

Nexus Copy Number Training

Soheil Shams, Ph.D.
BioDiscovery, Inc.

If you need an accessible version of this presentation, please contact the presenter.



What can you accomplish with Nexus Copy Number?

- Incorporate samples from various array platforms into a single project
- Identify common areas of aberrations across your data set
- Identify genes and GO terms in aberrant regions



Additional Features in Nexus Copy Number Professional

- Incorporation of external data such as miRNA and gene expression
- Statistical class comparisons
- Enrichment analysis
- Identification of natural grouping of samples via clustering
- Identify regions correlated with continuous valued phenotypes (e.g. survival)



What Will Be Covered?

- Copy number analysis basics; Segmentation and Allele Specific computation
- Basics of loading, processing, and viewing results
- Settings
- External Data
- Clustering
- Comparisons



DNA Copy Number Analysis Workflow

ImaGene®

Image Processing

- Convert image intensity values (pixels) into signal strength
- Perform basic QC test and remove “bad” spots

Nexus Copy Number™

Single Sample Analysis

- Pre-process data and identify regions of copy number change (segmentation)
- Measure QC based on probe-to-probe variations
- Visualize the result for a single sample

Population Analysis

- Integrate data from many arrays
- Identify areas of common aberration
- Identify genes and biological pathways being disrupted

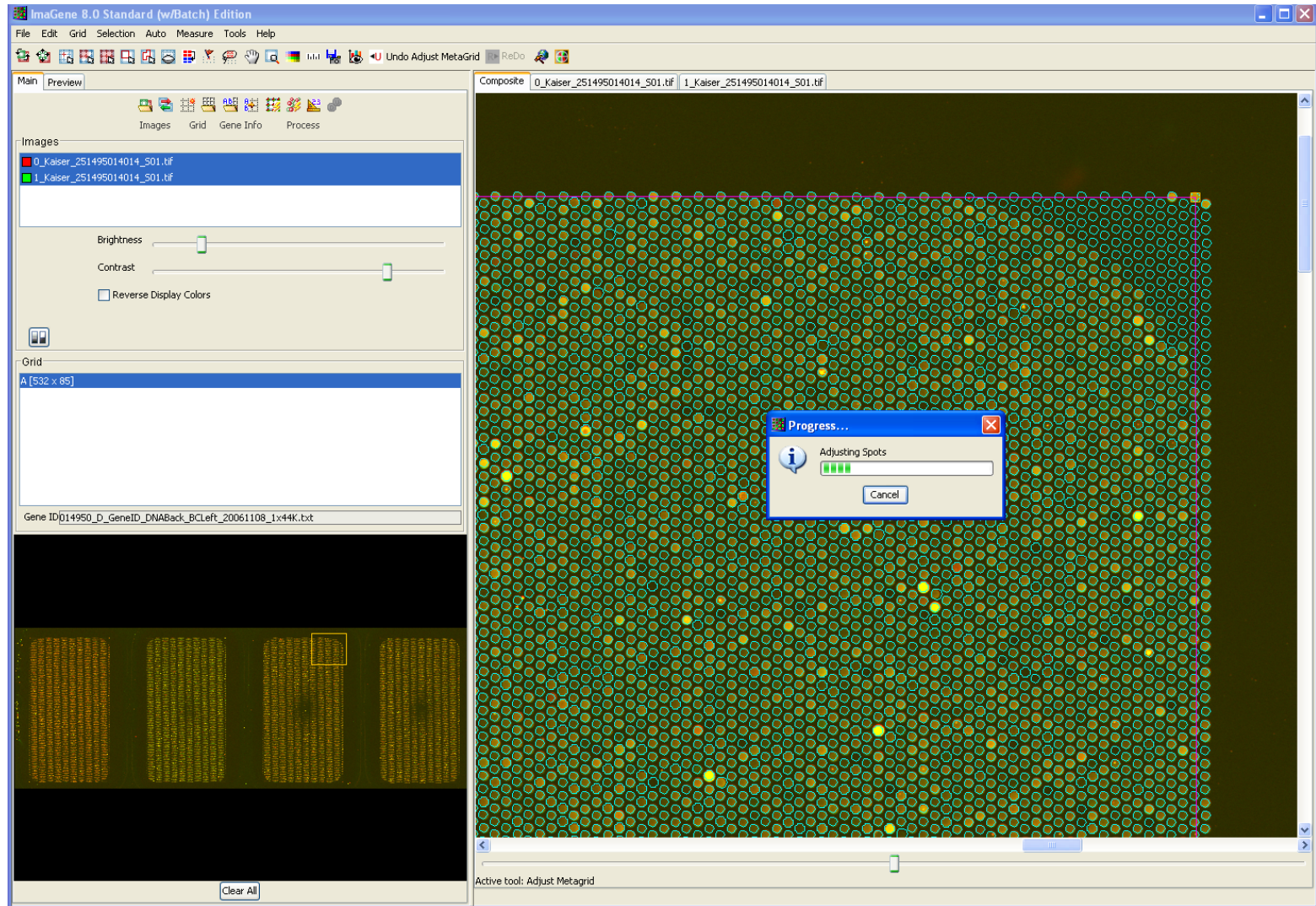


Image Processing

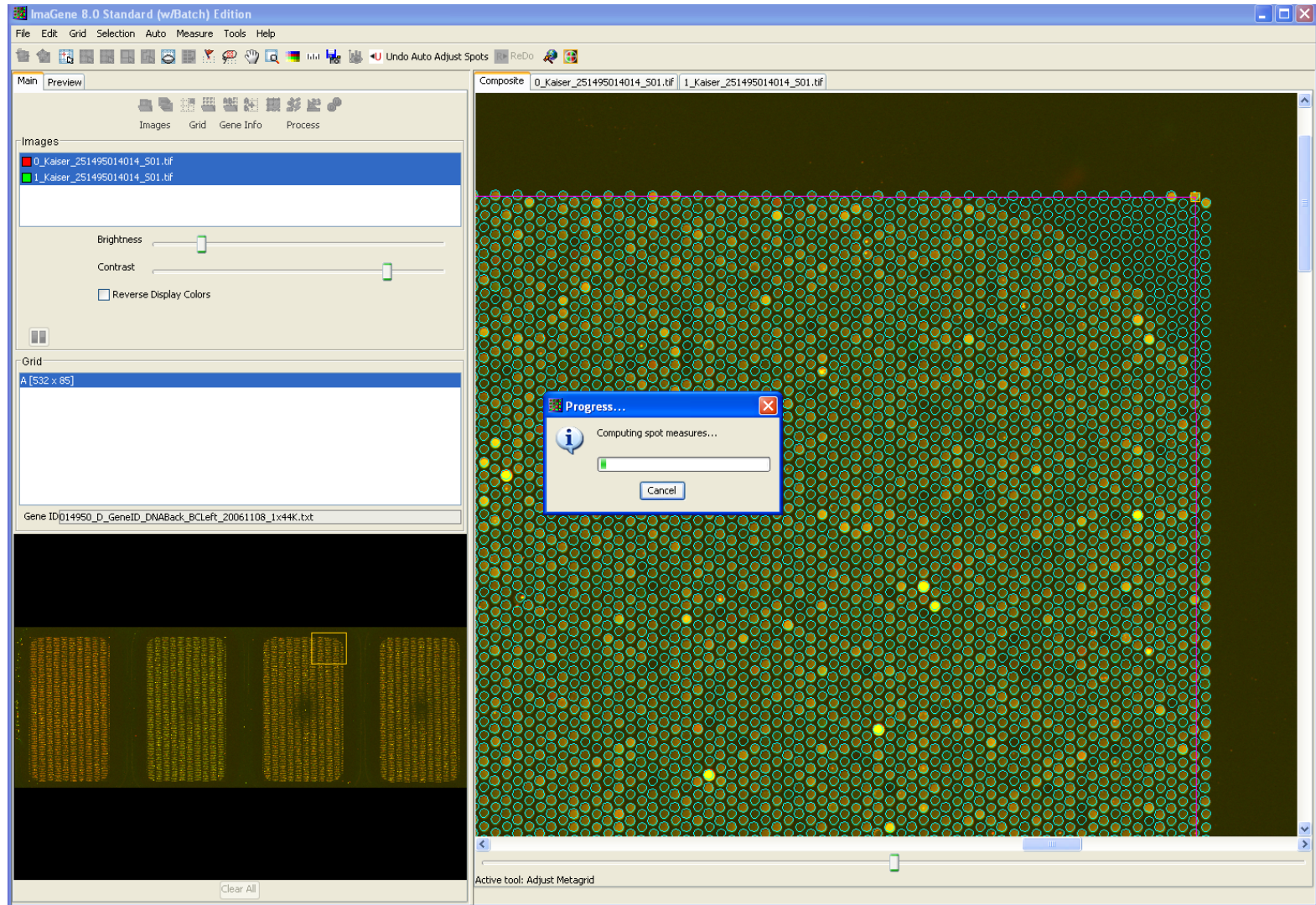
Demonstration using ImaGene version 8



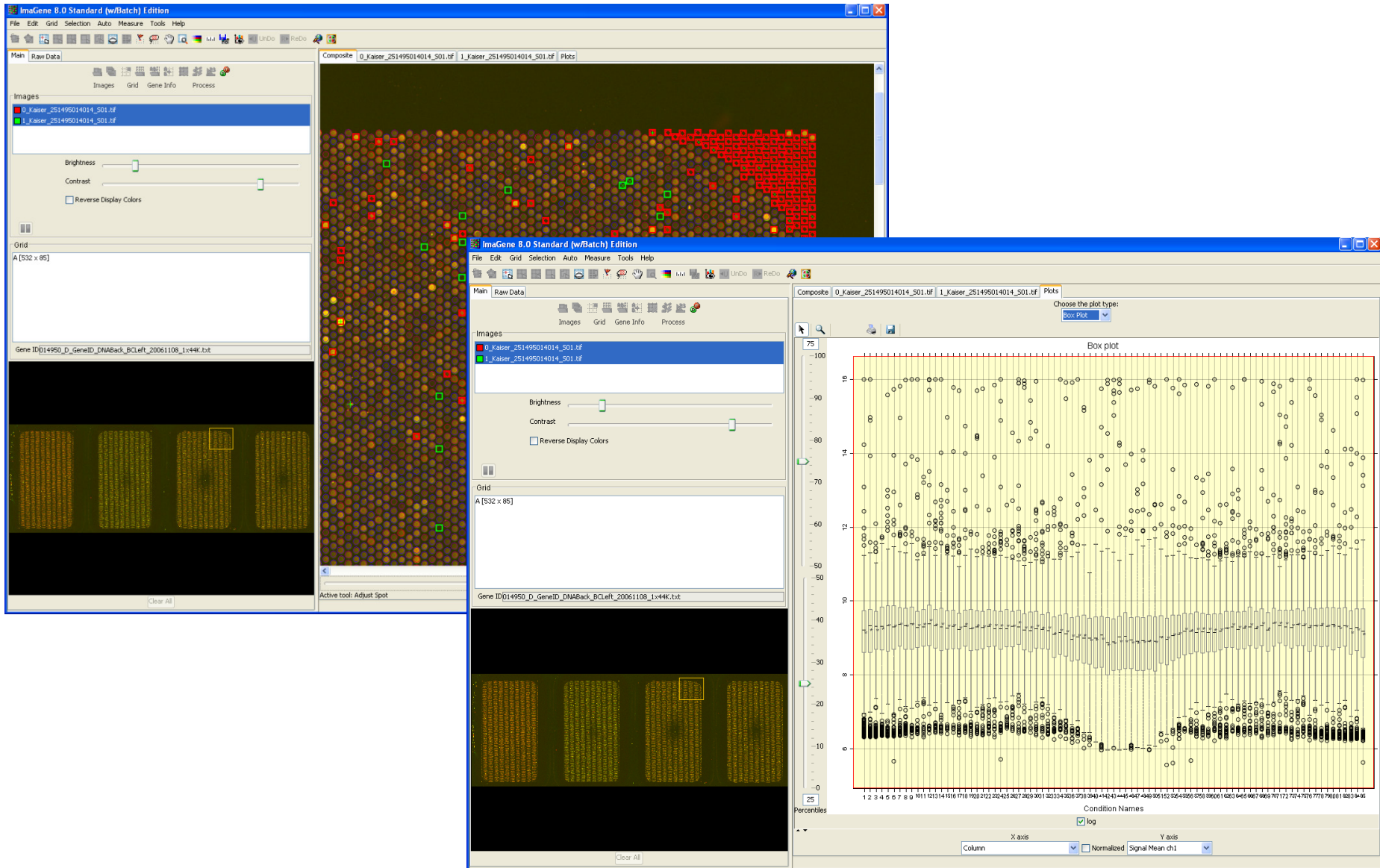
Load Image and Place Grid



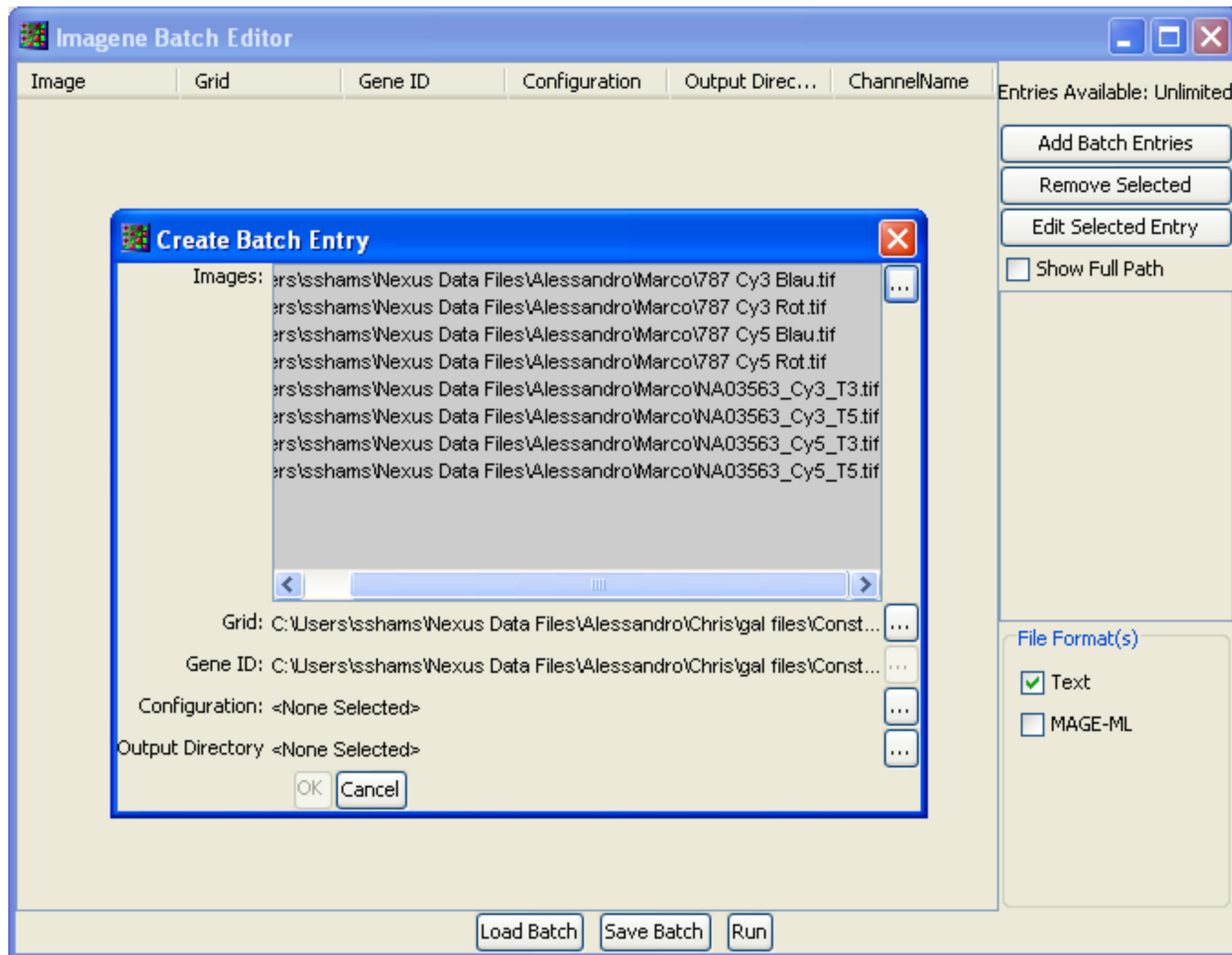
Quantify Spots & Perform QC



Save & Visualize Data



Walk-Away Batch Processing



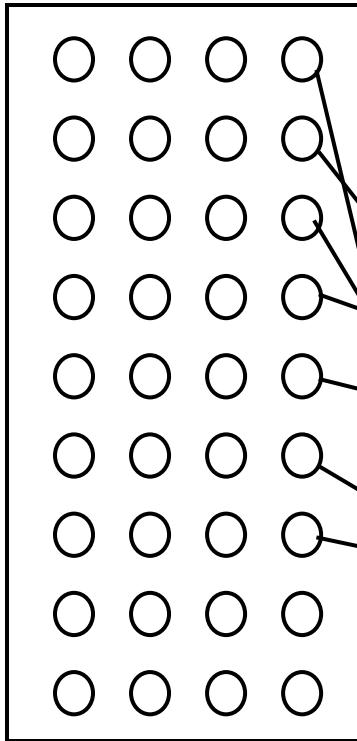
Section 1: Identification of Copy Number Change Events

Segmentation Algorithm

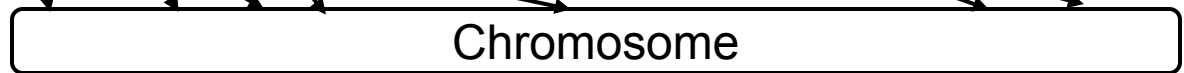


Mapping Spots to Regions on Chromosome

Array



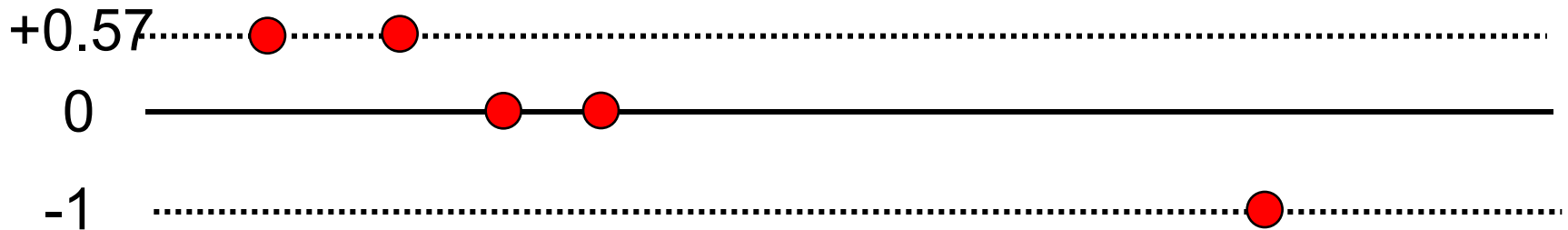
Probes can be BACs or Oligos



Computational Challenges

- Experiment results in a table:

Probe Location	Expr.	Control	Ratio	Log Ratio
Chr1:10-20	150	100	3/2	+0.57
Chr1:50-60	300	200	3/2	+0.57
Chr1:70-90	500	500	2/2	0
Chr1:100-120	60	60	2/2	0
Chr1:250-300	500	1000	1/2	-1



Segmentation Algorithms

- Many different algorithms have been proposed by industry and academics
- Range from simple approaches based on z-score to model based statistical approaches
- BioDiscovery has developed its proprietary rank segmentation algorithm based on a well accepted Circular Binary Segmentation (CBS) algorithm
- CBS was found to be a superior method to a number of other popular methods in paper by Willenbroc & Fridlyand, *Bioinformatics*, 2005
- Rank Segmentation performed very well in a number of recent comparison papers



Recent Comparisons

- Paper suggest a new algorithm, Birdsuite, for Affy SNP6 processing involving creating probe-specific models

Integrated genotype calling and association analysis of SNPs, common copy number polymorphisms and rare CNVs

Joshua M Korn^{1-5,10}, Finny G Kuruvilla^{1,4-6,10}, Steven A McCarroll^{1,4,5}, Alec Wysoker¹, James Nemesh¹, Simon Cawley⁷, Earl Hubbell⁷, Jim Veitch⁷, Patrick J Collins⁷, Katayoon Darvishi⁸, Charles Lee⁸, Marcia M Nizzari¹, Stacey B Gabriel¹, Shaun Purcell^{1,5}, Mark J Daly^{1,5,9} & David Altshuler^{1,4,5,9}

NATURE GENETICS VOLUME 40 | NUMBER 10 | OCTOBER 2008

Probes spanned by CNV	Total in Category	Birdsuite	Partek	Nexus	Nexus (relaxed)
<= 1 probes	325	3.7%	1.2%	0.6%	1.2%
2-5 probes	256	32.8%	0.4%	0.8%	5.5%
6-10 probes	112	61.6%	1.8%	13.4%	42.0%
11-20 probes	71	64.8%	4.2%	47.9%	69.0%
> 20 probes	129	93.8%	11.6%	72.9%	74.4%



Comments on Paper

- The birdsuite algorithm used here was optimized for SNP6 processing and optimal parameters were used
- Algorithm took advantage of genotype calls and B-allele values not just log-ratios
- Nexus parameters were default values and not optimized
- Nexus only looked at log-R values
- Settings reported requiring a minimum of 5 probes per call



Another Recent Comparison Paper



Evaluation of Seven CNV Detection Methods using Whole Genome SNP arrays from Myopia Samples

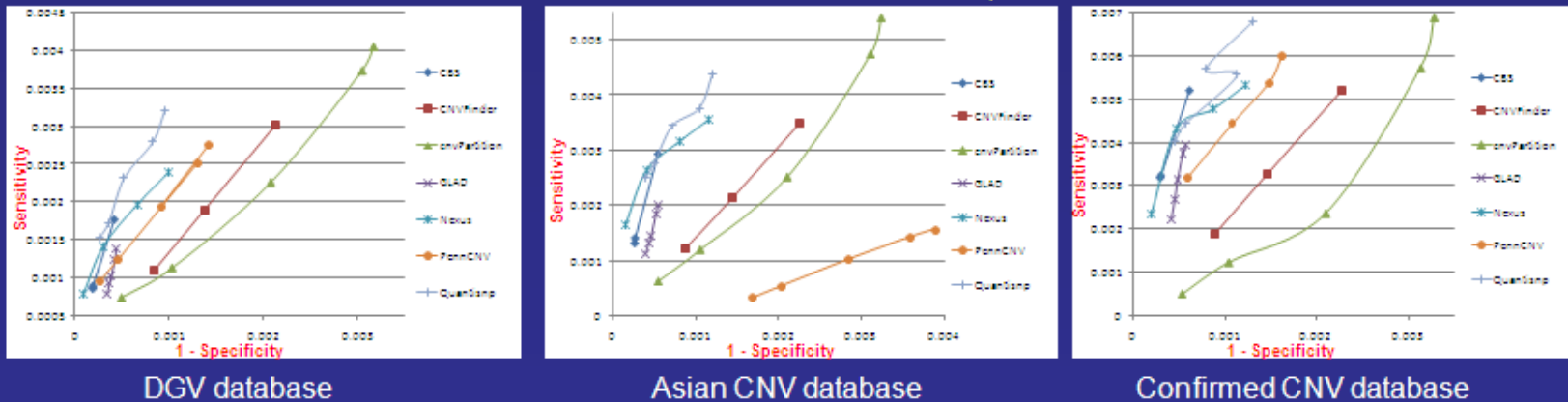
Yi-Ju Li^{1,2}, Andrew Dellinger¹, Mark Seielstad³, Liang Goh⁴, Terri L. Young^{1,4}, Seang Mei Saw^{5,6}.

¹Center for Human Genetics, Duke University Medical Center, Durham, NC; ²Department of Biostatistics and Bioinformatics, Duke University Medical Center, ³Genome Institute of Singapore, Singapore; ⁴Duke-NUS Graduate Medical School, Singapore; ⁵National University of Singapore, Singapore; ⁶Singapore Eye Research Institute, Singapore.



Poster presented at ASHG 2008

Figure 2. ROC Curves for Each Method and Database Comparison



Method comparison: ROC curves (sensitivity vs. 1-specificity) for each method on the 3 databases. Nexus and QuantiSNP are consistently best across datasets. CBS and GLAD are good, but limited in range of sensitivity. The performance of PennCNV was not consistent across datasets.

BioDiscovery's View on Segmentation Algorithms

- We strive to develop good quality and reliable segmentation algorithms
- Our goal is to create algorithms that work across many platforms
- We use real-world assumptions in the algorithm (e.g. mosaicism does exist in real samples)
- We do not intend to “compete” with segmentation algorithm developers
- Nexus' main goals are post segmentation analysis



Rank Segmentation Algorithm

Overview



Segmentation

1. Sort probes based on log-ratio values and give each probe their rank as the score
2. Recursively segment each chromosome to identify segment boundaries that separate two adjacent segments with a given significance level



Settings Window: Analysis Panel

- Significance Threshold
- Max Contiguous Probe Spacing
- Min Number of Probes per Segment
- High gain
- Gain
- Loss
- Big Loss

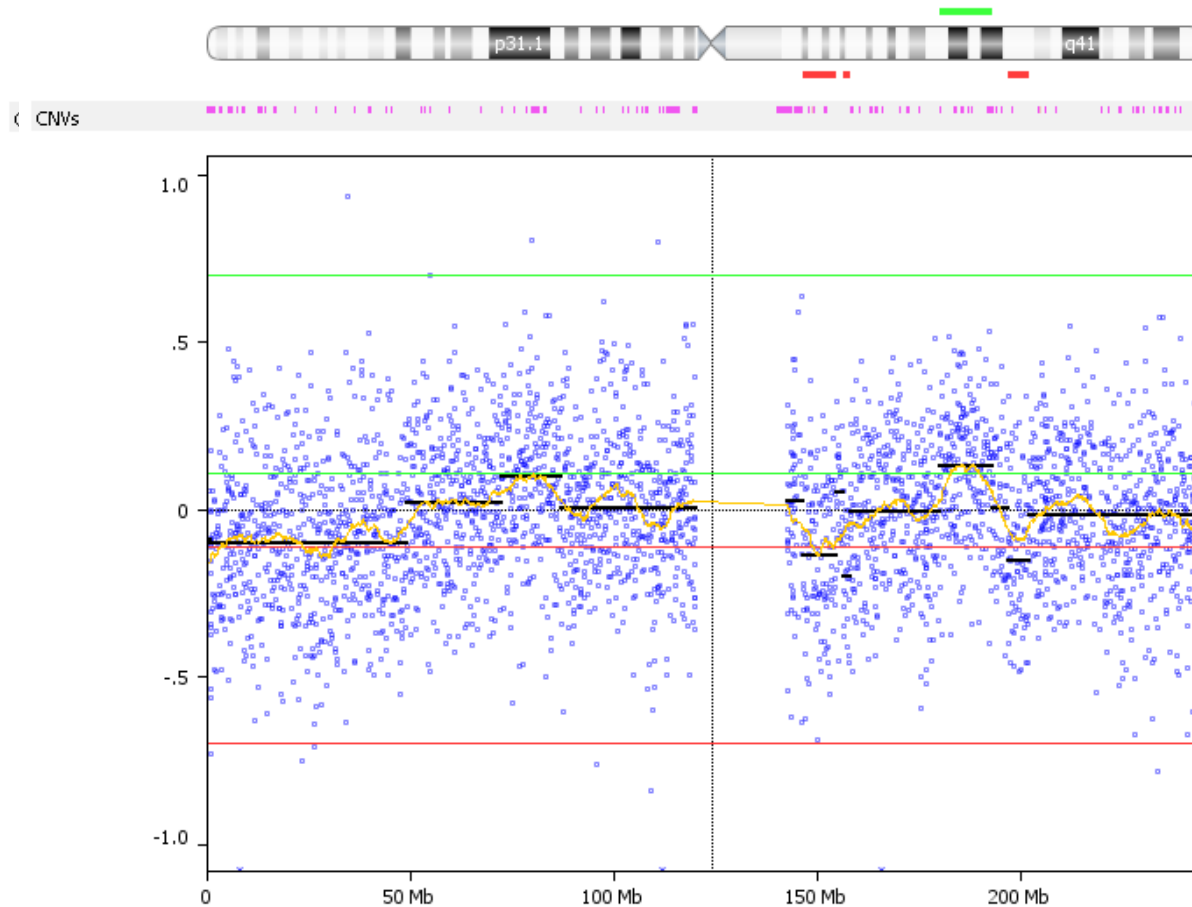


Significance Threshold

- For adjusting the sensitivity of the calling algorithm
- Smaller number = more stringent before creating a new clusters.
- Significance threshold should be set based on expected noise.
- For oligo arrays, we recommend using 1×10^{-6}
- For BAC array (lower density), we recommend 5×10^{-5}



Effect of Significance Threshold



Block1_B-23-2008: Chromosome 1

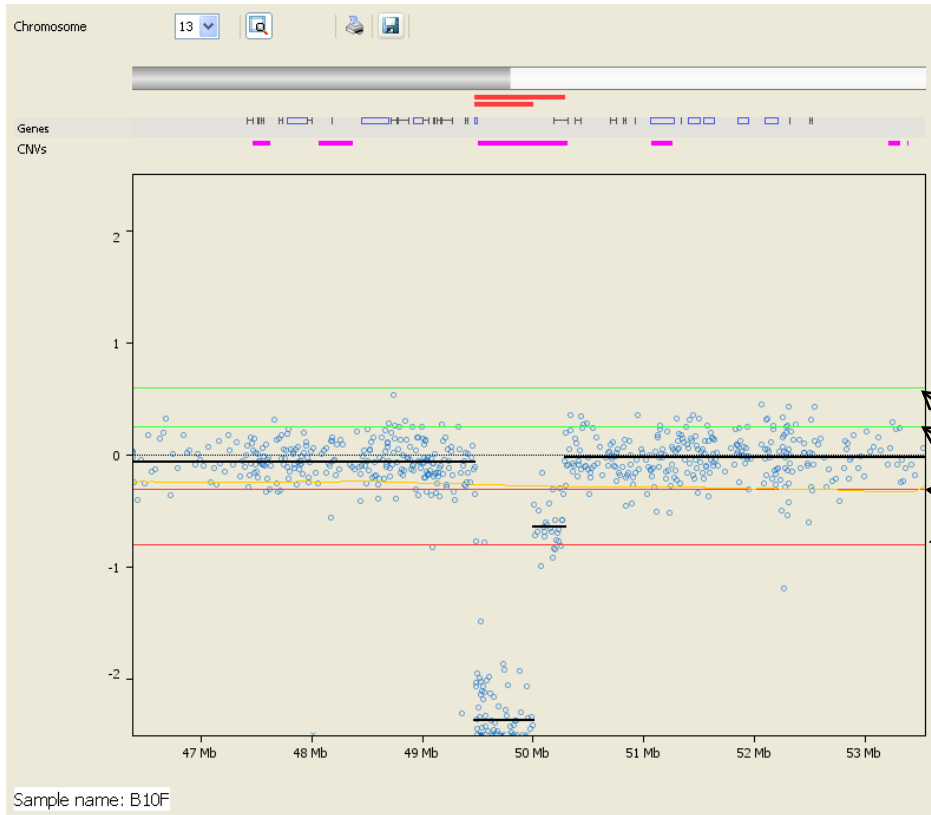
Significance = $1E-8$

Thresholds for Making Calls

- High gain – Two or more copy gain
- Gain – Single copy gain
- Loss – Hemizygous loss
- Big Loss – Homozygous loss



Calling Thresholds



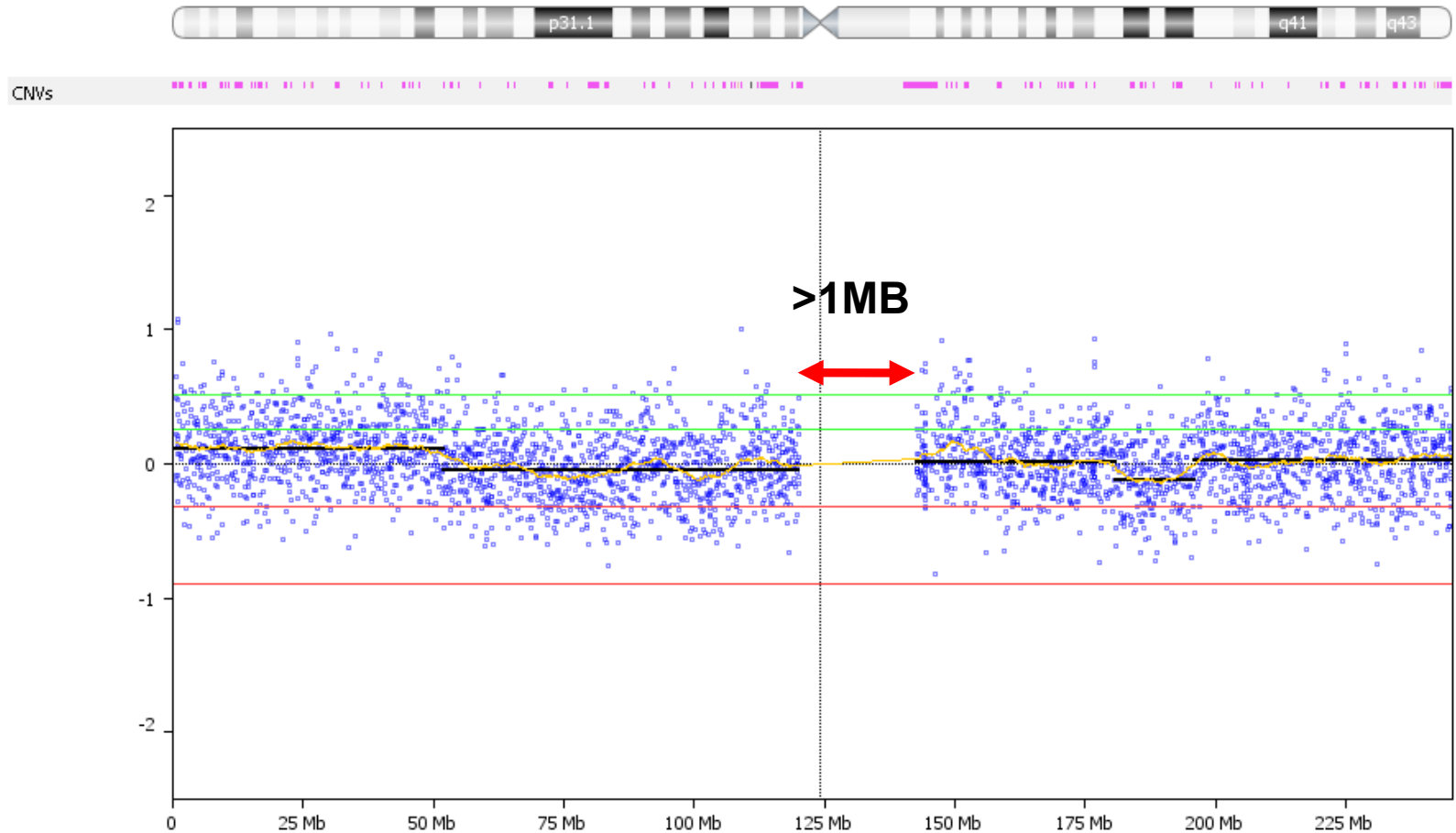
Analysis	
Type:	Rank Segmentation
Significance Threshold	1.0E-7
Max Contiguous Probe Spacing (Kbp)	1000
Min number of probes per segment	4
High Gain	0.5
Gain	0.2
Loss	-0.25
Big Loss	-0.7

Max Contiguous Probe Spacing

- Specifies the maximum spacing between adjacent probes before breaking a segment.
- Especially useful for
 - areas across the centromere where software will not make any calls where there are no probes
 - focused arrays where there are probes only in parts of the chromosome



Maximum Contiguous Spacing



a08-001: Chromosome 1



Min Number of Probes Per Segment

- Sets the minimum number of probes used to make a call
- Useful to prevent creation of a segment based only on a few probes in a region

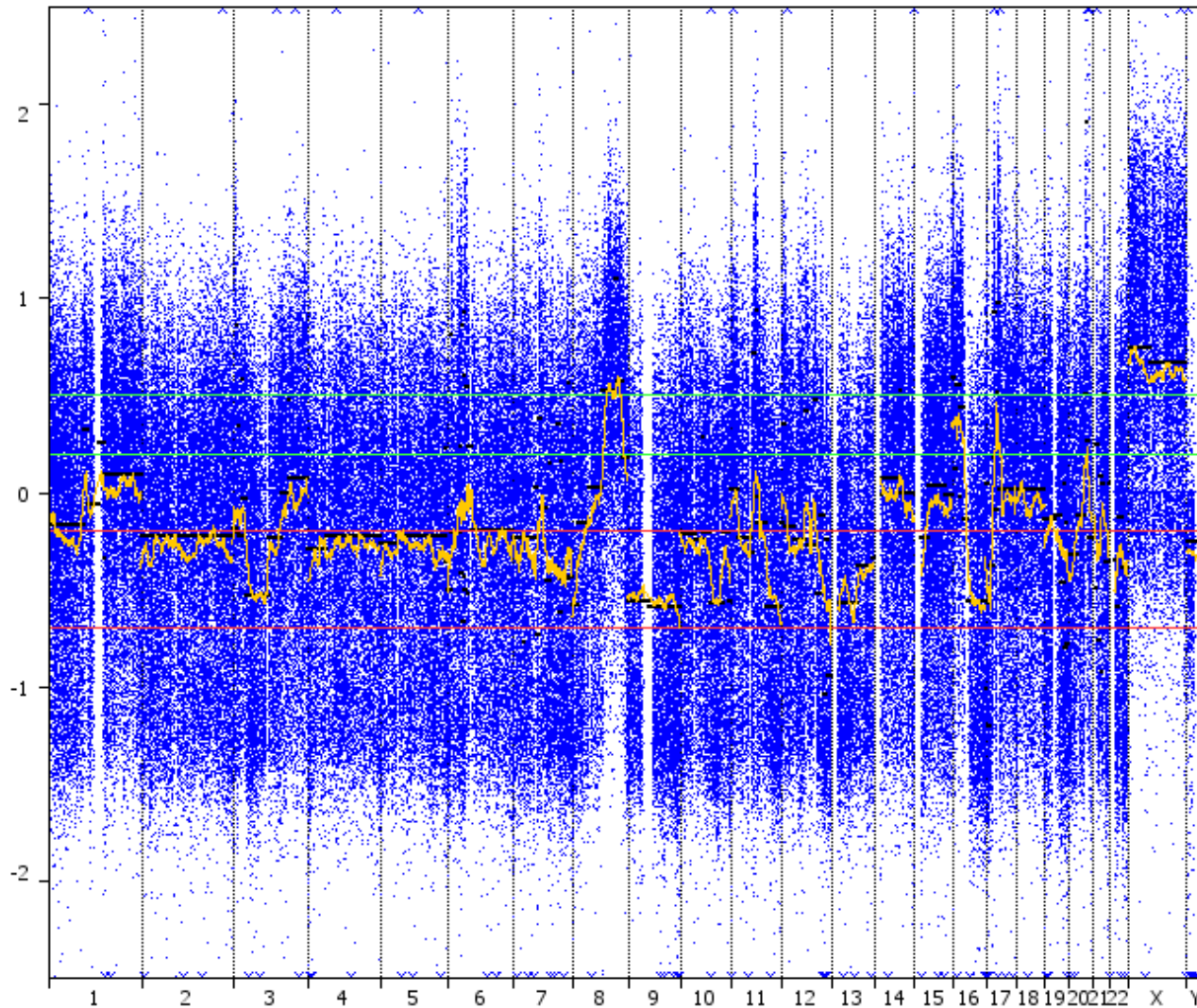


Robust Variance Sample QC

- Calculates the probe to probe variation across the genome
- A single parameter is used to remove extreme outliers that one would expect to be due to copy number breakpoints.



Example QC Scores



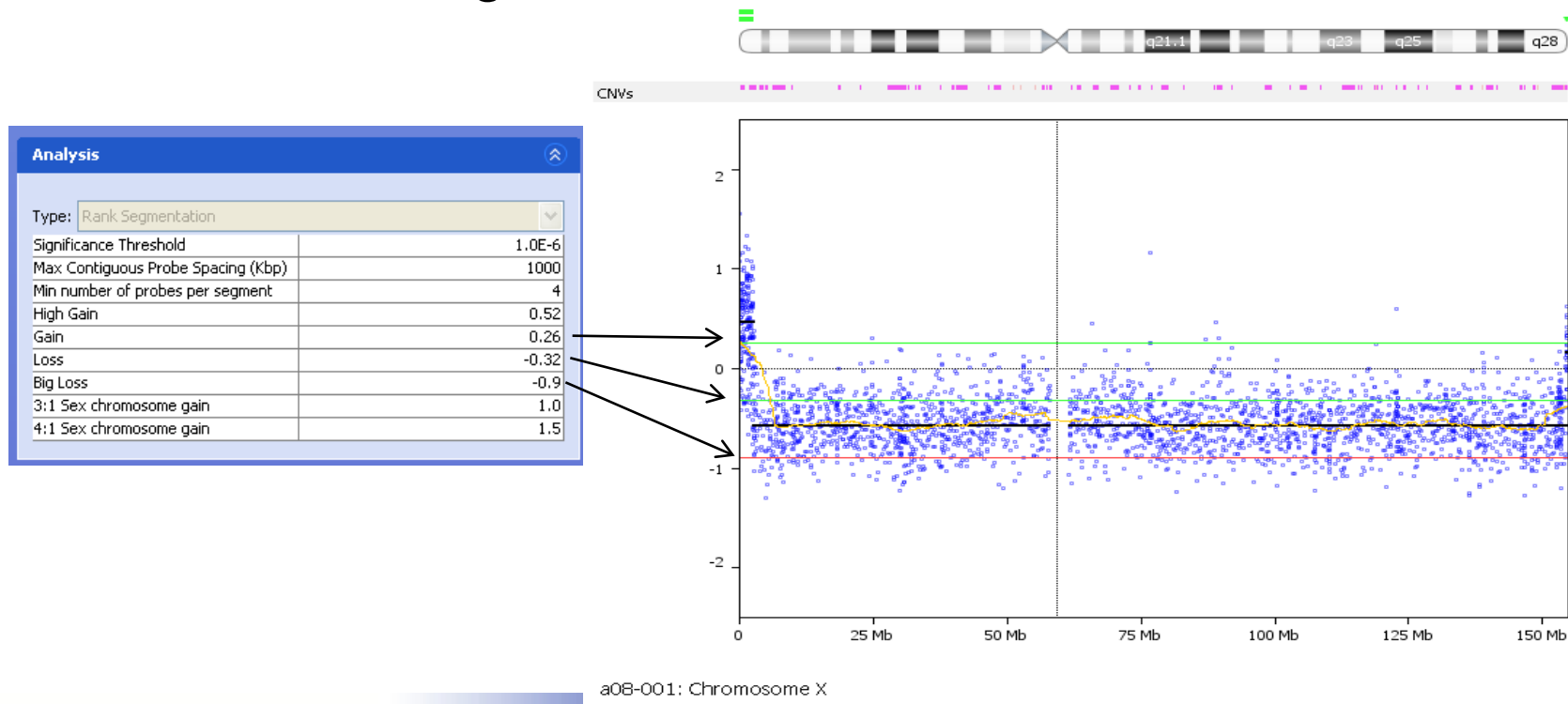
QC Score = 0.022

Sample: 7510 F7 (T)



Gender Specific Cutoffs

- New in version 4, if factors **Gender** and **Control Gender** are defined with values (Male or Female), the thresholds for the sex chromosomes are automatically set
- For example of sample is *Male* and reference is Female, the no change value for X chromosome



SNPRank Segmentation

Combining B-Allele Information with Log Ratio



Background

- Applicable only to SNP Arrays, e.g. Illumina and Affymetrix arrays
- Current support is provided for all Illumina arrays processed in Beadstudio so the column B-Allele Freq. is available in the output
- Affymetrix SNP 6 arrays processed through GTC yield .CNCHP files that contain B-Allele difference values that Nexus maps to B-Allele Freq
- Supported using custom data type to import pre-processed data



B-Allele Freq. Bands

Father:

AABABBBBABABBBBAAAABABB

Mother:

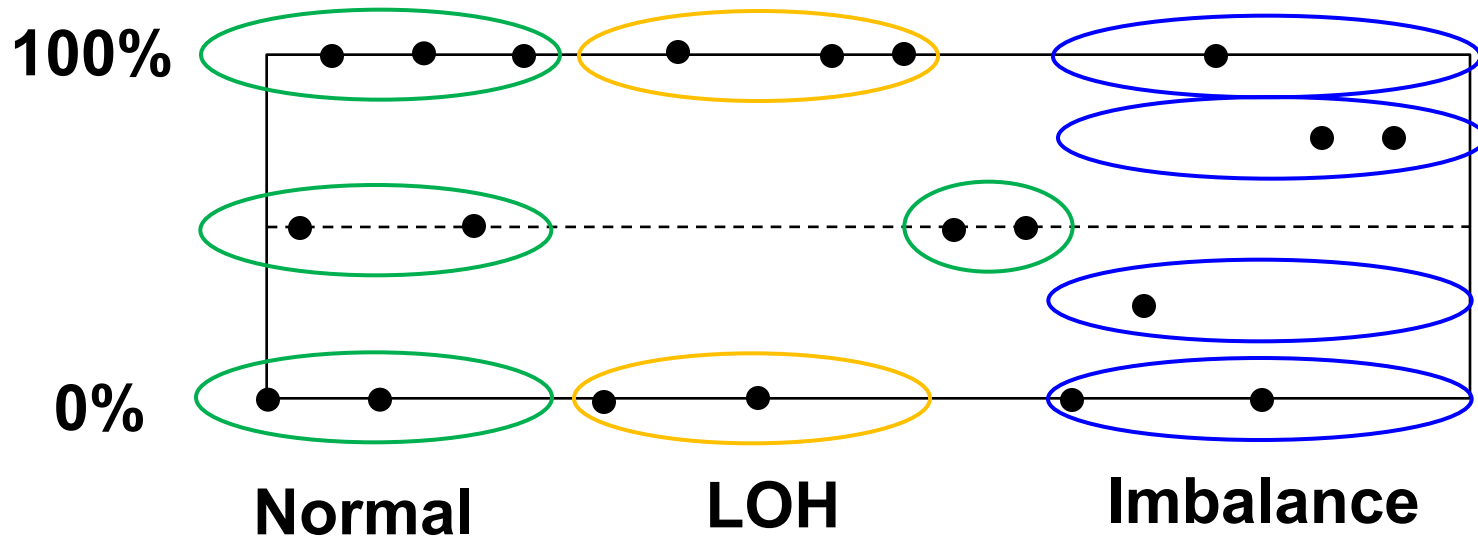
ABBABA

ABBABBAAA

AABABB



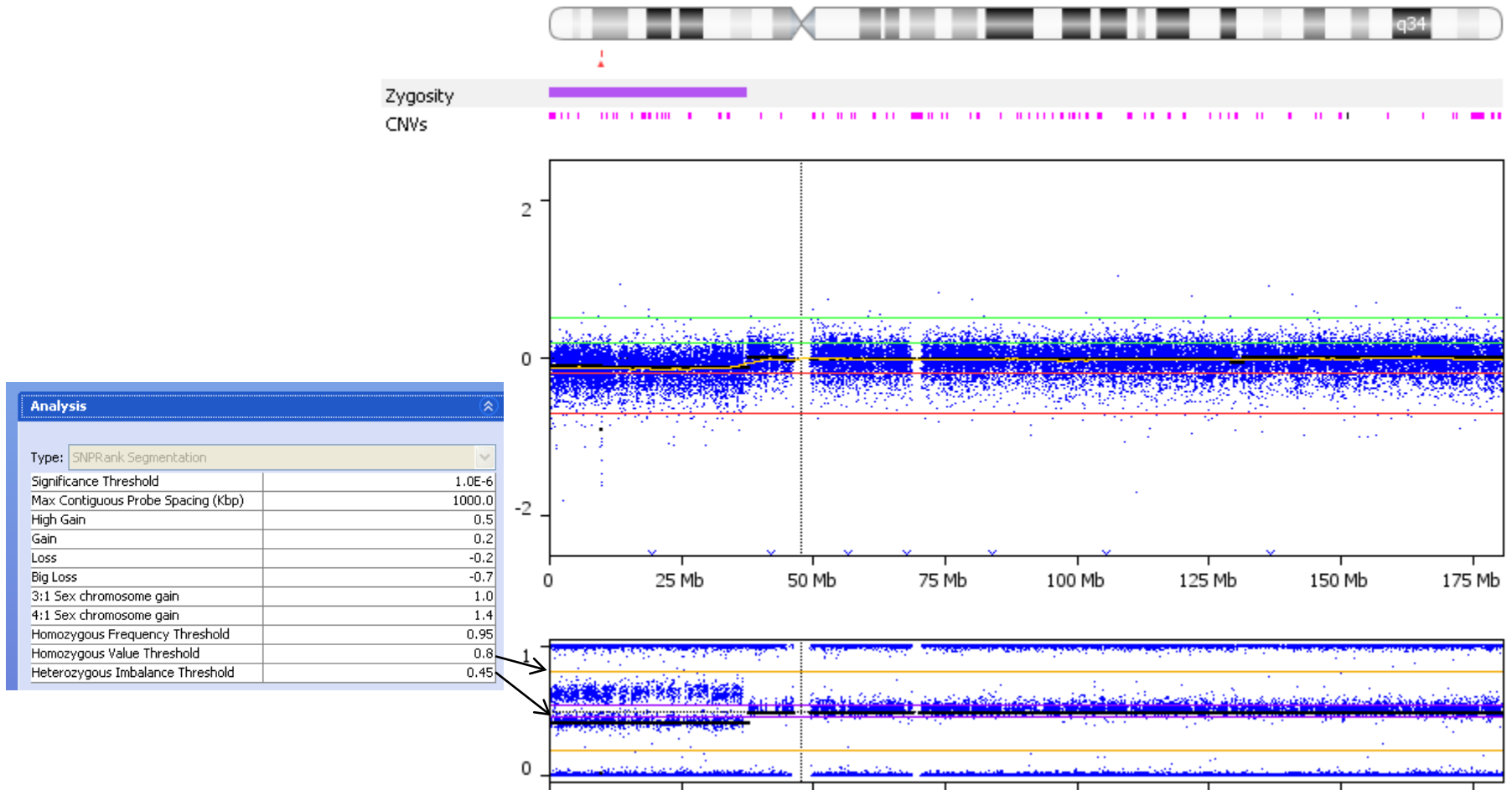
AA AB BB AA BB AB B A B A B B AB AB AAA AAB BBB AAA ABB ABB



		Allelic Loss		Allelic Imbalance	Normal
		LOH	Total Allelic Loss	Allelic Imbalance	
Gain	High Copy Gain	X		X	
	One Copy Gain			X	
Loss	One Copy Loss	X			
	Homozygous Loss		X		
Normal		X			

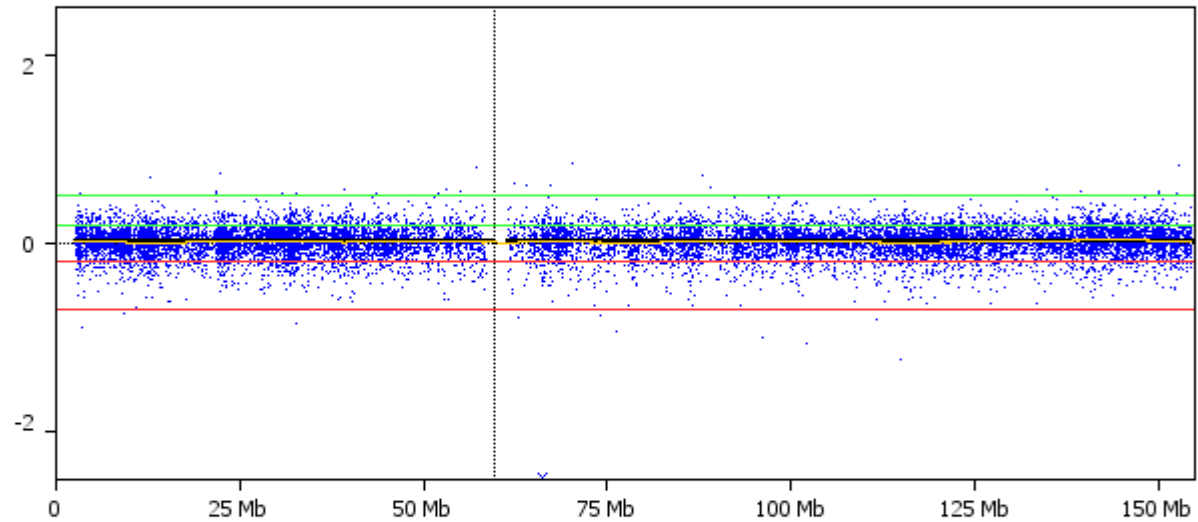
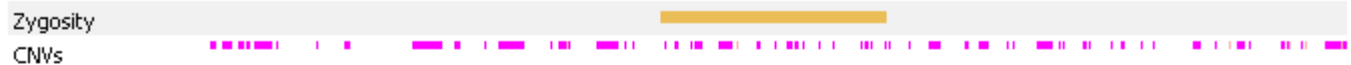


SNPRank Parameters

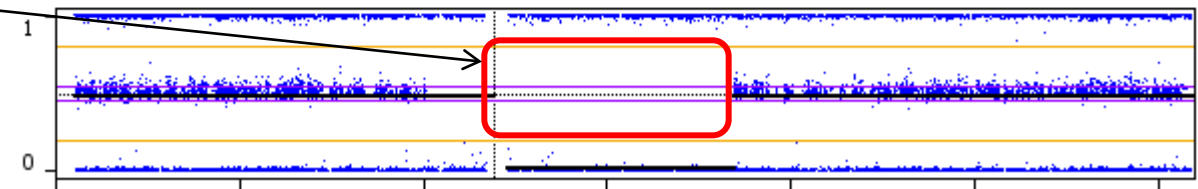


CG-3: Chromosome 5

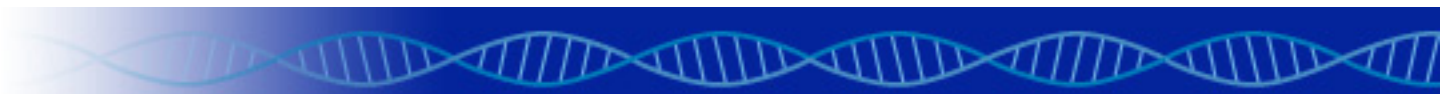
Detection of Homozygous Regions



Analysis	
Type:	SNPRank Segmentation
Significance Threshold	1.0E-6
Max. Contiguous Probe Spacing (Kbp)	1000.0
High Gain	0.5
Gain	0.2
Loss	-0.2
Big Loss	-0.7
3:1 Sex chromosome gain	1.0
4:1 Sex chromosome gain	1.4
Homozygous Frequency Threshold	0.95
Homozygous Value Threshold	0.8
Heterozygous Imbalance Threshold	0.45

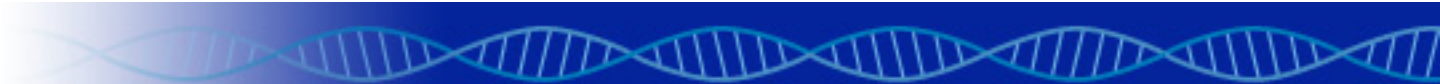


CG-3: Chromosome X



Section 2

Basics of Loading and Processing Data



What this Section Will Cover

- How to load data – demonstrate using a project containing samples from two array platforms
- Get an overview of the parameters for processing data
- Look at areas of common aberrations
- Generate graphical figures and reports



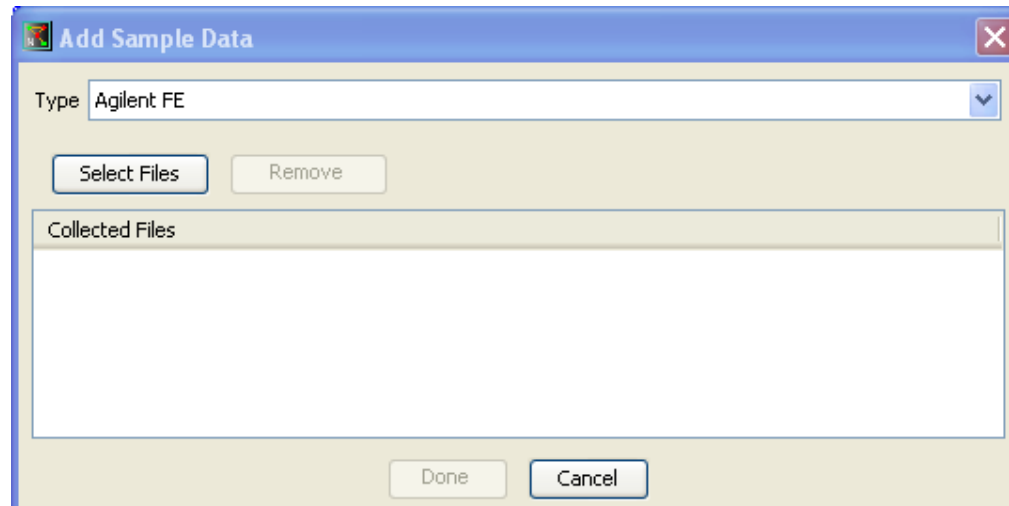
Some Common Data Types

- Affymetrix
- Agilent FE
- BlueFuse
- GenePix
- Illumina
- ImaGene
- NimbleGen



How to Load Data into Nexus Copy Number

- New in version 4, there are two methods for loading data
 - **Load Descriptor:** Select a file that contains all relevant information
 - **Load Data:** Use a file chooser to select the raw data files
- Load Data option is appropriate when no clinical factors are available



The Sample Descriptor

- A tab-delimited text file specifying sample names, any associated clinical data (Factors), and the location of the raw input files
- Needed to load raw data into Nexus Copy Number
- One sample descriptor file per data type
- To load multiple data types (e.g. Affymetrix, Illumina) into one project, create a sample descriptor for each data type



Factors

- Any clinical data associated with a sample. E.g. Sex, age, tumor type
- Each sample can have an unlimited number of Factors
- Optional
- Can specify in sample descriptor or enter manually in the Data Set tab



Sample Descriptor Templates

- Template files for the various Data Types are available in the 'Templates' folder in the installation directory
- Open the appropriate template in Excel
- Populate it and 'Save As' in another location



Support for Dye-Swap Experiments

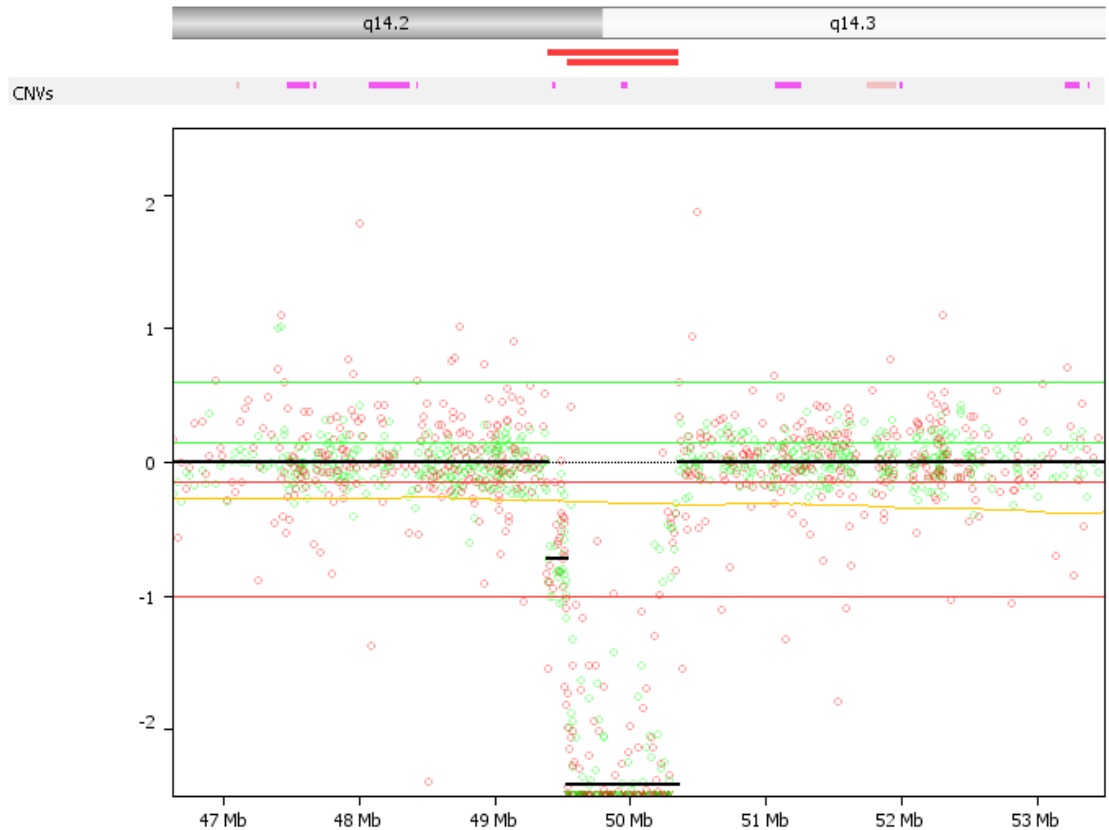
- New in version 4, you can indicate “replicate” arrays

	A	B	C	
1	Data Type:	Agilent FE		
2	Sample Name	File1	-File2	File
3	F1	F1cy5_251469319399.txt	F1cy3_251469319398.txt	
4	F2F	F2cy5_251469319401.txt	F2cy3_251469319400.txt	
5	F3F	F3cy5_251469319404.txt	F3cy3_251469319403.txt	
6	F4F	F4cy5_251469323249.txt	F4cy3_251469319405.txt	
7	F6F	F6cy5_251469321738.txt	F6cy3_251469321843.txt	45M
8	F7F	F7cy5_251469320894.txt	F7cy3_251469320857.txt	
9	F8F	F8cy5_251469323247.txt	F8cy3_251469320892.txt	
10	F9F	F9cy5_251469323249.txt	F9cy3_251469323248.txt	
11	F10F	F10cy5_251469321841.txt	F10cy3_251469321840.txt	F21

Combine Replicates Between Arrays

Type:

Analy:



F3F: Chromosome 13

Importing Other Data Types

- Open the file called customdatatypes.txt in the Nexus Copy Number installation folder and edit it

	A	B	C	D	E	F	G	H	I	J	K
1	Data Type	Probe	Full Location	Chromosome	BP Start	BP End	LogRatio	Flag	Structure	Header Starts	B Allele Freq
2	CNAG	ID		Chromosome	Position	Position	log2ratio_AB				
3	SPROC	Clone					Log2Rat	Bad_P			
4	Multi1	Name		Chromosome	Start	End			Multiple		
5	Unimi	Name		Chr	Position	Position	Log R Ratio				B Allele Freq
6	Quanti	Sample ID		Chromosome	Start (bp)	End (bp)	Copy Number		Segments		
7	Stanford550K	Name		Chr	Position	Position			Multiple		
8	TCGA_Mskcc_GBM	ProbeID		Chr	Pos.Start	Pos.End			Multiple		
9	TCGA_Harvard_GBM	CloneID		Ch	Pos	Pos	log2ratio				
10	Broad_SNP6seg	ID		chrom	loc.start	loc.end	seg.mean		Segments		
11	Broad_SNP6copynumber	Marker		Chromosome	PhysicalP	PhysicalP	log2ratio				
12	Stanford550K_Seg	Center		Chr	Start	Stop	Score		Segments		
13	Stanford550K_Combined	Name		Chr	Position	Position	Log R Ratio				B Allele Freq
14	TCGA_Mskcc_GBM_1	ProbeID		Chr	Pos	Pos			Multiple		
15	Mskcc_seg	ID		chrom	loc.start	loc.end	seg.mean		Segments		
16	Harvard_seg	ID		Chromosome	Start	End	Segment_Mean		Segments		
17	Harvard_multi	CloneID		Ch	Pos	Pos			Multiple		
18	Nimblegen GSE101089	ID_REF					VALUE			ID_REF	



Creating a Project

- Specify project name
- Select location of project
- Specify organism and build so that correct problemappings are loaded
- Click 'Create'



Defining New Organisms

- It is simple define new organisms by creating a new folder in the C:\Program Files\BioDiscovery\Nexus 4\Organisms folder
- Define probe locations and genome annotations (e.g., genes, miRNA, GO terms, etc.)



Loading Data

- Click on “Load Descriptor” and select the sample descriptor file
- Once data is loaded, each sample shows “Unprocessed” in the “Status” column
- Click “Load Descriptor” again to load additional samples in other sample descriptor files



Processing Data

- May need to adjust Settings
- Click “View” to process selected samples. Once processing is complete, results are displayed automatically



Results

- See graphical output in the Genome, Chromosome, and Summary tabs
- View numerical output in the Table, Aggregate, and Aggregate Participation tabs



Genome/Chromosome Tabs

- Sort samples
- Zoom in and out
- Sample Drill-down
- View annotations within the application or link out to external databases
- View by Factor Aggregates



Individual Sample Drill Down

- Ideograms, summary aberration plots, and numerical reports on individual samples
- Overview plot of gains and losses
- Annotation tracks
- Numerical reports with aberrant regions, cytoband location...



Saving Images and Exporting Reports

- Save images to file or copy to clipboard (options under Options menu)
- Print
- Export numerical data as tab-delimited text files



Additional Features in Nexus Copy Number Standard Edition

- New in version 4 – Query tool
 - Search for one or more genes, regions, etc. within the project
 - Returns the status of each sample for each query term as well as aggregate summary
- Aggregate table
 - Identify regions of change in a set percentage of population
 - Identify minimum common areas (Peaks only switch)
 - Identify “significant” regions using the frequency p-value calculated based on the STAC algorithm



STAC: A method for testing the significance of DNA copy number aberrations across multiple array-CGH experiments

Sharon J. Diskin, Thomas Eck, Joel Greshock, et al.

Genome Res. 2006 16: 1149-1158
Access the most recent version at doi:[10.1101/gr.5076506](https://doi.org/10.1101/gr.5076506)



Section 2

How to Integrate External Data into Nexus Copy Number



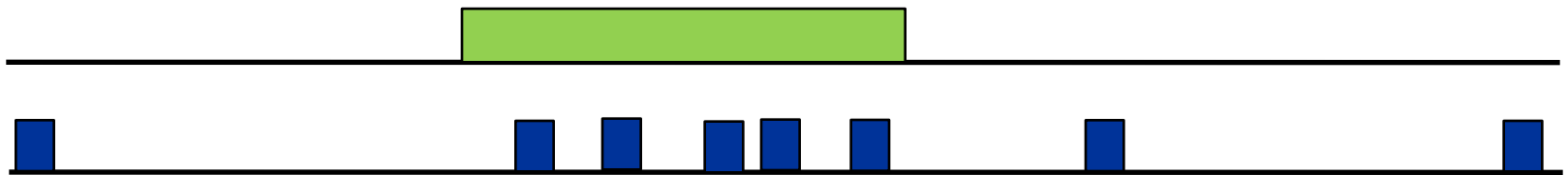
What is External Data?

- miRNA data
- Gene expression profile data



What can you do with it?

- Overlay expression results with copy number changes and view alongside each other
- Identify regions of “hot spots” where the correlation of differentially regulated genes falling in aberrant regions is significant (p-value)



What this Section Will Cover

- Expression Analysis using Nexus Expression
- Acceptable formats for loading data
- How to load expression data from an array or data from a published paper
- How to view and interpret the results



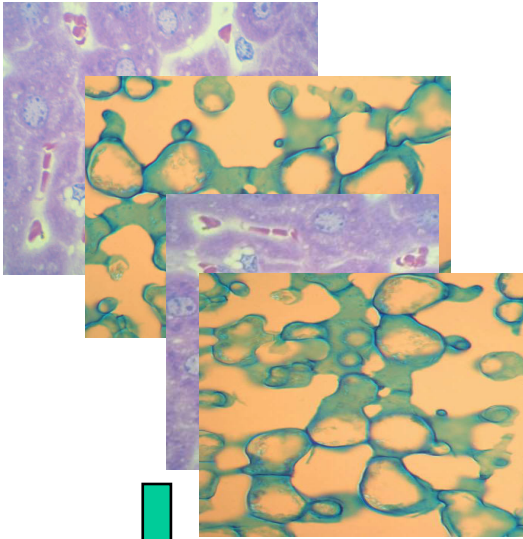
Gene Expression Analysis

General Overview



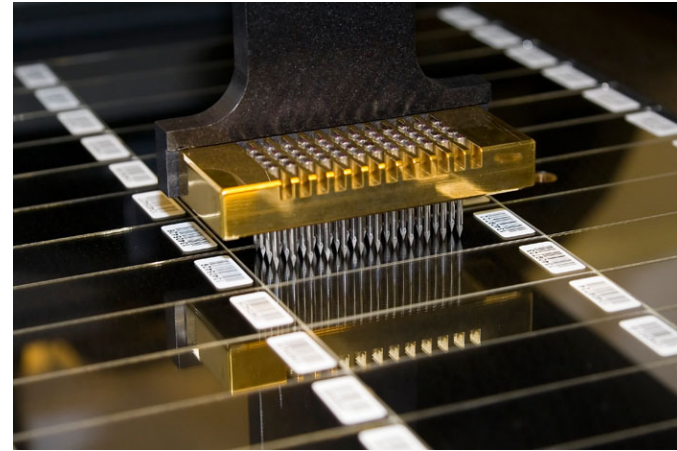
Information Flow

Biological Samples



Sample ID	Treatment	Time
Sample 1	Treated	0 hr
Sample 2	Untreated	0 hr
Sample 3	Treated	24 hr
Sample 4	Untreated	24 hr

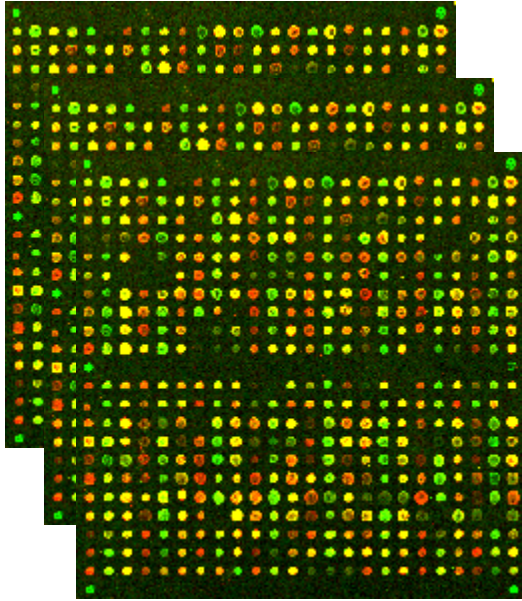
Arrays



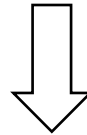
Probe	GeneSymbol	Bio Proc
Probe 1	ABC	Growth
Probe 2	SDF2	Decay
Probe 3	QRT	Growth
...



Hybridized Array Integrated Phenotypes with Genotype



Probe ID	Cy3 Strength	Cy5 Strength
P		
P		
P		
..		
P		
..		
...		
Probe 1	5000	2000
Probe 2	340	500
...



Model



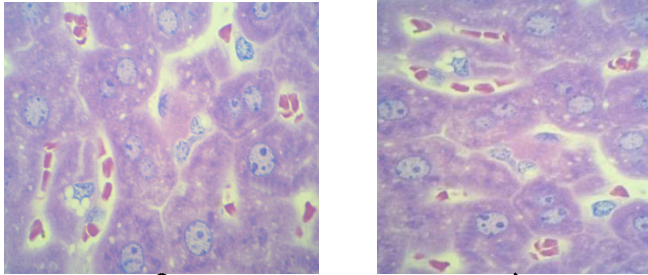
Objective of Gene Expression Analysis

- What genes are expressed differently between different samples
 - E.g. Genes that are significantly up regulated in lung cancer vs. breast cancer
- What do these genes do? Are there common biological themes



Experimental Setup

Grade 1

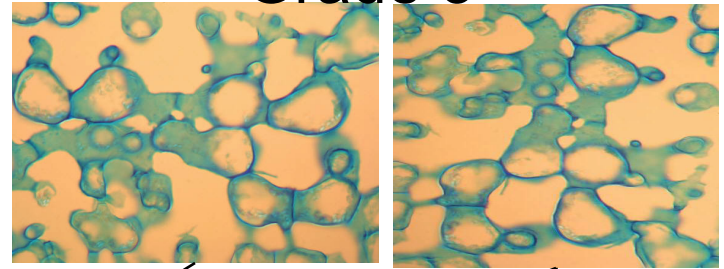


◦ Array 1

◦ Array 2

◦ Array 3

Grade 3



◦ Array 4

◦ Array 5

◦ Array 6



Nexus Expression Computation

- Transform measurements of probes on arrays to measurements of probes per sample
- Can remove batch effects in the process
- Estimate probe variance
- Comparisons are performed using a t-test of the results
 - FDR Correction
 - Pooling of variances based on intensity



Required Files

- Array Descriptor

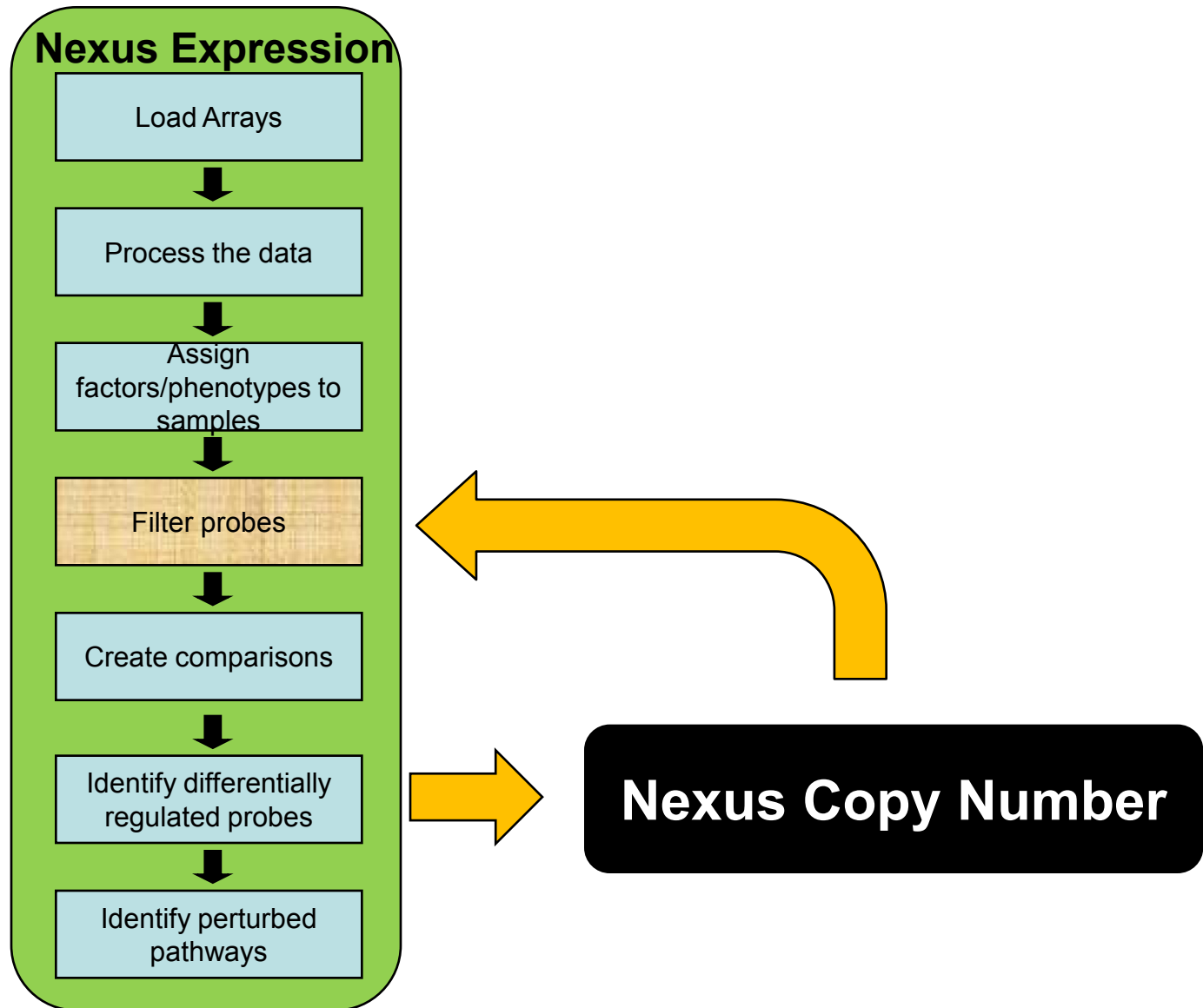
Data Type:	GenePix	Array Type:	Agilent/HG44K
Array	File	Channel 1 Sample	Hyb Month
Array 1	C:\Data\array1.gpr	Sample 1	June
Array 2	C:\Data\array2.gpr	Sample 1	July
Array 3	C:\Data\array3.gpr	Sample 2	June
Array 4	C:\Data\array4.gpr	Sample 3	June
Array 5	C:\Data\array5.gpr	Sample 4	July
Array 6	C:\Data\array6.gpr	Sample 4	July

- Sample Descriptor

Sample	Grade
Sample 1	1
Sample 2	1
Sample 3	3
Sample 4	3



Expression Analysis Data Flow



Input File Format

- **Gene Identifier**
 - Input file contains Gene Symbols
- **Probe Identifier**
 - Input file contains Probes
 - Extension of Gene Identifier format.
 - Not used commonly



Input File Containing Gene Identifier

- Required columns: **Gene Symbol, Regulation**
- Optional columns: **p-value, log-ratio**



Input File Containing Probe Identifiers

- Required columns: **Probe**, **Regulation**
- Optional columns: **p-value**, **log-ratio**
- Also need a probe to gene symbol mapping file placed in the **ProbeSymbolMappings** folder
- Need to edit **mappingdescriptors.txt** file in **ProbeSymbolMappings** folder to specify where to locate the mapping file and to indicate columns used in the mapping file



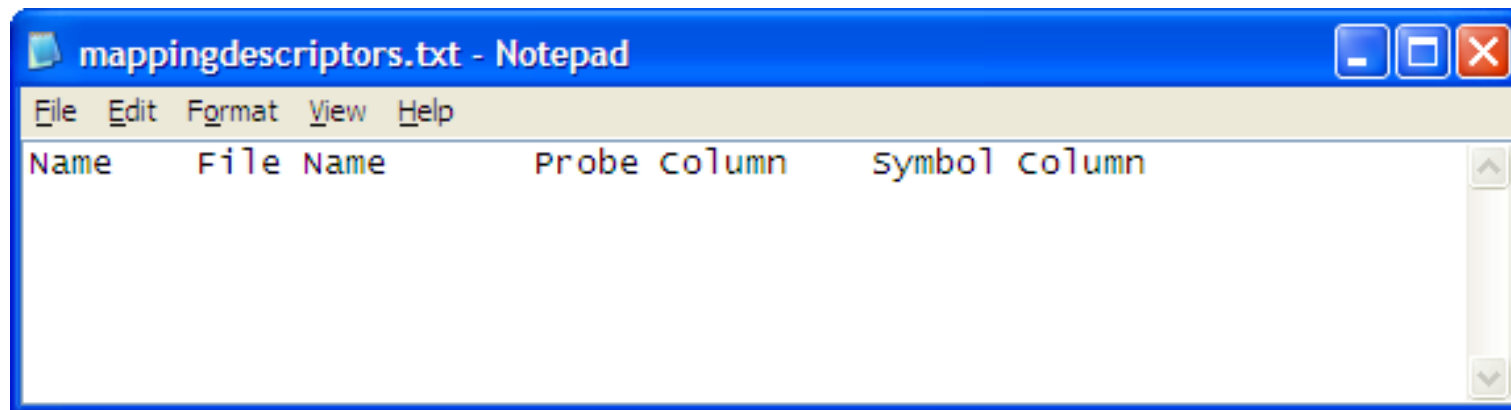
Input File Columns/Fields

- **Regulation** – possible values: **up** and **down**. Anything else in place of these values will mean no change.
- **p-value** and **log-ratio** – If present in input file, value will be displayed when moving the mouse over the expression data track tick marks in the Genome/Chromosome tabs



Probe Identifier File Format - Additional Requirements

- Probe to gene symbol mapping text file in **ProbeSymbolMappings** folder
- **mappingdescriptors.txt** file in **ProbeSymbolMappings** folder



Loading External Data

- Within **External Data** tab
- Click **Add** button
- Select files to load



External Data Tab

- Contains sub tabs, one for each external data type. E.g. **Expression, miRNA**
- Each sub tab
 - Lists all external data files loaded into project
 - User selects what to view by marking off “up” and “down” checkboxes
 - Can edit **Description** column to add a description for each external data file that was loaded



Viewing Results - Graphical

- In Genome/Chromosome tabs, select **Expression** from the **View** menu
- One track for each external data file
- Up-regulated in magenta
- Down-regulated in blue
- Move mouse over tick marks to display name and additional info. such as p-value, log-ratio when available



Viewing Results - Numerical

- **Expression p-value** column in Aggregate tab
- **Expression** tab in Enrichment results
- **Expression p-values** tab in Annotation window
- **Expression p-values** tab in Comparison window



Expression tab in Enrichment results

Term	External data file name
P-value	Significance of these genes being differentially regulated and appearing in this aberrant region.
Q-bound	FDR-corrected (for multiple testing) p-value
Present	The number of differentially regulated genes present in this external data set that are located in this aberrant region
Total	Total number of genes in the external data file



Expression p-values tab in Annotations Window

Comparison	External data file name
Total Genes	Number of genes in this region
Diff. Regulated Genes	Number of differentially regulated genes located in this region
P-value	Significance of these genes being differentially regulated and appearing in this aberrant region.



Section 3

Clustering



What is the Purpose of Clustering?

- To discover natural groupings in a set of samples without prior knowledge of any class labels
- Can help determine what factors are affecting certain phenotypes



What this Presentation Will Cover

- Available clustering algorithms in Nexus Copy Number
- Dendrogram
- Factor enrichments within clustering analyses



Clustering Settings

- Available from the Options menu (Options -> Clustering Settings)
- Clustering window has
 - Clustering algorithms
 - And associated parameters



Clustering Algorithms and Parameters

- K-means
 - Ignore sex chromosomes
 - Cluster count
- Average Linkage Hierarchical
 - Ignore sex chromosomes
- Single Linkage Hierarchical
 - Ignore sex chromosomes
- Complete Linkage Hierarchical
 - Ignore sex chromosomes



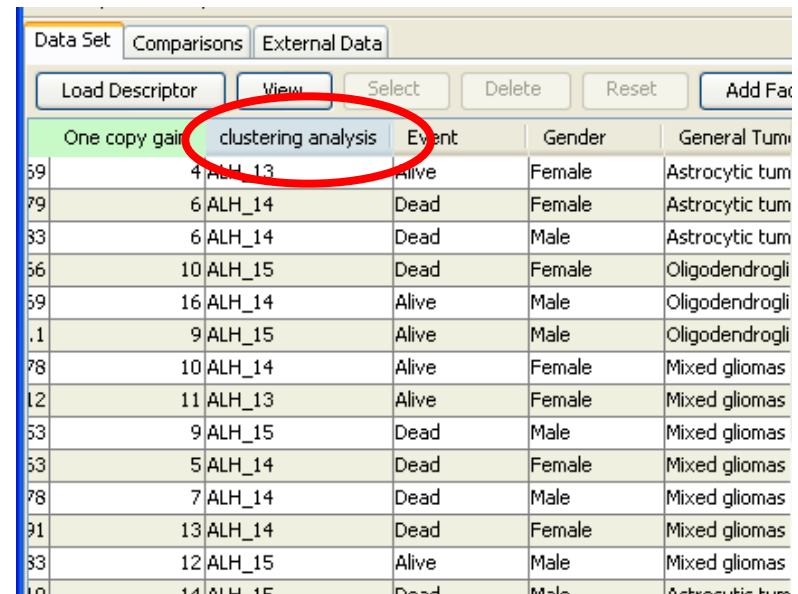
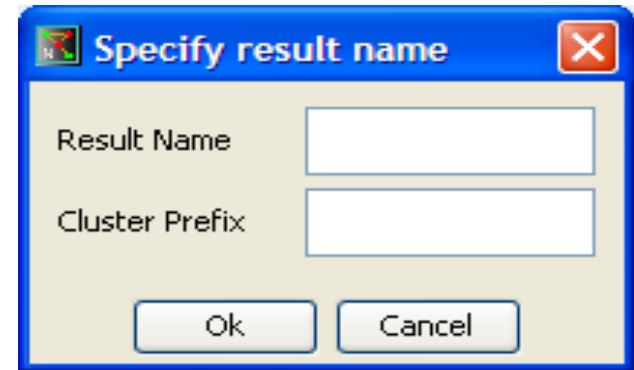
Dendrogram

- Drawn if using a hierarchical clustering algorithm
- Can adjust cluster count
- Can label ends of tree branches with factors
- Dendrogram image can be printed or saved as an image file
- Dendrogram for a previously performed clustering analysis cannot be re-drawn within Nexus Copy Number



Saving Clustering Results

- Can save clustering results by specifying a result name and cluster prefix
- A new column specifying which cluster each sample belongs to is added to Data Set tab
- Clustering column headers have blue background



	One copy gain	clustering analysis	Event	Gender	General Tum
59	4	ALH_13	Alive	Female	Astrocytic tum
79	6	ALH_14	Dead	Female	Astrocytic tum
83	6	ALH_14	Dead	Male	Astrocytic tum
56	10	ALH_15	Dead	Female	Oligodendrogl
59	16	ALH_14	Alive	Male	Oligodendrogl
.1	9	ALH_15	Alive	Male	Oligodendrogl
78	10	ALH_14	Alive	Female	Mixed gliomas
12	11	ALH_13	Alive	Female	Mixed gliomas
53	9	ALH_15	Dead	Male	Mixed gliomas
53	5	ALH_14	Dead	Female	Mixed gliomas
78	7	ALH_14	Dead	Male	Mixed gliomas
91	13	ALH_14	Dead	Female	Mixed gliomas
83	12	ALH_15	Alive	Male	Mixed gliomas
10	14	ALH_15	Dead	Male	Astrocytic tum

Numerical Clustering Results in Data Set Tab

- Right click on column header to get additional options
 - **Details**
 - shows algorithm and parameters used for this clustering analysis
 - **Factor Enrichment**
 - provides statistics on the number and percent of samples present in each factor value group within each cluster
 - provides a p-value for the likelihood of these samples clustering into the particular cluster by chance.



Graphical Results

- In Genome/Chromosome tabs
- **View -> Factor Aggregates**



Section 4

Comparisons



What Can We Learn From Performing Comparisons?

- See differences between two sets of samples in a project
- Find regions that are significantly different between the two groups



What this Presentation Will Cover

- Different comparison types
- How to create a comparison
- How to view and interpret the results



Comparison Types

- **Avg. of others** - One factor set is compared to the average of all the other factor sets
- **Sequential** - Compares each factor set in a list to the subsequent factor set in the list
- **Paired** – Pairs up specific factor sets for comparison
- **Selected** – Pairwise comparison of a factor set with each other factor set
- **Custom** – Compares one factor set to average of user-selected factor sets



Comparisons Tab

- Create/Add comparisons
- Delete comparisons
- Edit comparison names
- View details of which samples belong in each group being compared



Graphical Results

- Genome/Chromosome tabs in Comparisons window
- Frequency plot of each group in bottom panel
- Top panel shows frequency difference between the two groups in a Comparison
 - red = loss regions
 - green = gain regions



Frequency Difference Plot

- Gain magnitude of first group $>$ gain magnitude of second group, difference plotted in green above 0 baseline
- Gain magnitude of first group $<$ gain magnitude of second group, difference plotted in green below 0 baseline
- Loss magnitude of first group $>$ loss magnitude of second group, difference plotted in red below the 0 baseline
- Loss magnitude of first group $<$ loss magnitude of second group, difference plotted in red above the 0 baseline



Significant Track

- Green bars represent gain regions where difference is significant
- Red bars represent loss regions where difference is significant
- Based on p-value threshold and differential threshold (minimum percentage difference)



Numerical Results

- Table in the **Regions** tab
- Contains the regions which meet both the **Differential Threshold** and the **P-Value Threshold**
- P-value from Fisher's Exact test



Wrap Up

- Many other features not covered here
- Online training seminars on various topics
- Join User Forum for important announcements including training webinars:

<http://www.biodiscovery.com/index/siteforum-app>

