illumina

Quality Measures for CytoChip™ Microarrays

How to evaluate CytoChip Oligo data quality in BlueFuse[®] Multi software.

Data quality is one of the most important aspects of any microarray experiment. This technical note outlines the ways that users can check the quality of their microarray data, to assess laboratory procedures, before proceeding with downstream analyses. Checking the data quality enables users to assess whether the genomic changes seen in a sample are real or artefacts.

All CytoChip Oligo, CytoChip Focus Constitutional, and SNP microarrays come with BlueFuse Multi software. Quality control (QC) metrics are displayed in the report for each experiment. The composite image for each array can also be viewed to check for any major artefacts (**Subarray QC** shortcut). In addition, a confidence call is given to the imbalance calling of the chromosomes. It is also possible to view QC metrics across a database to assess data quality over time (detailed in the *BlueFuse Multi Reference Manual*).

This document should be used as a guide only. It remains the responsibility of the laboratory personel to make the final decision on the usability of array data.

For any further assistance, contact Illumina technical support at techsupport@illumina.com.

Overview of QC Metrics

QC Measures

In BlueFuse Multi, the report displays QC measures that are relevant to the selected experiment. These measures should be consulted together with the **Subarray QC** shortcut and the confidence values for calling when reviewing results.

SD Autosome/Robust—is a measure of the dispersion of log₂ ratio of all clones on the array. A small value indicates that the log₂ ratios of the autosome clones are tightly clustered around zero, which is their expected value. It is calculated on the normalized but unsmoothed data. SD autosome is calculated using all autosomal clones on the array; SD robust is calculated using the middle 66% of the data, allowing the removal of outliers. The SD robust value is used by the software for region calling to determine gain and loss thresholds. An SD value that is too low suggests a problem with washing or hybridization, resulting in lots of non-specific signals left on the array.

% Included Clones—percentage of all clones that were not excluded due to inconsistency between clone replicates on an array. This measure is used for bacterial artificial chromosome (BAC) arrays (CytoChip Focus Constitutional and 24sure). If the level is low, it may indicate a major artefact on the array. There is no exclusion on oligo arrays; therefore, this value is always 100%.

Mean Spot Amplitude Ch1/Ch2—the mean fluorescent signal intensities for the two channels; channel 1 = sample (typically Cy3) and channel 2 = reference (typically Cy5). This metric is variable due to the differences between available scanners. The mean spot amplitude metric can give an indication of how well the DNA has labeled with fluorescent dyes. More importantly, extremely high values can indicate over-scanning of the microarray image or poor washing so that there is lots of non-specific signal left on the array. The balance between channels can be assessed but the Cy5 signal tends to give a higher intensity than Cy3. Major differences in the channels may indicate a labeling or a scanner problem.

SBR Ch1/Ch2—the overall signal-to-background ratio for each fluorescent channel. This value is calculated by dividing the background-corrected fluorescent signal intensity from the DNA clones by the raw background signal level. SBR indicates how clearly the spots can be detected above the background. The SBR metric gives an indication of how well the labeling and washing steps have been performed.

DLR Raw/Fused — derivative log ratio (DLR) measures the spread of the difference in log₂ ratios between all the pairs of adjacent clones along the genome. For a simple profile, this statistic will be very similar to autosome standard deviation. The presence of many step changes in a complex profile will cause the autosome standard deviation to become inflated, whereas the DLR metric is much less sensitive to this problem. DLR fused is the DLR calculated on the normalized but unsmoothed data.

Additional SNP QC Metrics

CytoChip Oligo SNP and CytoChip Cancer SNP array designs include DNA probes that measure the presence of biallelic SNPs in a sample, in addition to those that assess copy-number imbalance. In the report for SNP arrays, three metrics relate to the performance of SNP calling in the experiment. These metrics are:

SNPs on Array—the total number of SNPs on the array design (CytoChip Oligo SNP = 26,541; CytoChip Cancer SNP = 59,593).

Reference SNPs—the number of SNP loci where the reference sample has a known genotype and which is not homozygous for the digested allele. All other SNP loci are excluded from analysis. For any particular reference DNA, approximately one-third of loci are expected to be homozygous for the digested allele.

% High-Confidence SNPs—the percentage of Reference SNPs where the genotype call has a confidence value of 95% or more. This measure indicates the overall quality of the SNP data and mainly corresponds to how well the peaks of the SNP histogram are resolved from each other. Overlapping peaks leads to low-confidence SNP calls and a low value for the percentage of high-confidence SNPs.

Oligo Microarray QC Guidance

Oligo QC Metrics

Basic QC criteria settings can be used as a guide to assess the performance of an oligo microarray experiment before analyzing the results (Table 1). Illumina oligo arrays include: CytoChip Oligo, CytoChip ISCA, CytoChip Cancer, CytoChip Focus 8x60K, 4x180K,

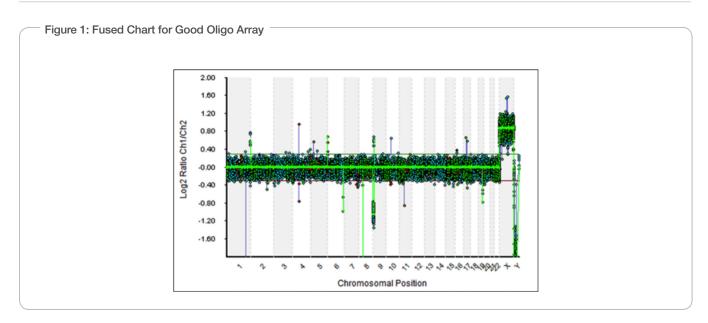
- Table 1: Basic QC Metrics ----

SD Autosome/Robust	Good	0.08-0.15
	Too low	< 0.07
	Too high	> 0.17
6 Included Clones	100% for oli	go designs
Vlean Spot Amplitude	Good	400-2,000
(signal)*	Too low	< 400
	Too high	> 2,500
SBR	Good	5–20
	Too low	< 3
	Too high	> 30
DLR Raw/Fused	Good	< 0.15
	OK	< 0.20
	Fail	> 0.24

and oligo SNP array designs. The CytoChip Focus 4x180K platform requires less DNA (200 ng) for hybridization; as a result, the QC metrics for mean spot amplitude are different (Table 2).

These values are based on preliminary results and may be modified once there is more experience using the platform.

SD Autosome/Robust	Good	0.08-0.15
	Too low	< 0.07
	Too high	> 0.17
% Included Clones	100% for oli	go designs
Mean Spot Amplitude	Good	200+
(signal)*	Too low	< 200
	Too high	> 2,500
SBR	Good	5–20
	Too low	< 2
	Too high	> 30
DLR Raw/Fused	Good	< 0.15
	OK	< 0.20
	Fail	> 0.24



A Good Quality Oligo Array Result

Figure 1 shows a fused chart plot from BlueFuse Multi and is an example of a result from a good experiment. There is clear separation of normal regions and regions of imbalance. For this type of result, there is no need to rely on QC metrics.

Figure 2 shows the composite image for the array, which can be viewed using the **Image QC** shortcut in the software. Oligo slides have QC measures incorporated during microarray manufacture and are supposed to have dark corners with a few bright spots in the extreme corners. Figure 3 is an image of a scan that shows the corner of the microarray. The bright spots are visible, with clear dark corners and no spots visible in this region. There is also a very even overall background on this image. If the dark corners are visible, it is indicative of non-specific fluorescent signal, perhaps due to a washing issue.

Figure 2: Composite Image for Good Oligo Array

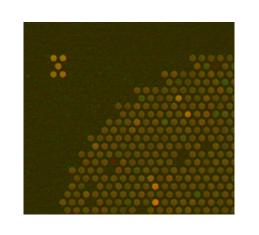
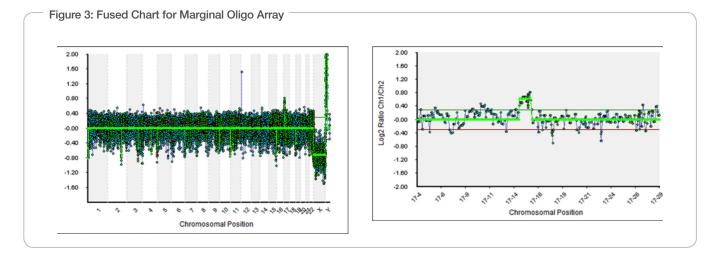


Table 3 shows the QC metrics for a good oligo array.

Table 3: QC Metrics for Good Oligo Array

Parameter	Value	Comment
SD Autosome/Robust	0.12/0.09	Good robust SD score; below 0.1 is a very good result
% Included Clones	100.00	No replicates on the array-can be ignored
Mean Spot Amplitude Ch1/Ch2	325.87/429.86	Low side of mean signal but acceptable
SBR Ch1/Ch2	8.02/9.62	SBRs fall within the optimal range
DLR Raw/Fused	0.15/0.14	Good fused DLR value; any value below 0.15 is a very good result



A Marginal Quality Oligo Array Result

Figure 3 shows a fused chart plot from BlueFuse Multi and is an example of a result from a difficult-to-call oligo array. QC metrics must be taken into account when analyzing this array. On closer examination (right panel), large changes are called but would require extra care when checking through the imbalances.

Figure 4 is a scan image that shows the corner of this microarray. It is clear that the image has been over-scanned, as there are very bright signals. It is not likely to be a wash problem as spots in the dark corners are still not visible.

Table 4 shows the QC metrics for the marginal oligo array.

 Figure 4: Composite Image for Marginal Oligo Array

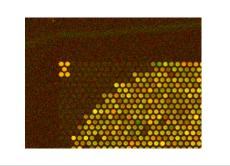
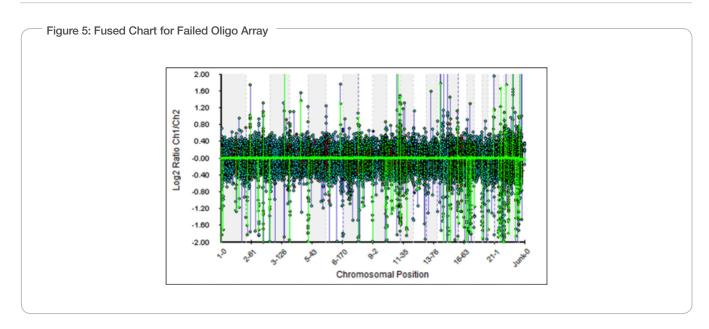


Table 4: QC Metrics for Marginal Oligo Array

Parameter	Value	Comment
SD Autosome/Robust	0.19/0.15	Anything below 0.15 should be interpretable for large changes; this result is borderline
% Included Clones	100.00	No replicates on the array-can be ignored
Mean Spot Amplitude Ch1/Ch2	6250.15/7278.70	Massively over-scanned and will cause suppression of some ratios, which can result in under-calling
SBR Ch1/Ch2	63.63/22.52	SBRs are artificially high due to the over-scanning of the arrays. This result would have been enhanced with a reduced scanner setting.
DLR Raw/Fused	0.21/0.21	Any value below 0.2 should be interpretable for large changes; this result is borderline



A Failed Quality Oligo Array Result

Figure 5 shows a fused chart plot from BlueFuse Multi and is an example of a result from a failed oligo array. The data are very noisy, which makes it difficult to determine the true calls.

Figure 6 is a scan image that shows the corner of this microarray. The array looks reasonable, the background looks low, and the dark corners are acceptable.

Table 5 shows the QC metrics for the failed oligo array.

Figure 6: Composite Image for Failed Oligo Array

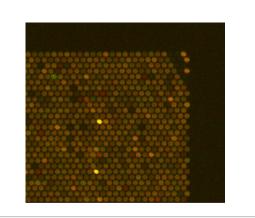
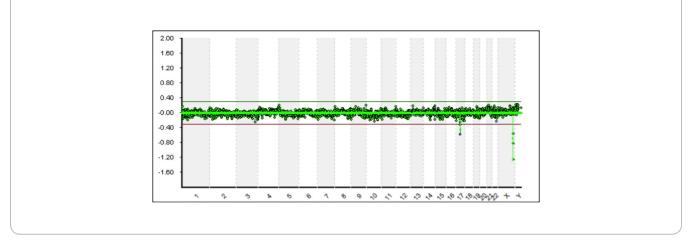


Table 5: QC Metrics for Failed Oligo Array

Parameter	Value	Comment
SD Autosome/Robust	0.27/0.19	Very poor robust SD score; any value above 0.17 is a bad result and almost impossible to call accurately
% Included Clones	100.00	No replicates on the array-can be ignored
Mean Spot Amplitude Ch1/Ch2	711.75/1151.61	Signal is around optimal, so not the cause of the noise
SBR Ch1/Ch2	8.55/7.70	SBRs within acceptable range
DLR Raw/Fused	0.28/0.28	Very high DLR, probably a DNA-related issue





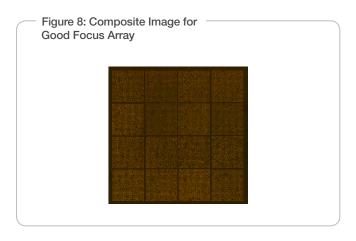
CytoChip Focus BAC QC Guidance

CytoChip Focus Constitutional (BAC Array) QC Metrics

These basic QC criteria settings can be used as a guide to assess CytoChip Focus Constitutional microarray results (Table 6).

A Good Quality CytoChip Focus Array Result

Figure 7 shows a fused chart plot from BlueFuse Multi and is an example of a result from a good experiment. In normal regions, the spots are tight to the log_2 ratio 0 axis. There is clear demarcation of



regions of imbalance. For this type of result, there is no need to rely on QC metrics.

Figure 8 is the scan image for this CytoChip Focus Constitutional array. It shows a clean image, and even dark background with no wash-related artefacts or other flaws.

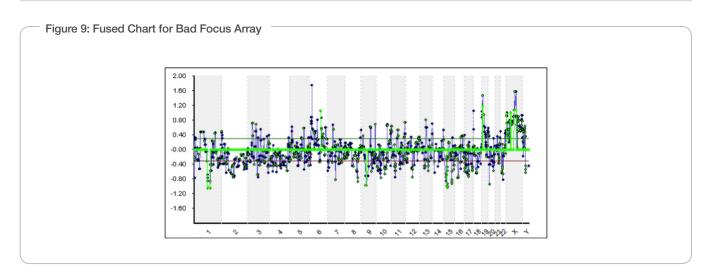
Table 7 shows the QC metrics for the good Focus array.

— Table 6: Basic QC Metrics

SD Autosome/Robust	Good	0.05-0.11
	Too low	< 0.05
	Too high	> 0.17
% Included Clones	Good	> 95%
Mean Spot Amplitude	Good	700–3,000
(signal)*	Too low	< 500
	Too high	> 3,000
SBR	Good	3–12
	Too low	< 3
	Too high	> 15
DLR Raw/Fused	Good	< 0.2
	OK	< 0.22
	Fail	> 0.24

Table 7: QC Metrics for Good Focus Array

Parameter	Value	Comment
SD Autosome/Robust	0.06/0.06	Very good robust SD score; below 0.1 is a good result
% Included Clones	99.29	Near-perfect clone inclusion, less than 1% excluded
Mean Spot Amplitude Ch1/Ch2	1453.22/2713.96	Good signal
SBR Ch1/Ch2	5.91/6.05	SBRs well within the acceptable range
DLR Raw/Fused	0.05/0.04	Excellent fused DLR value; any value below 0.15 is a very good resul



A Bad Quality CytoChip Focus Array Result

Figure 9 shows a fused chart plot from BlueFuse Multi and is an example of a result from a bad or failed experiment. In general, the array results look very noisy; the resulting called imbalances cannot be trusted.

The scan image of this CytoChip Focus Constitutional array (Figure 10) shows that the slide has not been washed effectively. There is lots of background and wash artefacts. It is possible that the array dried out during hybridization. There are also speckles on the array that indicate the labeled DNA pellet was not resuspended properly in hybridization buffer. Also, the image looks quite red.

Table 8 shows the QC metrics for the bad Focus array.

Figure 10: Composite Image for Bad Focus Array

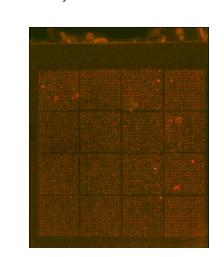
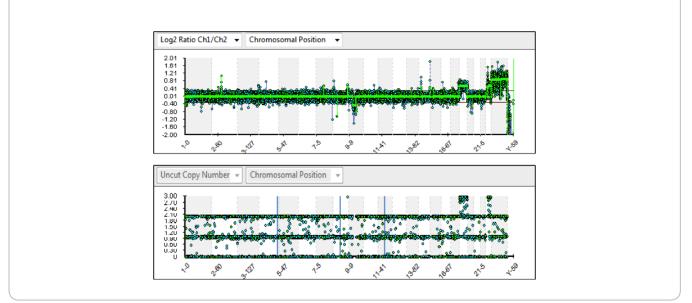


Table 8: QC Metrics for Bad Focus Array

Parameter	Value	Comment
SD Autosome/Robust	0.35/0.32	Very poor robust SD score; any value above 0.17 is a bad result and almost impossible to call accurately
% Included Clones	48.46	Over half the clones have been excluded, leading to uninterpretable results
Mean Spot Amplitude Ch1/Ch2	39.51/1152.80	Signal far too low in Ch1 (Cy3) and may be one of the causes of the noise
SBR Ch1/Ch2	0.18/4.16	SBR far too low in Ch1 (Cy3) due to low signal
DLR Raw/Fused	0.55/0.49	DLR far too high

Figure 11: Fused Charts for Good SNP Data



SNP Array QC Guidance

For CytoChip Oligo SNP and CytoChip Cancer SNP, use the oligo QC metrics detailed in this document to assess the copy-number aspect of each experiment. Use this section to assess the quality of the SNP probe performance and the resulting confidence in the UPD/LOH regions called.

In addition to the general oligo QC metrics and standard BlueFuse Multi views, examine the Fused SNP Copy Number Chart. Additional QC metrics and the SNP histogram provide essential QC information.

Good Quality SNP Data

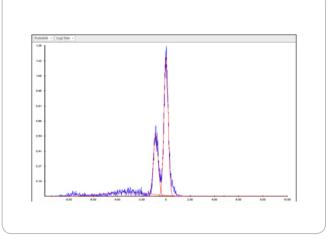
Both the fused chart and fused SNP copy-number chart (Figure 11) show clear array results. There is good separation of SNPs in the chart, shown by the clearly separated SNP data points at copy-number values 0 (AA, uncut), 1 (AB, one allele cut), and 2 (BB, both alleles cut), with few spots plotted in between.

The QC metrics in the report show a good percentage of highconfidence SNPs:

% High-confidence SNPs:	90.88
SNPs on array:	9,911
Reference SNPs:	6,088

Finally, viewing the SNP Histogram (Figure 12) shows clearly defined peaks that represent 0, 1, and 2 copies of SNPs.

- Figure 12: Histogram for Good SNP Data





Poor Quality SNP Data

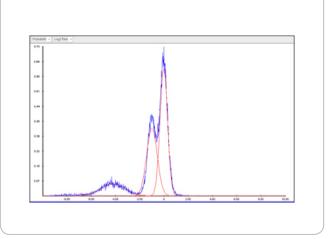
Both the fused chart and fused SNP copy-number chart (Figure 13) show noisy/borderline array results. In the SNP chart, there is good separation of SNP data points between 0 and 1 copy number, but not between 1 and 2.

The QC metrics in the report show a borderline percentage of high-confidence SNPs:

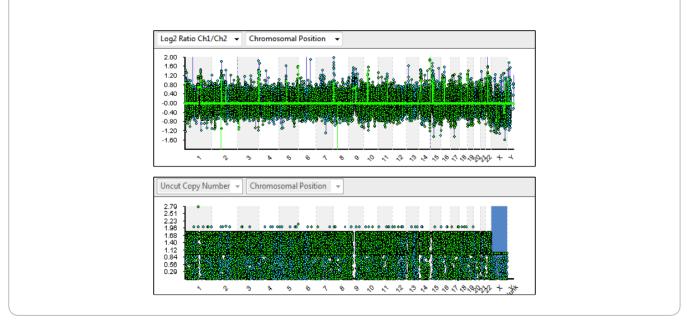
% High-confidence SNPs:	66.14
SNPs on array:	59,593
Reference SNPs:	36,436

Finally, viewing the SNP Histogram (Figure 14) shows clearly separated peaks that represent 0 and 1, but not 1 and 2 copies of SNPs. This result may indicate incomplete digestion of the DNA.

- Figure 14: Histogram for Poor SNP Data







Failed Quality SNP Data

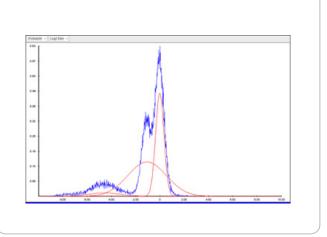
Both the fused chart and fused SNP copy-number chart (Figure 15) show noisy array results. In the SNP chart, there is no clear separation at 0, 1, or 2.

The QC metrics in the report show a low percentage of high-confidence SNPs:

% High-confidence SNPs:	22.67
SNPs on array:	59,593
Reference SNPs:	36,436

Viewing the SNP Histogram (Figure 16) shows no separation of the peaks and BlueFuse Multi has failed to fit the peaks. This result indicates poor or failed digestion of the genomic DNA prior to labeling and hybridization.

- Figure 16: Histogram for Failed SNP Data



Illumina • 1.800.809.4566 toll-free (U.S.) • +1.858.202.4566 tel • techsupport@illumina.com • www.illumina.com

FOR RESEARCH USE ONLY

© 2014 Illumina, Inc. All rights reserved.

Illumina, BlueFuse, CytoChip, the pumpkin orange color, and the Genetic Energy streaming bases design are trademarks or registered trademarks of Illumina, Inc. All other brands and names contained herein are the property of their respective owners. Pub. No. 1570-2014-038 Current as of 09 December 2014

