Skyline & Panorama:
Key Tools for Establishing a Targeted LC/MS Workflow

Genentech
A Member of the Roche Group

Kristin Wildsmith
Scientist, Biomarker Development
LabKey User Conference
October 6, 2016

©2012, Genentech

Biomarkers enable drug development

Drug development

Research → Pre-IND → Phase I → Phase II → Phase III (Pivotal) → Clinical Practice

Biomarker development

Biomarker Discovery → **Clinical Practice**

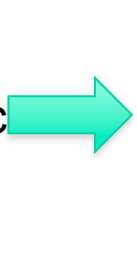
Clinical validation

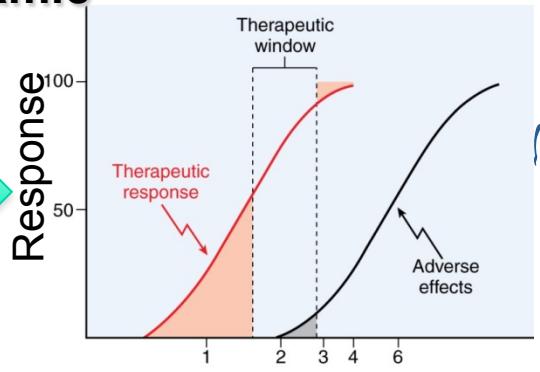
Technical development

Preclinical validation

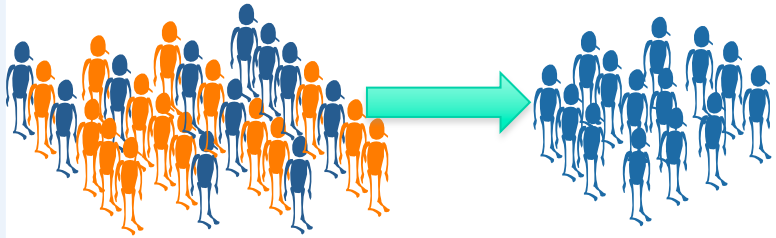
Pharmacodynamic

Therapeutic





Diagnostic



©2012, Genentech

Mass spectrometry is a quantitative tool for biomarker development and Labkey server is used for data management

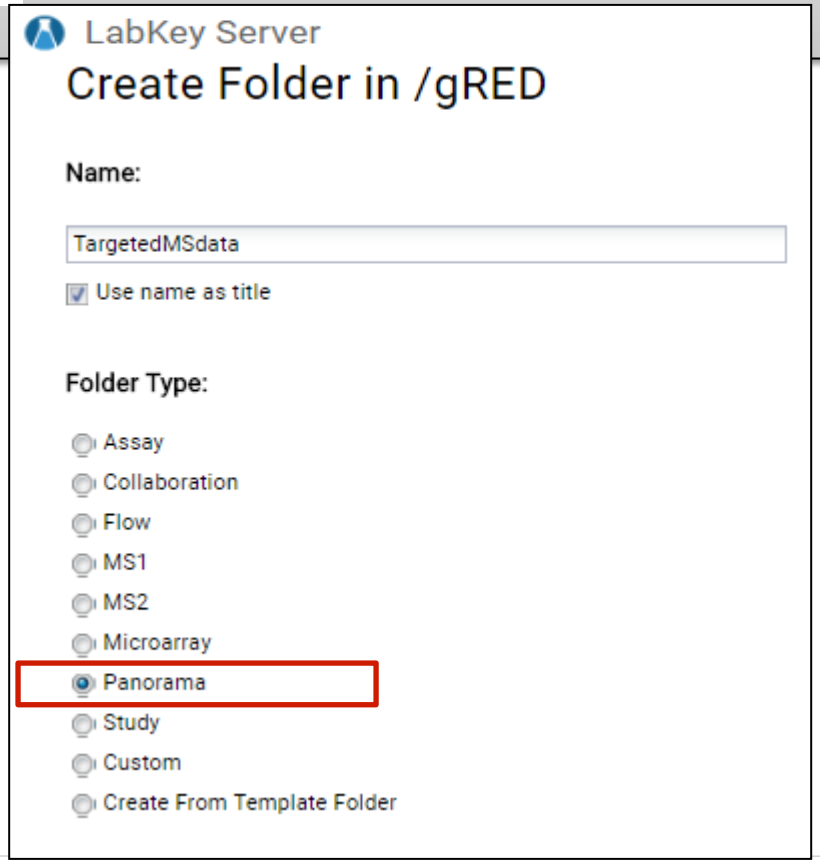
Biomarker Discovery | Clinical validation | Clinical Practice

LTQ-Orbitrap
(High resolution)

- Discovery (unbiased)
- Targeted-discovery

- Multiplex panel
- Single analyte quantitation

Triple Quad
(High sensitivity)

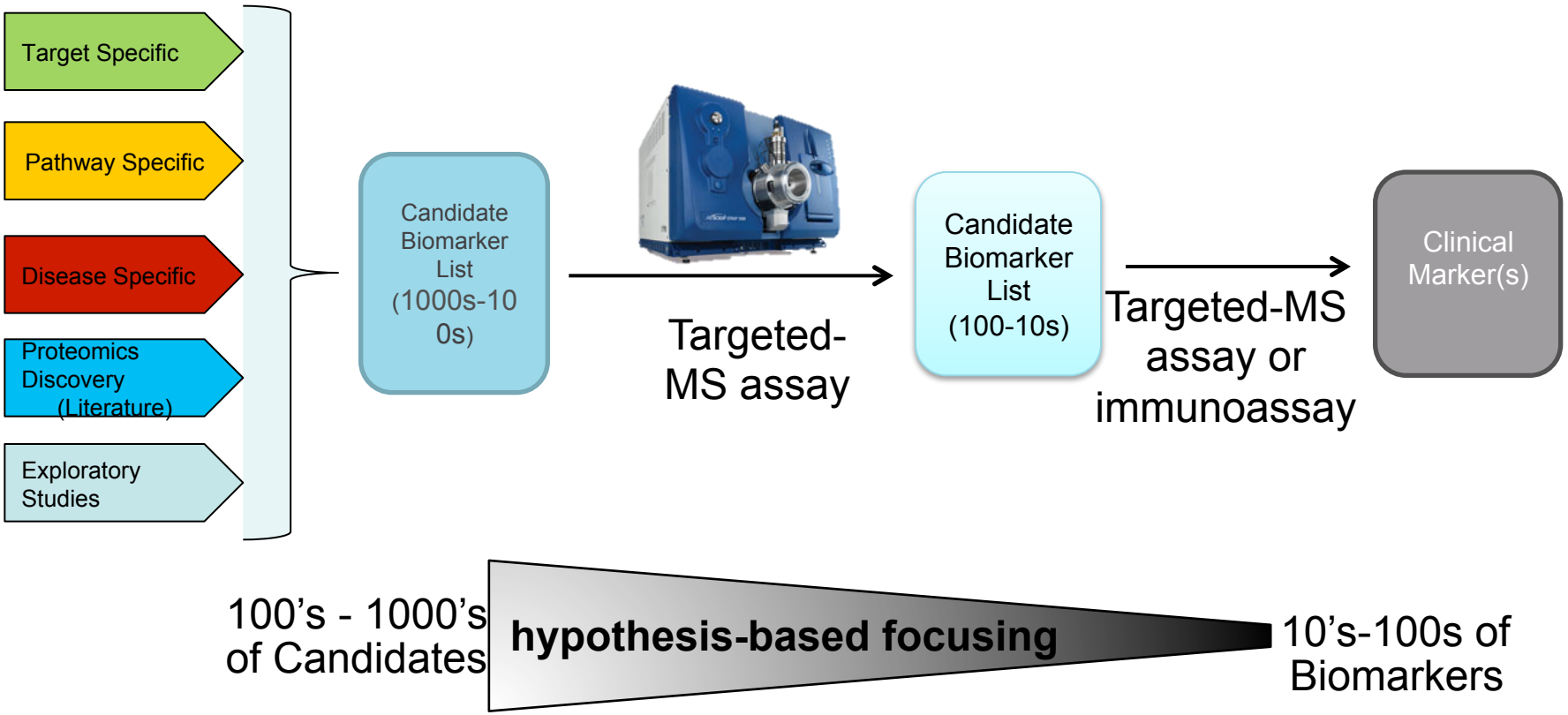


Labkey is a data management solution for MS data

Development of customized folders accelerated Targeted MS assay development

©2012, Genentech

New model for developing protein clinical biomarkers: Hypothesis driven multiplexing with mass spectrometry



MS can provide a faster and more successful translation of protein biomarkers to the clinic

©2012, Genentech

Genentech
A Member of the Roche Group

Multiplexed MRM assay developed for candidates

Candidate proteins in CSF

protease digestion

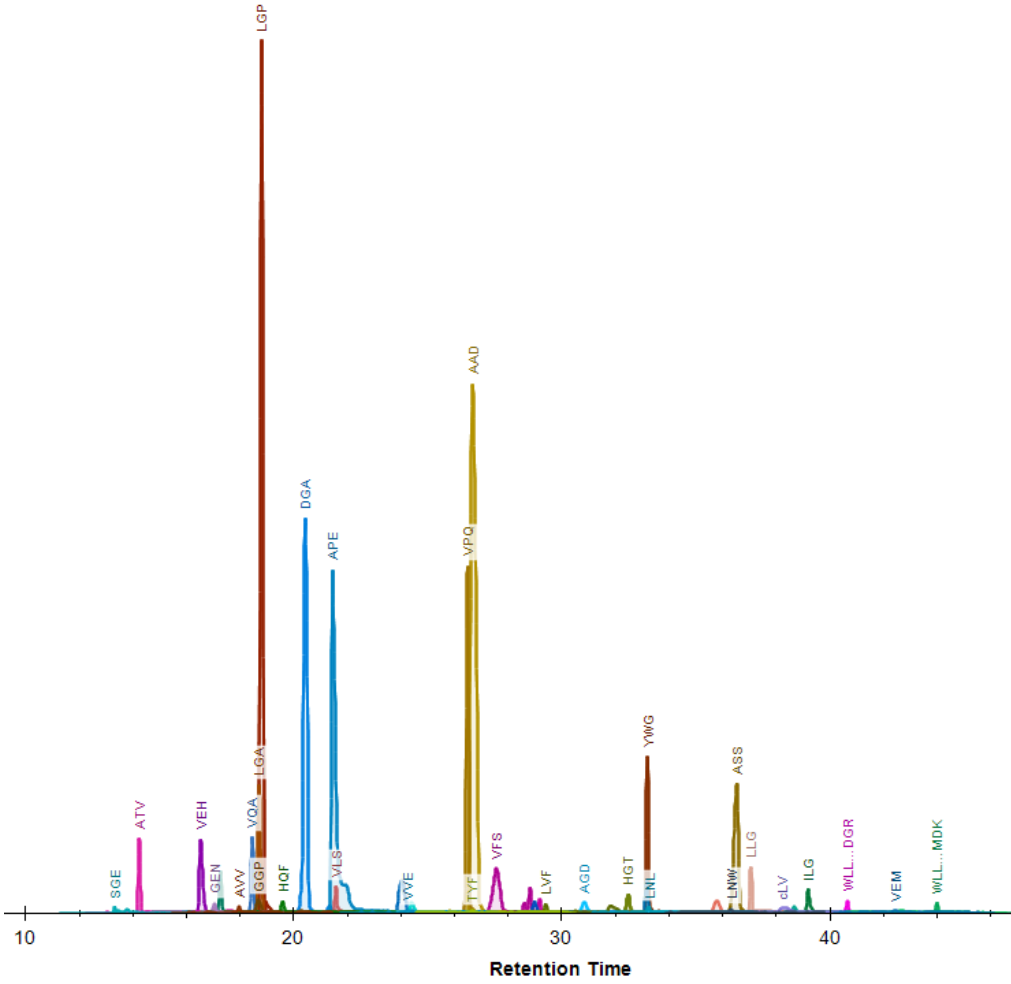
Signature peptides

+internal standards (heavy peptides)

LC-MS/MS

Quantitation

Multiplex



©2012, Genentech

Genentech
A Member of the Roche Group

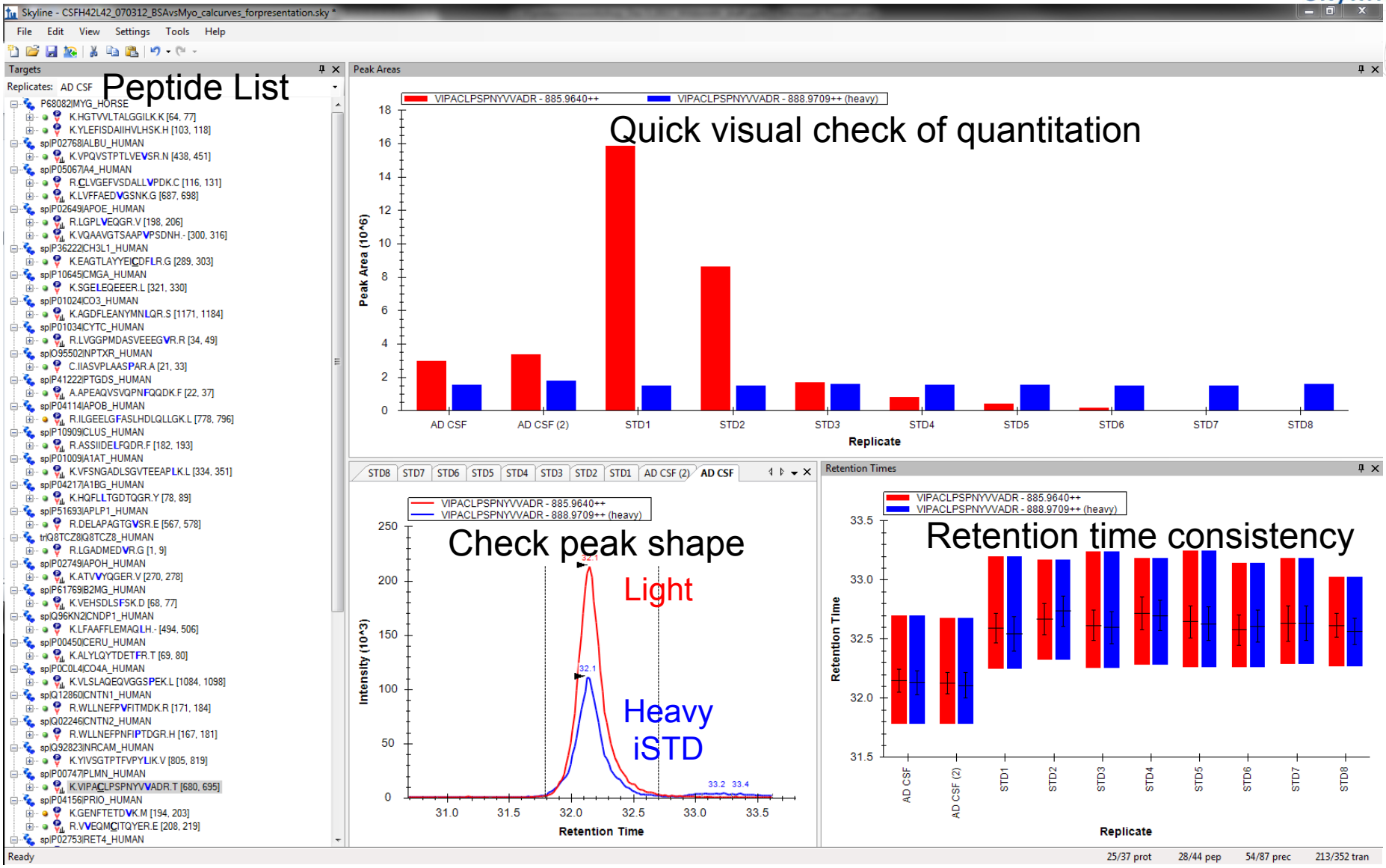
Skyline supports targeted MS techniques

- Skyline is an open-source software from MacCoss lab at University of Washington
- Vendor neutral
- Supports method development and data analysis:
  - Targeted MS/MS
    - Selected reaction monitoring (SRM, aka MRM)
    - Parallel reaction monitoring (PRM)
  - Data-independent-acquisition (DIA) / SWATH
  - Targeted discovery with MS^1 quantitative methods
  - Discovery proteomics with MS^2 methods

©2012, Genentech

Open-source software available from MacCoss lab at University of Washington:
https://brendanx-uw1.gs.washington.edu/labkey/project/home/software/Skyline/begin.view

Multiplexing enabled by targeted-proteomics software



Open-source software available from MacCoss lab at University of Washington:
https://brendanx-uw1.gs.washington.edu/labkey/project/home/software/Skyline/begin.view

©2012, Genentech

Genentech
A Member of the Roche Group

Old workflow for multiplexed MRM creates data management challenges

Target Proteins

Select Peptides & build MRM method

Skyline

Generate Data

MQ MultiQuant 2.1

Peak integration and calibration curve generation

QC sample data

TIBCO Spotfire

- Iterative method development
- Complex, MANUAL data analysis
- Large data files from multiple platforms

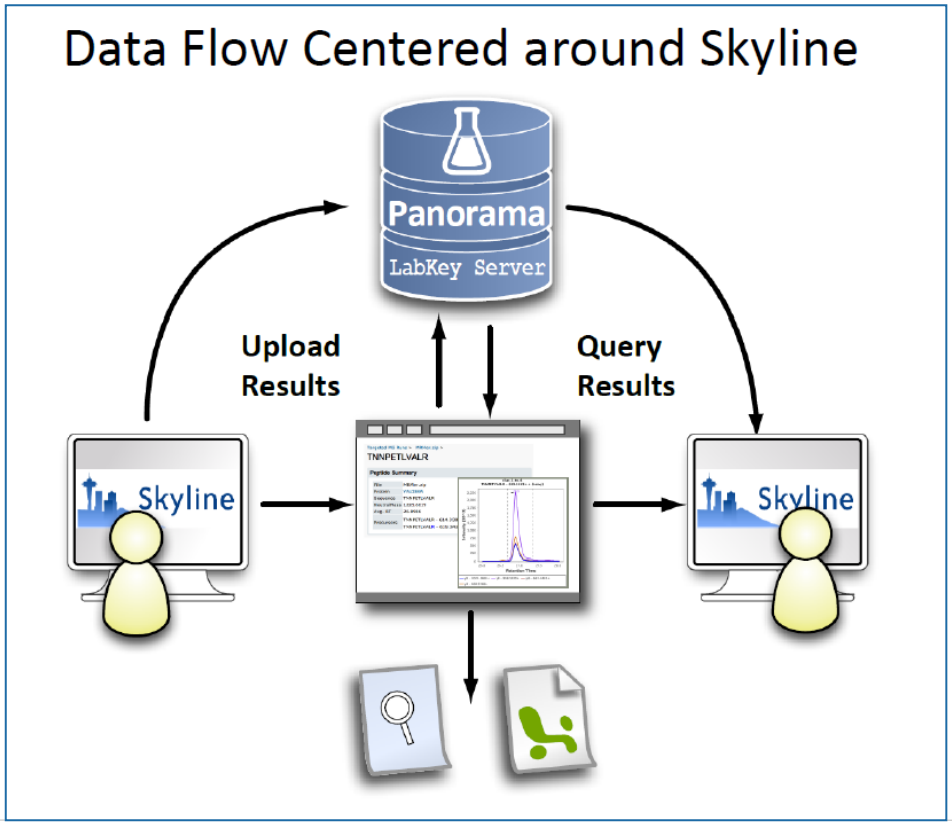
Statistical analysis and data visualizations

©2012, Genentech

Data management simplification with Panorama

Panorama is a repository server application supporting Skyline

- Module within LabKey server
- Open-source bioinformatics data management platform
- Searchable database storing Skyline data files

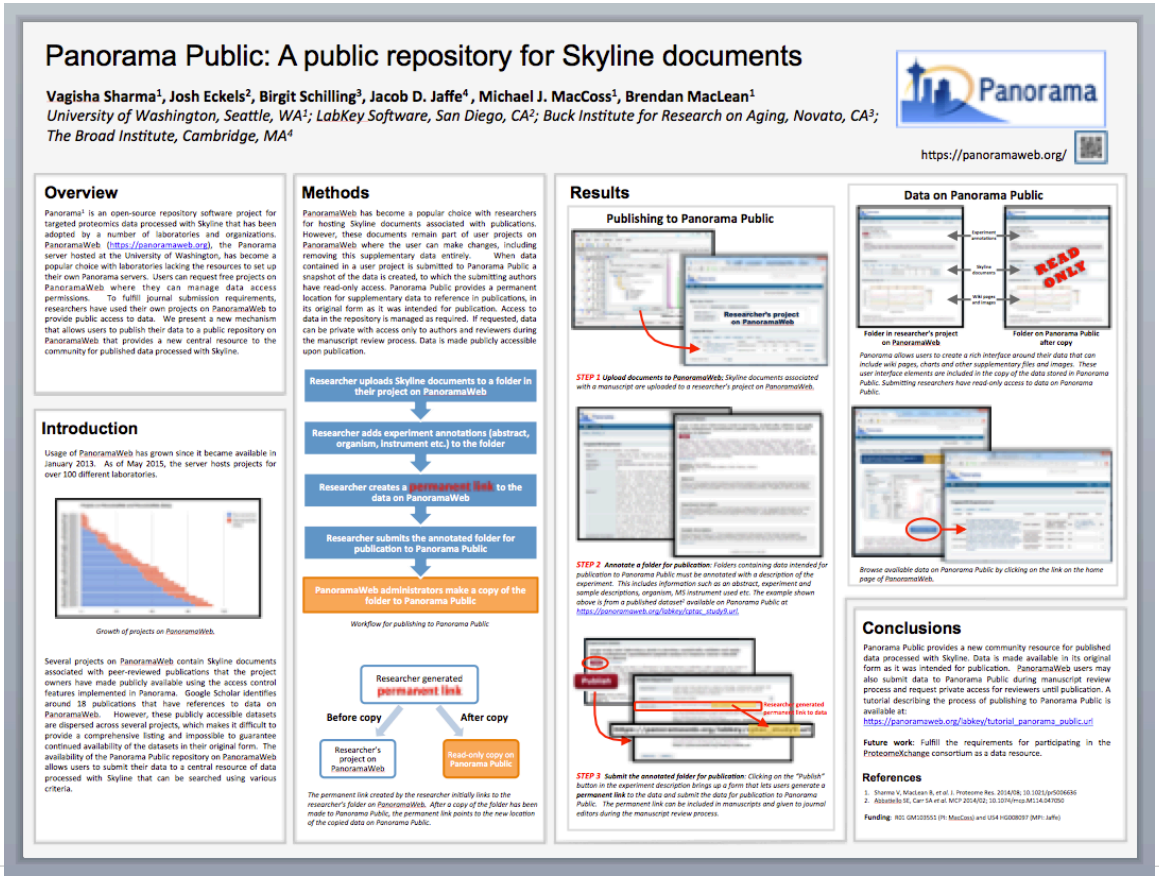


©2012, Genentech

Genentech
A Member of the Roche Group

Panorama increases rate of method development

- MacCoss-lab hosted instance (includes CPTAC, ABRF shared libraries)
- Shares read-only data to support publications
- Workflow for sharing data with reviewers and then general public



Panorama Public: A public repository for Skyline documents

Vagisha Sharma[1], Josh Eckels[2], Birgit Schilling[3], Jacob D. Jaffe[4], Michael J. MacCoss[1], Brendan MacLean[1]
University of Washington, Seattle, WA[1]; LabKey Software, San Diego, CA[2]; Buck Institute for Research on Aging, Novato, CA[3]; The Broad Institute, Cambridge, MA[4]

https://panoramaweb.org/

Overview

Panorama[1] is an open-source repository software project for targeted proteomics data processed with Skyline that has been adopted by a number of laboratories and organizations. PanoramaWeb (https://panoramaweb.org), the Panorama server hosted at the University of Washington, has become a popular choice with laboratories lacking the resources to set up their own Panorama servers. Users can request free projects on PanoramaWeb where they can manage data access permissions. To fulfill journal submission requirements, researchers have used their own projects on PanoramaWeb to provide public access to data. We present a new mechanism that allows users to publish their data to a public repository on PanoramaWeb that provides a new central resource to the community for published data processed with Skyline.

Introduction

Usage of PanoramaWeb has grown since it became available in January 2013. As of May 2015, the server hosts projects for over 100 different laboratories.

Growth of projects on PanoramaWeb.

Several projects on PanoramaWeb contain Skyline documents associated with peer-reviewed publications that the project owners have made publicly available using the access control features implemented in Panorama. Google Scholar identifies around 18 publications that have references to data on PanoramaWeb. However, these publicly accessible datasets are dispersed across several projects, which makes it difficult to provide a comprehensive listing and impossible to guarantee continued availability of the datasets in their original form. The availability of the Panorama Public repository on PanoramaWeb allows users to submit their data to a central resource of data processed with Skyline that can be searched using various criteria.

Methods

PanoramaWeb has become a popular choice with researchers for hosting Skyline documents associated with publications. However, these documents remain part of user projects on PanoramaWeb where the user can make changes, including removing this supplementary data entirely. When data contained in a user project is submitted to Panorama Public a snapshot of the data is created, to which the submitting authors have read-only access. Panorama Public provides a permanent location for supplementary data to reference in publications, in its original form as it was intended for publication. Access to data in the repository is managed as required. If requested, data can be private with access only to authors and reviewers during the manuscript review process. Data is made publicly accessible upon publication.

Workflow for publishing to Panorama Public

The permanent link created by the researcher initially links to the researcher's folder on PanoramaWeb. After a copy of the folder has been made to Panorama Public, the permanent link points to the new location of the copied data on Panorama Public.

Results

Publishing to Panorama Public

STEP 1 Upload documents to PanoramaWeb: Skyline documents associated with a manuscript are uploaded to a researcher's project on PanoramaWeb.

STEP 2 Annotate a folder for publication: Folders containing data intended for publication to Panorama Public must be annotated with a description of the experiment. This includes information such as an abstract, experiment and sample descriptions, organism, MS instrument used etc. The example shown above is from a published dataset[2] available on Panorama Public at https://panoramaweb.org/labkey/cptac_study9.url

STEP 3 Submit the annotated folder for publication: Clicking on the "Publish" button in the experiment description brings up a form that lets users generate a permanent link to the data and submit the data for publication to Panorama Public. The permanent link can be included in manuscripts and given to journal editors during the manuscript review process.

Data on Panorama Public

Panorama allows users to create a rich interface around their data that can include wiki pages, charts and other supplementary files and images. These user interface elements are included in the copy of the data stored in Panorama Public. Submitting researchers have read-only access to data on Panorama Public.

Browse available data on Panorama Public by clicking on the link on the home page of PanoramaWeb.

Conclusions

Panorama Public provides a new community resource for published data processed with Skyline. Data is made available in its original form as it was intended for publication. PanoramaWeb users may also submit data to Panorama Public during manuscript review process and request private access for reviewers until publication. A tutorial describing the process of publishing to Panorama Public is available at: https://panoramaweb.org/labkey/tutorial_panorama_public.url

Future work: Fulfill the requirements for participating in the ProteomeXchange consortium as a data resource.

References

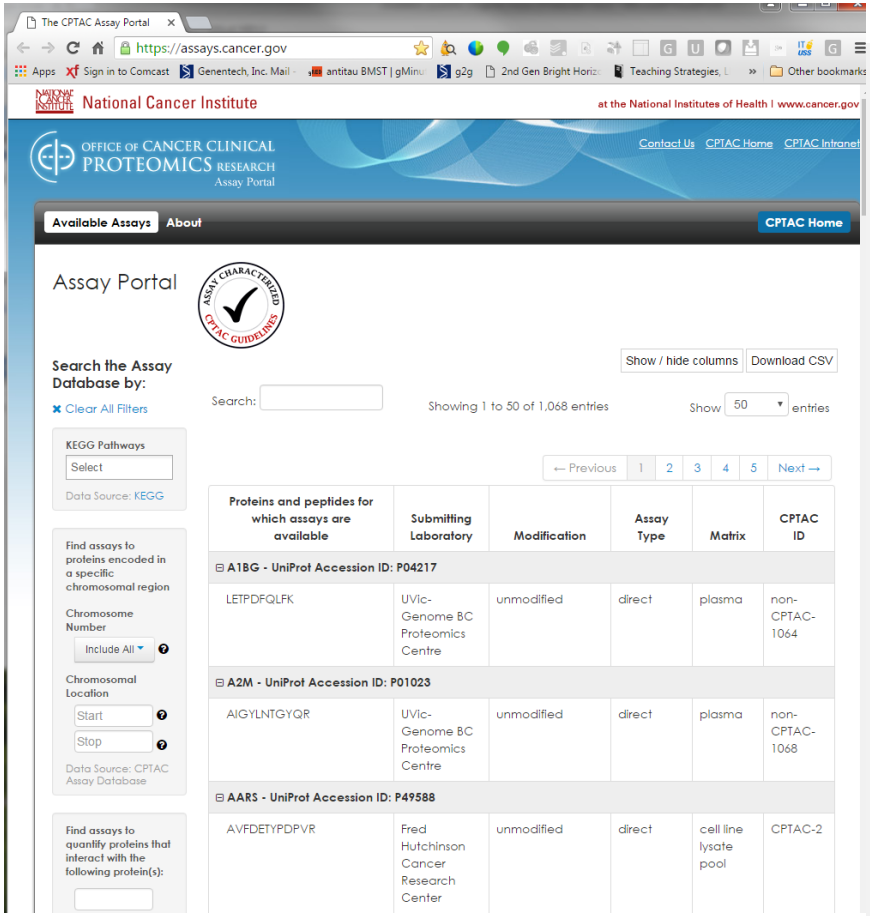
1. Sharma V, MacLean B, *et al.* J. Proteome Res. 2014/08; 10.1021/pr5006636
2. Abbatiello SE, Carr SA *et al.* MCP 2014/02; 10.1074/mcp.M114.047050

Funding: R01 GM103551 (PI: MacCoss) and U54 HG008097 (MPI: Jaffe)

©2012, Genentech

Genentech
A Member of the Roche Group

Enabling data sharing



©2012, Genentech

Panorama Partnership enables use of Panorama in secure environment and expand capabilities

- Collaborate directly with MacCoss lab and LabKey teams
- Installed Labkey server on-site
- Guide Skyline & Panorama software development priorities
- All partners contribute to common development fund
- All partners give input on priorities and design decisions

- Genentech/Roche partners since Oct 2013 (1 of 4)

Field strongly supporting Skyline and Panorama development

Academic/Government:
2 R01s
1 P41
1 U01

Industry:

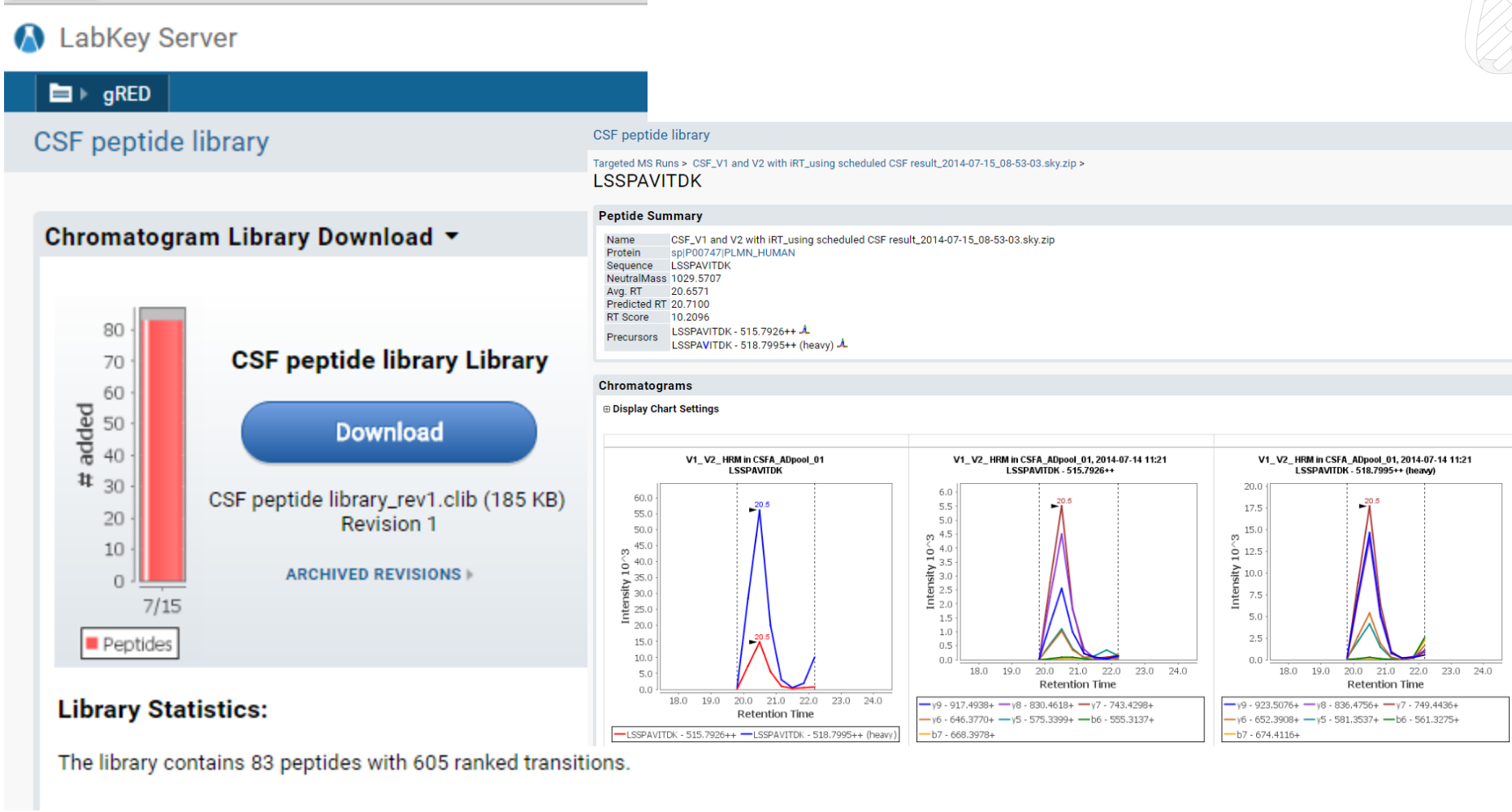
https://skyline.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=funding

©2012, Genentech

Roche Genentech
A Member of the Roche Group

Sharing data internally

Accelerating method building: Chromatogram libraries



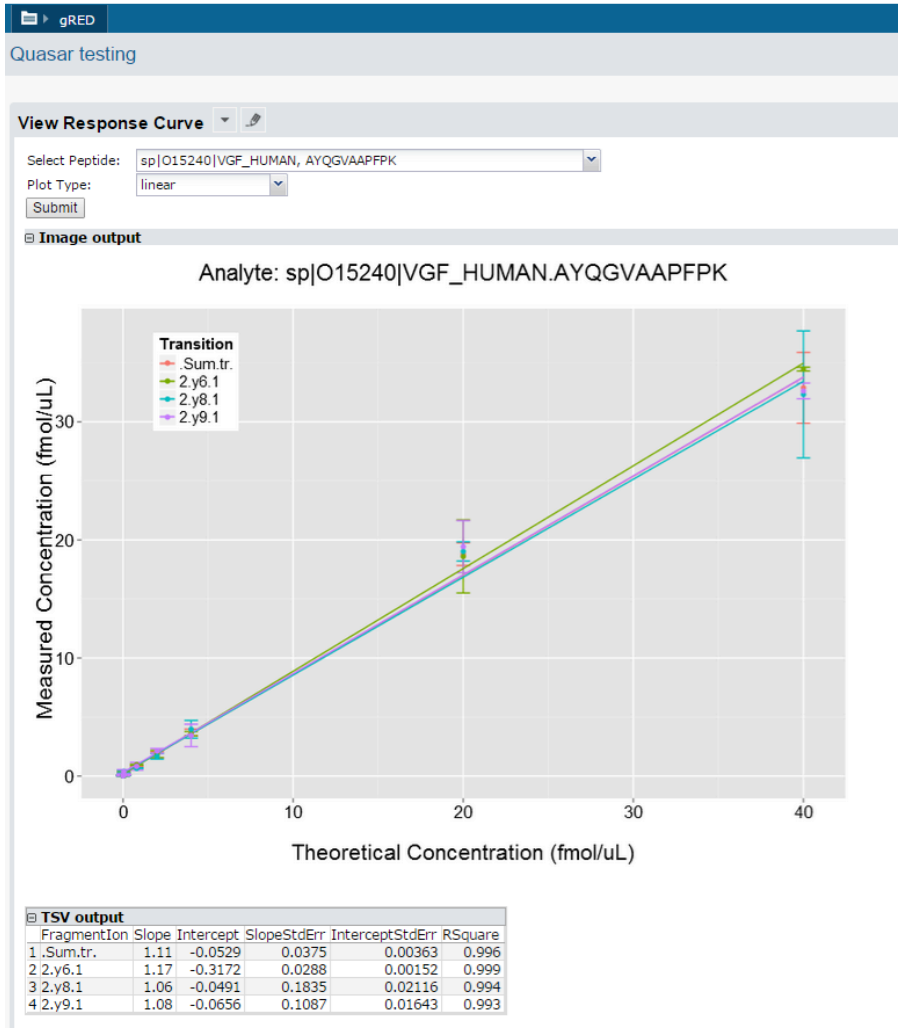
©2012, Genentech

Genentech
A Member of the Roche Group

Key areas of focus for Panorama Partnership

Vision: Data analysis directly in Panorama

- Tracking method development (versioning)
- Implementing data analysis tools directly in Panorama
  - Quality Control folder
  - AutoQC
  - Statistics



	FragmentIon	Slope	Intercept	SlopeStdErr	InterceptStdErr	RSquare
1	.Sum.tr.	1.11	-0.0529	0.0375	0.00363	0.996
2	2.y6.1	1.17	-0.3172	0.0288	0.00152	0.999
3	2.y8.1	1.06	-0.0491	0.1835	0.02116	0.994
4	2.y9.1	1.08	-0.0656	0.1087	0.01643	0.993

Adaptation of CPTAC Response Curve R script by
Laura Harris on Roche Panorama test server

©2012, Genentech

Panorama Partners progress

Tracking method development: Version Chaining

LabKey Server

TargetedMS

Method Build

Targeted MS Runs >

MethodBuildThird_2016-03-24_08-16-29.sky.zip

Document Summary

Name	MethodBuildThird_2016-03-24_08-16-29.sky.zip RENAME DOWNLOAD		
Protein Count	3	Peptide Count	30
Precursor Count	30	Transition Count	155

Document Versions

Document Name	Imported	Imported By	Note	Proteins	Precursors	Transitions
MethodBuildThird_2016-02-23_12-21-54.sky.zip	02/23/2016 12:22 PM	JEckels		4	49	313
MethodBuildThird_2016-02-23_12-22-33.sky.zip	02/23/2016 12:22 PM	JEckels		4	45	287
MethodBuildThird_2016-02-23_12-22-47.sky.zip	02/23/2016 12:22 PM	JEckels		4	45	279
MethodBuildThird_2016-02-23_13-41-10.sky.zip	02/23/2016 1:41 PM	JEckels		3	36	208
MethodBuildThird_2016-02-24_11-00-03.sky.zip	02/24/2016 11:00 AM	JEckels		3	33	180
MethodBuildThird_2016-02-24_11-30-52.sky.zip	02/24/2016 11:30 AM	JEckels		3	31	169
MethodBuildThird_2016-02-26_14-28-57.sky.zip	02/26/2016 2:29 PM	JEckels		3	31	169
MethodBuildThird_2016-03-23_22-00-40.sky.zip	03/23/2016 10:00 PM	JEckels		3	30	159
MethodBuildThird_2016-03-24_08-16-29.sky.zip	03/24/2016 8:16 AM	JEckels		3	30	155

©2012, Genentech

Genentech
A Member of the Roche Group

Panorama Partners QC folder

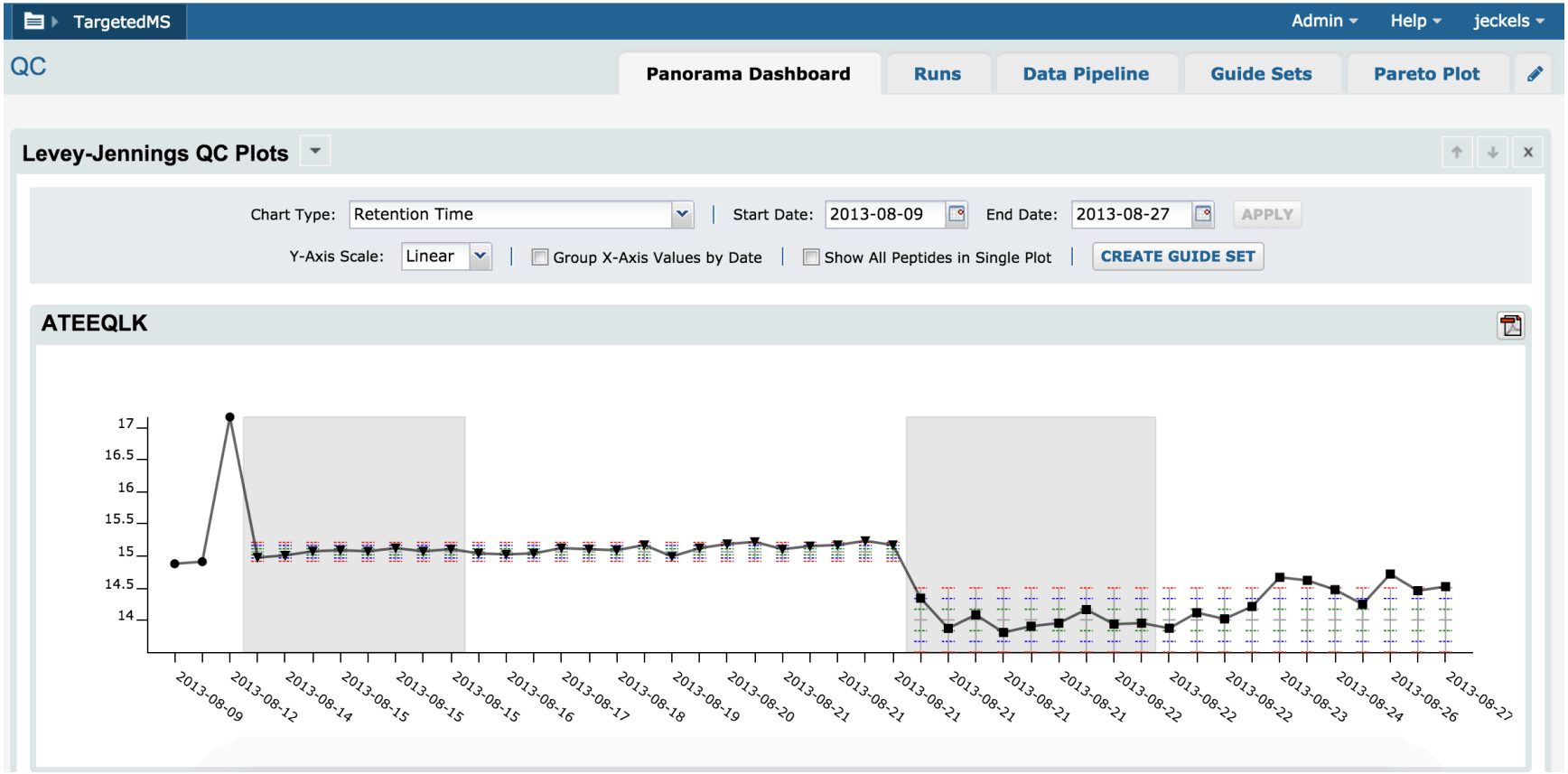
Development of QC folder to monitor instrument performance

©2012, Genentech

Roche Genentech
A Member of the Roche Group

Define reference regions for QC statistics

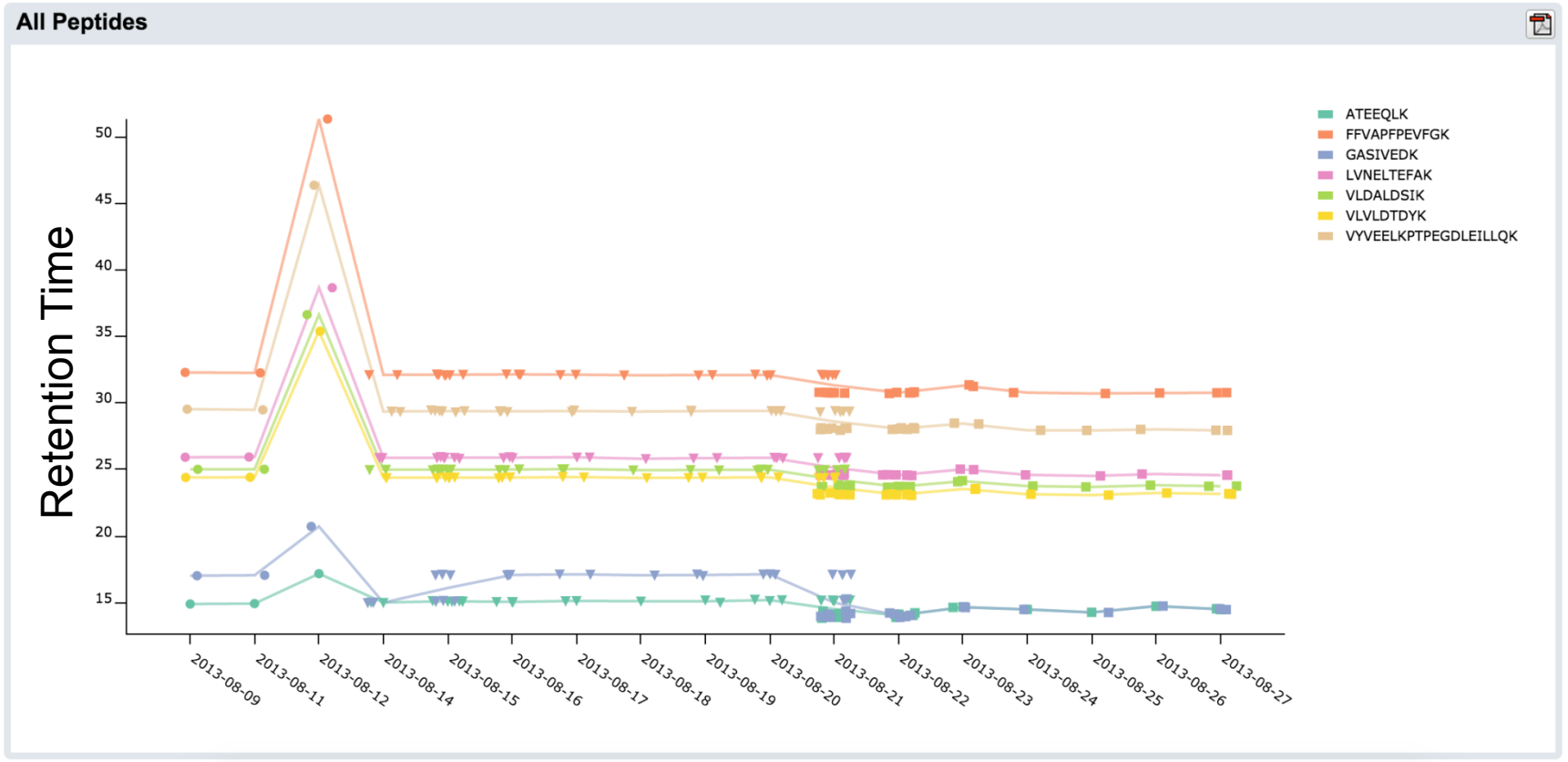
QC – Levey-Jennings Plots



©2012, Genentech

Panorama Partners QC folder

Evaluate performance of multiple peptides



©2012, Genentech

Panorama Partners QC folder

Track instrument performance

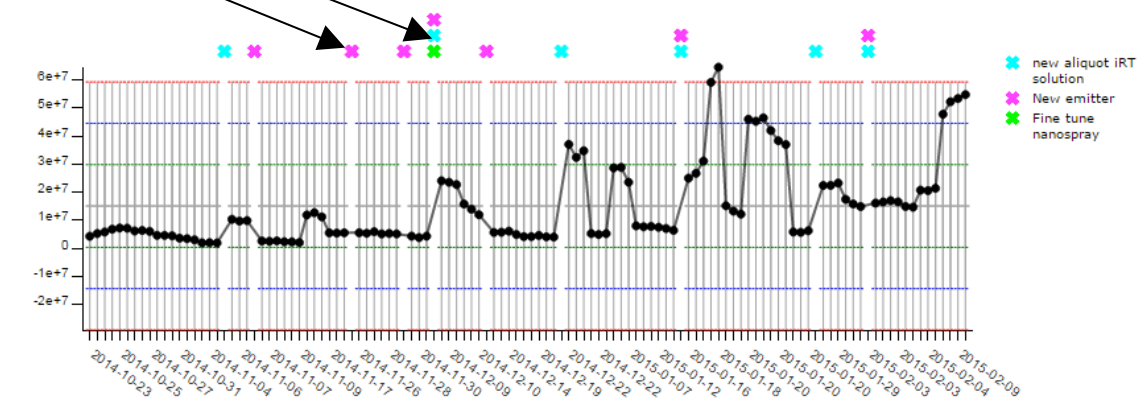
QC Annotation Type

VIEWS | CHARTS | INSERT NEW | IMPORT DATA | DELETE | EXPORT | PRINT | PAGE SIZE | 1 - 6 of 6

	Name	Description	Color	Folder	Created	Modified
EDIT	Fine tune nanospray			QC_for L and H_ratio	2015-02-10	2015-02-12
EDIT	Instrumentation Change			Shared	2015-01-30	2015-01-30
EDIT	new aliquot iRT solution			QC_for L and H_ratio	2015-02-09	2015-02-12
EDIT	New emitter			QC_for L and H_ratio	2015-02-09	2015-02-12
EDIT	Reagent Change			Shared	2015-01-30	2015-01-30
EDIT	Technician Change			Shared	2015-01-30	2015-01-30

Y Axis Scale: Linear | Start Date: 2014-10-23 | End Date: 2015-02-09 | APPLY | Chart Type: Peak Area

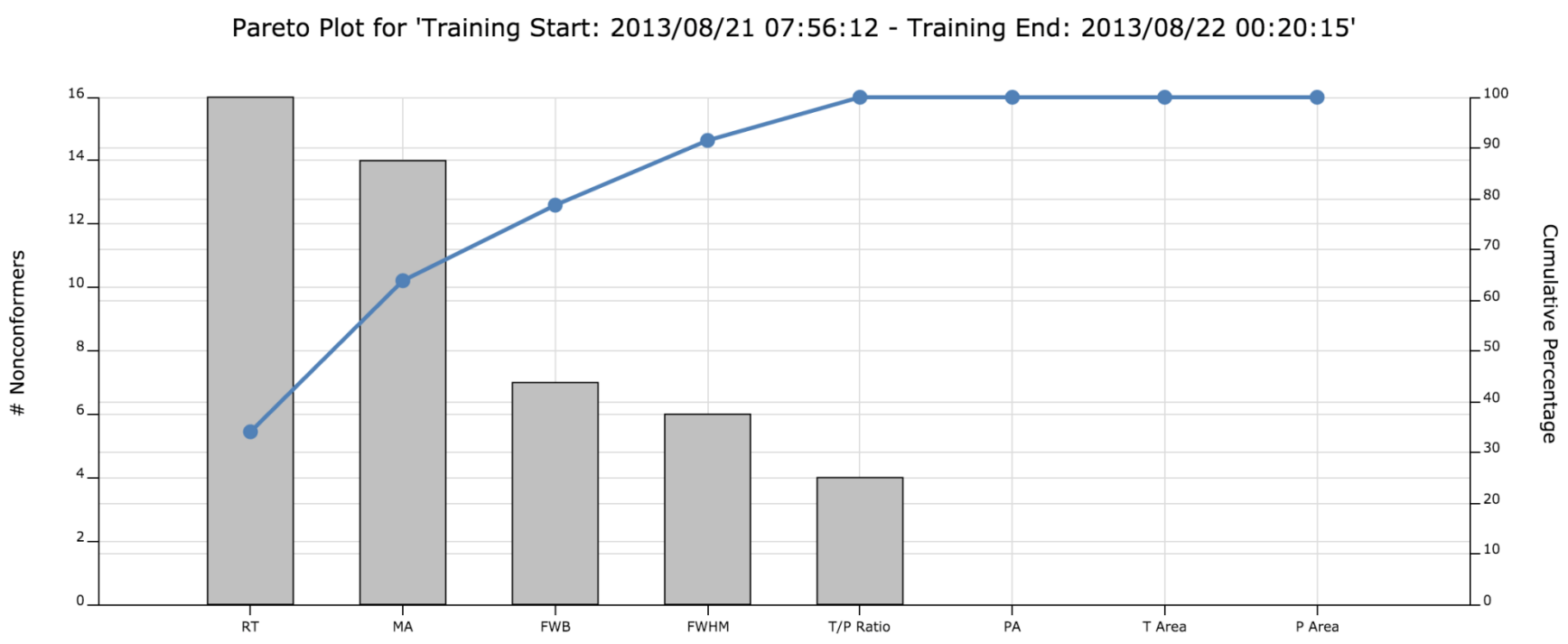
ADVTPADFSEWSK



©2012, Genentech

Determine which feature drives variability in method

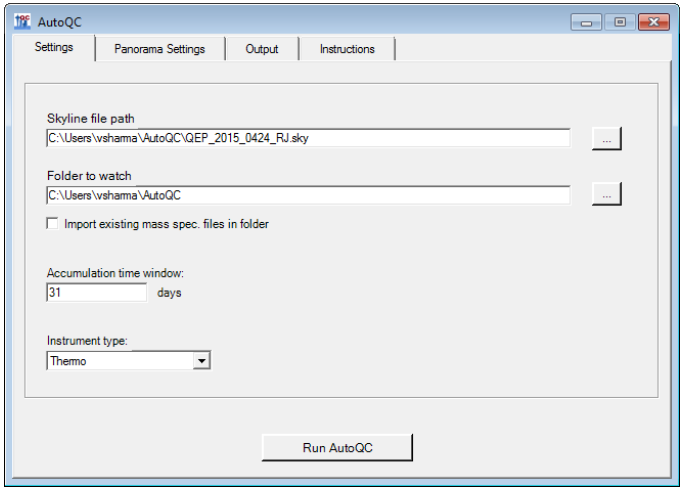
Guide Set 2

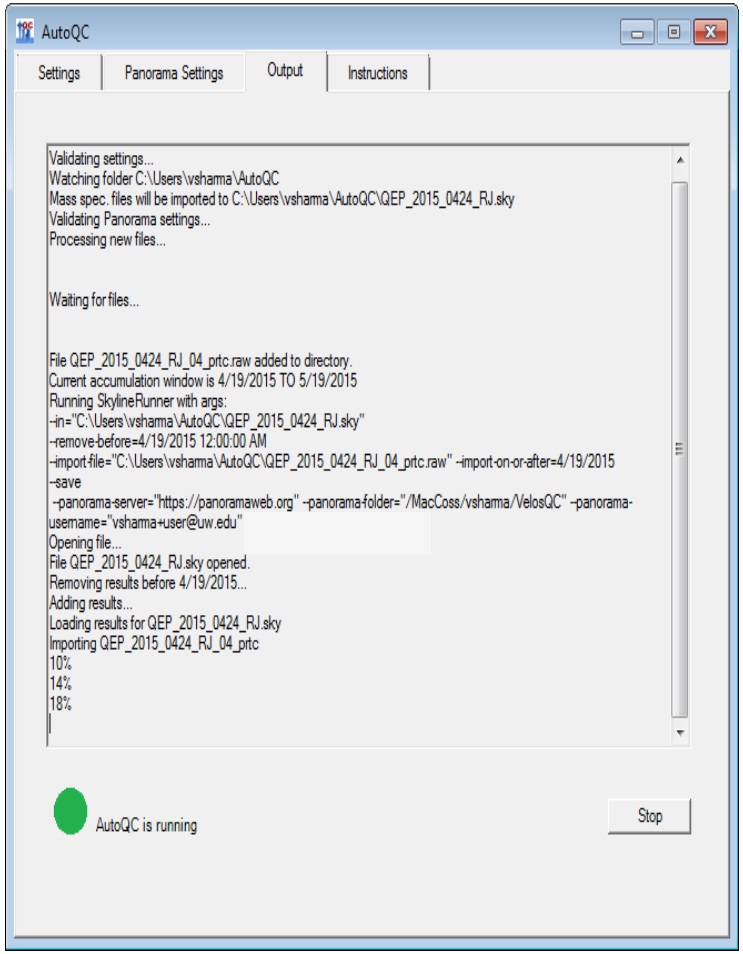


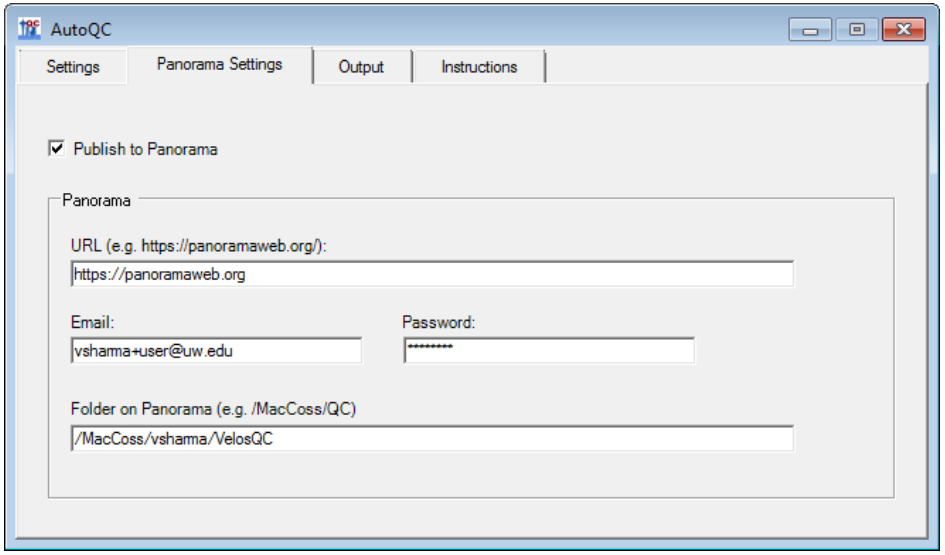
Retention time variable

©2012, Genentech

Automatically upload QC runs to Panorama



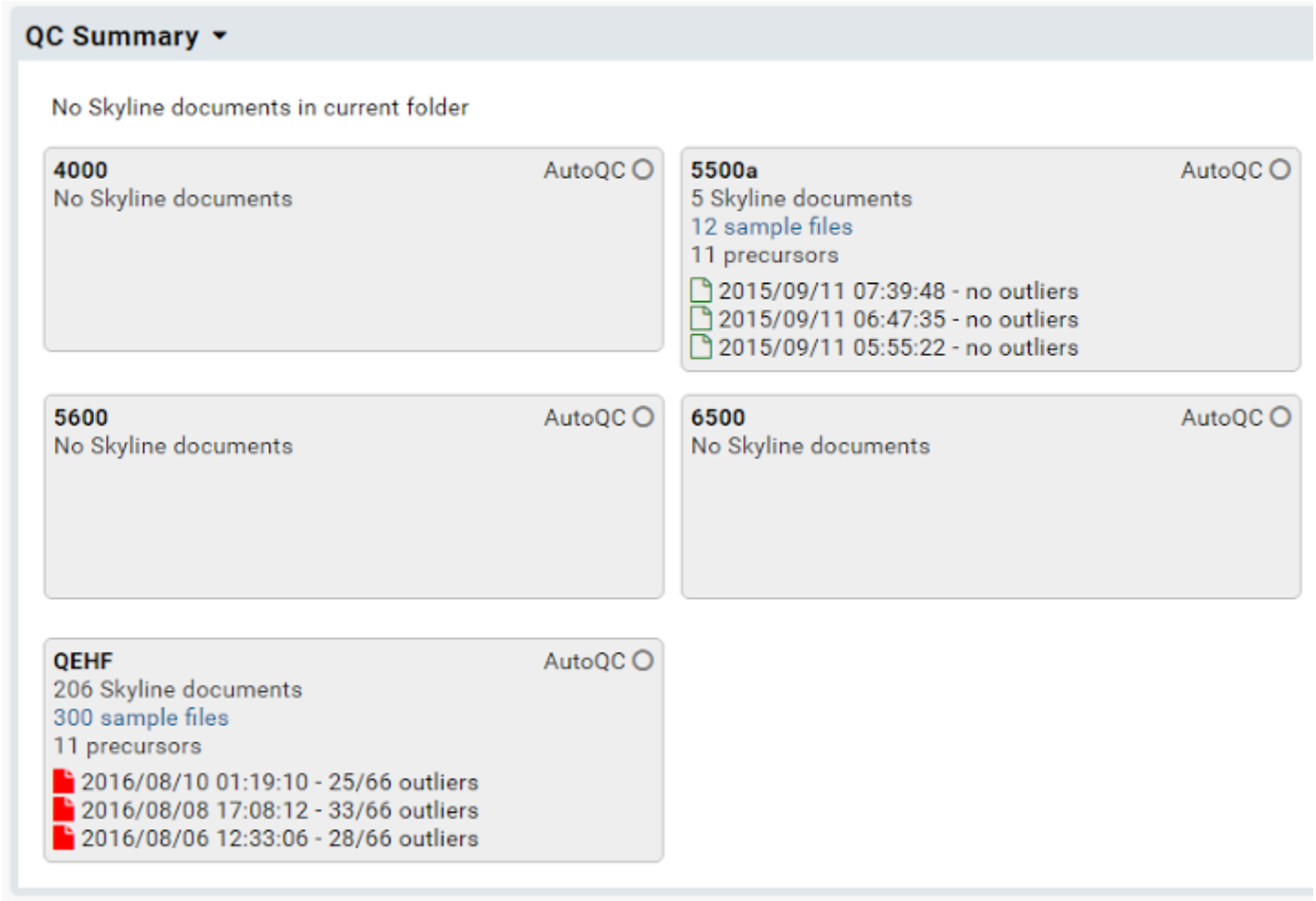




©2012, Genentech

Genentech
A Member of the Roche Group

Monitor active instruments using QC summary view



©2012, Genentech

Small molecule support

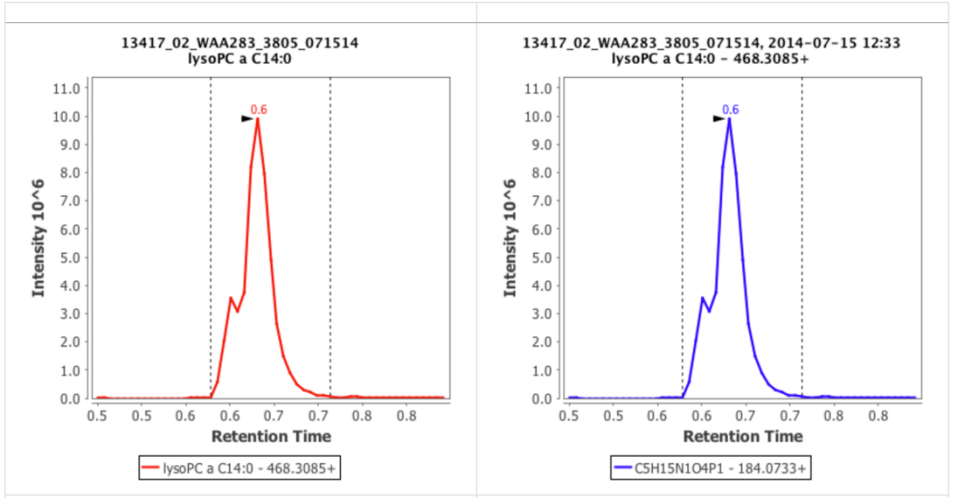
Panorama

Small Molecule Precursor List

GRID VIEWS ▾ REPORTS ▾ CHARTS ▾ EXPORT ▾ PRINT PAGING ▾

Molecule / Label	Description
PC	
LysoPC	

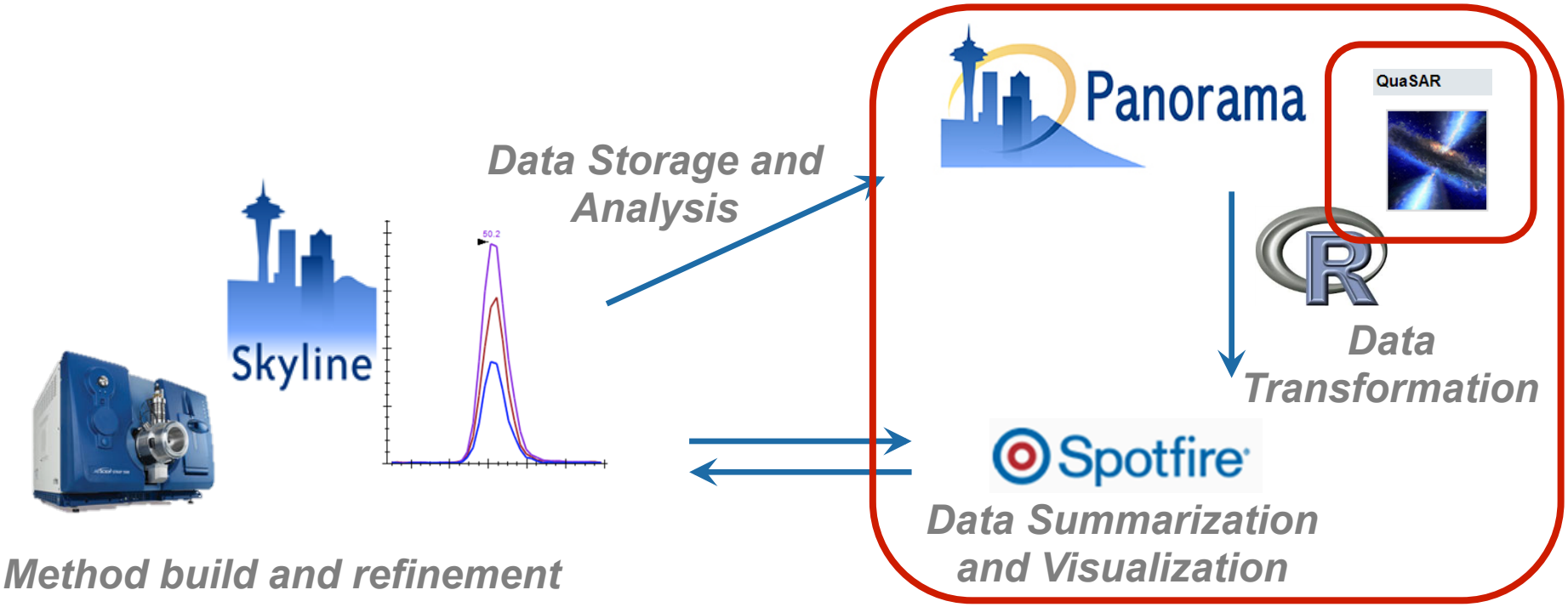
Custom Ion Name	Ion Formula	Mass Average	Mass Monoisotopic	Precursor	Q1 Z	Q1 m/z	Mass Average	Mass Monoisotopic
lysoPC a C14:0	C22H47N1O7P1	468.5881	468.3090	lysoPC a C14:0	1+	468.3085	468.5881	468.3090
lysoPC a C16:0	C24H51N1O7P1	496.6416	496.3403	lysoPC a C16:0	1+	496.3398	496.6416	496.3403
lysoPC a C16:1	C24H49N1O7P1	494.6257	494.3247	lysoPC a C16:1	1+	494.3241	494.6257	494.3247
lysoPC a C17:0	C25H53N1O7P1	510.6683	510.3560	lysoPC a C17:0	1+	510.3554	510.6683	510.3560
lysoPC a C18:0	C26H55N1O7P1	524.6951	524.3716	lysoPC a C18:0	1+	524.3711	524.6951	524.3716
lysoPC a C18:1	C26H53N1O7P1	522.6792	522.3560	lysoPC a C18:1	1+	522.3554	522.6792	522.3560
lysoPC a C18:2	C26H51N1O7P1	520.6633	520.3403	lysoPC a C18:2	1+	520.3398	520.6633	520.3403
lysoPC a C20:3	C28H53N1O7P1	546.7009	546.3560	lysoPC a C20:3	1+	546.3554	546.7009	546.3560
lysoPC a C20:4	C28H51N1O7P1	544.6850	544.3403	lysoPC a C20:4	1+	544.3398	544.6850	544.3403
lysoPC a C24:0	C32H67N1O7P1	608.8554	608.4655	lysoPC a C24:0	1+	608.4650	608.8554	608.4655
lysoPC a C26:0	C34H71N1O7P1	636.9089	636.4968	lysoPC a C26:0	1+	636.4963	636.9089	636.4968
lysoPC a C26:1	C34H71N1O7P1	636.9089	636.4968	lysoPC a C26:1	1+	636.4963	636.9089	636.4968
lysoPC a C28:0	C36H75N1O7P1	664.9624	664.5281	lysoPC a C28:0				
lysoPC a C28:1	C36H73N1O7P1	662.9465	662.5125	lysoPC a				
Acylcarnitine								



©2012, Genentech

Genentech
A Member of the Roche Group

Connecting Labkey's Panorama to data analysis tools



©2012, Genentech

Data analysis within Panorama: QuaSAR

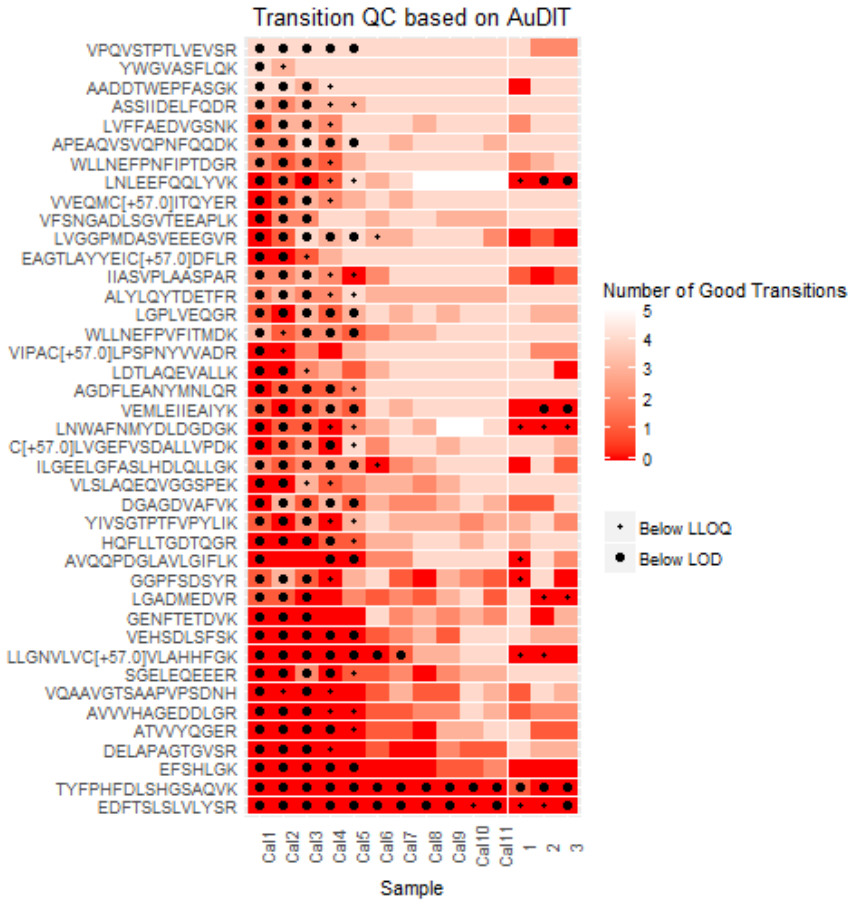
Accelerating method development using automated tools

QuaSAR processes SRM/ MRM-MS data calculating:

- limits of detection and quantification (LOD and LOQ),
- Mean and Coefficient of variation (CV) for all transitions of each peptide in a set of samples
- Peptide analyte concentration in unknown samples
- Runs the AuDIT algorithm to evaluate the transition quality and detect interference

Reports

Transitions QC Plot: AuDIT



©2012, Genentech

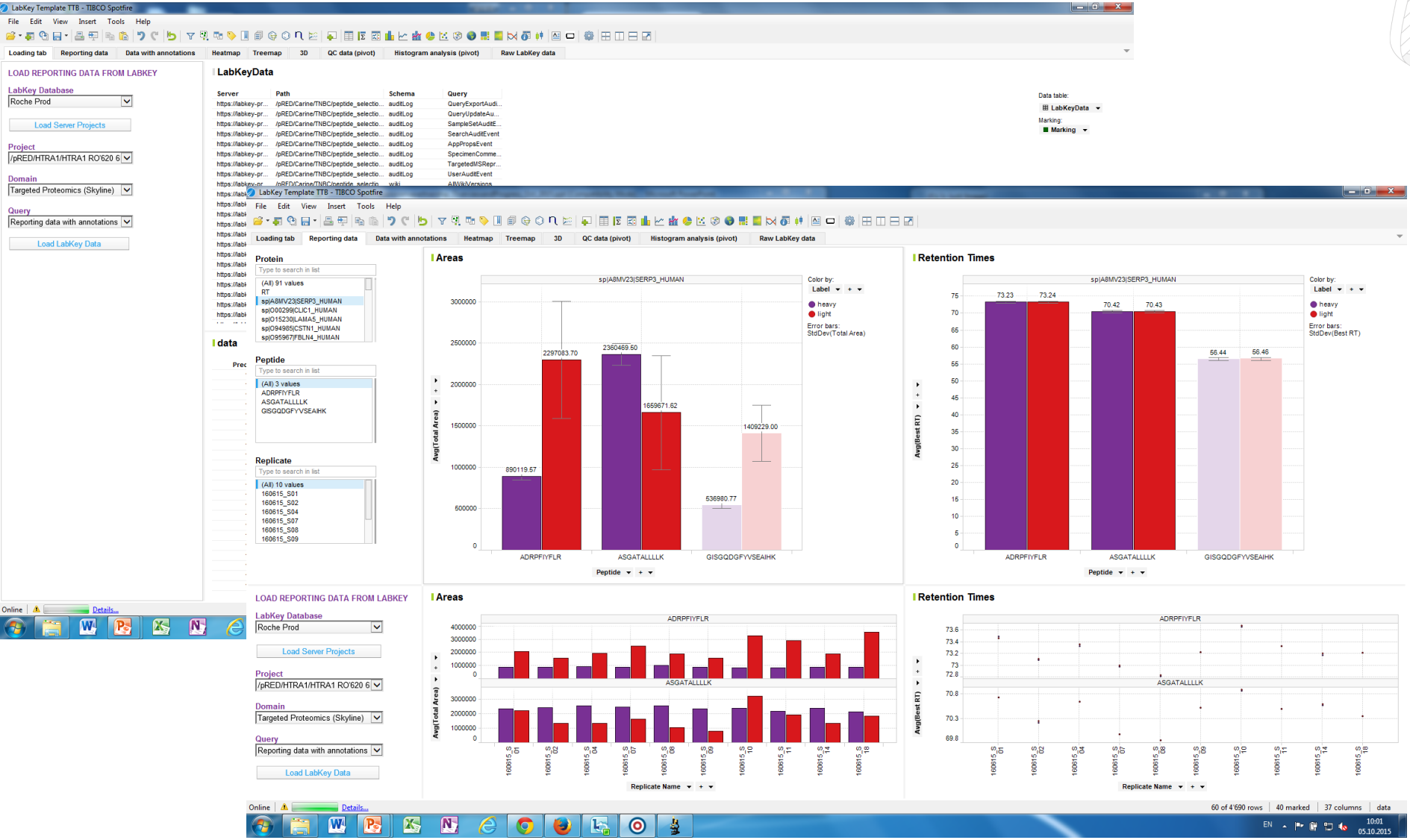
Shadi Eshghi

Modified from QuaSAR script at http://www.broadinstitute.org/~manidr/QuaSAR-1_2.zip

Genentech
A Member of the Roche Group

Panorama Partners Progress: Data Summarization

Connecting LabKey's Panorama to Spotfire for visualization



©2012, Genentech

Roche Genentech
A Member of the Roche Group

Labkey helping to accelerate biomarker development through support of MS data management

Biomarker Discovery → Clinical validation → Clinical Practice

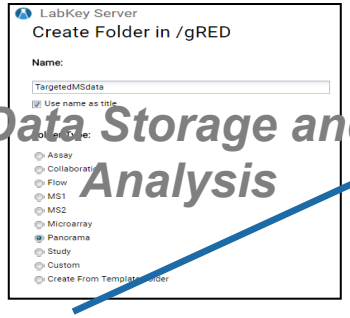
LTQ-Orbitrap
(High resolution)

- Discovery (unbiased)
- Targeted-discovery

- Multiplex panel
- Single analyte quantitation

Triple Quad
(High sensitivity)

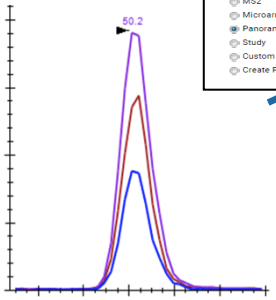
Developing customized features within Labkey to simplify method development and data analysis



Data Storage and Analysis

Panorama

Skyline



Data Transformation

Spotfire

Method build and refinement

Data Summarization And Visualization

©2012, Genentech

Genentech
A Member of the Roche Group

Acknowledgements

Genentech

OMNI Biomarker Development

- Paul Auger
- Veronica Anania
- Shadi Eshghi
- Rod Mathews
- Lee Honigberg

gRED IT

- Jon Sasano

Spotfire

- Laura Harris
- Joshua Haznedar

Panorama Partners

- Josh Eckels (Labkey)
- Brendan MacLean (UW)
- Vagisha Sharma (UW)
- Tom Dunkley (Roche)
- Michel Petrovic (Roche)

UW Targeted Proteomics Course 2014

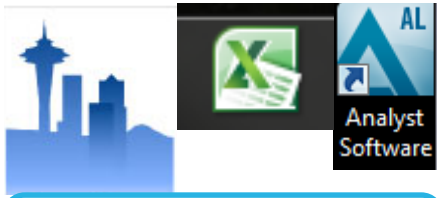
©2012, Genentech

Genentech
A Member of the Roche Group

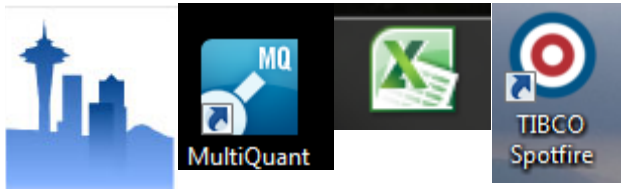
Appendix

©2012, Genentech

Large data files from multiple platforms creates data management challenges



Method and data generation



Data QC and analysis

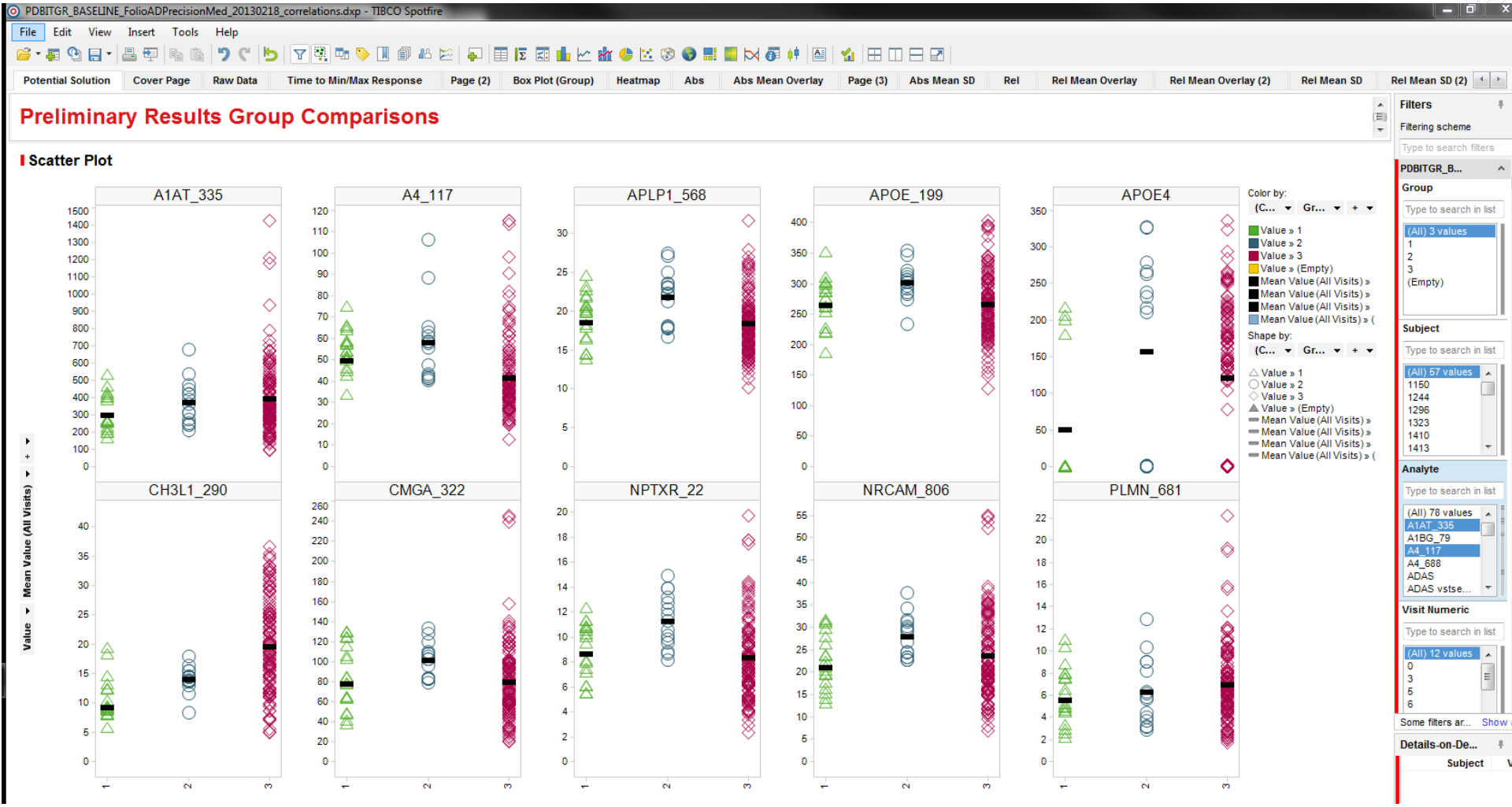
Types of files generated:

- Skyline method and data visualization files
- Instrument method file
- Raw MS data
- Quantitation file
- Excel spreadsheets for import to Spotfire
- Spotfire QC file
- Spotfire biomarker result file
- Additional (externally generated) biostats *.csv files and R plots

©2012, Genentech

Peptide levels by diagnostic group

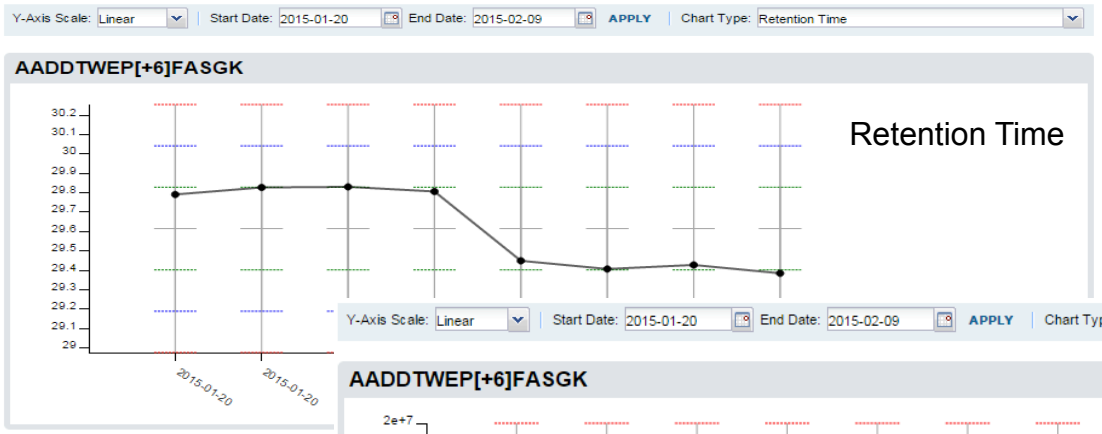
Spotfire visualization of multiplex results

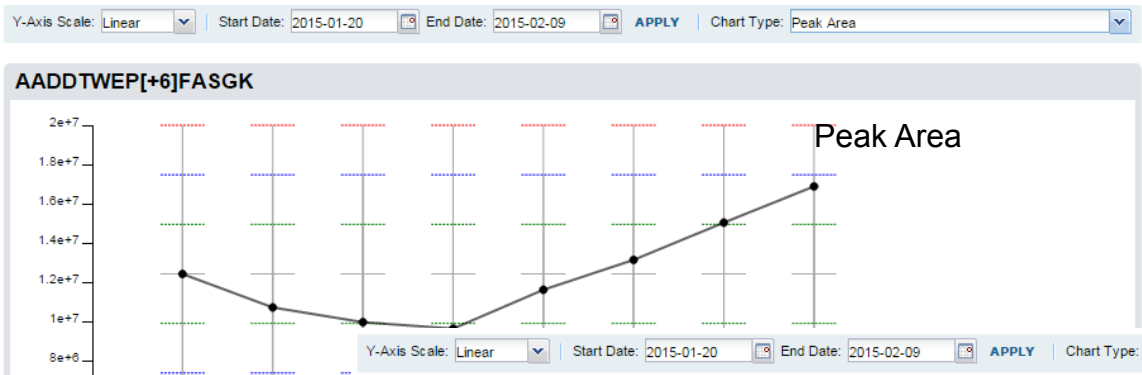


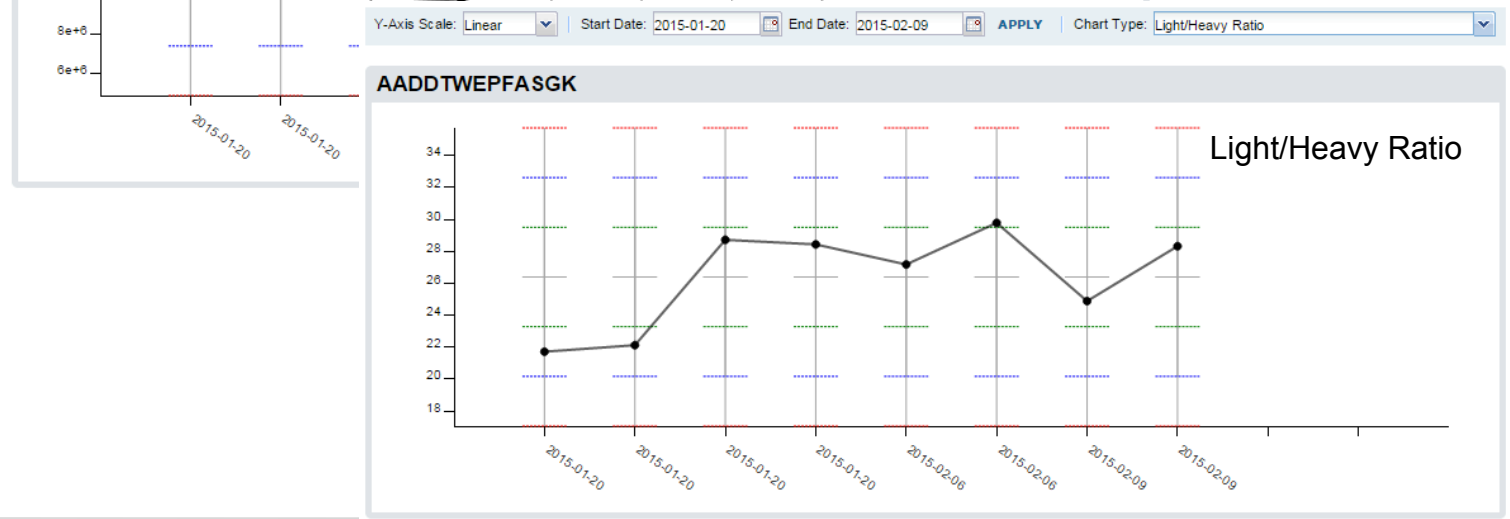
©2012, Genentech

Joshua Haznedar PKPD Genentech
A Member of the Roche Group

Development of QC folder to monitor instrument performance



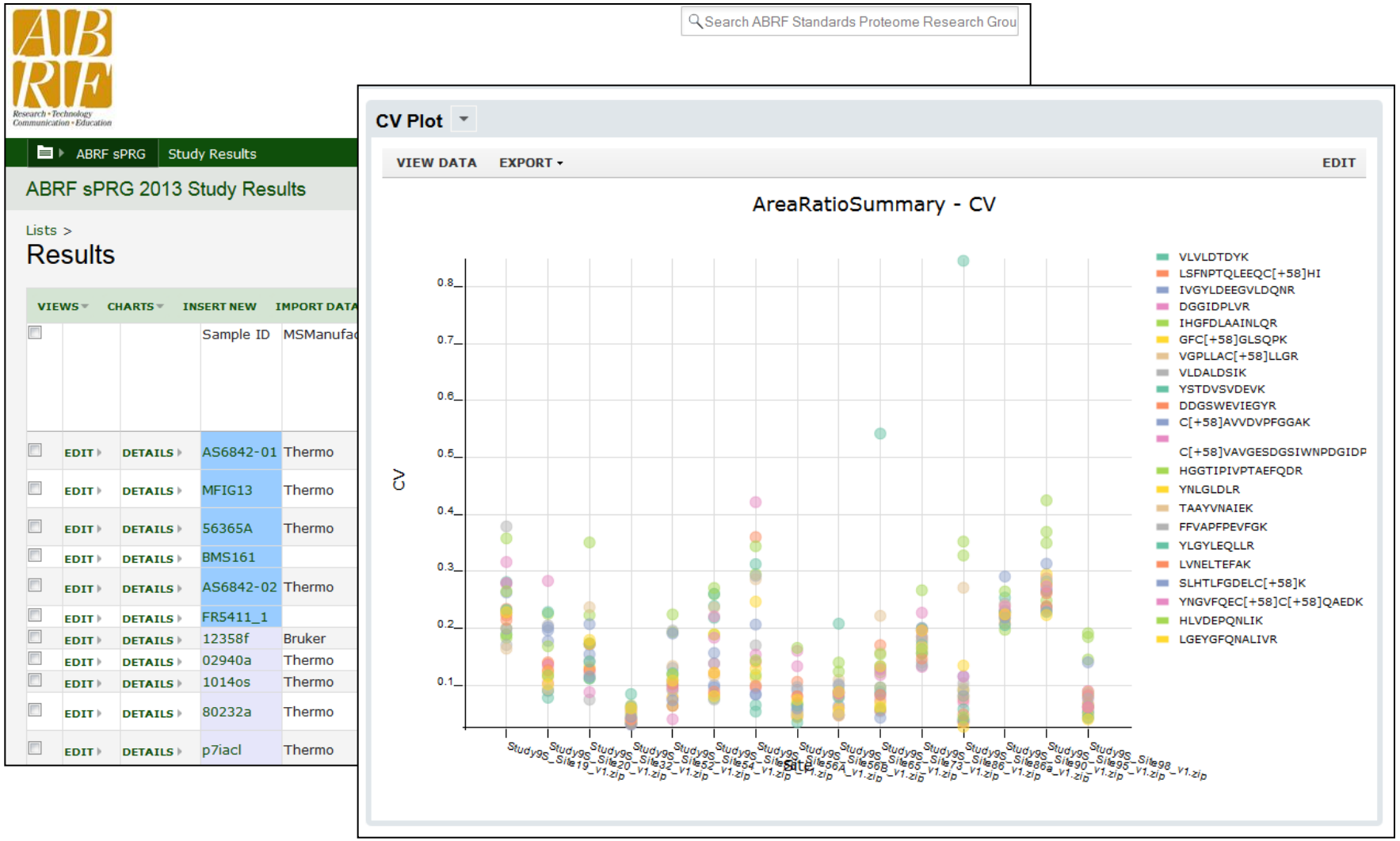




©2012, Genentech

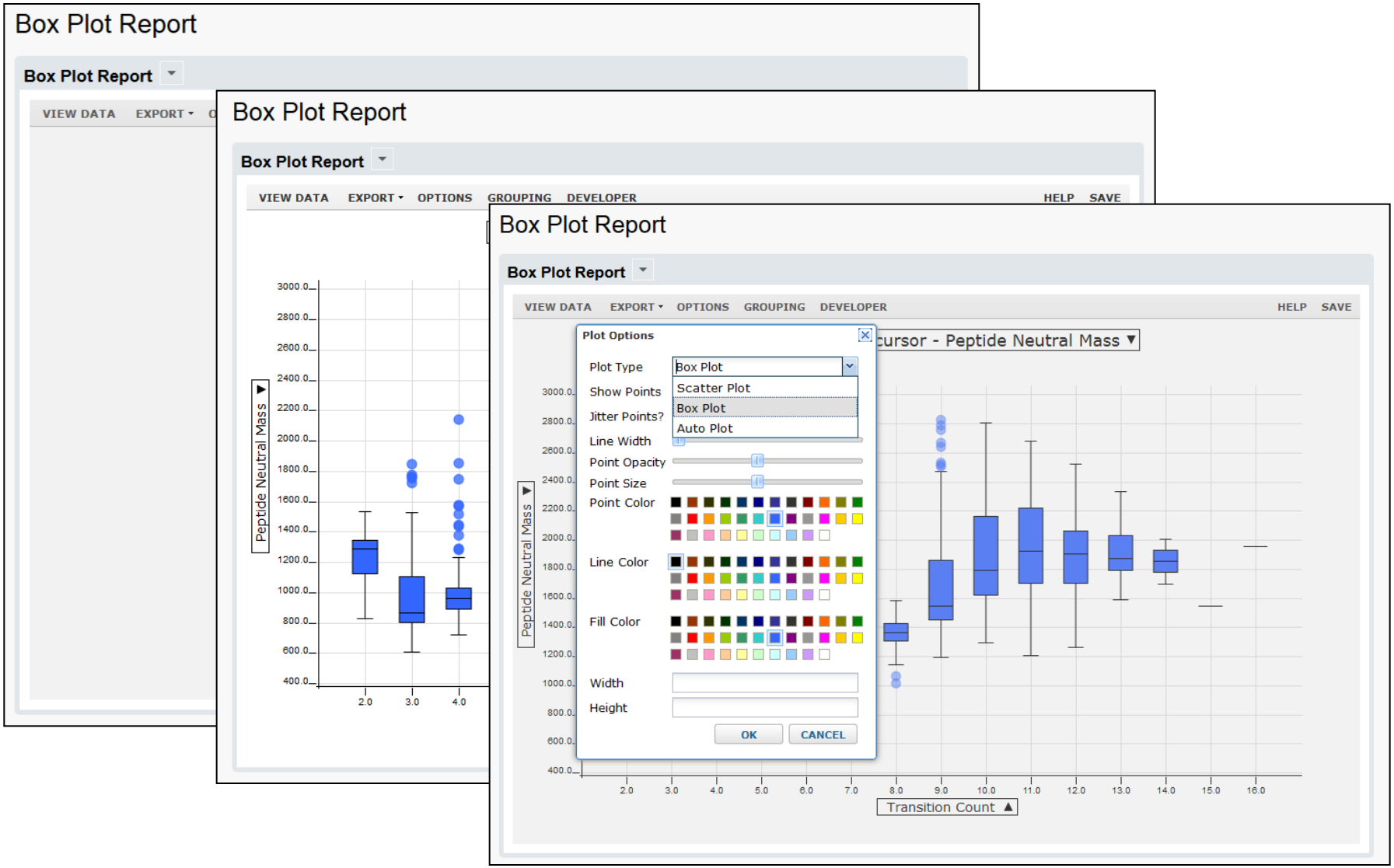
Roche Genentech
A Member of the Roche Group

Custom Reporting



©2012, Genentech

Point-and-Click Visualizations



©2012, Genentech

Panorama Software

- Module in LabKey Server
  - Open source (Apache 2.0 license)
  - Easy to install
  - Tested and in production over 5 years
  - Extensive core features
  - Projects and folders
  - Powerful security model
- Relational database supports full Skyline document model
- Integration with Skyline
  - Publish files directly to a Panorama server
- Web browser interface
  - Several data views familiar to Skyline users

Aggregating and Publishing

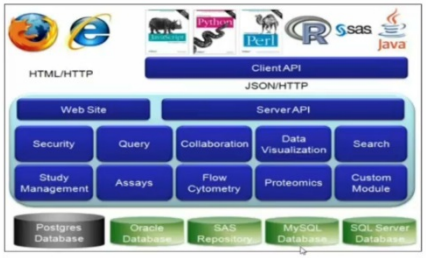
- Publish fully annotated Skyline documents
- Build chromatogram libraries
- Aggregate lab QC data (future)

Panorama

- Free hosted version (http://panoramaweb.org)
  - 25 separate projects so far (CPTAC and ABRF sPRG)
  - User controlled security
- Locally installable server application
- Free and open source (Apache 2.0)

Proteomics Standards Research Group (sPRG)

Back End Panorama (Labkey)



- Final datasets (estimate ~40-60)
- Perform Java/SQL Queries for downstream analyze data
- Generate report for 2014 ABRF Meeting and Manuscripts
- Build and Host Spectral Library directly through Panorama for future proteomic experiments

CPTAC MRM Assay Web Portal

- Goals
  - Promote the development and dissemination of MRM-based proteomic assays with high quality standards
  - Facilitate user community access to CPTAC assay development efforts
- Integration with Panorama
  - Use Panorama as a data repository
  - Use LabKey's Client APIs to pull data out of Panorama for viewing in the web portal
  - Provide access to chromatogram libraries built in Panorama

©2012, Genentech