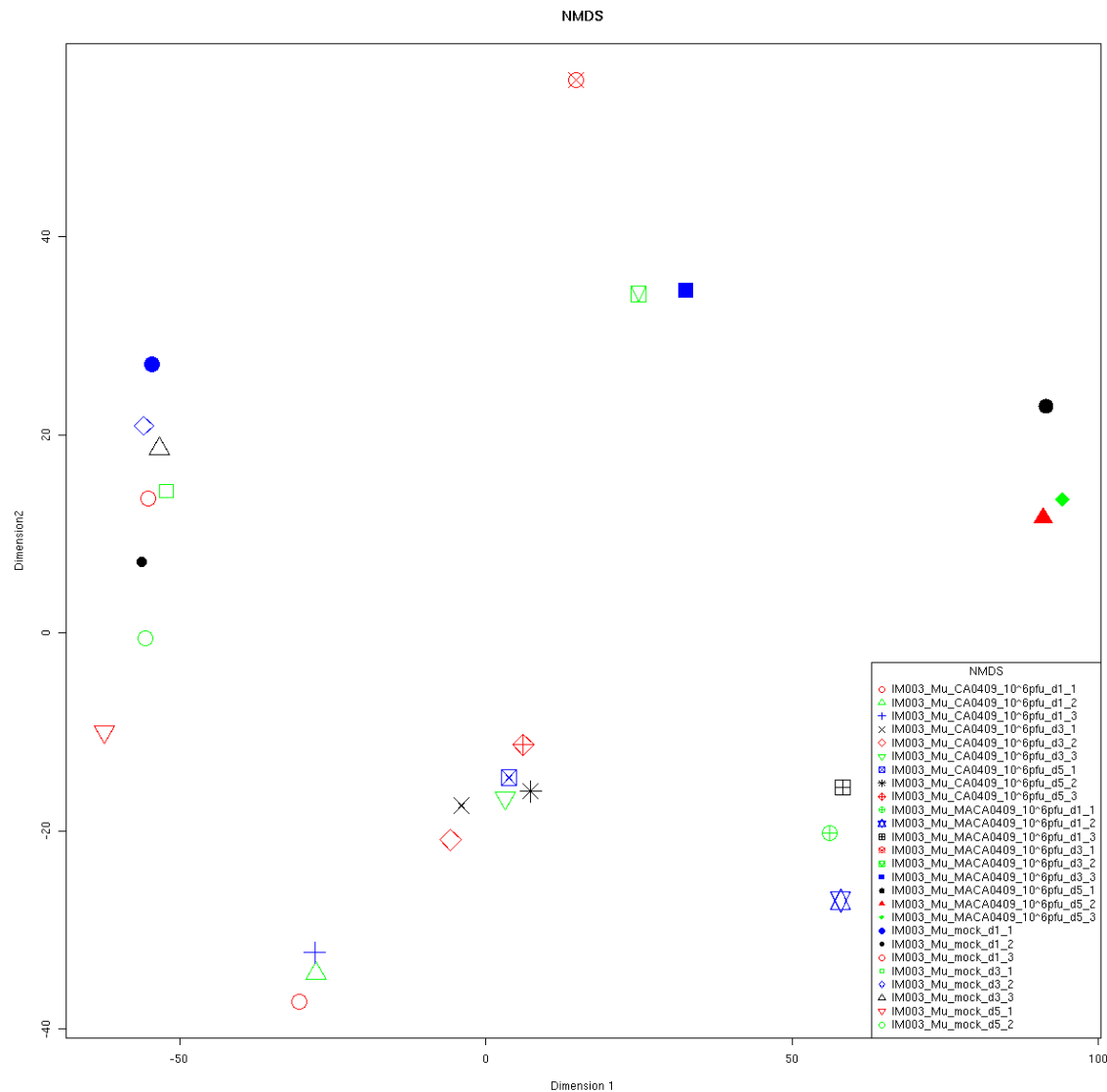
Cluster Dendrogram

QC Analysis

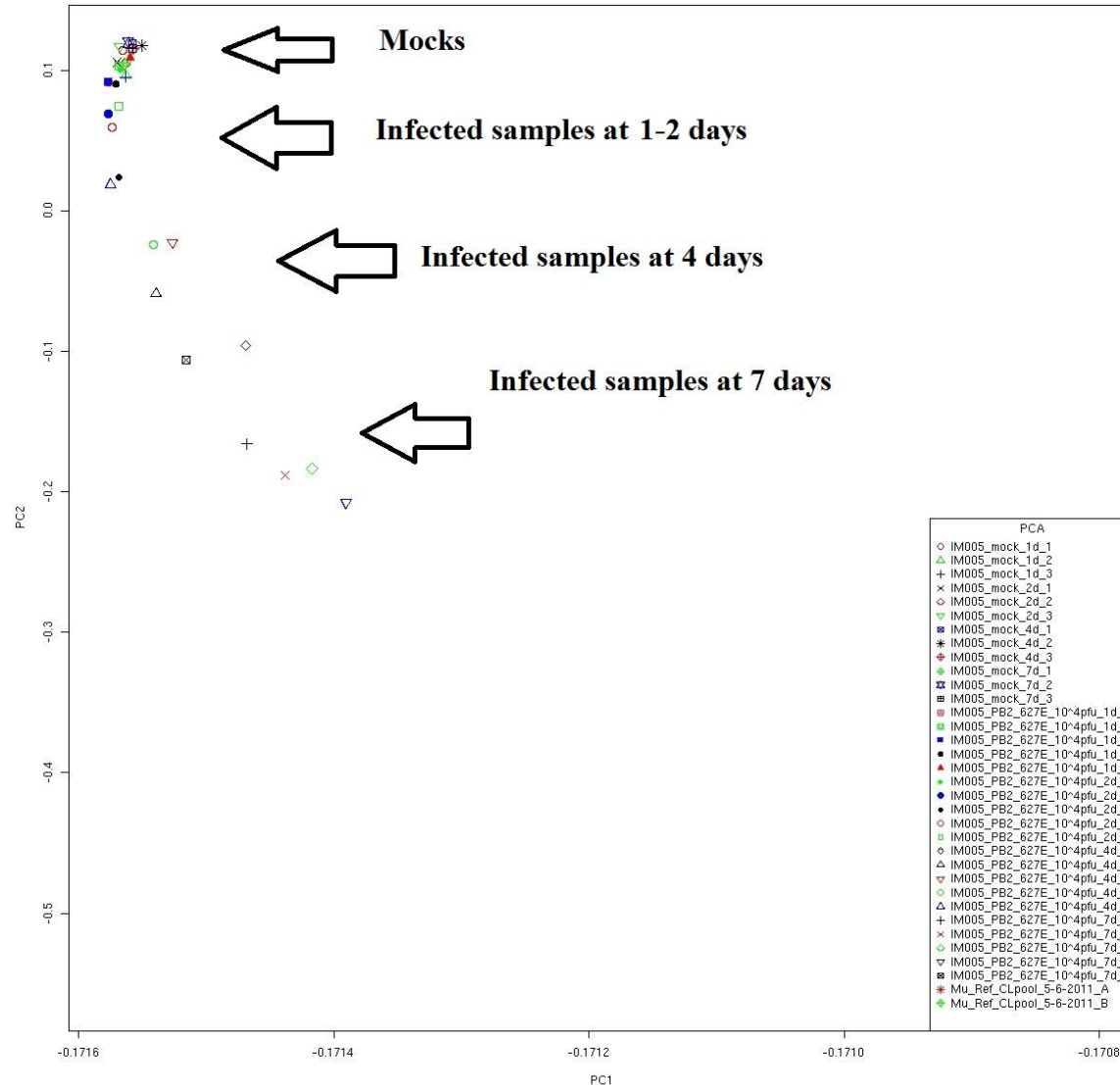
- Hierarchical Cluster Analysis



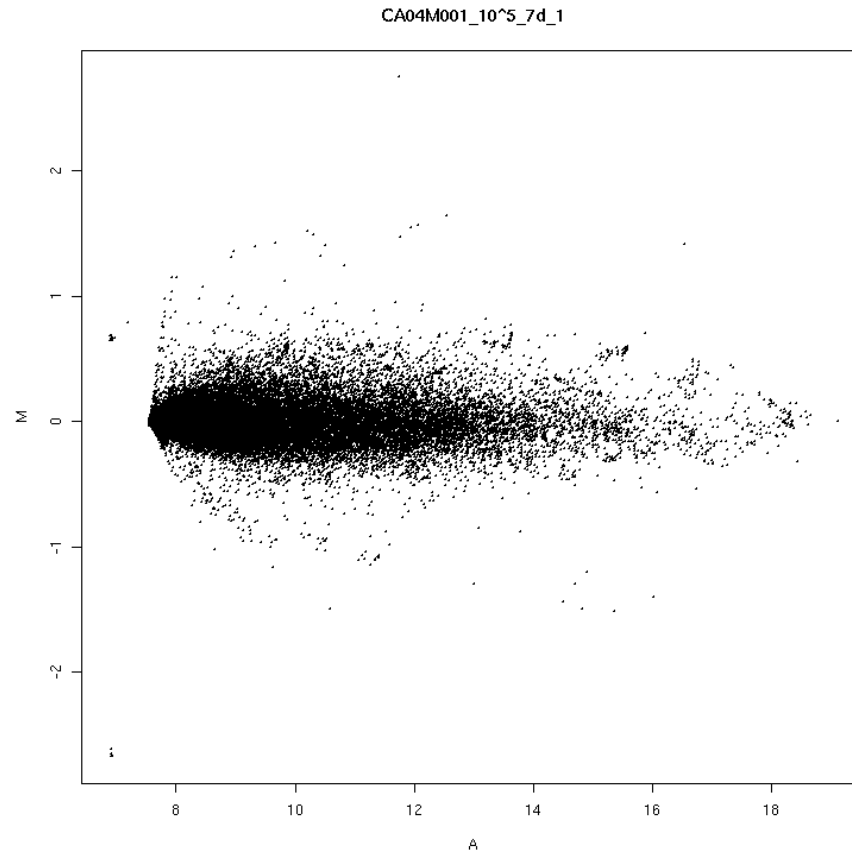
QC Analysis: NMDS



QC analysis: Principal Component Analysis



QC Analysis: MA Plots



To determine whether normalization is needed, the MA-plot is generated to plot of the distribution of the red/green intensity ratio ('M') plotted by the average intensity ('A').

GEO Submission

- Prepare the microarray data to be pushed into the public repository

dose and time course experiment

IM007 - VN1203-NS1 trunc

IM008 - NGS Study 1 - VN1203 and VN1203 PB2 627E

IM009 - Repeat of IM003

IM010: TNFSFR1b and IDO1 KOs

IM011: LTB4R1 KO

Macaque

siRNA

Members Only

Shared Information

Miscellaneous

PROJECTS

IMAGE PROJECT

IMAGE SITE

ome

ermanent Link

upport

elp

PROCESS AND IMPORT DATA **SETUP**

Status	Created	Description
No data to show.		

P default

- All fields
- for GDC file
- QC view
- Run View

A

- QC AnyColor PrntFeat NonUnifOL N/A-1.00
- QC eQC OneColor LinFitSlope 0.90-1.20
- QC eQC OneColorSpike DetectionLimit 0.0.-2.00
- QC geQC MedPrntCV ProcSignal N/A-8.0
- QC gMedPrnt CVProcSignal N/A-8.00
- QC gNegCtrl AveBG SubSig -10.00-5.00
- QC gNegCtrl AveNetSig N/A-40.00
- QC gNegCtrl AveNetSig N/A-40.00-by-hyb-date
- QC gNegCtrl SDevBG SubSig N/A-10.00
- QC gSpatialDetrend RMSFiltered MinusFit N/A-15.00
- QC IsGoodGrid
- QC_metrics1
- write_normalized_matrix2GEO

Create

Customize View

Manage Views

☒ Apply View Filter

FOR MAGEML FILES

Data File Url

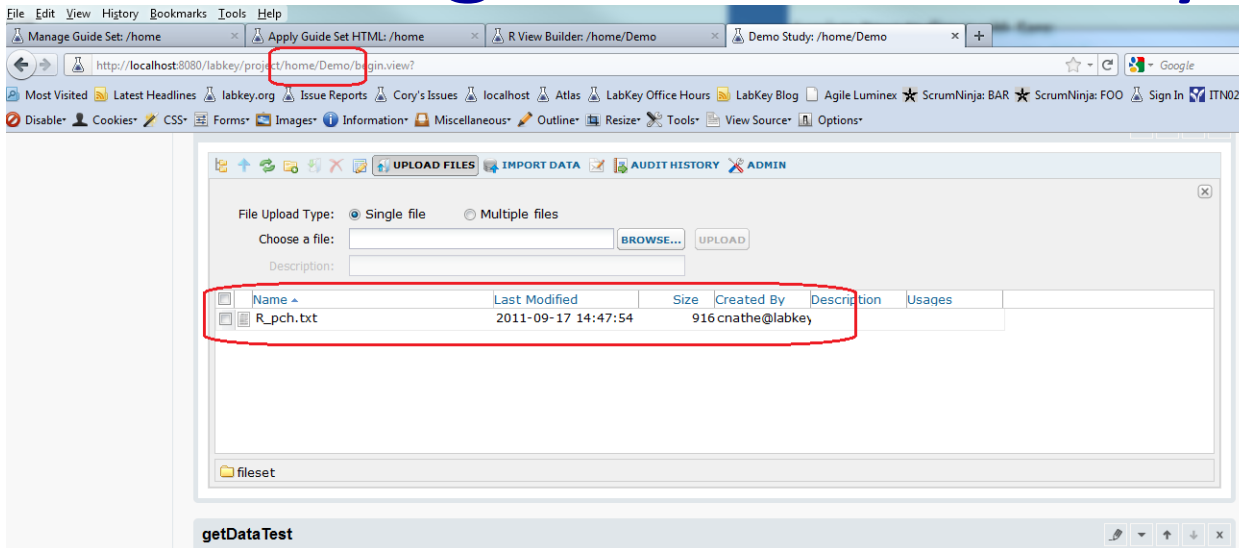
Runs

g protocol specific for Katze lab processing procedures and built for the Labkey Server v10.1 platform.

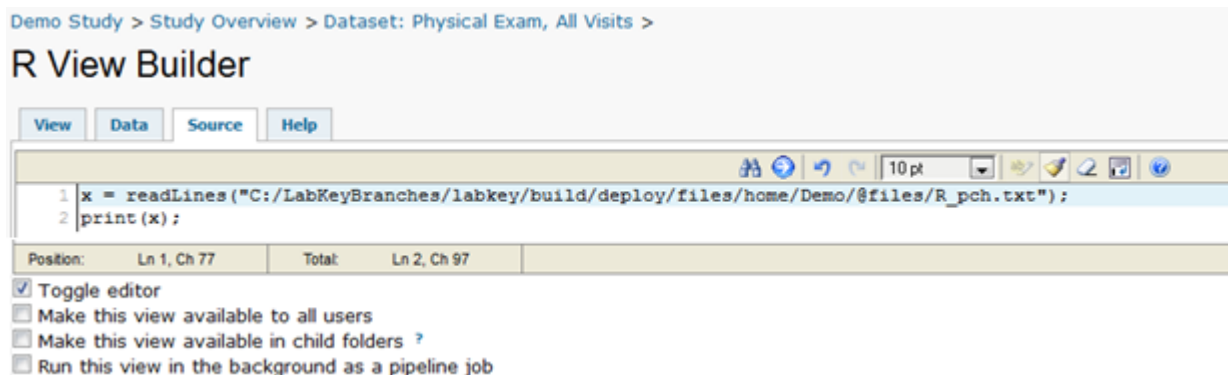
[/ RUNS](#) [VIEW RESULTS](#) [VIEW COPY-TO-STUDY HISTORY](#)

DELETE	ADD TO RUN GROUP	SHOW RESULTS	COPY TO STUDY	IMPORT DATA											
					Barcode	Batch #	Hyb Date	QC IsGoodGrid	QC AnyColor PrntFeat NonUnifOL N/A-1.00	QC eQC OneColorSpike DetectionLimit 0.0.-2.00	QC eQC OneColor LinFitSlope 0.90-1.20	QC geQC MedPrntCV ProcSignal N/A-8.0	QC gNegCtrl AveBG SubSig -10.00-5.00	QC gNegCtrl AveNetSig N/A-40.00	
					09_10^6pfu_d3_1	251486831424_1_2	2	2010-06-08		0.093296	1.2069	0.963964	3.305	-2.0379	19.4121
					10^6pfu_d1_3	251486831424_1_1	2	2010-06-08		0.137723	1.11533	0.970117	3.55831	-2.81099	19.5687
					10^6pfu d5 1	251486831423 1 4	2	2010-06-08		0.006664	1.11597	0.953432	3.44419	-1.43184	20.1616

Uploading a Configuration File Can Strengthen Our Analysis Power



- Preprocess 44x4K Agilent (target file)
- Limma (contrast matrix)



How Analysis Has Moved Forward: LabKey's Support

- Constant development on new tools
- Pushing integrated tools into a open source platform that is available to all
- Consistent support : labs of 4 to 43, either through a service contract or on the support boards online

Future Steps

- Additional GEO customization
- Adapting additional Bioconductor tools into our automated pipeline
- Further NGS integration into labkey
- NGS pipeline: (Galaxy server, QC methods)
- Visualizing libra output through R/bioconductor
- R-cytoscape
- Kineta: Integration of chemistry schema into postgres database

Acknowledgements

- Cory Nathe
- Josh Eckels
- Tim Owens
- Katze lab