

# Agilent CytoGenomics 2.0

For Detection of CNC, LOH and UPD



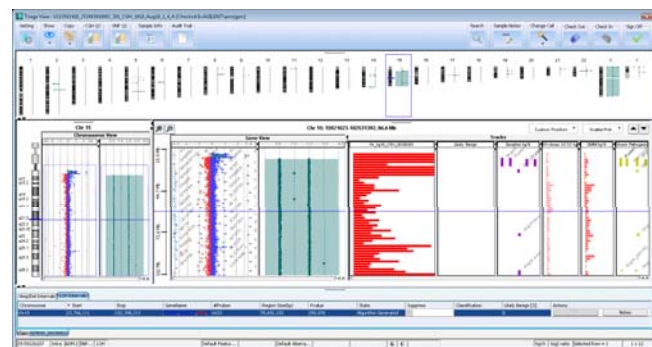
**New algorithms for CGH+SNP analysis of hematological tumor and constitutional samples**

**Arne Ijpma, Ph.D.  
Product Manager**



# Agilent CytoGenomics 2.0

## Value Proposition



Sample Information									
Array ID	US2102418_1_2								
Global Display Name	US2102418_1_2_02022011_EU_Male_Agilent_1_2_02022011								
Guest Sample	European Male (NA12891_1)								
Red Sample									
Test Summary Report for Sample									
US2102418_1_2_02022011									
Location	Size	Chromosome	#Probes	Copy	Copy	Copy	Copy	Copy	Copy
chr1:100,000,000-100,000,000	100,000,000	1	100,000	2	2	2	2	2	2
chr2:200,000,000-200,000,000	200,000,000	2	200,000	2	2	2	2	2	2
chr3:300,000,000-300,000,000	300,000,000	3	300,000	2	2	2	2	2	2
SNP Table View for Sample									
US2102418_1_2_02022011									
Location	Size	Chromosome	#Probes	Copy	Copy	Copy	Copy	Copy	Copy
chr1:100,000,000-100,000,000	100,000,000	1	100,000	2	2	2	2	2	2
chr2:200,000,000-200,000,000	200,000,000	2	200,000	2	2	2	2	2	2
chr3:300,000,000-300,000,000	300,000,000	3	300,000	2	2	2	2	2	2
Analysis Settings									
Alteration Algorithm	AZIM 2	Threshold	0.0						
GC Correction	ON	Window Size	200						
Centralization	ON	Bin Size	10						
Centralization Threshold	0.0	SNP Copy Number	ON						
SNP CN Confidence Level	0.95	LOH	ON						
LOH Threshold	0.0	Fuzzy Zone	OFF						
Resolving Level	0	Combine Replicates/Array	ON						
Genotype	Agilent	Alteration Filters	Agilent						
Feature Level Filters	gdcThreshold = true OR mcdThreshold = true OR gdcThreshold = true OR mcdThreshold = true OR	Alteration Filters	gdcThreshold = 0.0 AND mcdThreshold = 0.0 AND mcdThreshold = 0.0 AND mcdThreshold = 0.0						
Array Level Filters	NONE	Design Level Filters	NONE						
Genomic Boundaries	Not Applied	Metric Test Filters	NONE						

To enable the Cytogenetic Researcher to analyze a normal CGH (or CGH+SNP) sample in less than 5 minutes, including report generation, while at the same time providing the tools needed for analysis of complex samples such as hematological cancers.

# Agilent CytoGenomics 2.0

## Result of the cytogenetic analysis: Cyto Report

Cyto Report

Sample Information

Array ID : 252983010002\_1\_1

Global Display Name : US23502418\_252983010002\_S01\_CGH\_1010\_Aug 10\_1\_1

Green Sample : European Male (NA12891\_v1)

Red Sample : NA09208

Polarity : 1

DerivativeOfLogRatioSD : 0.167981

Amp/Del Intervals Table

Location	Size	Cytoband	#Probes	Amp/Del	P-value	Annotations
chr17:24457-5901054	5,876,598	p13.3 - p13.2	340	-0.866466	0.00E00	DOC2B, RPH3AL, C17orf97, FAM101B, VPS53, FAM57A, GEMIN4, ELP2P, GLOD4, RNMTL1, NXN, TIMM22, ABR, BHLHA9, TUSC5, YWHAE, CRK, MYO1C, INPP5K, LOC100306951, PITPNA, SLC43A2, SCARF1, RILP, PRPF8...

Amp=Amplification  
Del=Deletion

Total Amp/Del Intervals: 1

LOH Intervals Table

Location	#Probes	LOH Score	Annotations
chr17:72083-5853691	84	8.03E00	RPH3AL, C17orf97, FAM101B, VPS53, FAM57A, GEMIN4, ELP2P, GLOD4, RNMTL1, NXN, TIMM22, ABR, BHLHA9, TUSC5, YWHAE, CRK, MYO1C, INPP5K, LOC100306951, PITPNA, SLC43A2, SCARF1, RILP, PRPF8, TLCD2...

Total LOH Intervals: 1

Analysis Settings

Aberration Algorithm : ADM-2

Threshold : 6.0

Window Size : 2Kb

Bin Size : 10

SNP Copy Number : ON

LOH : ON

Fuzzy Zero : OFF

Combine Replicates(Intra Array) : ON

Aberration Filters : minProbes = 3 AND minAvgAbsLogRatio = 0.25 AND maxAberrations = 100000 AND percentPenetrance = 0

Design Level Filters : NONE

Metric Set Filters : NONE

Design : 029830\_20100921

GC Correction : ON

Centralization : ON

Centralization Threshold : 6.0

SNP CN Confidence : 0.95

Level

LOH Threshold : 6.0

Nesting Level : OFF

Genome : hg19

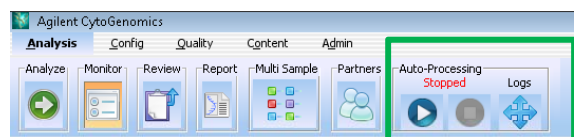
Feature Level Filters : glsSaturated = true OR rlsSaturated = true OR glsFeatNonUnifOL = true OR rlsFeatNonUnifOL = true

Array Level Filters : NONE

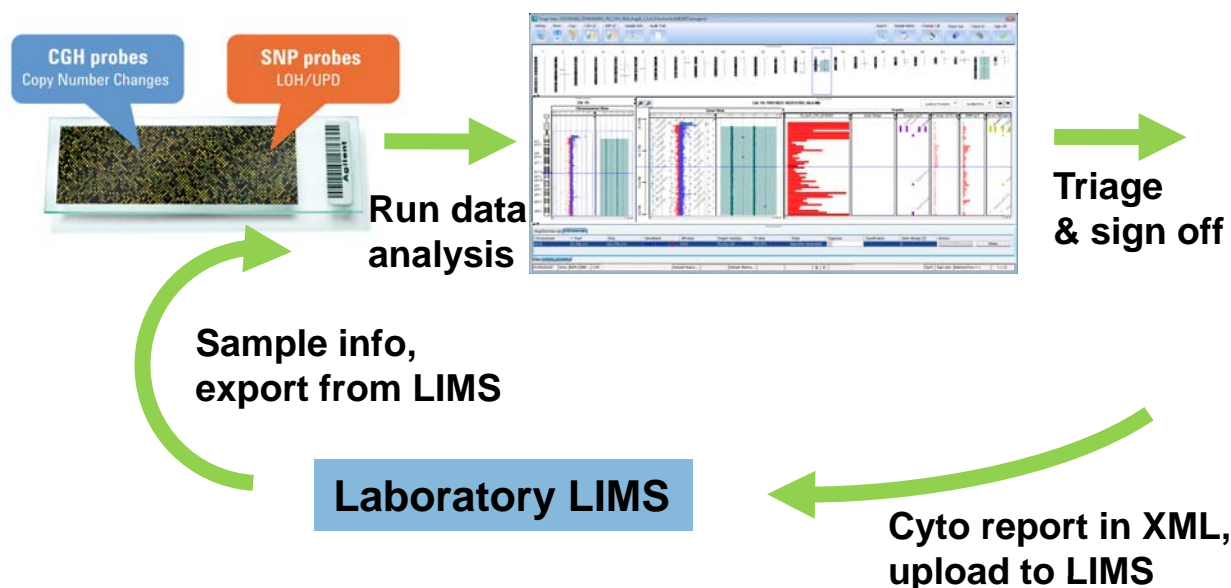
# Agilent CytoGenomics 2.0

## Streamlined data analysis for Cytogenetic research labs

- Simple and quick workflow for data analysis: Automation of data upload and analysis



- Convenient input and output support allowing for integration with LIMS



### Cyto Report

Sample Information					
Array ID	252983010001_1_2				
Global Display Name	US23502418_252983010001_S01_CGH_1010_Aug10_1_2_02022011				
Green Sample	European Male (NA12891_v1)				
Red Sample					

Text Summary Report for Sample US23502418_252983010001_S01_CGH_1010_Aug10_1_2_02022011					
Location	Size	Cytoband	#Probes	CopyCat	Probes
chr5:68123300-70636424	1,024,294	q11.2	37	0.296018	0.00000
LOC1007210C, GTF2H2D, LOC10072216, GUSBP3, SERP1A, SERP1B, SRN2, SRN1, LOC10072009, LOC1007210B, NBP, GTF2H2, LOC1007210C					
chr5:259315-362290	102,972	p25.3	14	0.400728	0.00000
chr5:31973370-32005082	35,912	p21.33	35	0.389420	0.00000
GUSP22, CYP21A2, TNXA, TNXB					

SNP Table View for Sample US23502418_252983010001_S01_CGH_1010_Aug10_1_2_02022011					
Location	#Probes	LOH Score	Annotations		
chr1:76226484-7970067	87	7.81E30	ACAC8, BARO07B, SNORD40C, SNORD45A, SNORD45B, MSN4, A5817, STGALNAC3, STGALNAC5, PGK, AK5, ZZZ3, USP3, FAM75A, NEXN, PUSP1, DNAH4, GPC2, MUC27B2, ITGFB		
chr1:172013001-197915297	517	4.60E01	C10orf1, FASLG, TNFSF18, TNFSF4, PRDM1, SLC5A11, ANKRD45, RPL20, CENPL, DARS2, GARS, SNORD1, SNORD47, SNORD80, SNORD79, SNORD78, SNORD44, SNORD77, SNORD76, SNORD75		
chr1:233464687-245963058	189	1.84E01	KALY10A, KIF11, SLC6F3, C10orf1, TAGBP1, RFBP2, TOMM20, SNORA14B, RSKA, ARID4B, GGP31, TCCE, SODALNT2, GACA, LYST, MBR1507, NDI1, GPR137B, ERO1LB, EDARADD		
chr2:211536392-221570900	245	2.30E01	CPS1, ERBB4, MIR548F2, KIF22, SPAG16, VWZL2, BARD1, ARCA12, ATRC, FHL, MRE6, PEG3, TMEM189, KRCOS, MARCH4, SMARCA1, RPL37A, KIFBP2, KIFBP5, TNF1		

Analysis Settings					
Alteration Algorithm	ADRA-2	Threshold	6.0		
GC Correction	ON	Window Size	2Kb		
Centralization	ON	Bin Size	10		
Centralization Threshold	6.0	SNP Copy Number	ON		
SNP CN Confidence Level	0.95	LOH	ON		
LOH Threshold	6.0	Fuzzy Zero	OFF		
Nesting Level	0	Combine Replicates/Intra Array	ON		
Genome	hg19	Aberration Filters	minProbes = 6 AND minAvgAbsLogRatio = 0.0 AND maxAberrations = 100 AND percentPenetration = 0		
Feature Level Filters	ghoSaturated = true OR rfsSaturated = true OR ghFeatNonIntOL = true OR rfsFeatNonIntOL = true	Design Level Filters	NONE		
Array Level Filters	NONE	Metric Set Filters	NONE		
Genomic Boundaries	Not Applied				

# Agilent CytoGenomics 2.0

## Workflow setup and user roles

- Preloaded default workflows
- Initial set up: Configuring the Workflow

1. Configure analysis method

2. Configure CytoReport

3. Configure Workflow

- Day to day usage:
  - **Technician (manual) / Auto-processing**
    1. Upload TIFF images (or FE files)
    2. Map array to sample attributes (manual or upload SAF file)
    3. Engage **configured workflow** to run FE (TIFF only) and analysis
    4. Initial sample Triage
  - **Scientist**
    1. Customize analysis method and workflows
    2. Final Triage of sample and sign off
    3. Generate the CytoReport

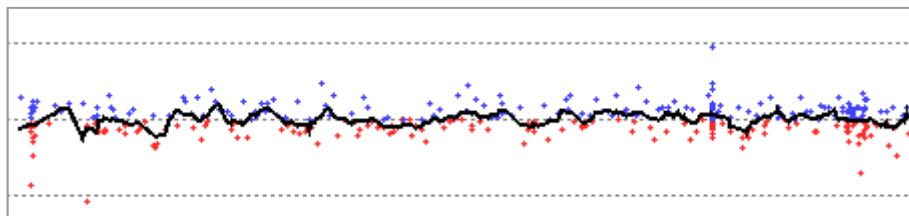


# Agilent CytoGenomics 2.0

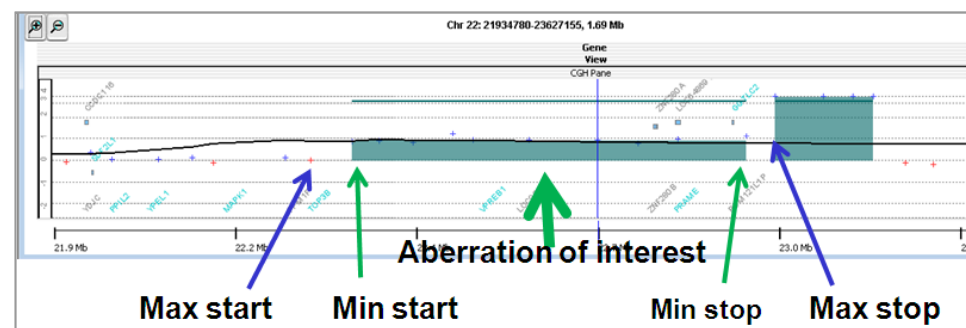
## New Features

- Most changes are ‘under the hood’ algorithm changes that provide support for CGH+SNP analysis for hematological cancer samples

- Moving Average



- Provide Min/Max Start and Stop coordinates for aberrations In reports



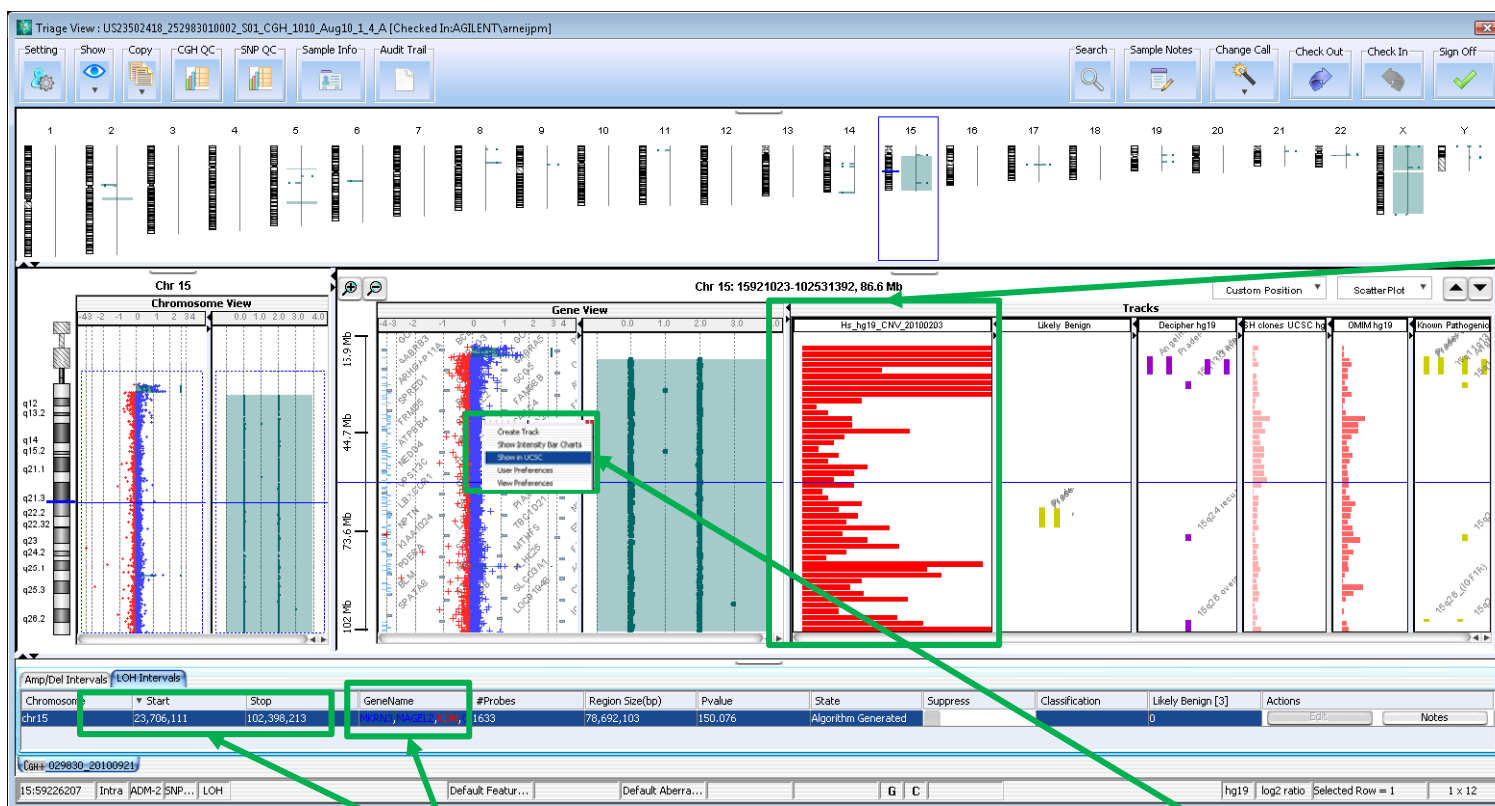
- Easy access to CGH&SNP QC metrics (added SNP metric thresholds)

No.	Array ID	Global Display Name	FE QC Report	CGH&SNP Fit	CGH&SNP QC Flag	Status	Green Sample	Red Sample
1	253333810001_1_4	US90603641_253333		<a href="#">View</a>	Evaluate(1/20)	Analyzed	European Male (NA109-0008 (Male))	
2	253333810002_1_4	US90603641_253333		<a href="#">View</a>	Evaluate(3/20)	Analyzed	European Male (NA109-0032)	
3	253333810001_1_3	US90603641_253333		<a href="#">View</a>	Evaluate(1/20)	Analyzed	European Male (NA109-0022 (Male))	



# Agilent CytoGenomics 2.0

## Sample Triage – Access to external databases



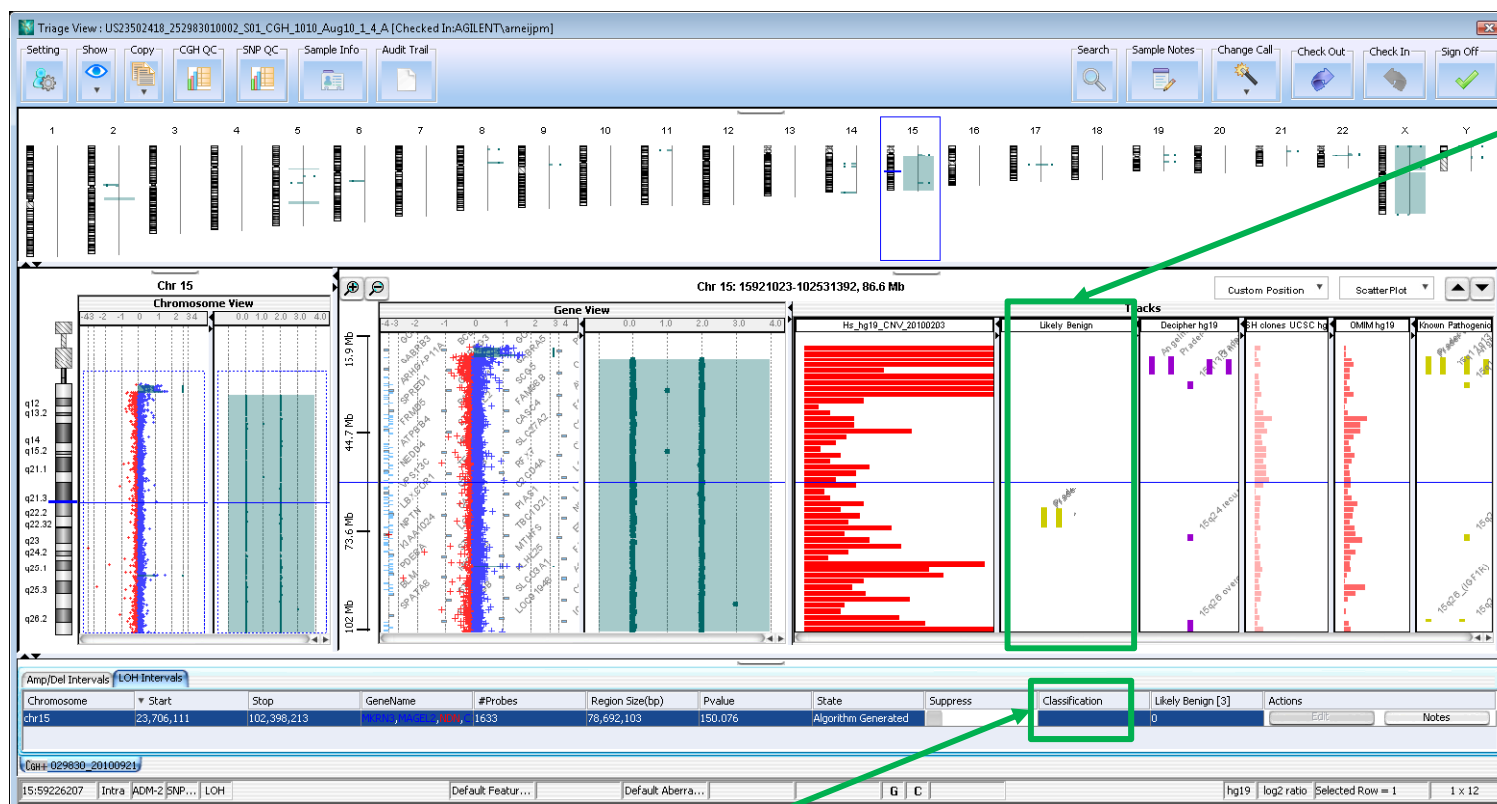
Preloaded track:  
DGV-CNV

Link out from aberrations and associated genes  
to external Databases: UCSC, DGV, OMIM,  
Entrez

View and compare  
aberrations in DGV and  
UCSC genome browser

# Agilent CytoGenomics 2.0

## Sample Triage – Using Classification and CytoGenomics Database



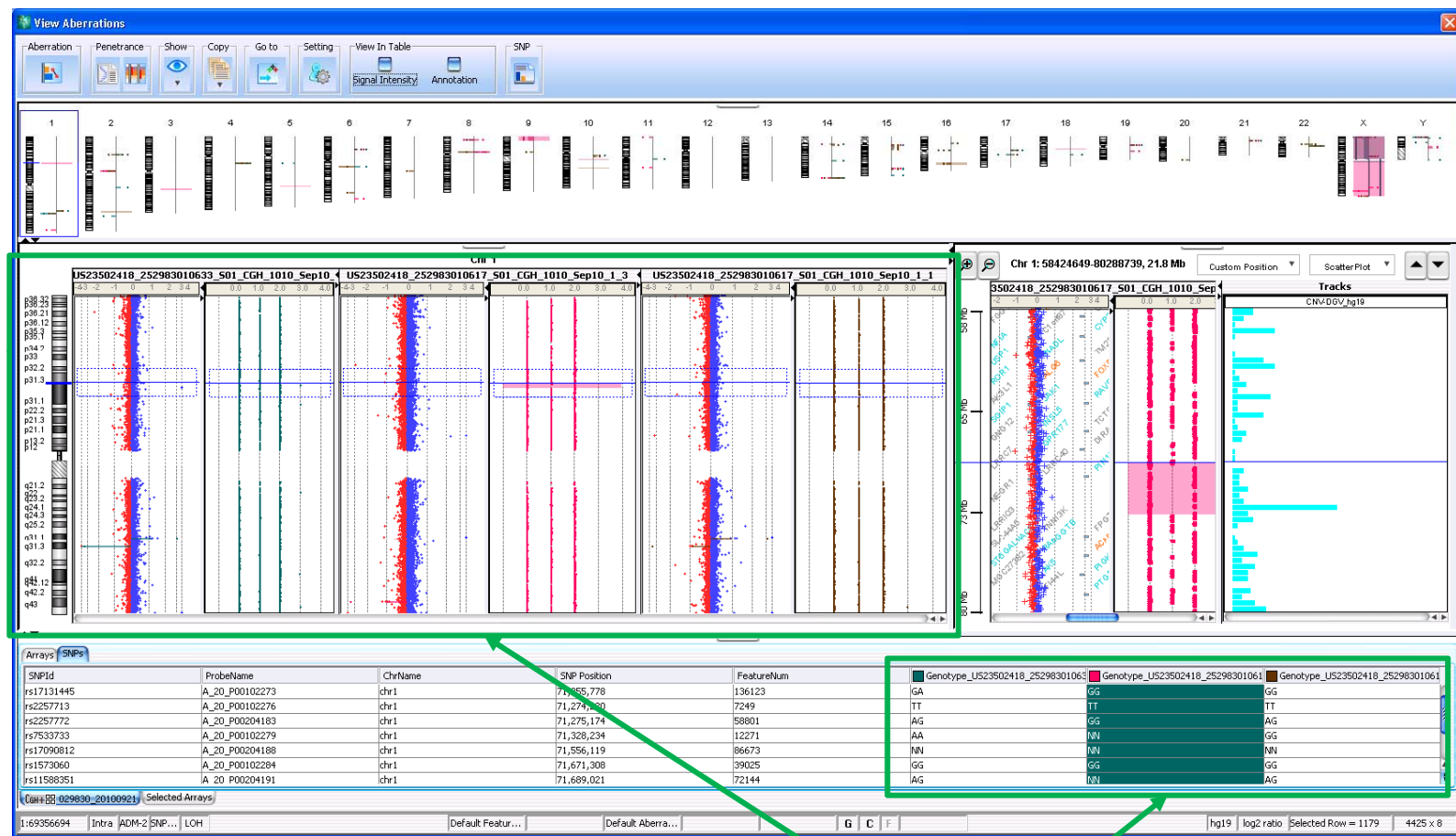
Auto updated tracks based on aberration classification in CytoGenomics Database

Right click on 'Classification' allows query for overlapping aberrations in CytoGenomics database: results will be shown as new track along the gene view



# Agilent CytoGenomics 2.0

## Multi Sample View

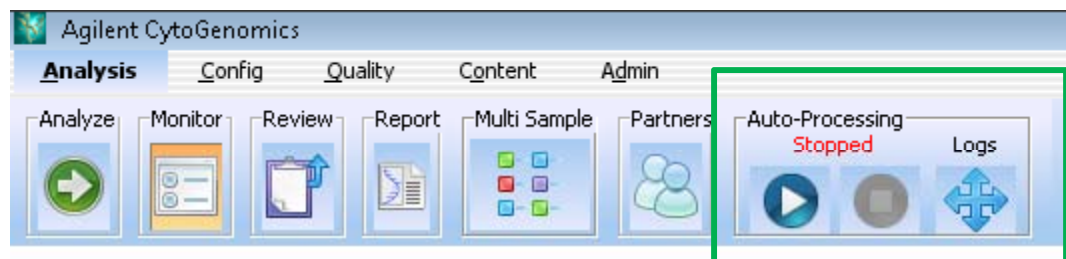


Comparing 2 or more samples side by side

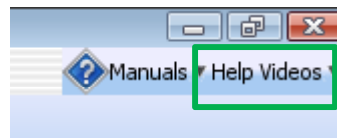
# Agilent CytoGenomics 2.0

## Additional features

- Workflow Automation Mode for automatic processing of TIFF images and report generation.

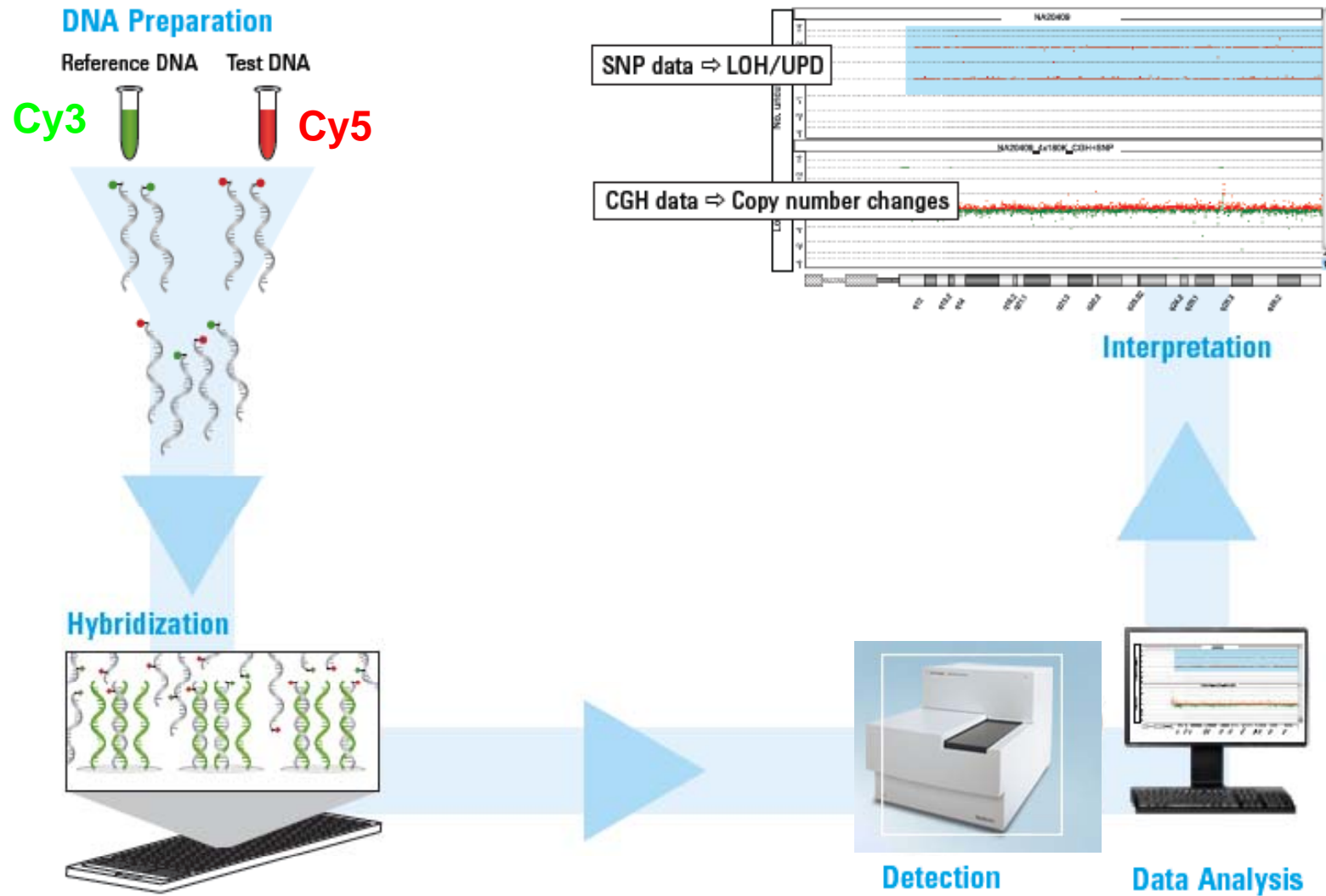


- Data upload to ISCA DB via Cartagenia Bench integration.
- Short 5 minute video tutorials



**Download CytoGenomics free trial @ <https://earray.chem.agilent.com/earray/>**

# CGH and CGH+SNP Workflow



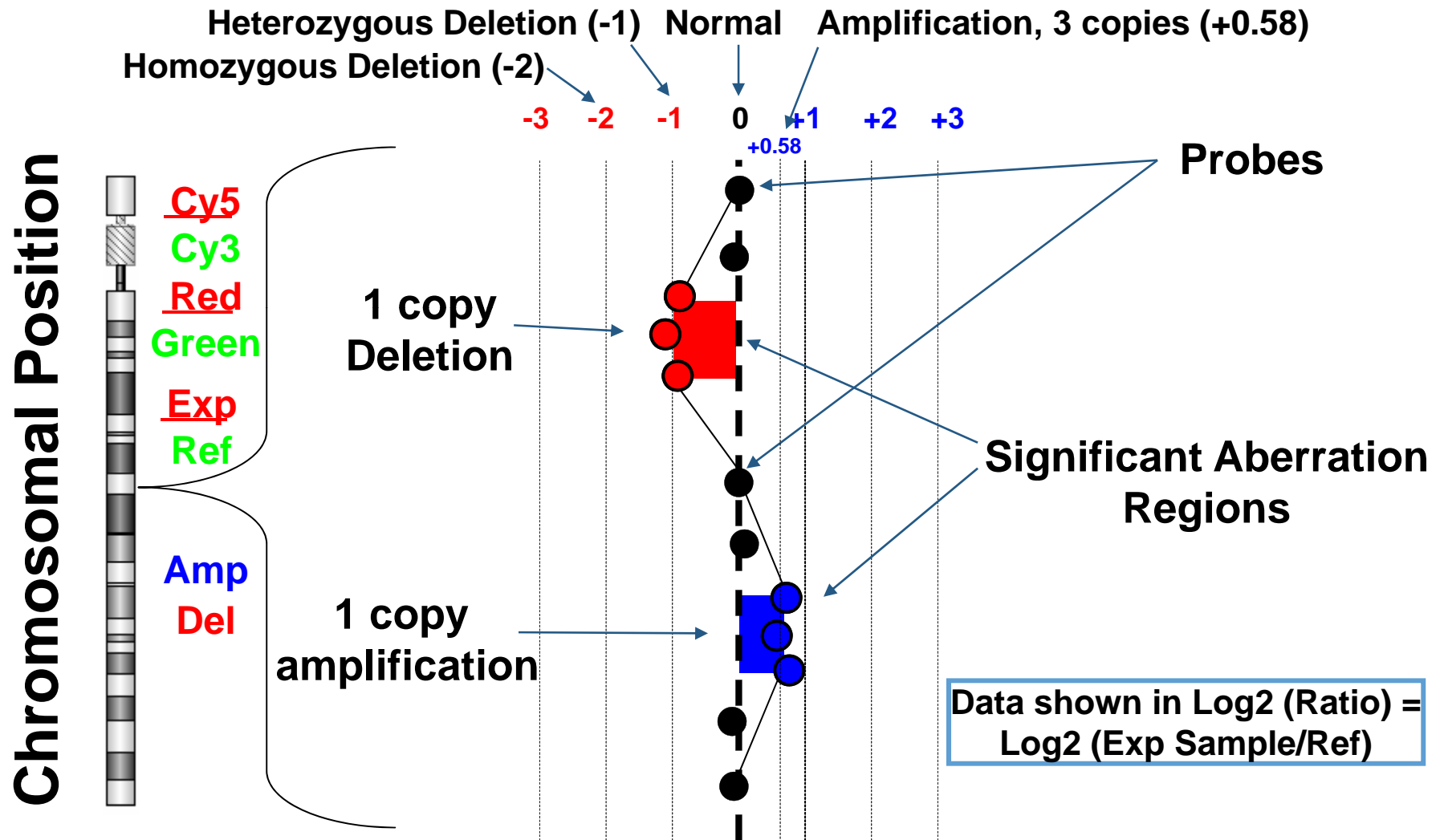
Not approved  
for use in  
diagnostic  
procedures

The Measure of Confidence



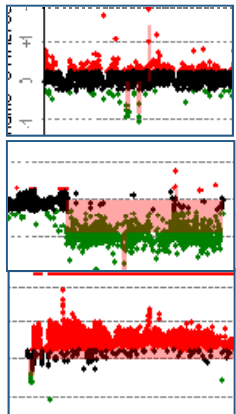
Agilent Technologies

# Schematic view of CGH Data Output



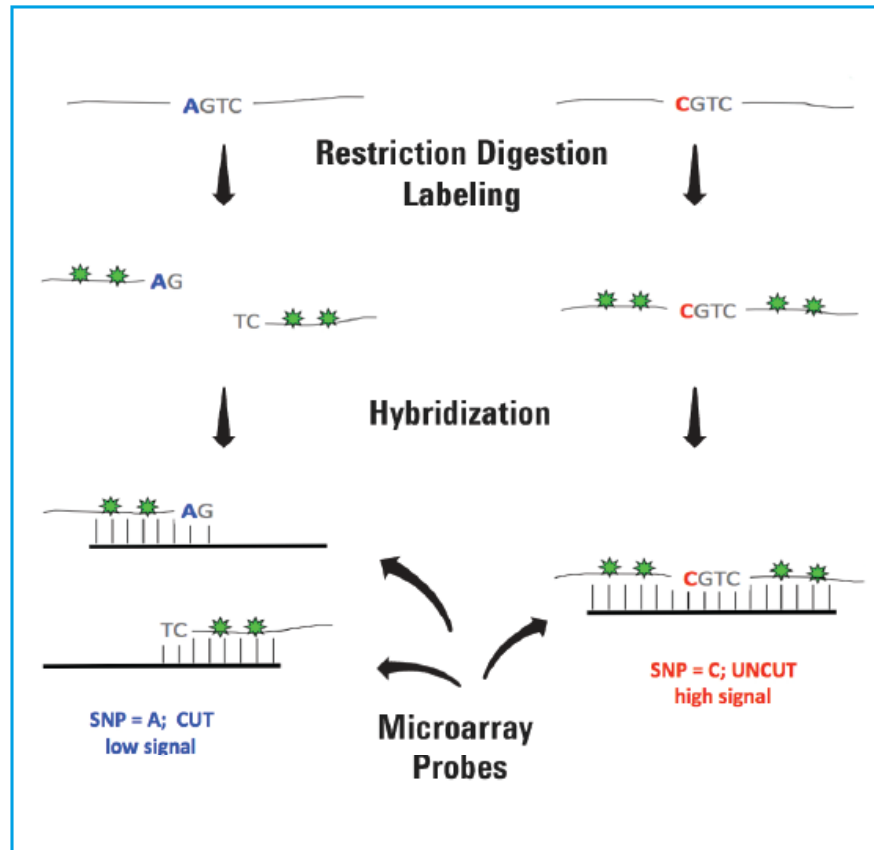
# Overview of 2-color CGH “math”

We measure the difference between 1 sample and 1 reference DNA, usually expressed as  $\log(2)$  (Sample/reference). Actual data are slightly compressed from ideal values.



Description	Average Sample CN	Ref CN	Ratio (S/R)	Ideal Log2(Ratio)	Actual data
Diploid	2	2	1	0	0
Deletion	1	2	0.5	-1	-0.9
Trisomy	3	2	1.5	+0.58	+0.53
50% mosaic deletion	1.5	2	0.75	-0.41	-0.37
50% mosaic trisomy	2.5	2	1.25	+0.32	+0.29
20% mosaic deletion	1.8	2	0.9	-0.15	-0.13
20% mosaic trisomy	2.2	2	1.1	+0.14	0.12

# Measuring SNPs using restriction enzymes

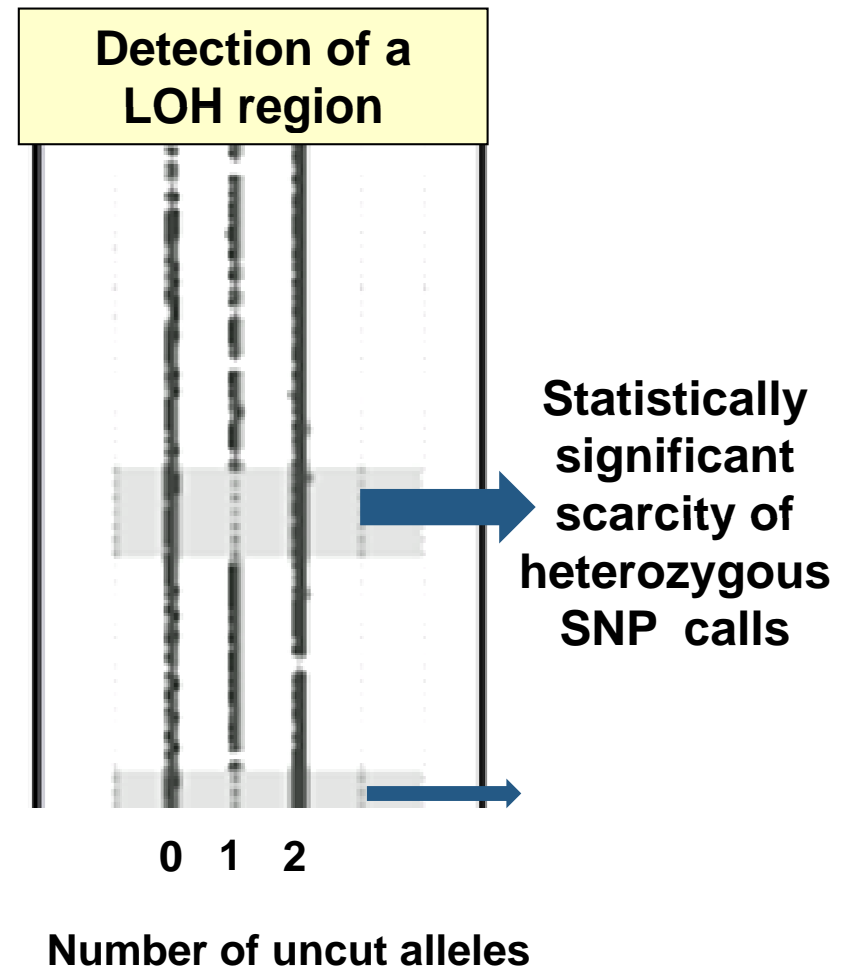
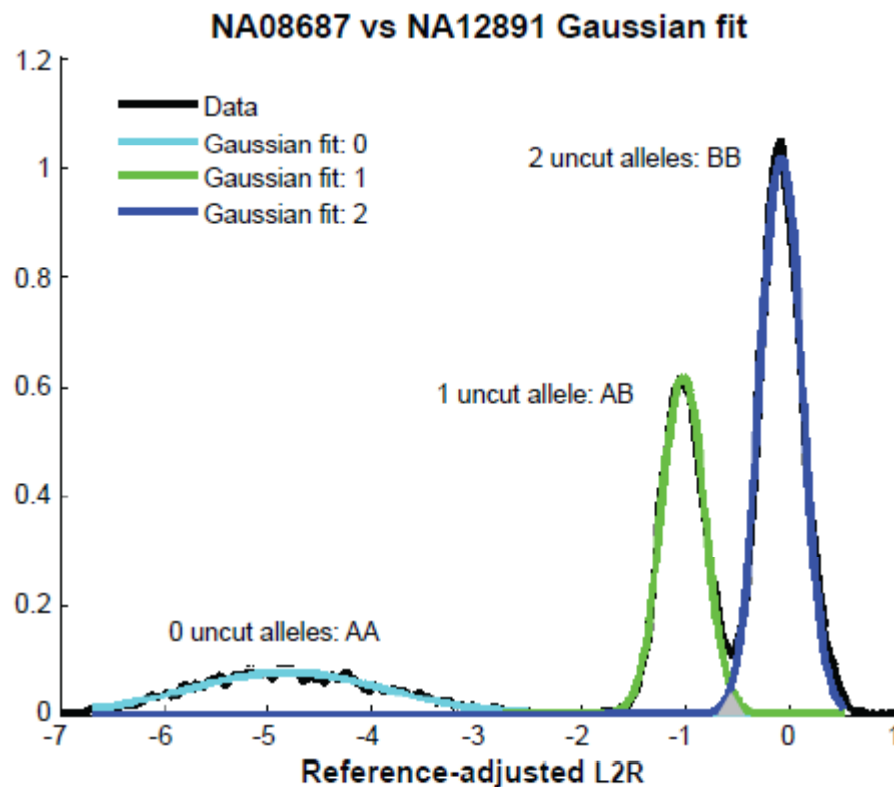


- A subset of probes on a CGH array measures SNPs in parallel to copy number, **on a single array**
- Genotype SNPs using **restriction digestion** (Alu/Rsa)
- Measure the copy number of one allele at each SNP site relative to a **known reference**
- Regions of LOH are located by finding genomic regions with a **statistically significant scarcity of heterozygous calls**
- **~5-10 Mb LOH/UPD resolution** across the entire genome

## How it works (cont'd)

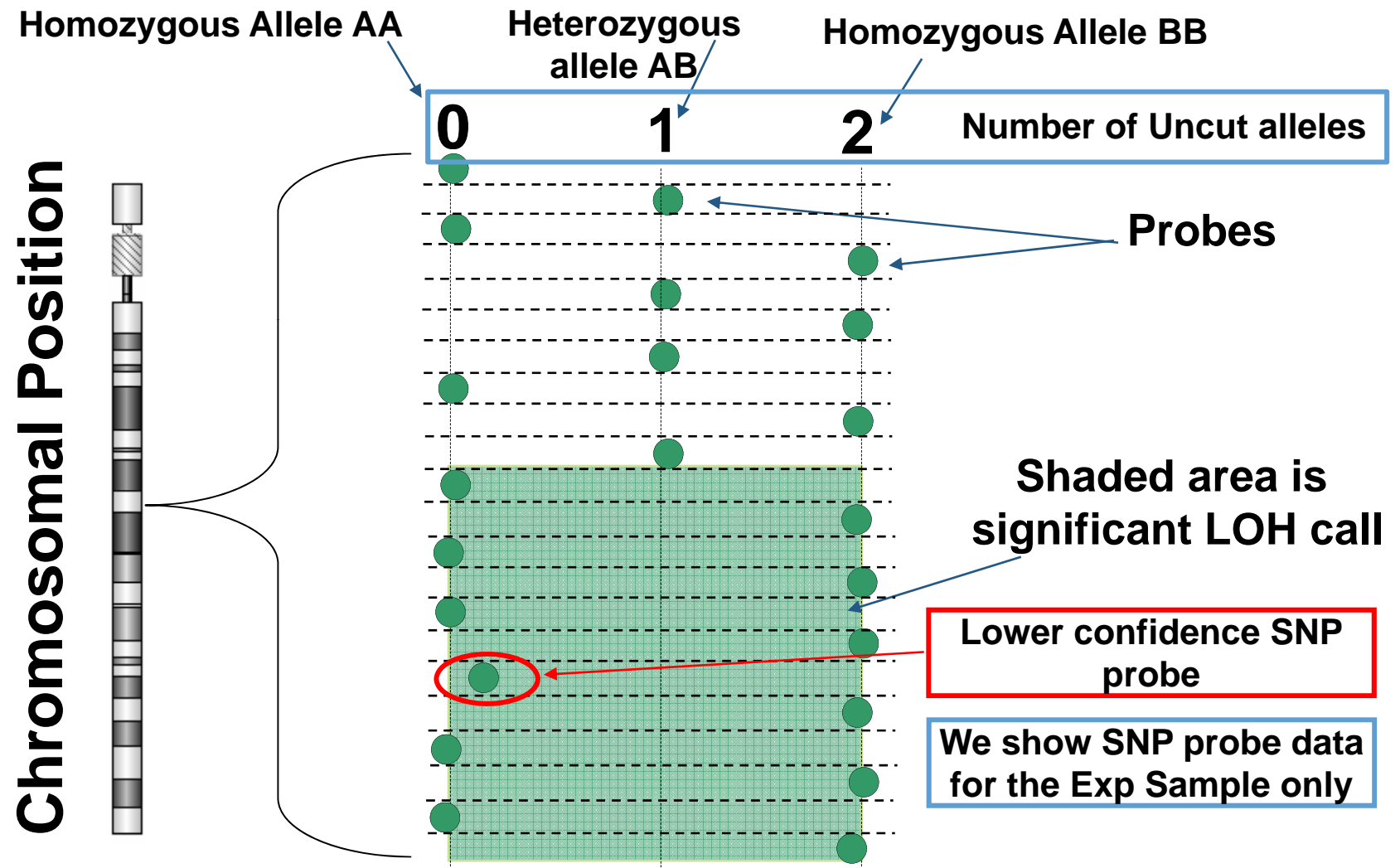
... the SNP calls are corrected for the reference

### Distribution of the SNP Probe reference-corrected log2 ratios





# Schematic View of SNP Data Output



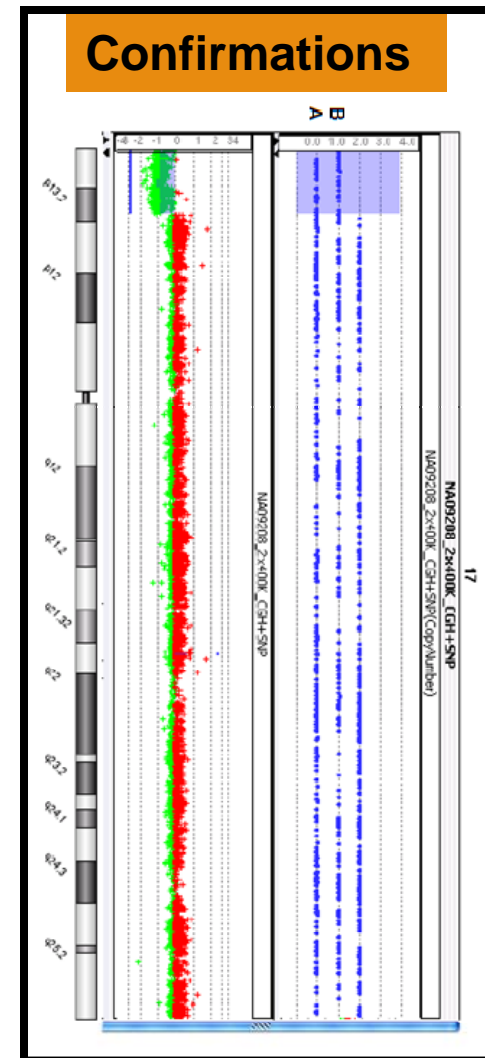
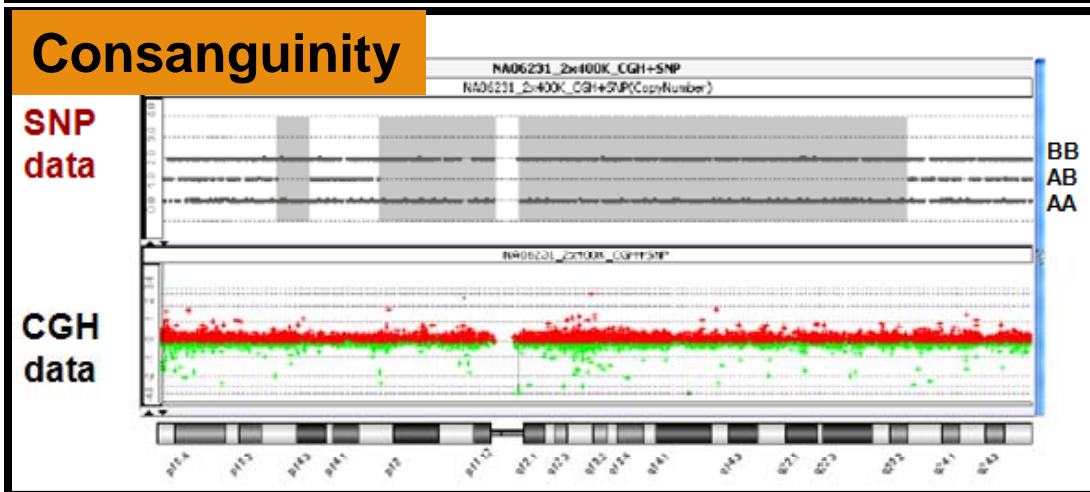
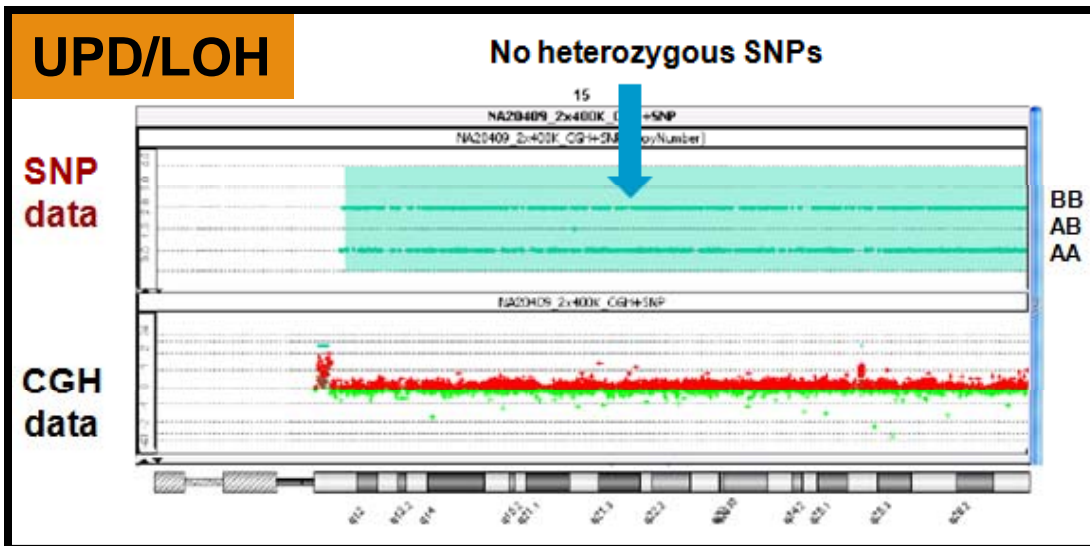
# Relationship of genotype to SNP status

- number of uncut alleles

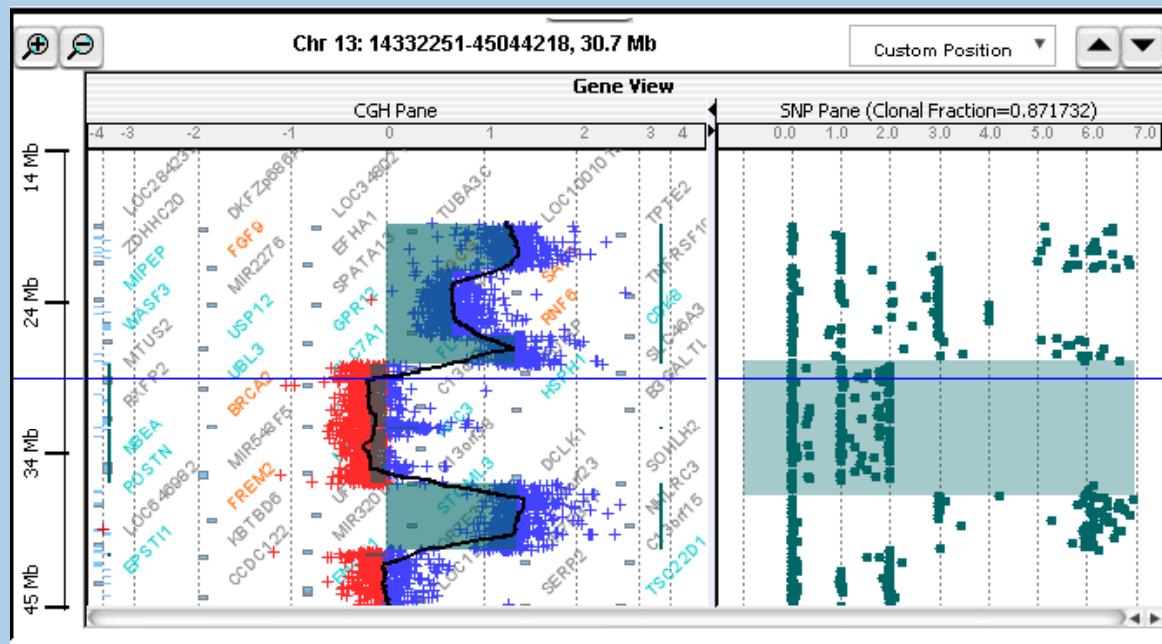
Genomic status	Genotype	No. uncut allele	
Normal diploid genome	AA, AB, BB	0, 1, 2	
Diploid genome with copy-neutral LOH or UPD	AA, BB	0, 2	
Hemizygous LOH	A, B	0,1	
Amplification: e.g. trisomy	AAA, AAB, ABB, BBB	0, 1, 2, 3	
Amplification: e.g. Tetraploidy -Unbalanced	AAAA, AAAB, ABBB, BBBB	0, 1, 3, 4	
Amplification: e.g. Tetraploidy -Balanced	AAAA, AABB, BBBB	0, 2, 4	

In this example, B is uncut allele

# CGH+SNP: Copy Number and LOH



# Support for CGH+SNP analysis of Hematological Cancer samples



# What's different for analyzing cancer samples?

**Constitutional samples are usually homogeneous, eg all cells are the same**

**“Mosaic” refers to a mixed, or non-clonal population of cells. Cancer samples are often mosaics of tumor and normal cells. Constitutional samples are rarely mosaic, but it can happen.**

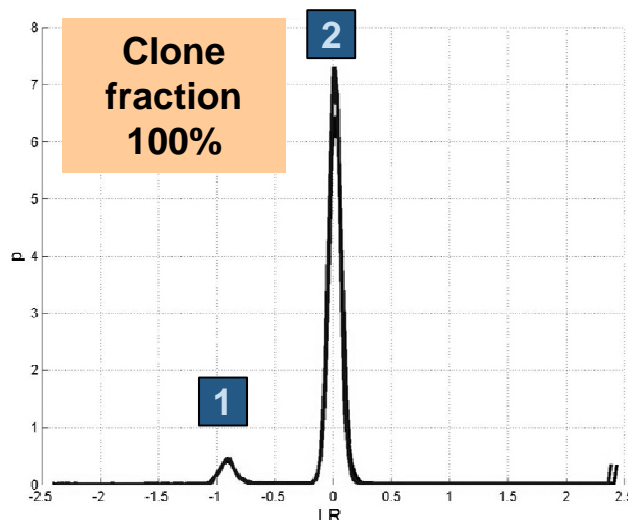


# DNA from Cancer samples adds complexity to data analysis

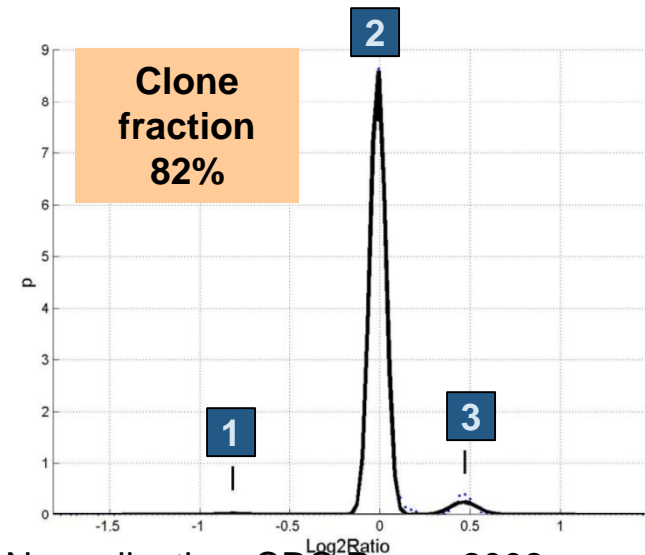
- DNA from cancer cells is typically mixed with an unknown fraction of normal DNA
- Multiple clones may be present
- Samples are not necessarily diploid, but may have 0, 1, 2, 3, 4... copies

## Distributions of Log2Ratios of CGH probes

Normal XY sample vs.  
XX ref



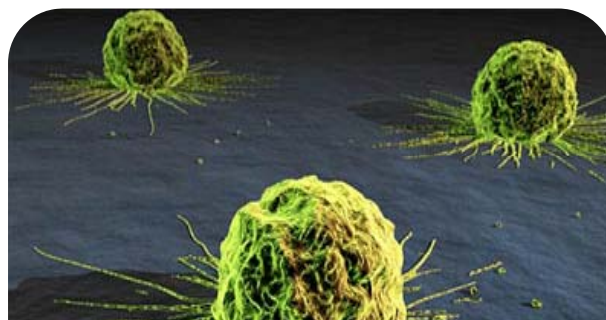
Chronic Lymphocytic  
Leukemia sample



Bo Curry, Jayati Ghosh, Charles Troup. Methods in Microarray Normalization, CRC Press, 2008

# Agilent CytoGenomics 2.0

Supports constitutional and cancer sample analysis



## Cancer

- Blood
- Bone marrow
- Solid (Frozen)
- Solid (FFPE)

## CGH

Yes  
Yes  
Yes  
Yes

## CGH+SNP

Yes  
Yes  
Yes\*  
No

\* data interpretation might be complex  
(manual peak assignment, etc.)



# Agilent CytoGenomics 2.0

## Hematological cancer CGH +SNP analysis

- CGH data plot based on complete sample (potentially mix of tumor clone and normal) ability to detect mosaicism as low as 10%
- SNP data plot based on aberrant clone(s) only
- Ability to detect mosaic LOH (based on SNP data) as low as 20%
- Mosaic fraction of major aberrant clone is reported by the software
- Easy access to CGH&SNP QC metrics for sample evaluation

# Agilent CytoGenomics 2.0

Improved access to CGH and SNP QC metrics

- High level info available on number of metrics that need evaluation

No.	Array ID	Global Display Name	FE QC Report	CGH&SNP Fit	CGH&SNP QC Flag	Status	Green Sample	Red Sample
1	253333810001_1_4	US90603641_253333		<input type="button" value="View"/>	Evaluate(1/20)	Analyzed	European Male (NA109-0008 (Male)	
2	253333810002_1_4	US90603641_253333		<input type="button" value="View"/>	Evaluate(3/20)	Analyzed	European Male (NA109-0032	
3	253333810001_1_3	US90603641_253333		<input type="button" value="View"/>	Evaluate(1/20)	Analyzed	European Male (NA109-0022 (Male)	
4	253333810001_1_2	US90603641_253333		<input type="button" value="View"/>	Evaluate(3/20)	Analyzed	European Male (NA109-0061 (Male)	

# Agilent CytoGenomics 2.0

Improved access to CGH and SNP QC metrics

- Metrics to inform on quality of the sample, but also on proper CN2 peak assignment

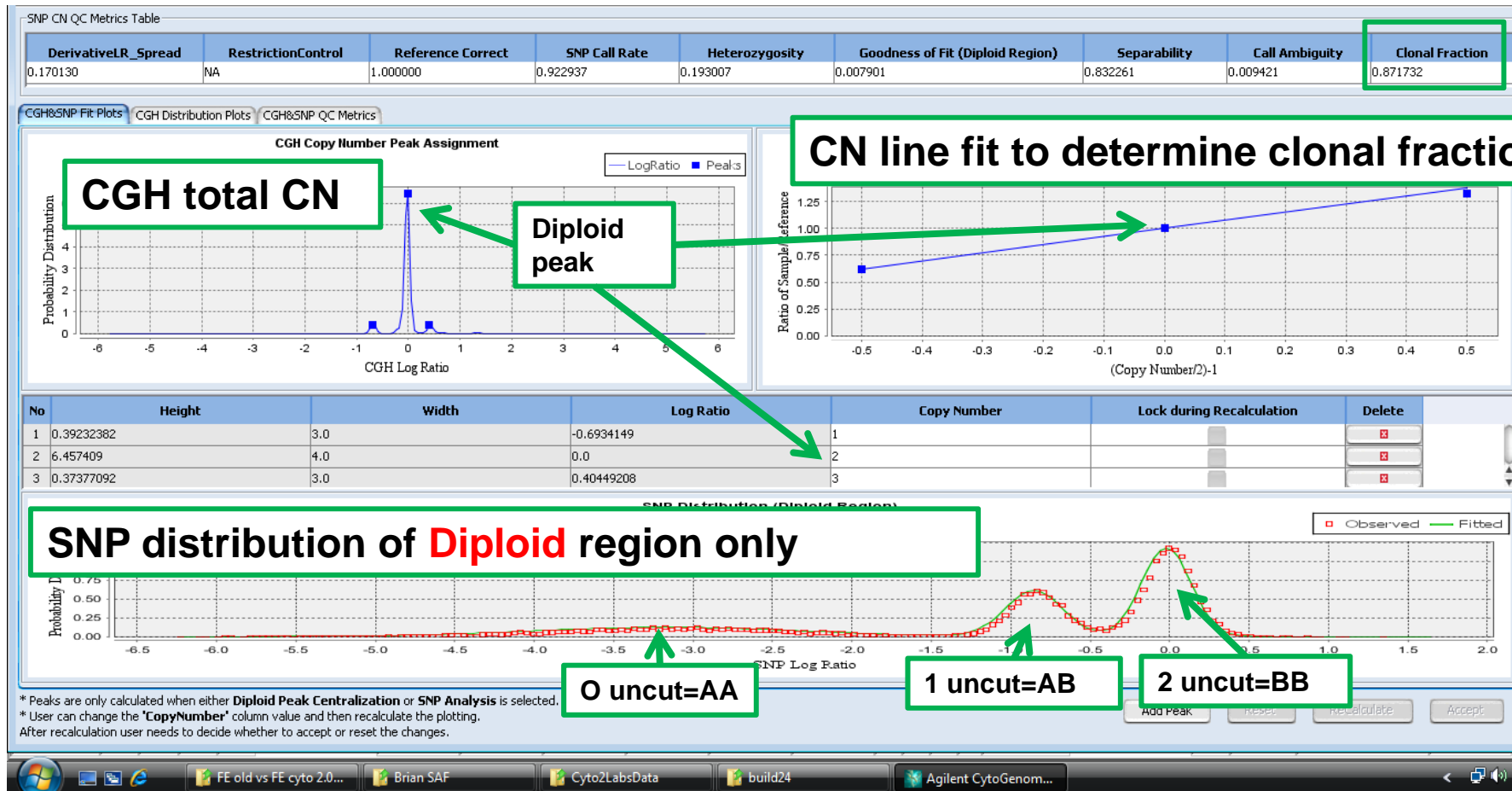
**CGH QC  
metrics**

**SNP QC  
metrics**

QC Metrics		
Metric Name	Value	Preferred Range
IsGoodGrid	1.000000	$\geq 1.0$
AnyColorPrntFeatNonUnifOL	0.087904	$< 5.0$
DerivativeLR_Spread	0.205153	$< 0.3$
gRepro	0.095980	$< 0.2$
g_BGNoise	2.650810	$< 15.0$
g_Signal2Noise	97.534716	$> 30.0$
g_SignalIntensity	258.546000	$> 50.0$
rRepro	0.090995	$< 0.2$
r_BGNoise	6.134300	$< 15.0$
r_Signal2Noise	135.547495	$> 30.0$
r_SignalIntensity	831.489000	$> 50.0$
RestrictionControl	NA	$> 0.8$
Reference Correct	1.000000	$> 0.8$
SNP Call Rate	0.530957	$> 0.6$
Heterozygosity	0.128273	$> 0.15$ and $< 0.35$
Goodness of Fit (Diploid Region)	U.U13/61	NA
Separability	0.873145	$> 0.8$ and $< 1.05$
Call Ambiguity	0.073391	$< 0.07$
Clonal Fraction	0.397550	NA
Call Accuracy	NA	$> 0.98$

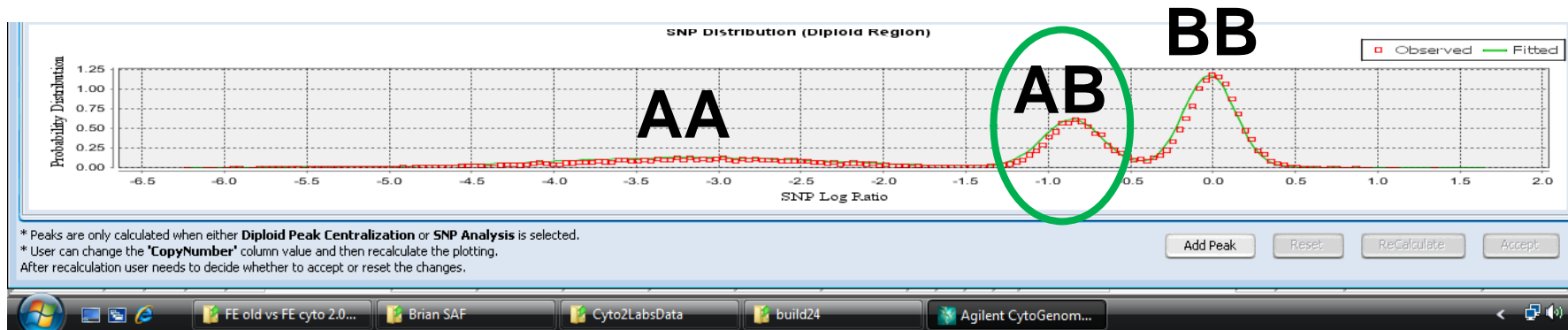
# Agilent CytoGenomics 2.0

Access to view of overall CGH&SNP probe distribution



# SNP QC metrics explained

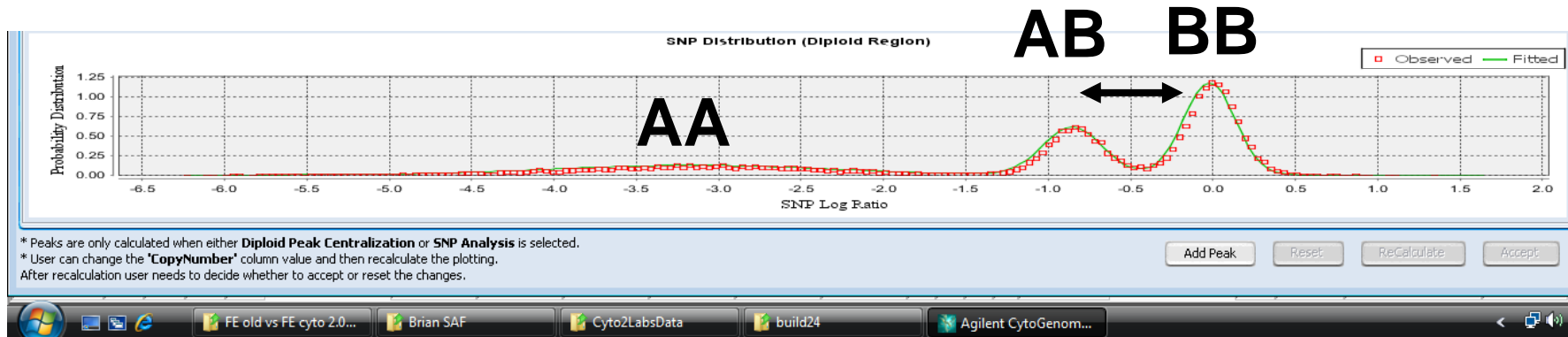
- Heterozygosity: Fraction of heterozygous SNPs (AB = 1 uncut allele peak)



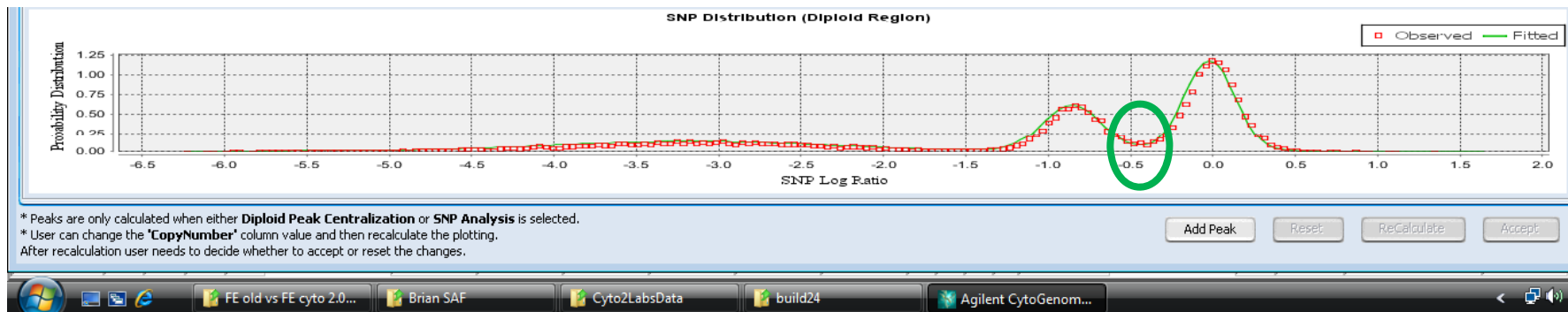
- Goodness of fit: how well did we fit the green 'fit' line to the red data points

# SNP QC metrics explained

- Separability: How well did the AB and BB peak separate

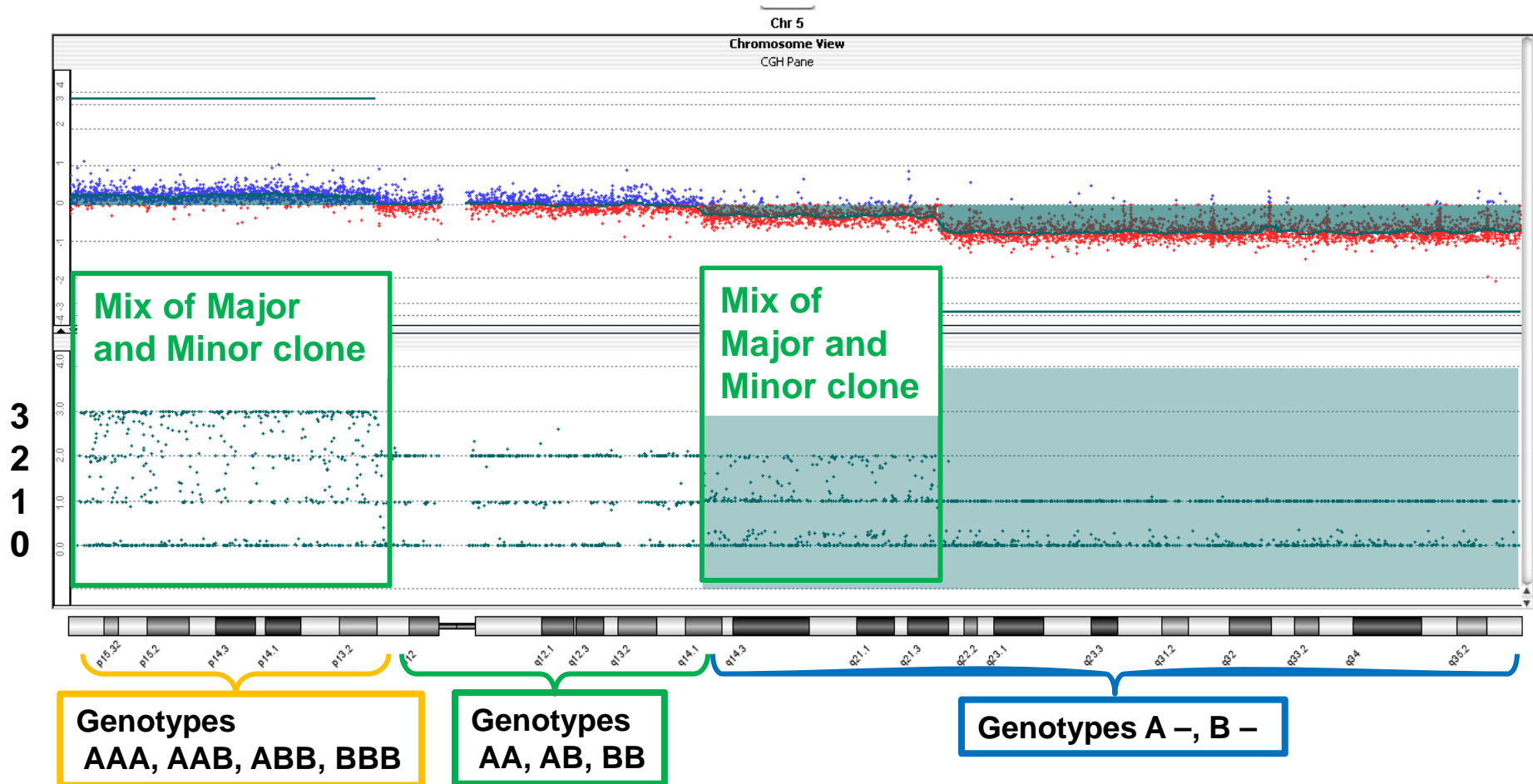


- Call Ambiguity: uncertain calls



# Example of CGH+SNP analysis from AML case

~85% major clone fraction

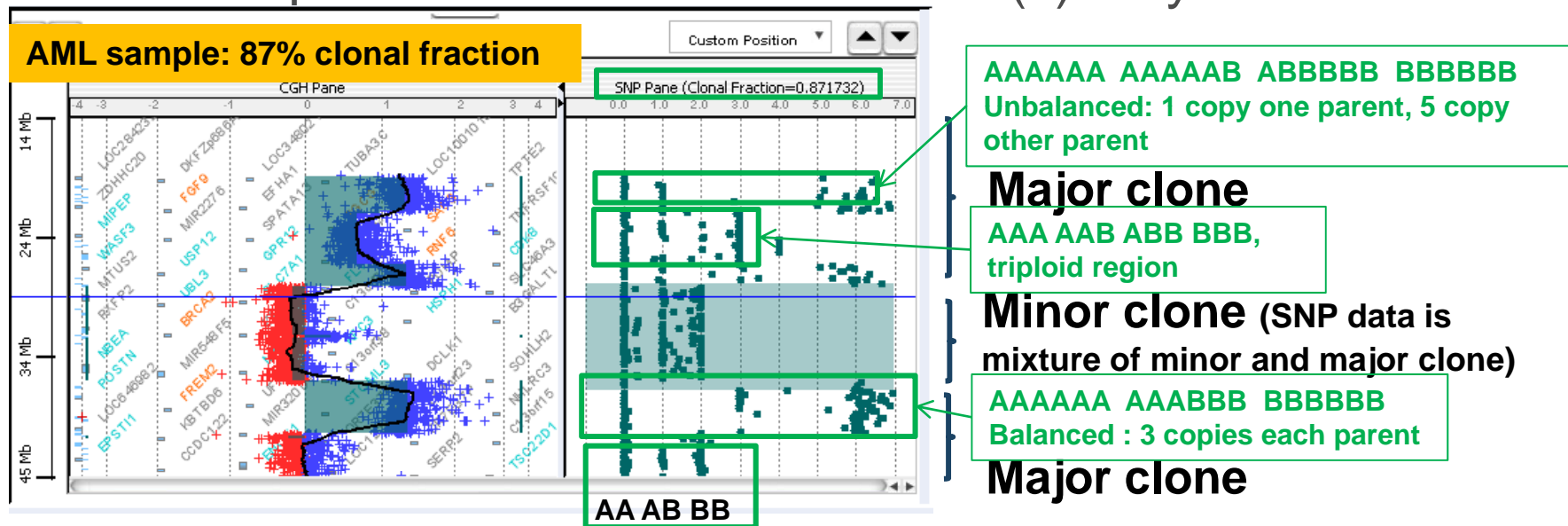




# Agilent CytoGenomics 2.0

## Hematological cancer CGH +SNP analysis

- CGH data plot based on complete sample (potentially mix of tumor clone and normal)
- SNP data plot based on aberrant clone(s) only



- Clonal fraction is reported by the software

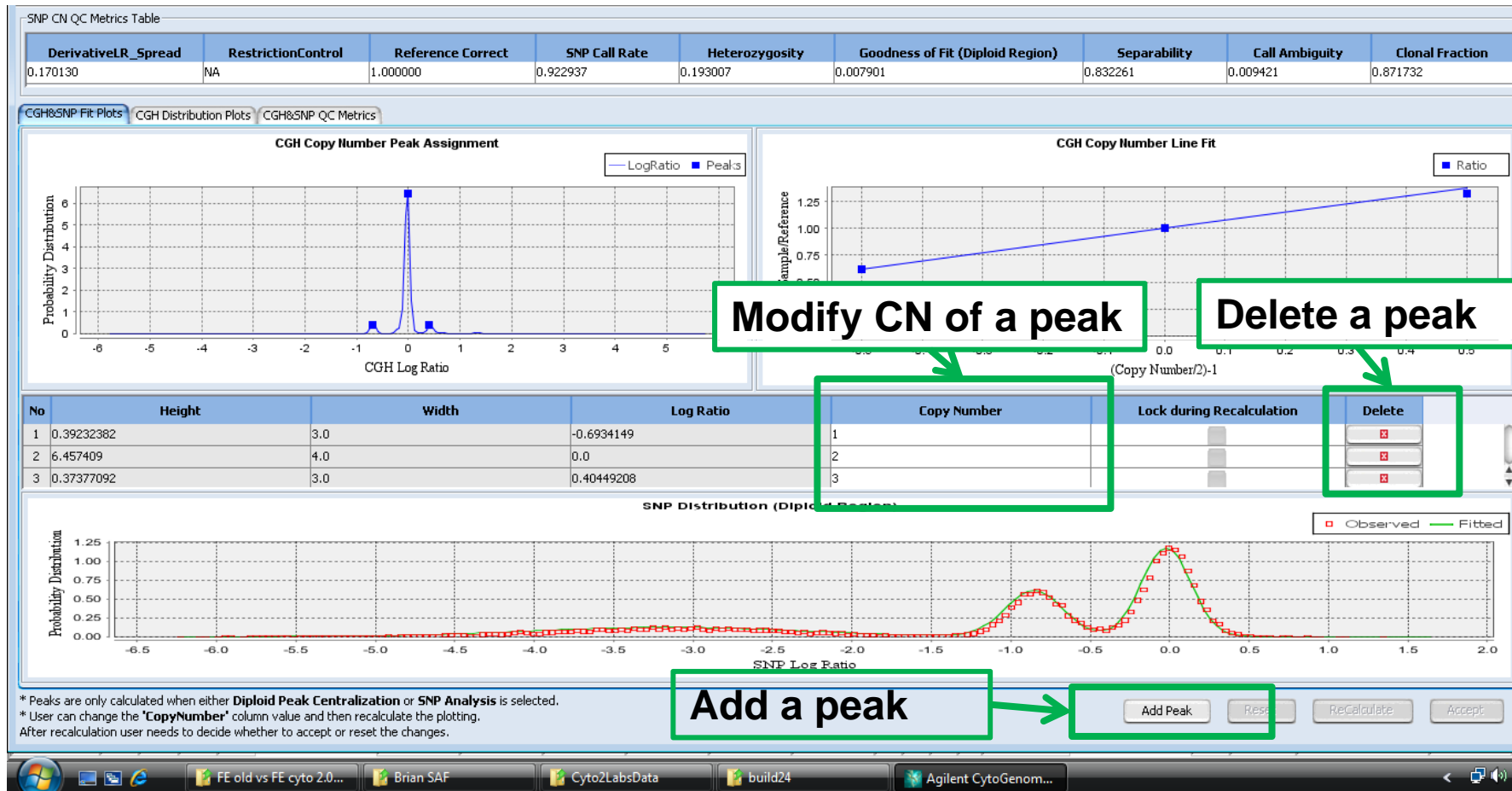
# Agilent CytoGenomics 2.0

## Hematological cancer CGH +SNP analysis

- New algorithms support CGH+SNP analysis in at least 90% of Hematological cancer samples without need for manual intervention
- Manual intervention is only needed in remaining 10% of difficult cases and these cases will often be flagged due to the many failed SNP QC metrics.
- Signs that manual intervention improved the analysis includes decrease in number of failed SNP QC metrics or simply better values of the metrics and better SNP plot fits

# Agilent CytoGenomics 2.0

## Manual Peak Reassignment



After accepting manual changes made to CGH CN peaks, a new analysis workflow will be launched (while keeping original analysis)

# Agilent CytoGenomics 2.0

## Hematological cancer CGH +SNP analysis

- CGH data plot based on complete sample (potentially mix of tumor clone and normal)
- SNP data plot based on aberrant clone(s) only
- Ability to detect mosaic LOH as low as 20%
- Mosaic fraction of major aberrant clone is reported by the software
- Easy access to CGH&SNP QC metrics for sample evaluation

# Agilent CytoGenomics 2.0

## System specs

---

- Client/server system with database (supports access to database of multiple users at the same time).
- Runs on PC (XP, Windows 7).
- Runs on Mac (OS X Leopard and OS X Snow Leopard).
- Supported on both 32 and 64 bit machines.
- Feature extraction is integrated component of CytoGenomics software.

**Download CytoGenomics free trial @ <https://earray.chem.agilent.com/earray/>**