Quality Control & Pre-processing Evaluation

of

E-GEOD-4600.raw.1_0

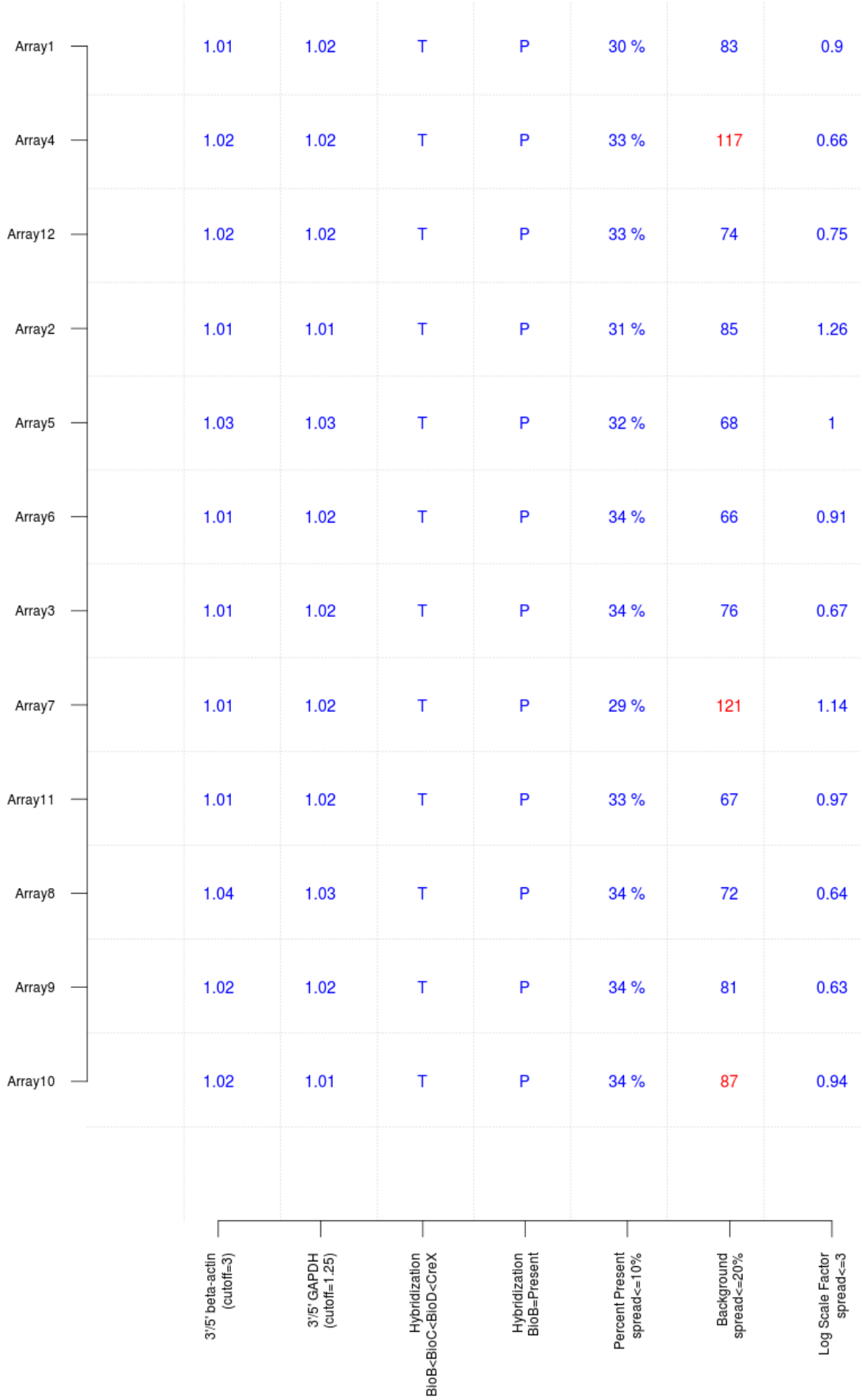
REPORT

Array names and grouping

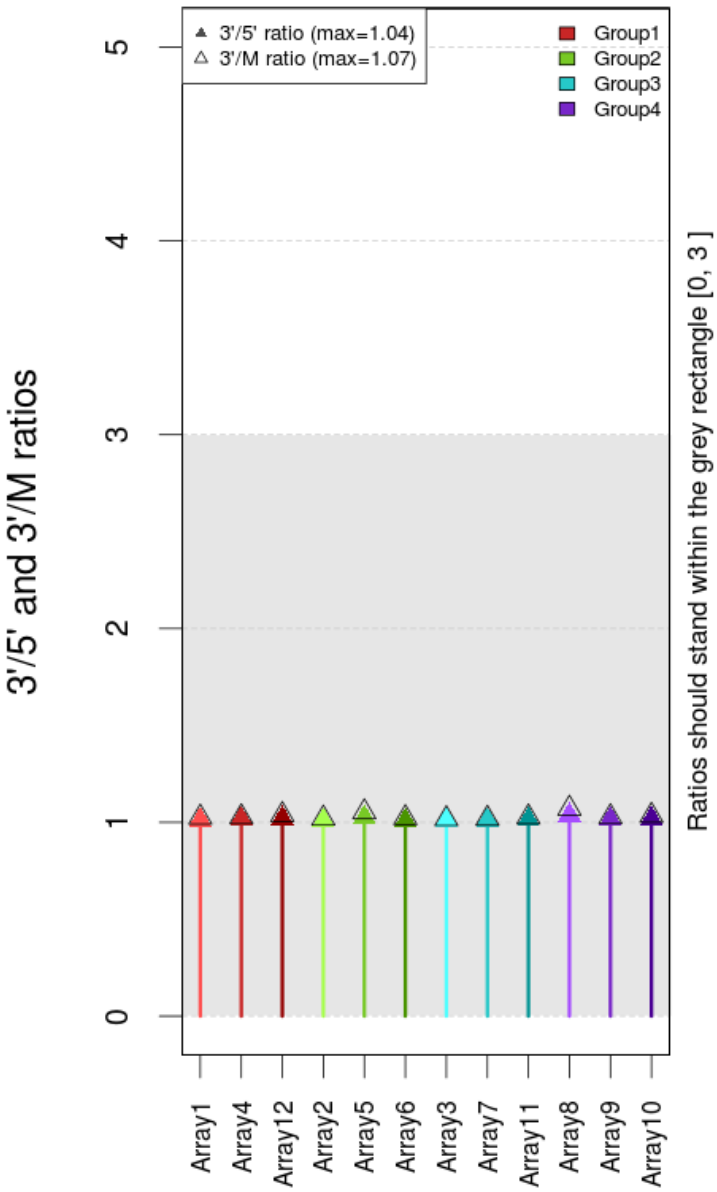
ArrayDataFile	SourceName	FactorValue
GSM102842.CEL	Array1	Group1
GSM102878.CEL	Array2	Group2
GSM102883.CEL	Array3	Group3
GSM102825.CEL	Array4	Group1
GSM102879.CEL	Array5	Group2
GSM102880.CEL	Array6	Group2
GSM102881.CEL	Array7	Group3
GSM102876.CEL	Array8	Group4
GSM102877.CEL	Array9	Group4
GSM102875.CEL	Array10	Group4
GSM102882.CEL	Array11	Group3
GSM102870.CEL	Array12	Group1

Summary of raw data quality indicators

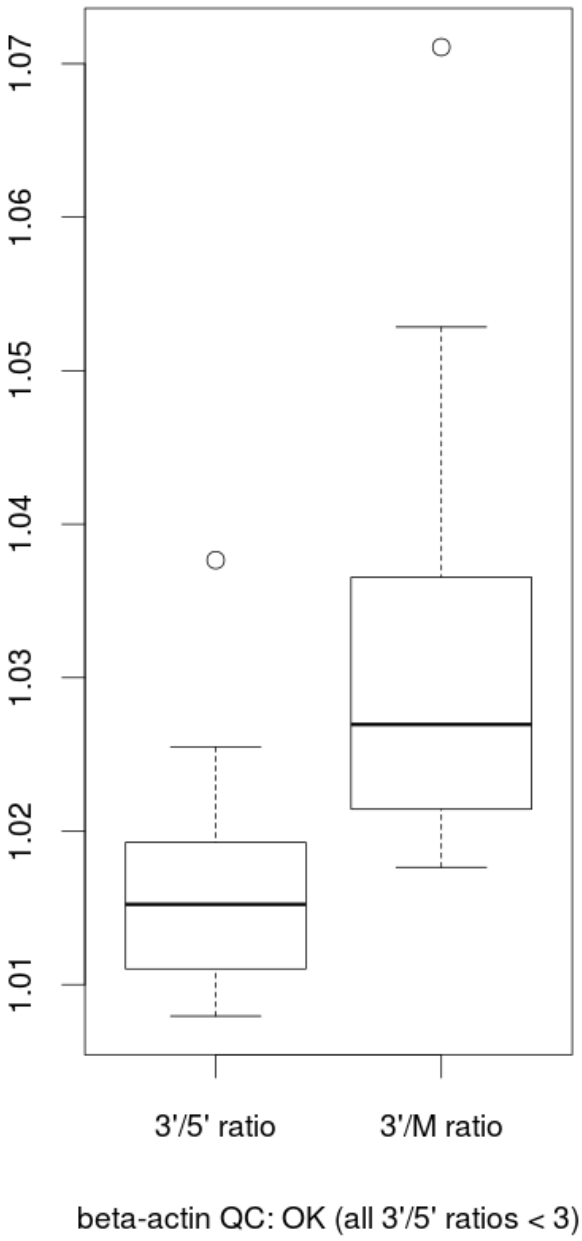
blue = "within" / red = "out of" recommended cut-off



RNA degradation of beta-actin

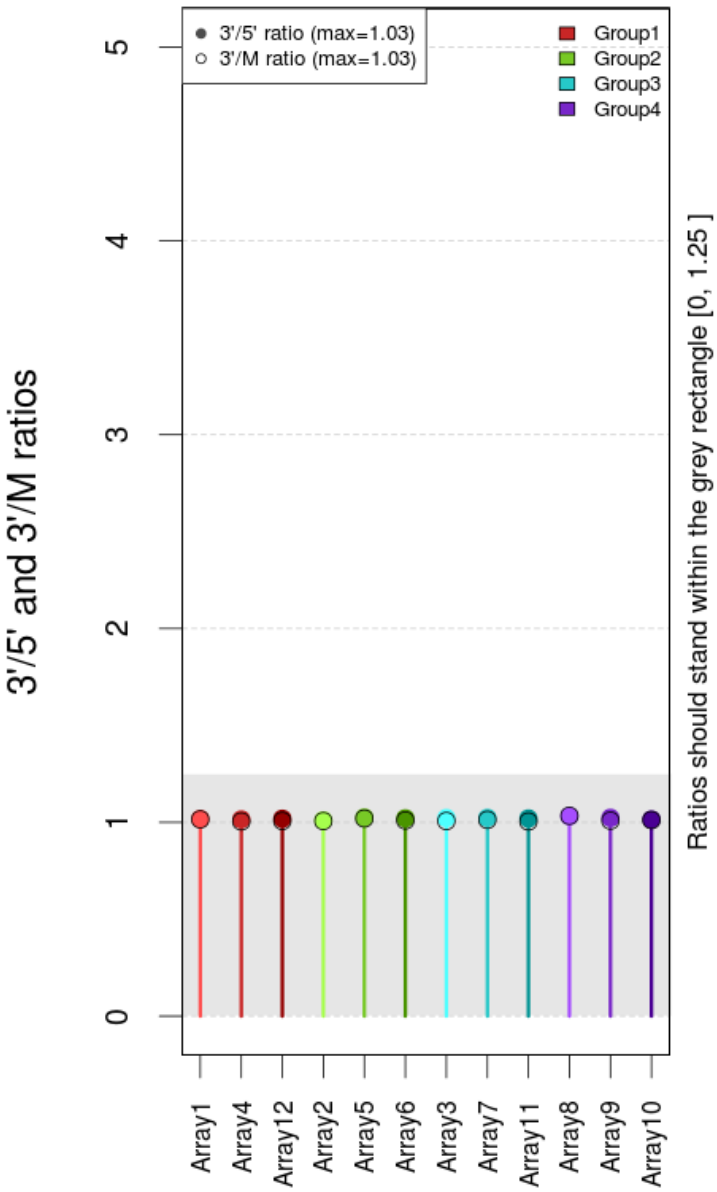


Boxplot of beta-actin ratios

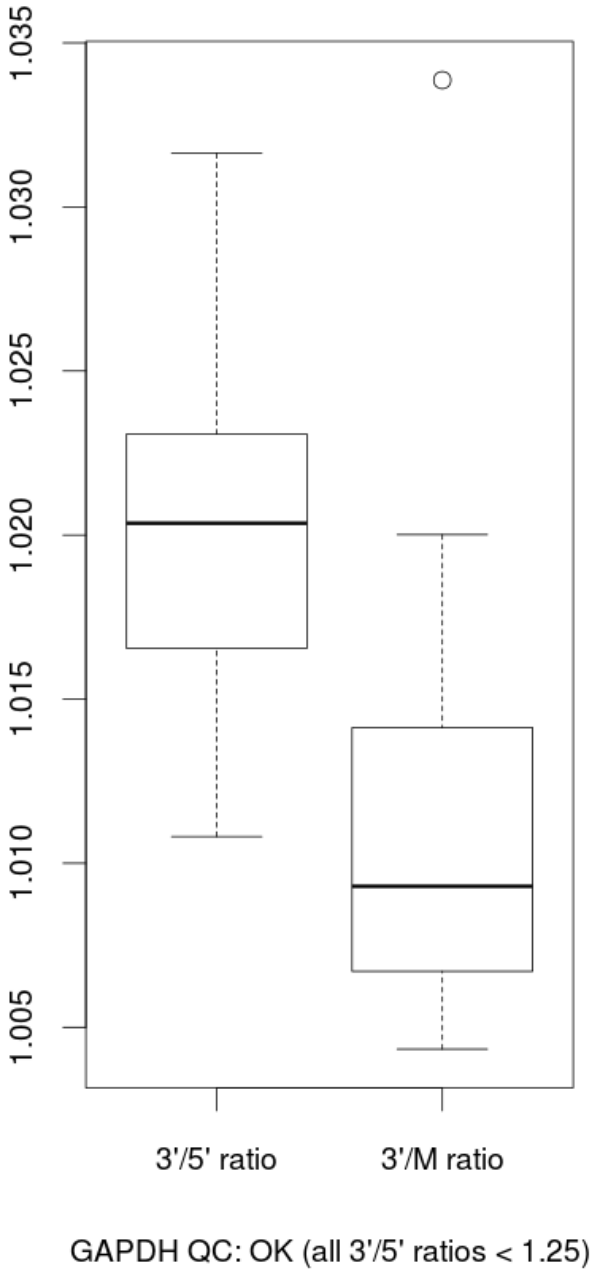


beta-actin QC: OK (all 3'/5' ratios < 3)

RNA degradation of GAPDH

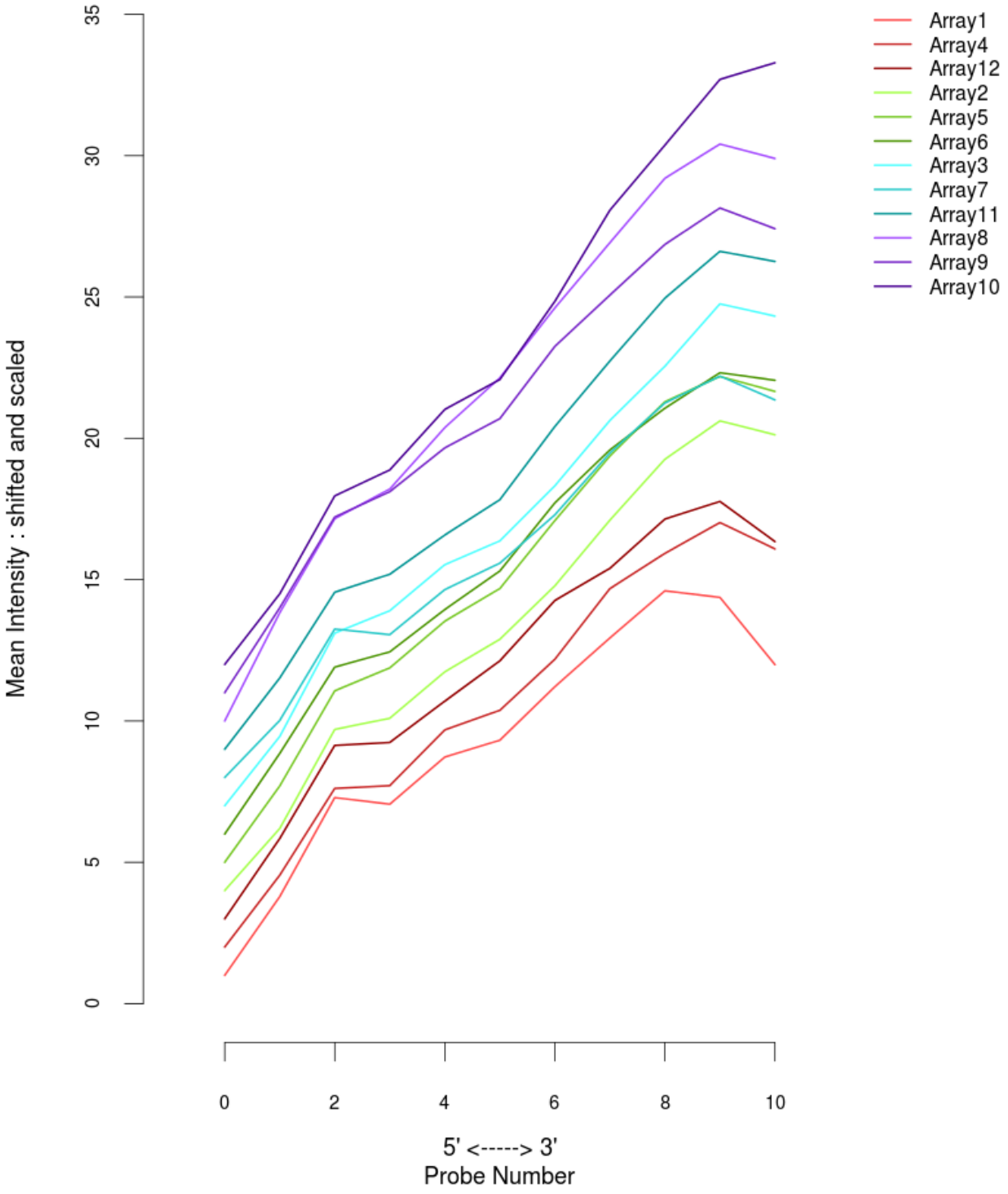


Boxplot of GAPDH ratios

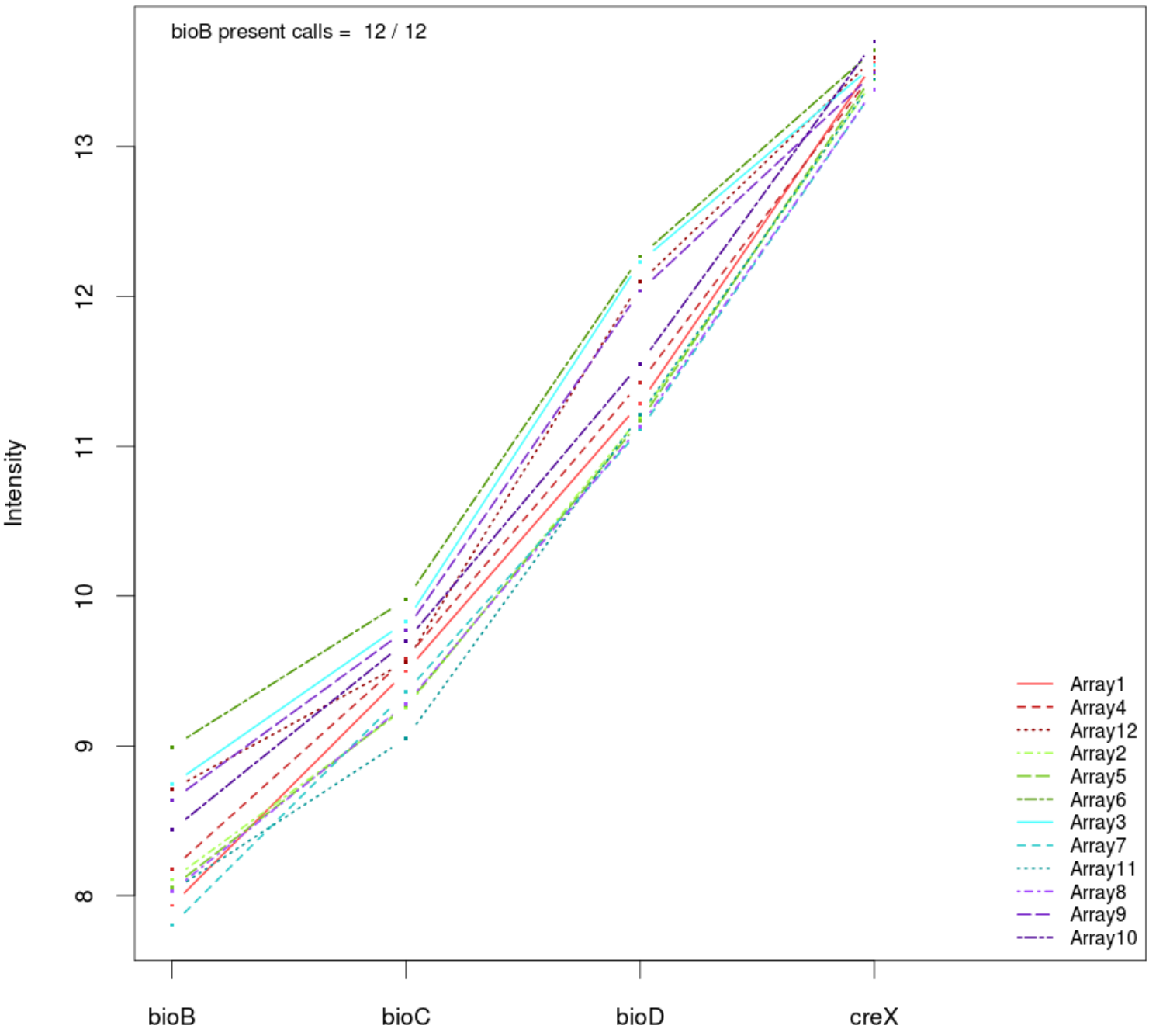


GAPDH QC: OK (all 3'/5' ratios < 1.25)

RNA degradation plot



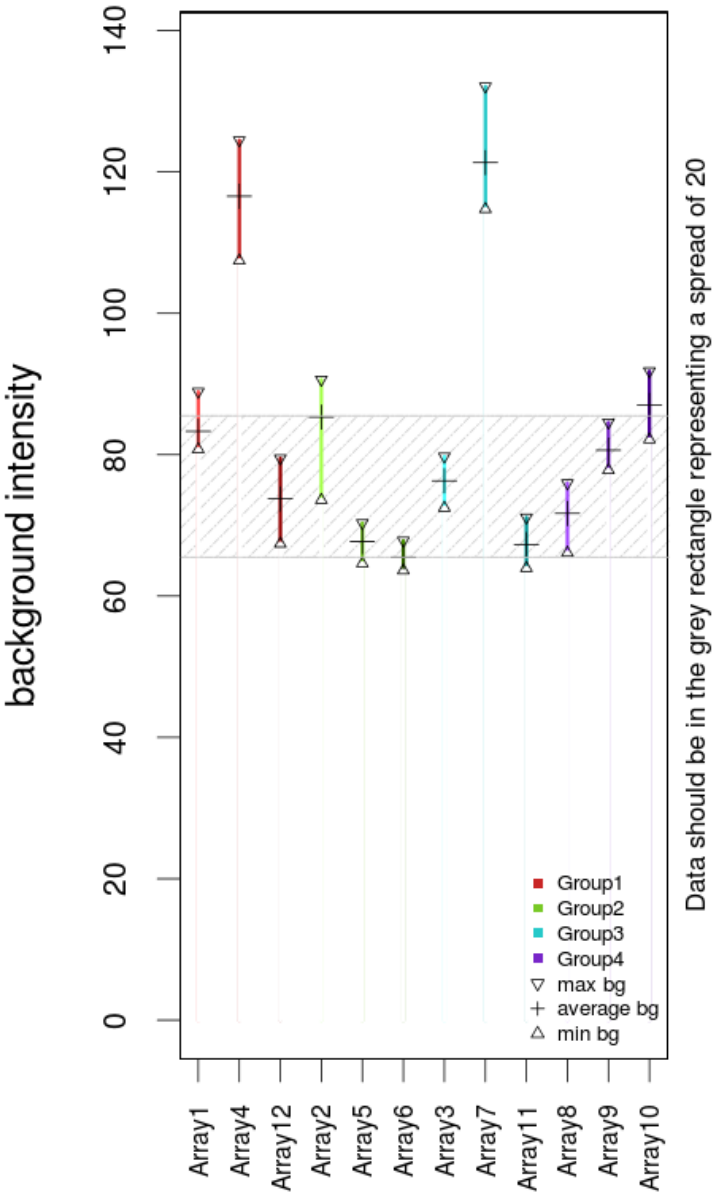
Spike-in Hybridization controls intensities and calls



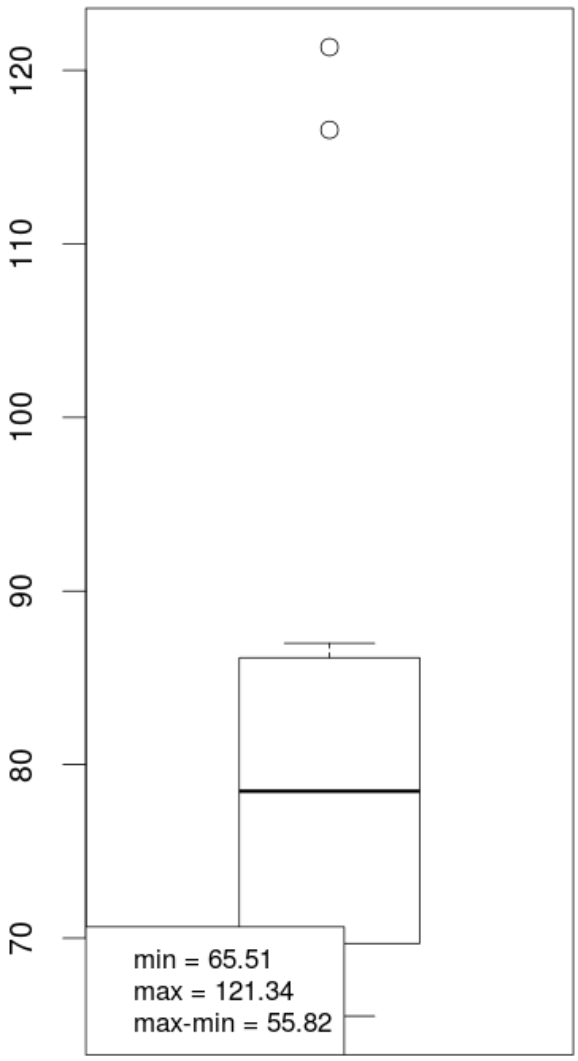
Intensities: OK (bioB < bioC < bioD < creX for all arrays)

BioB Present calls: OK (indeed all bioB are called present)

Plot of background intensity

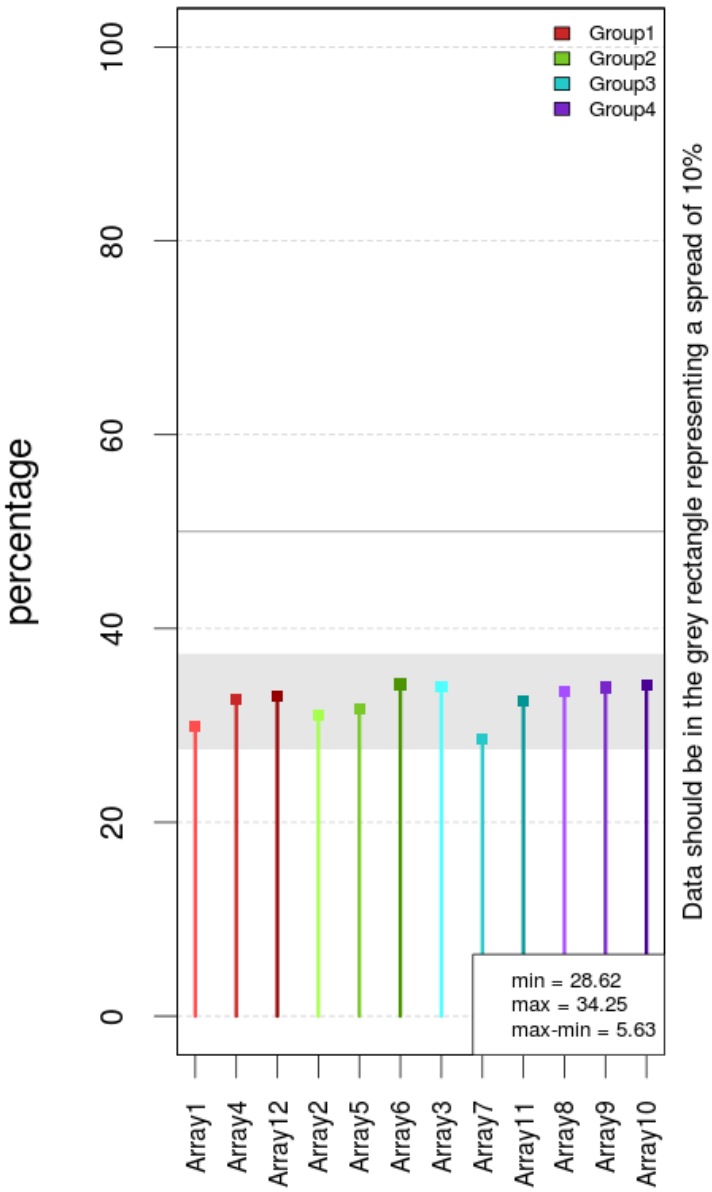


Average background intensity

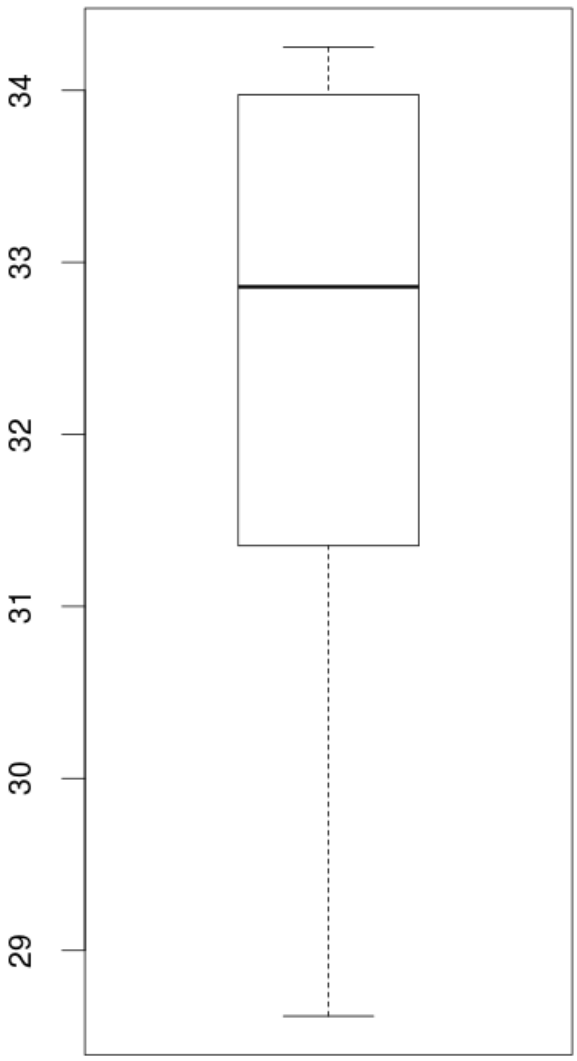


Background QC: not OK (spread > 20)

Plot of percent present

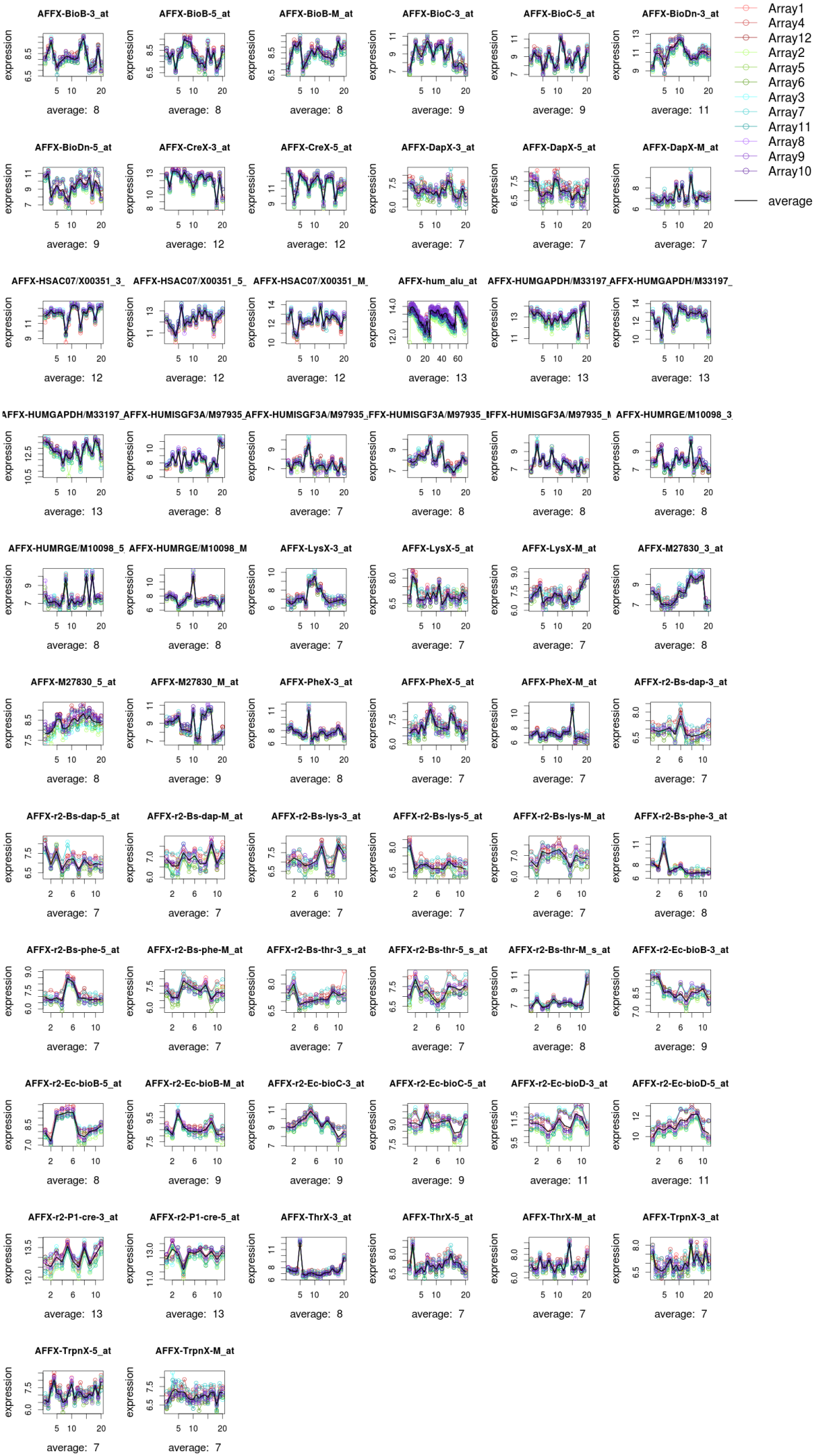


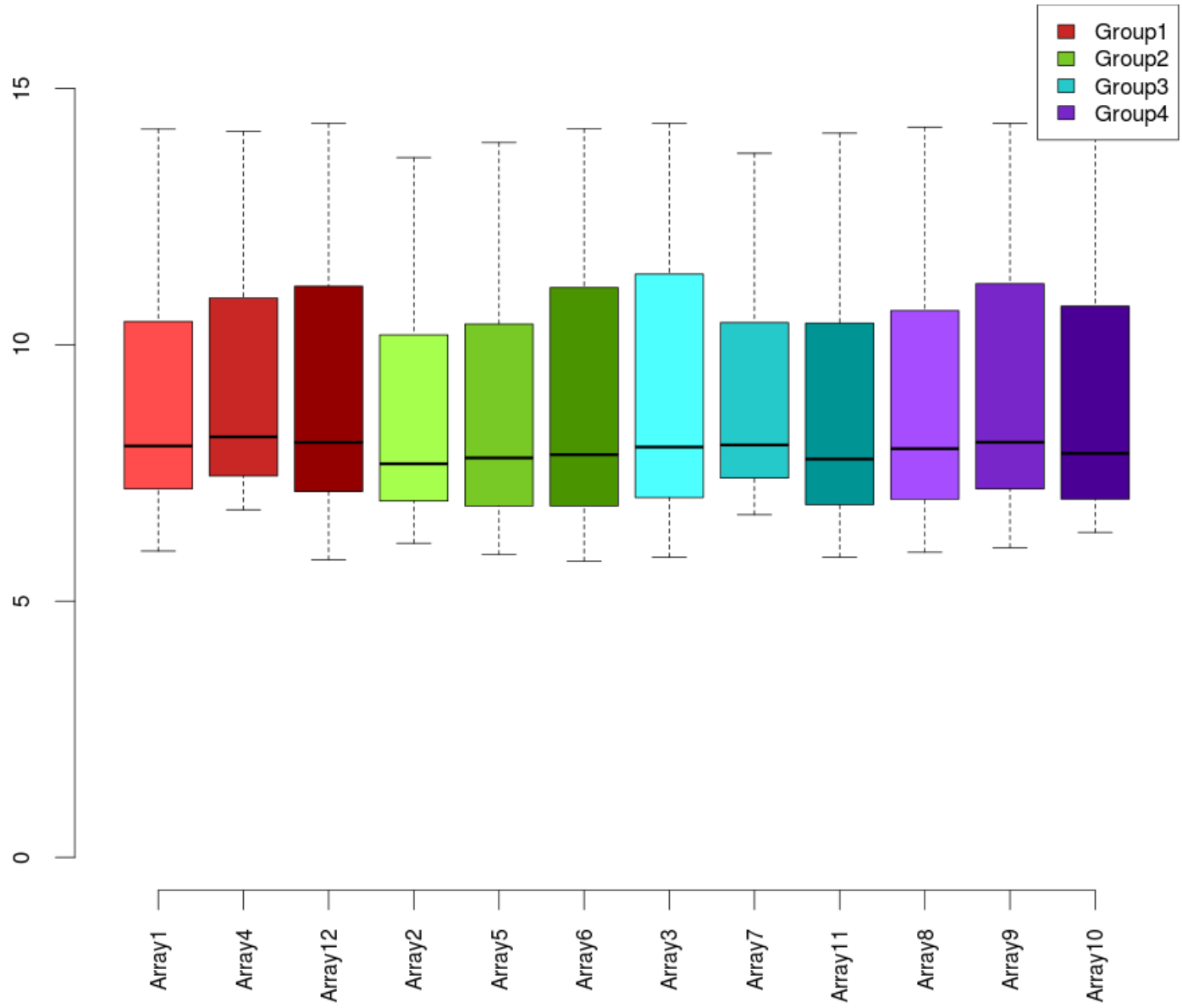
Boxplot of percent present

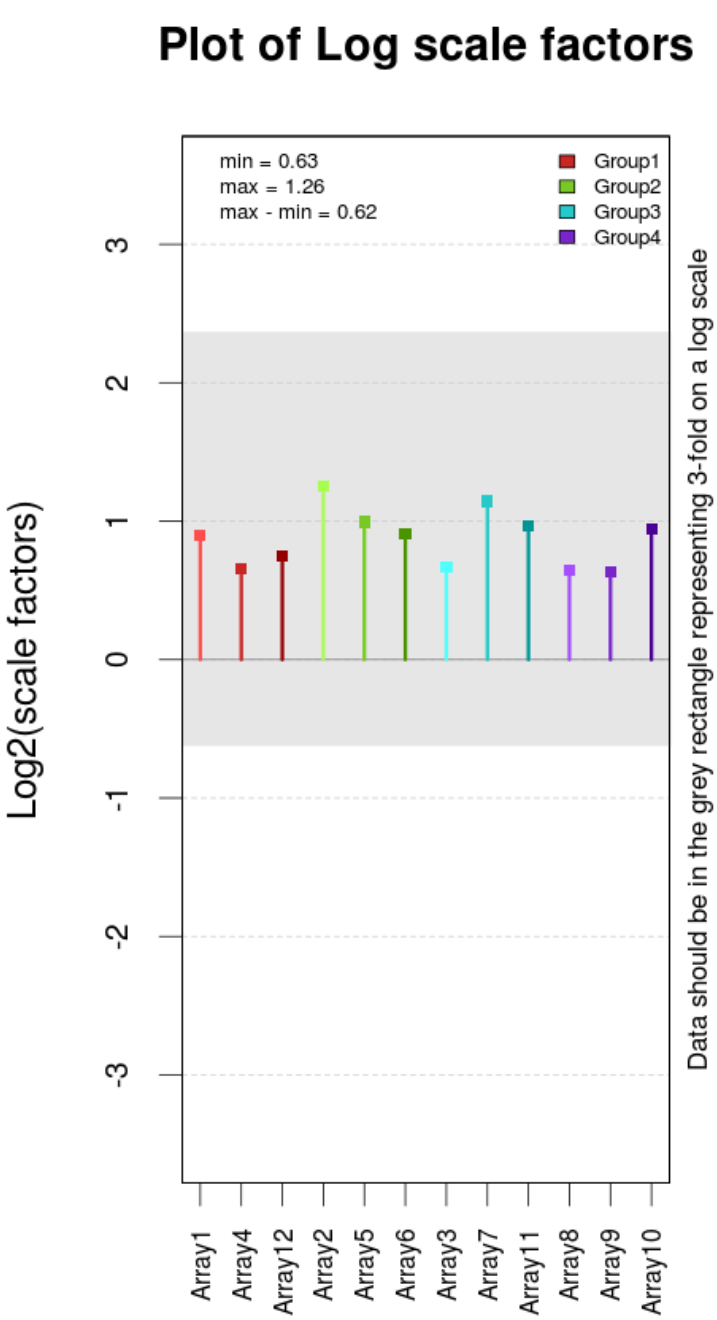


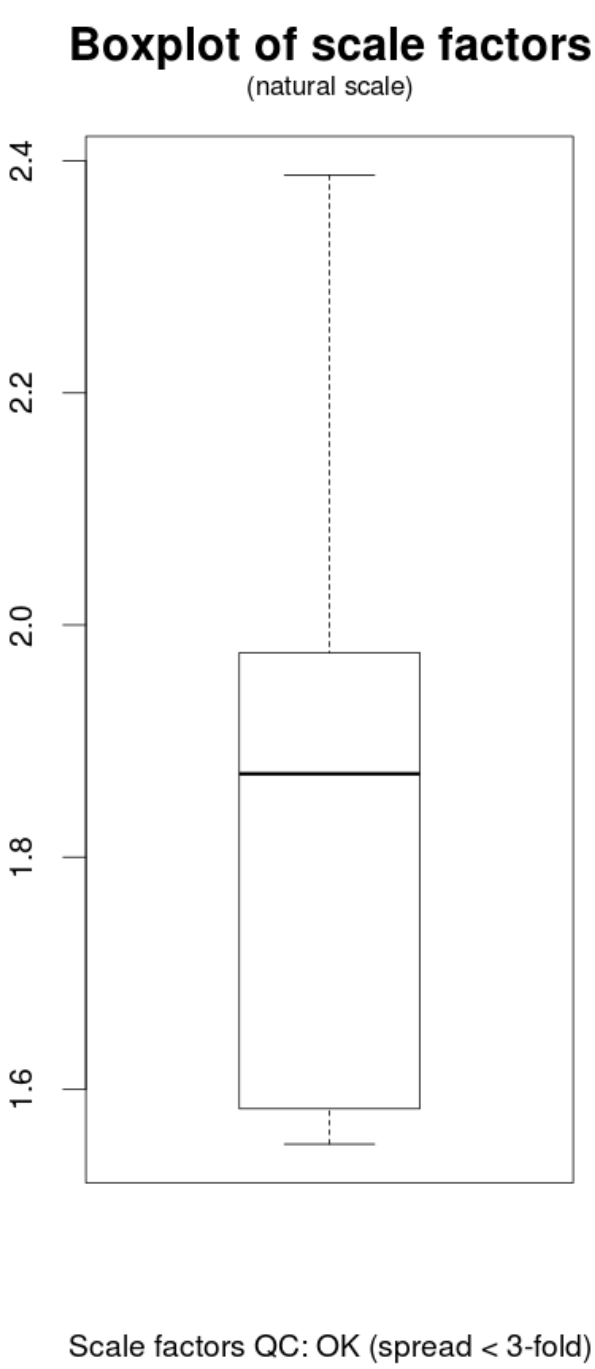
Percent present QC: OK (spread <= 10%)

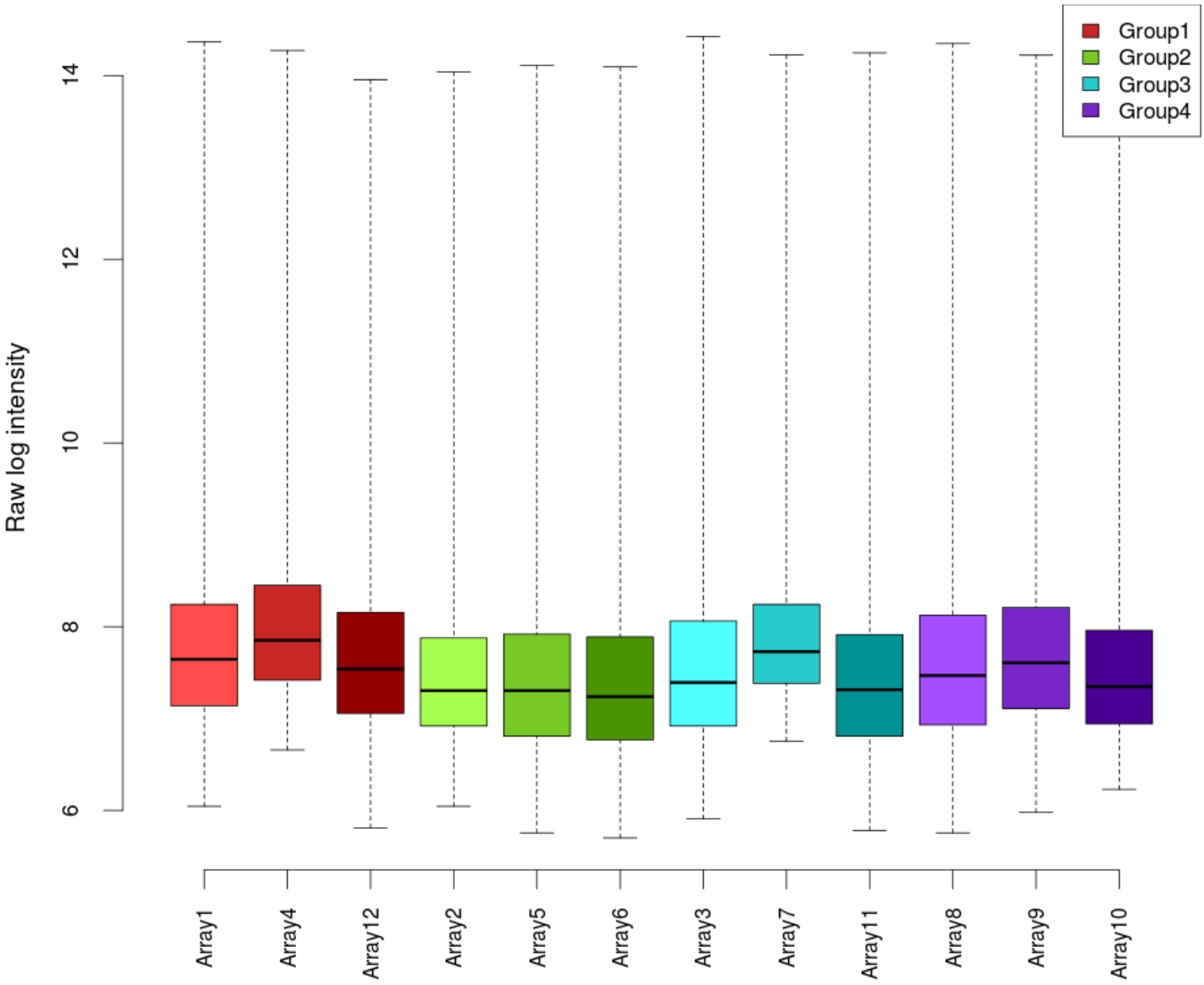
affx control profiles

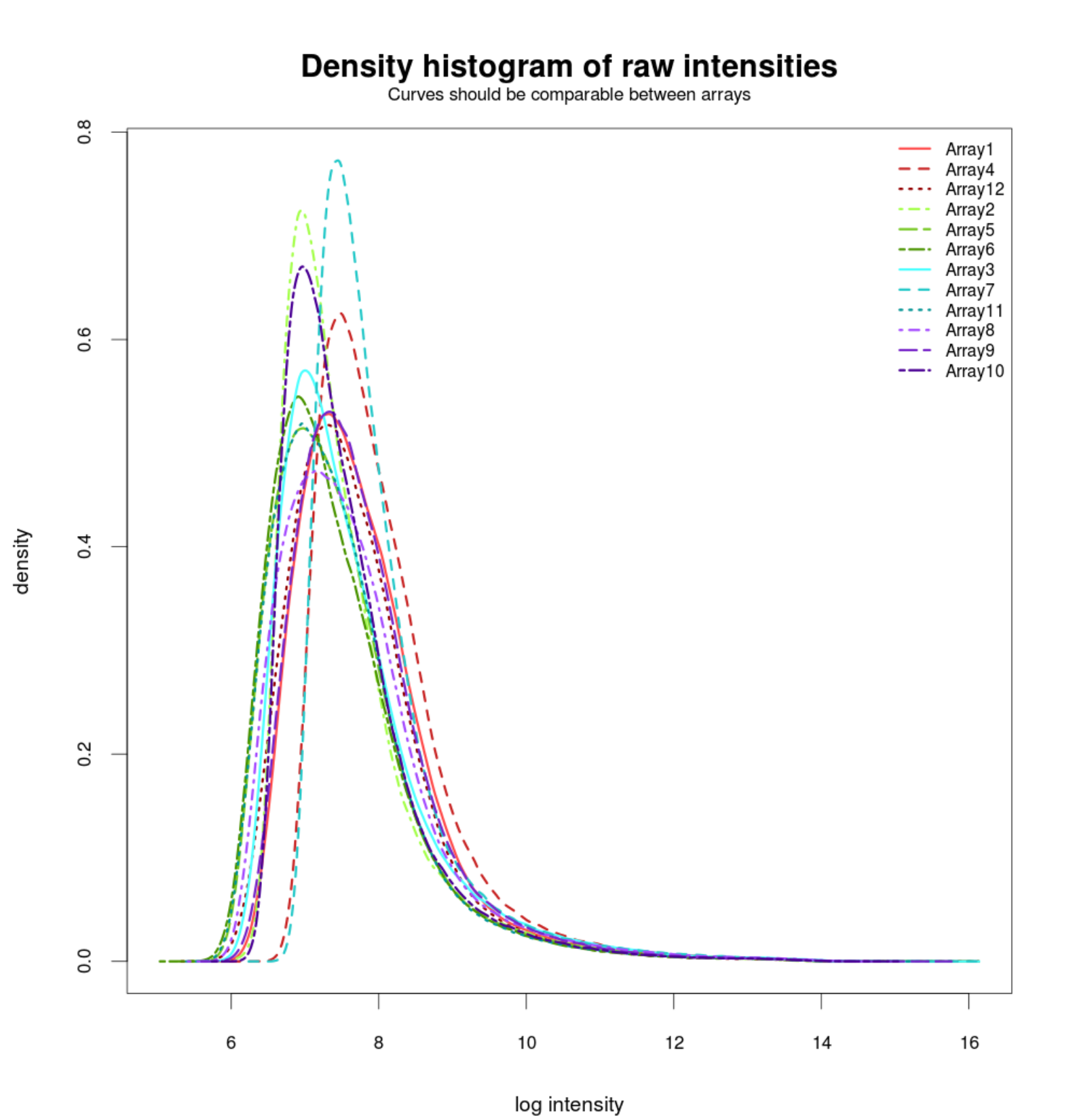



affx controls
Group1
Group2
Group3
Group4
0
5
10
15
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10


Plot of Log scale factors
min = 0.63
max = 1.26
max - min = 0.62
Group1
Group2
Group3
Group4
Log2(scale factors)
3
2
1
0
-1
-2
-3
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10
Data should be in the grey rectangle representing 3-fold on a log scale

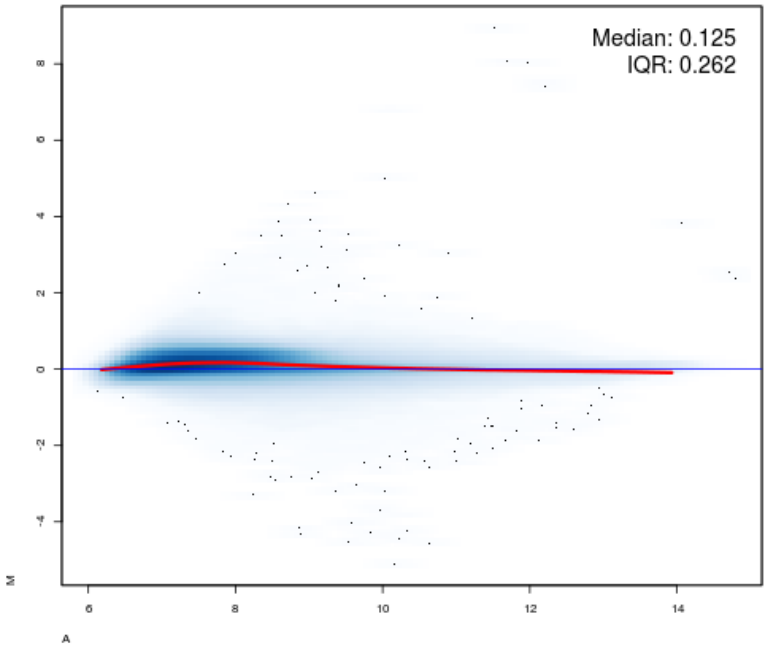

Boxplot of scale factors
(natural scale)
2.4
2.2
2.0
1.8
1.6
Scale factors QC: OK (spread < 3-fold)


Boxplot of raw intensities
Distributions should be comparable between arrays
Raw log intensity
6
8
10
12
14
Group1
Group2
Group3
Group4
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10

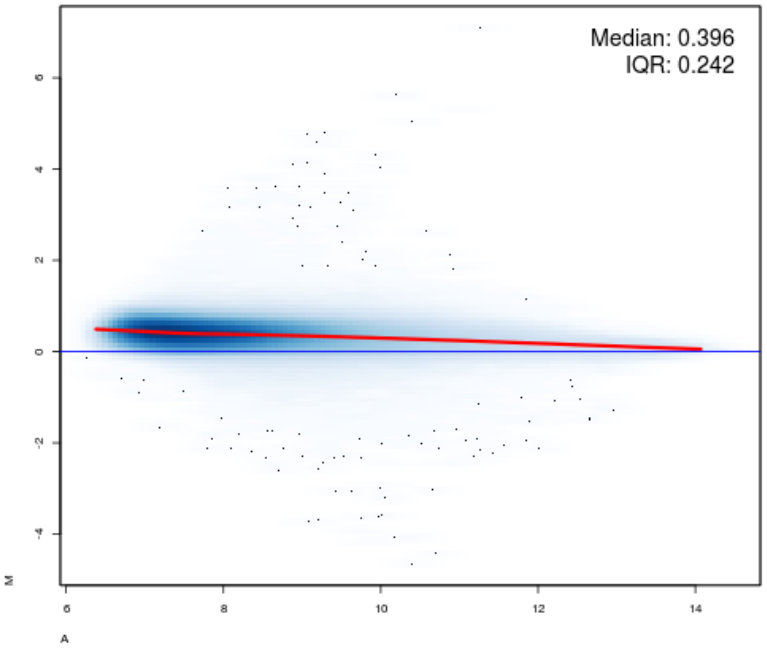

Density histogram of raw intensities
Curves should be comparable between arrays
density
log intensity
0.0
0.2
0.4
0.6
0.8
6
8
10
12
14
16
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10

MA plots of raw data 1 / 2

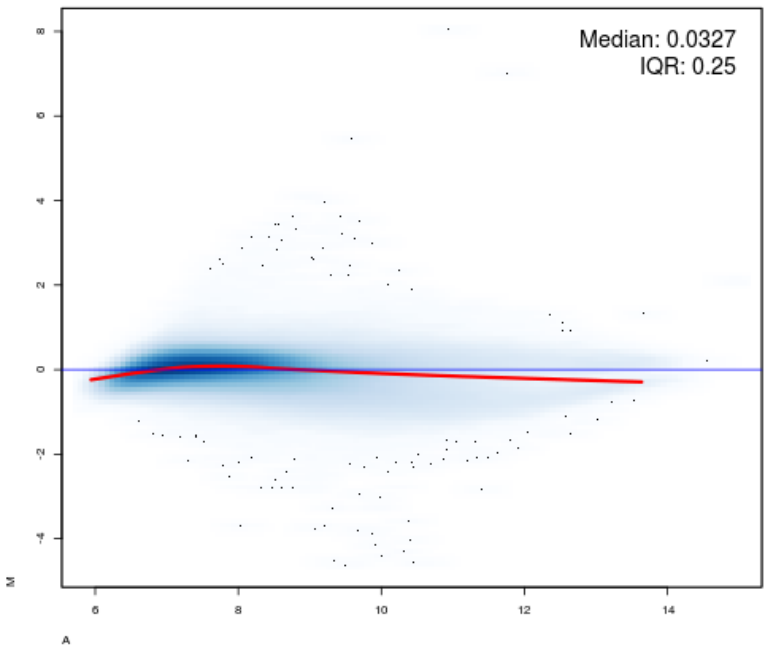
Array1 vs pseudo-median reference chip



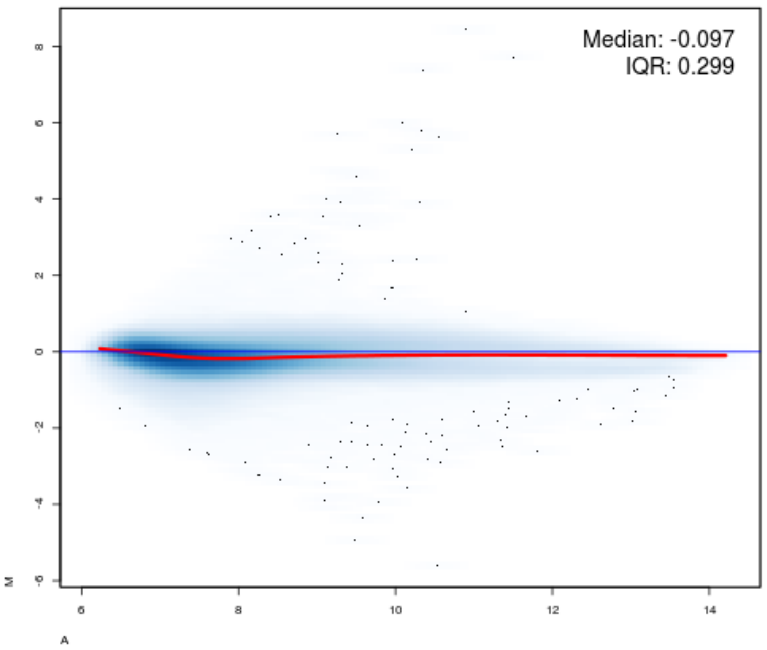
Array4 vs pseudo-median reference chip



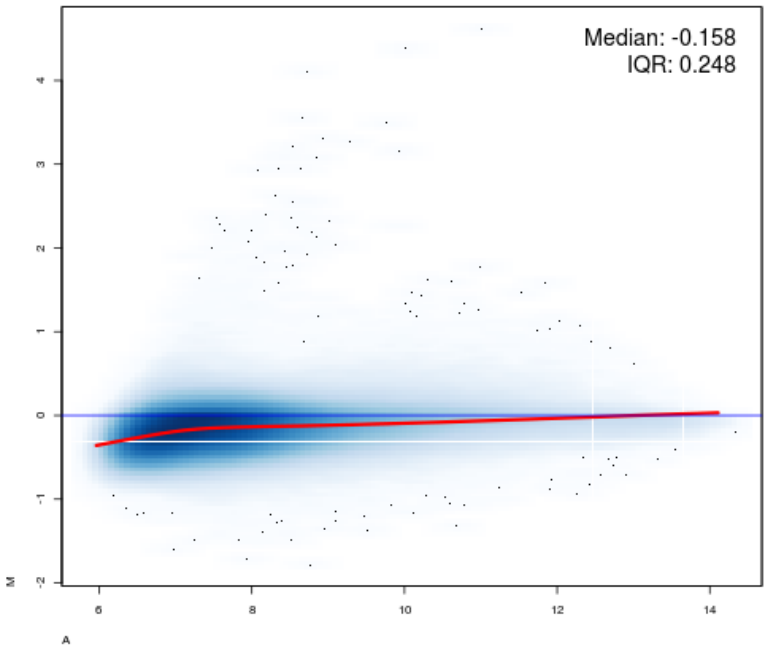
Array12 vs pseudo-median reference chip



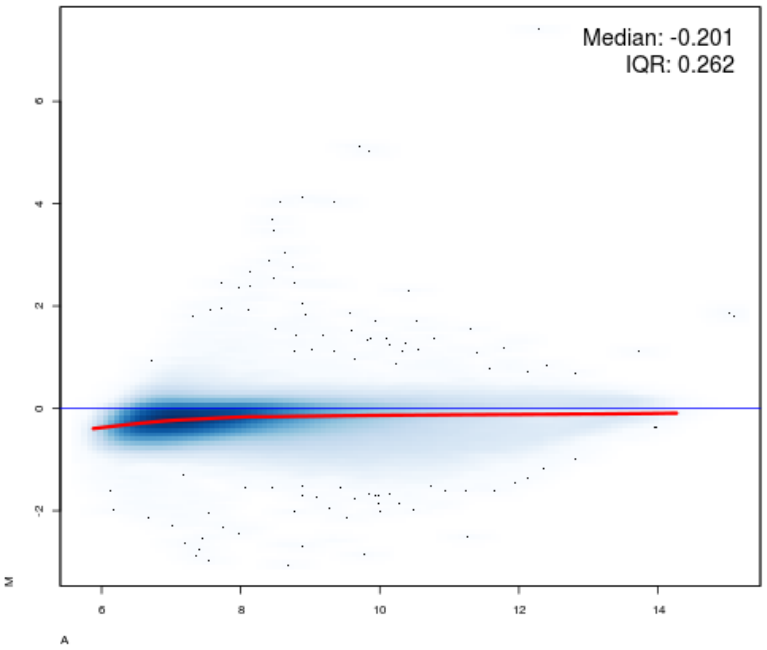
Array2 vs pseudo-median reference chip



Array5 vs pseudo-median reference chip

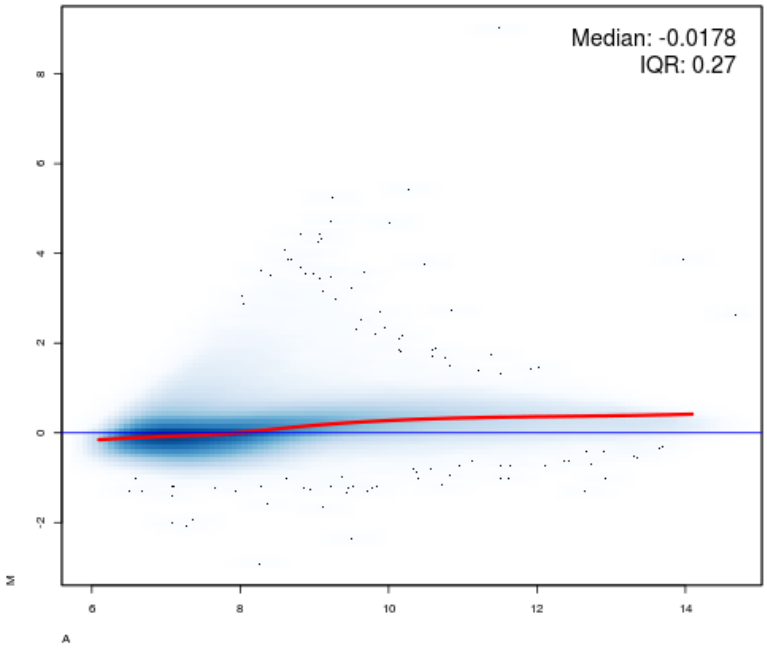


Array6 vs pseudo-median reference chip

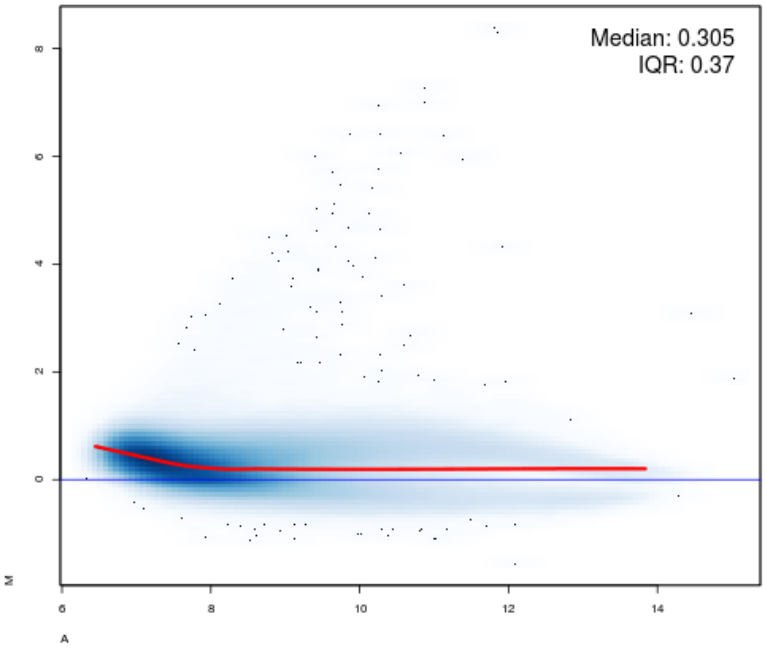


MA plots of raw data 2 / 2

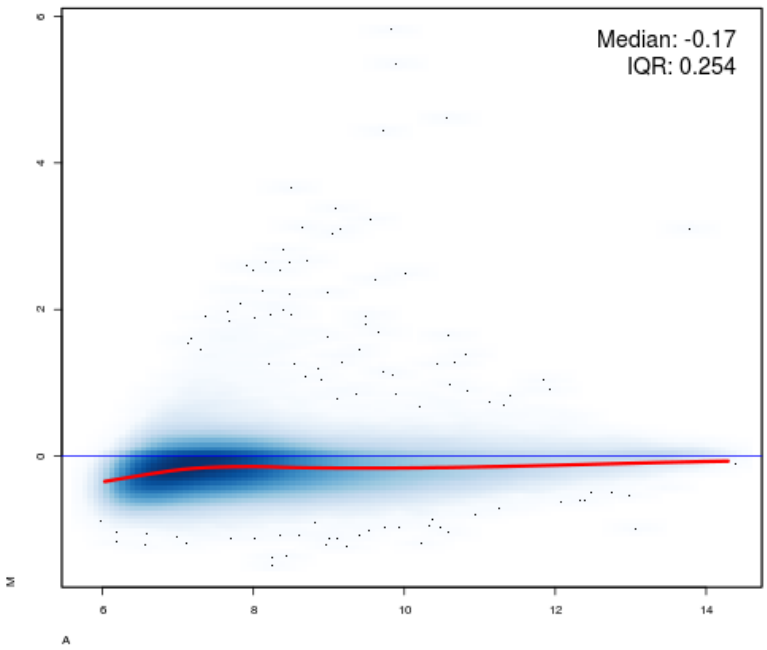
Array3 vs pseudo-median reference chip



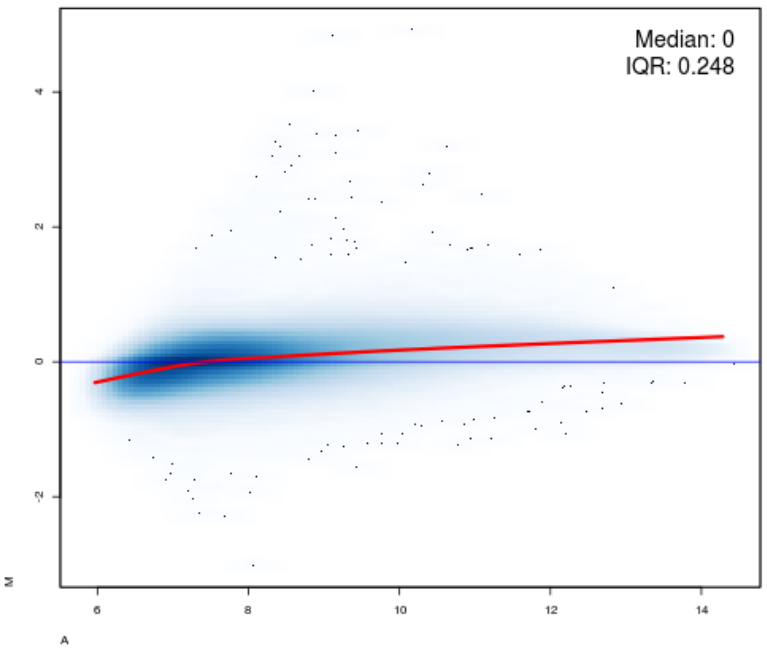
Array7 vs pseudo-median reference chip



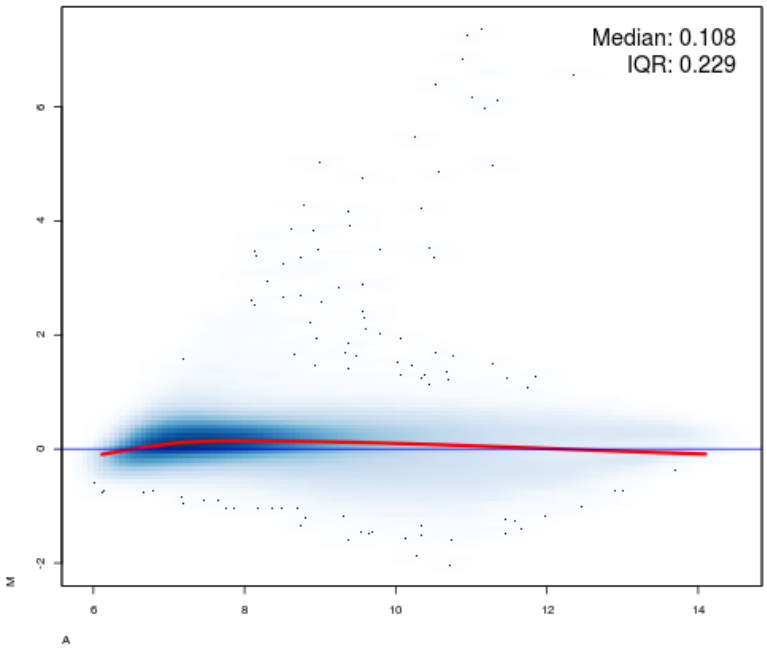
Array11 vs pseudo-median reference chip



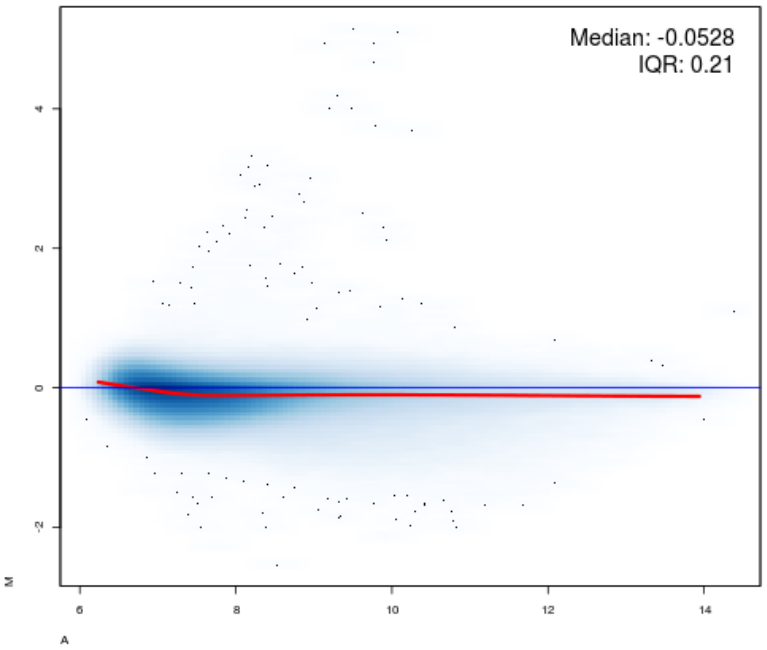
Array8 vs pseudo-median reference chip



Array9 vs pseudo-median reference chip

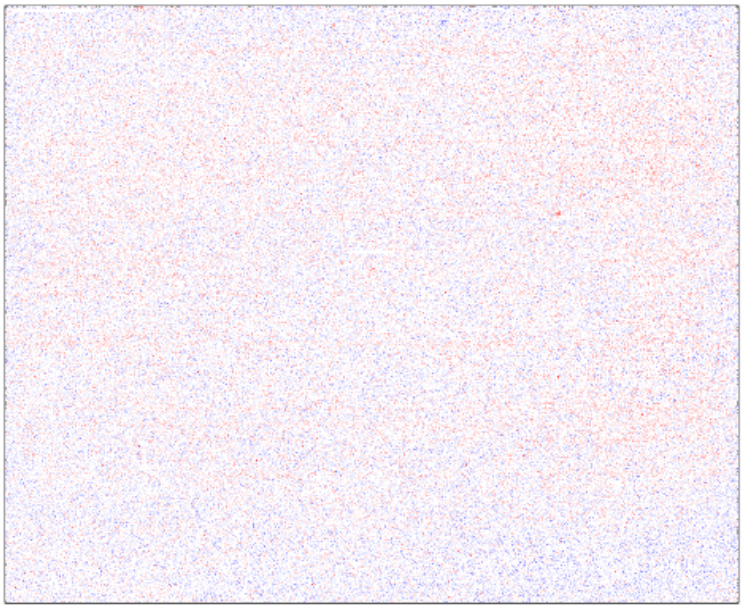


Array10 vs pseudo-median reference chip

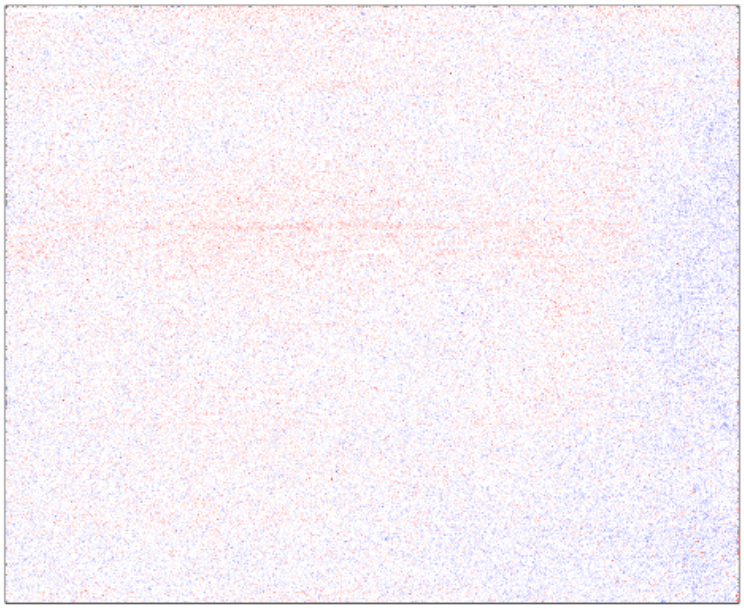


2D virtual PLM image for model characteristic: resids 1 / 2

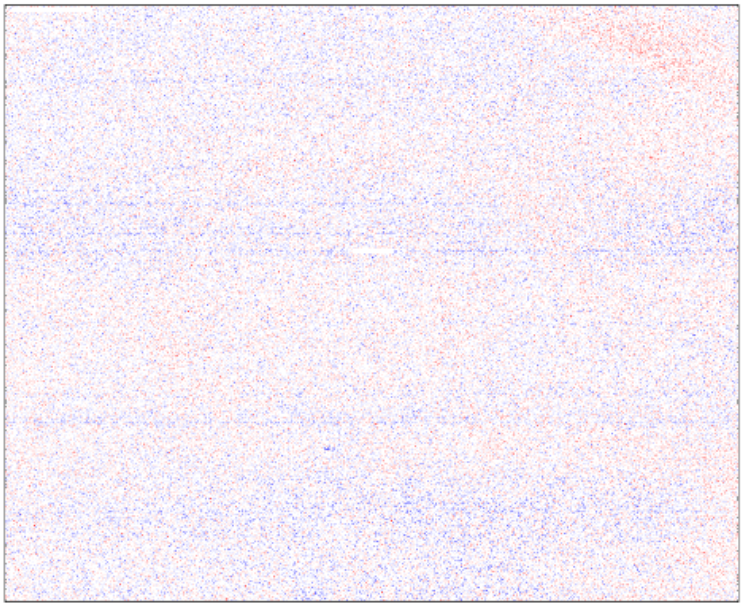
Array1



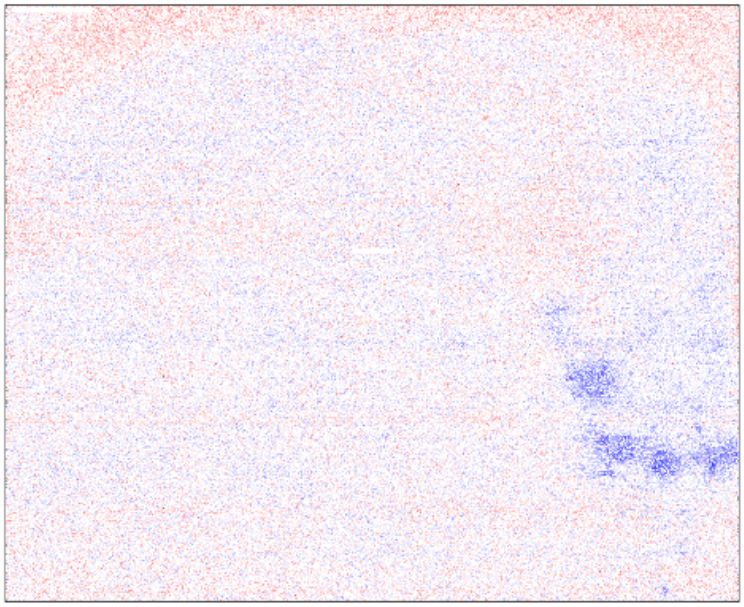
Array4



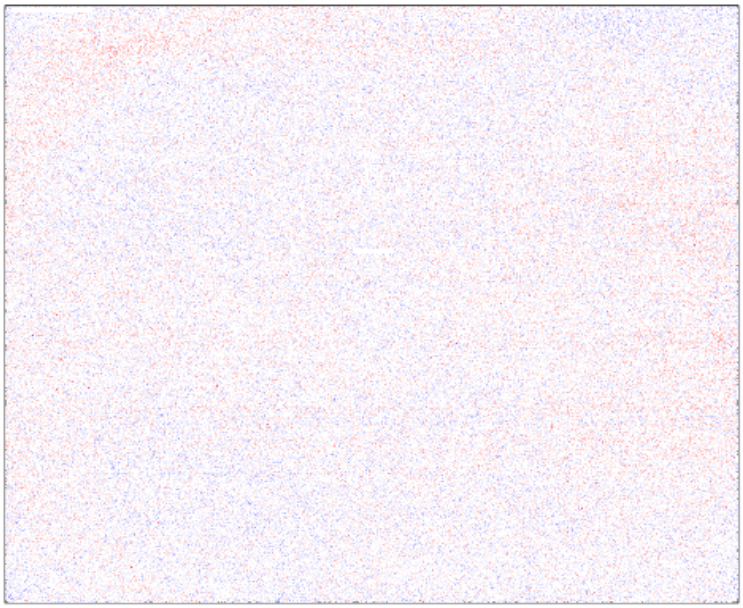
Array12



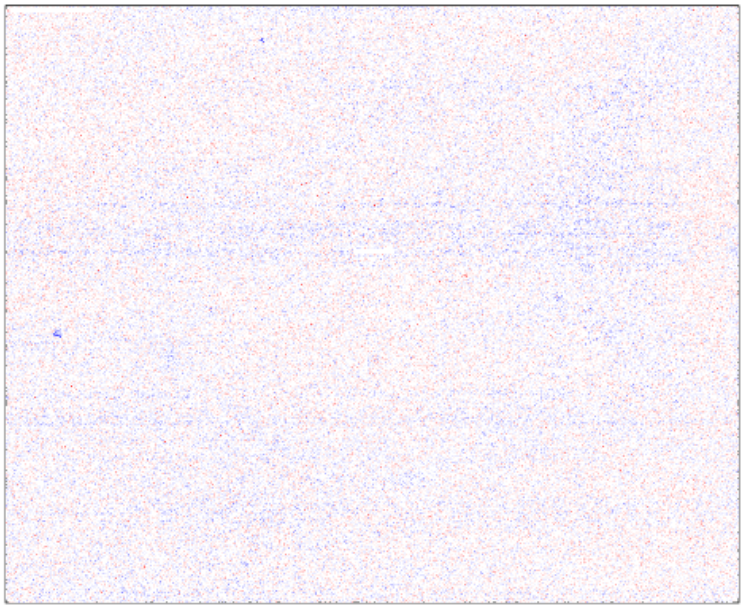
Array2



Array5

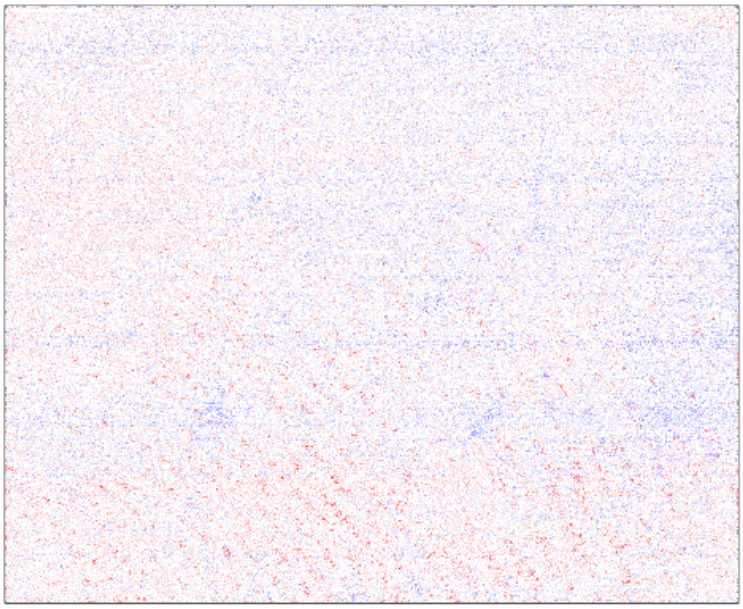


Array6

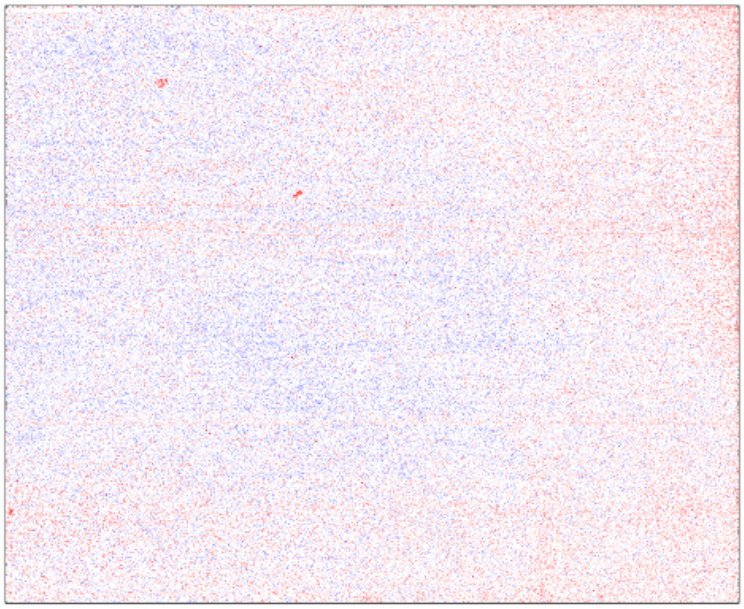


2D virtual PLM image for model characteristic: resids 2 / 2

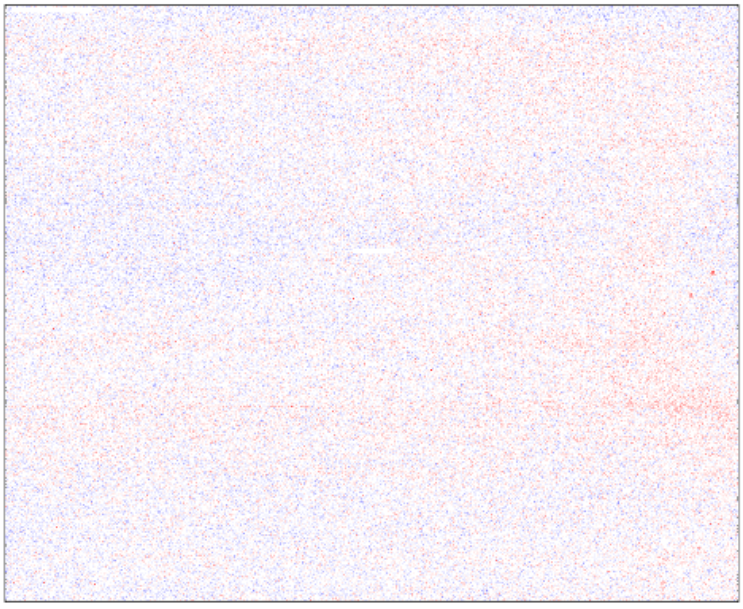
Array3



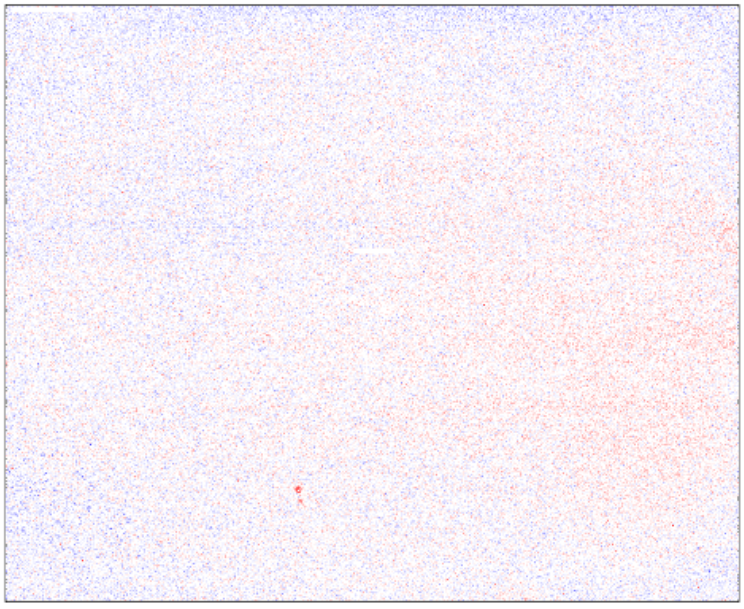
Array7



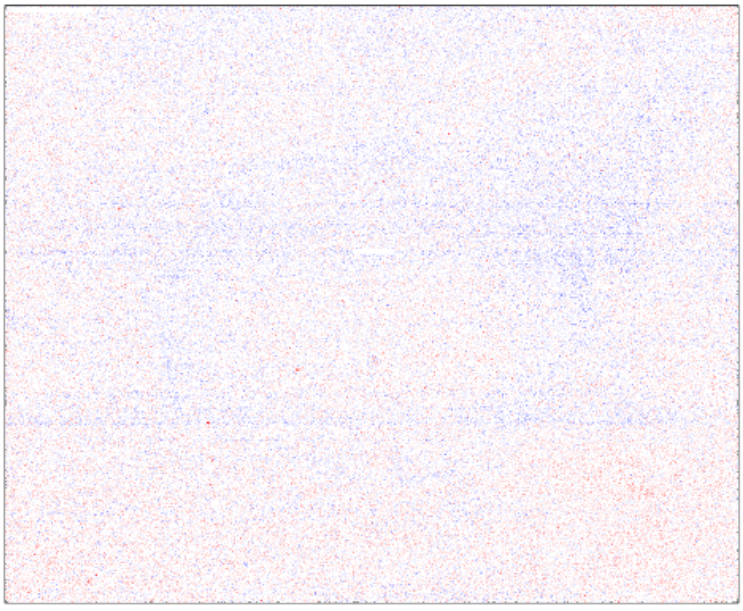
Array11



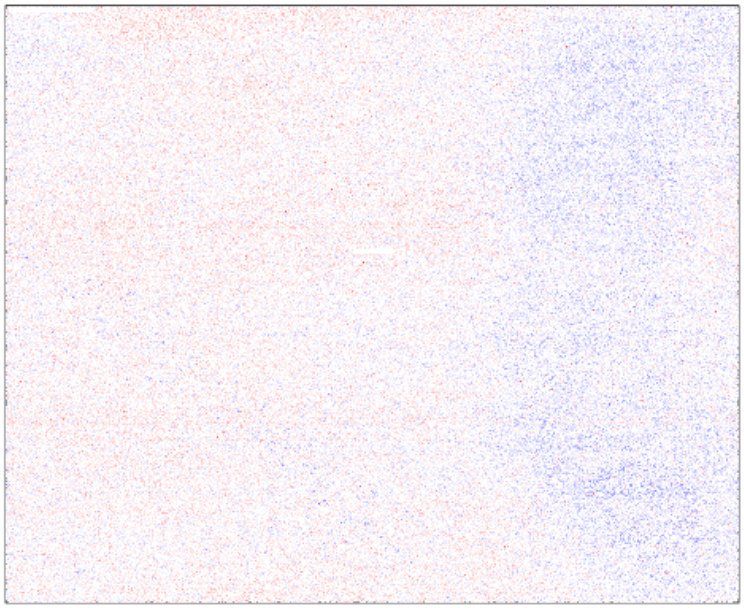
Array8



Array9

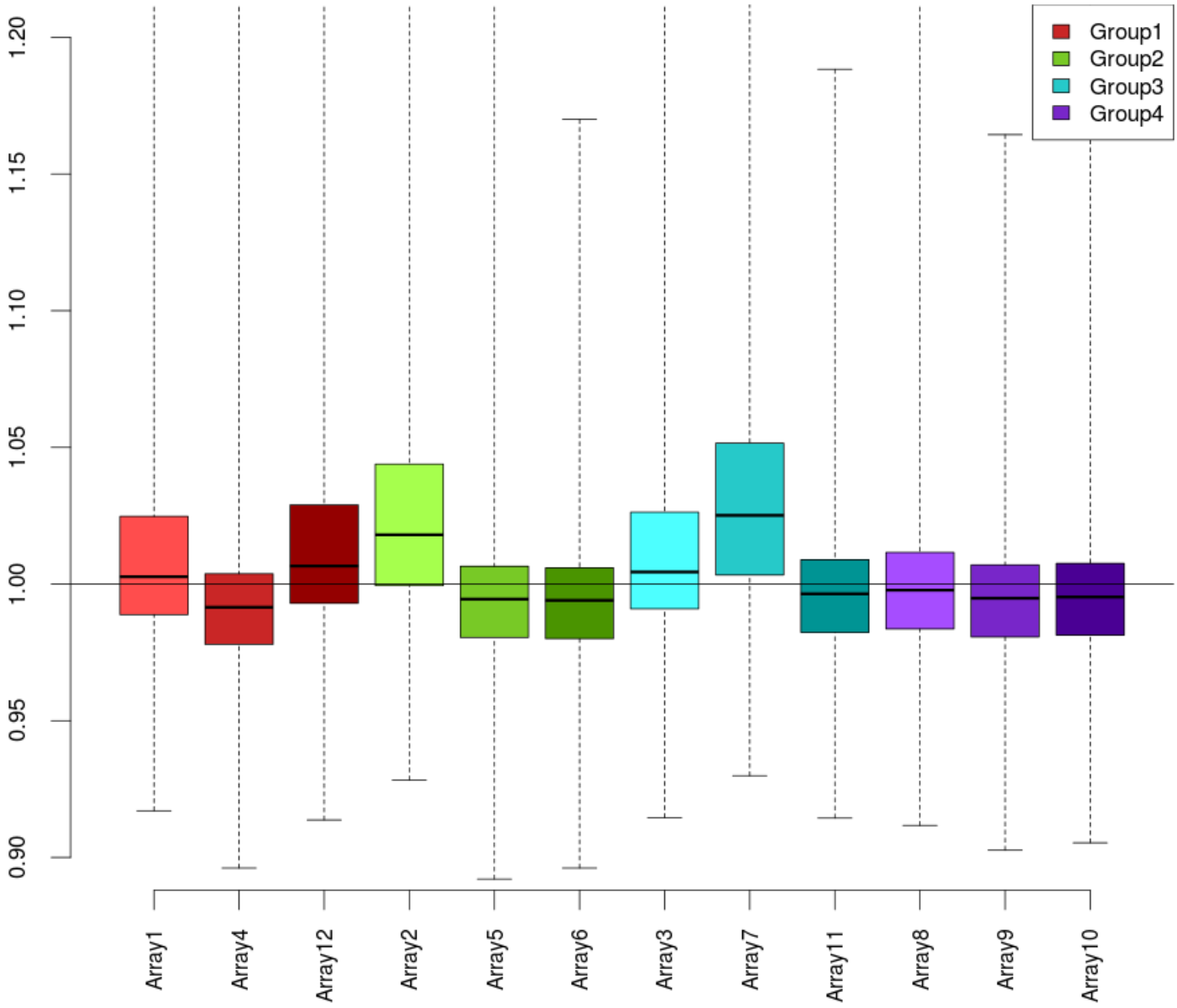


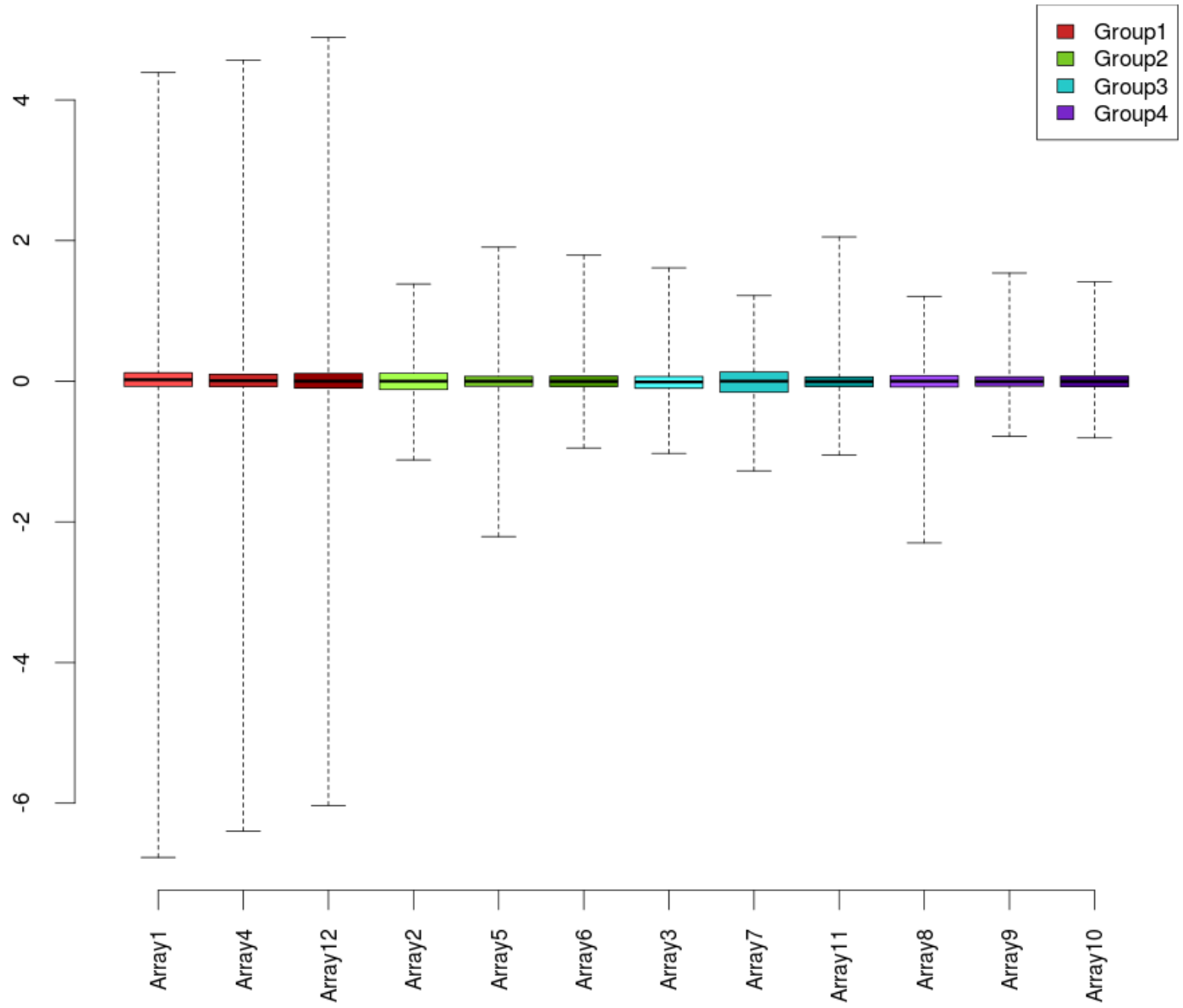
Array10

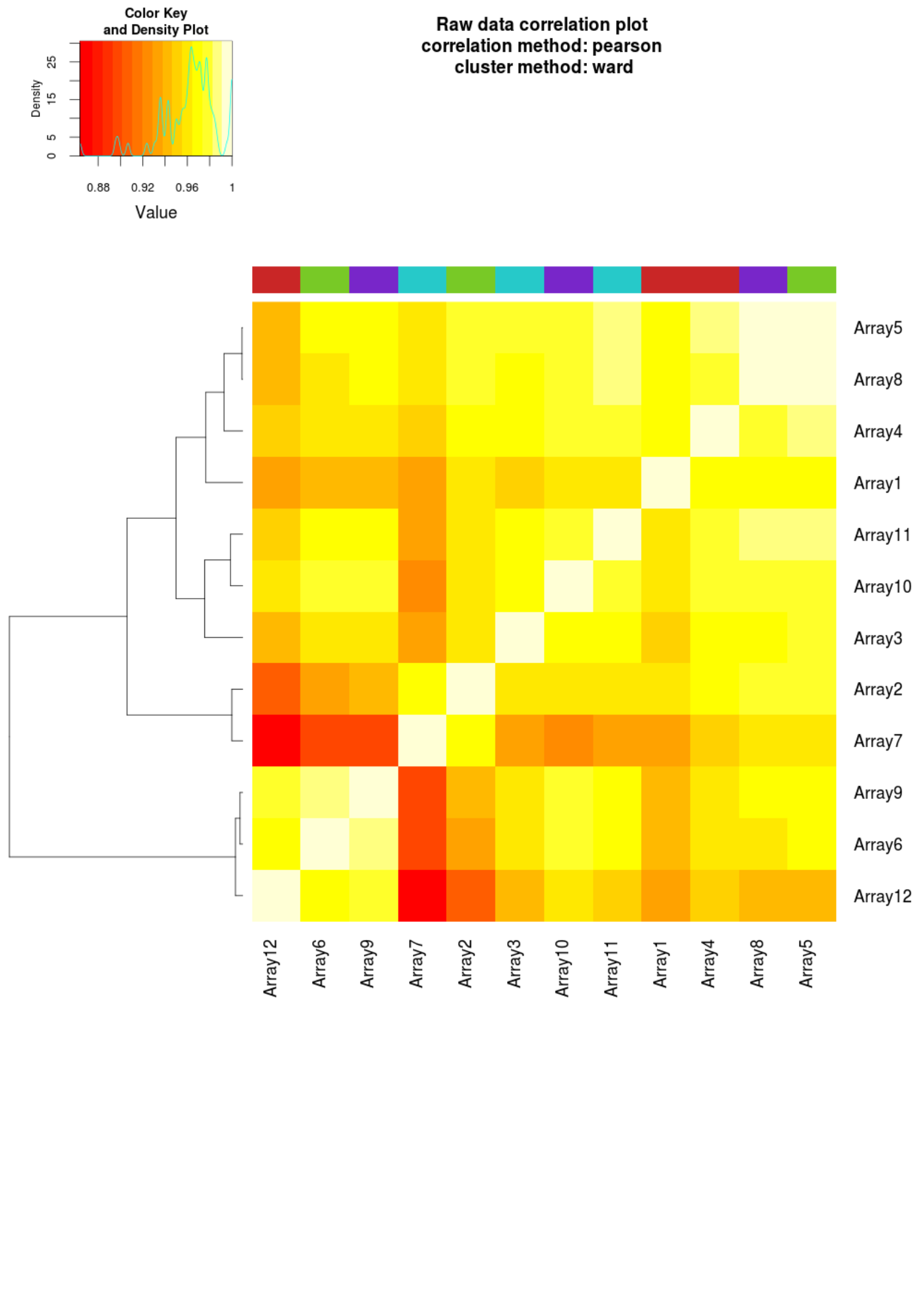


Normalized Unscaled Standard Errors (NUSE)

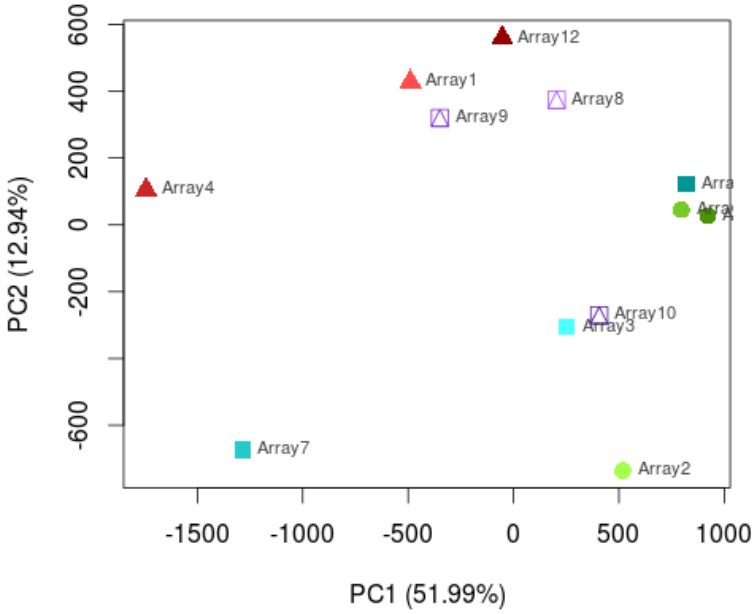
NUSE median value should be < 1.1

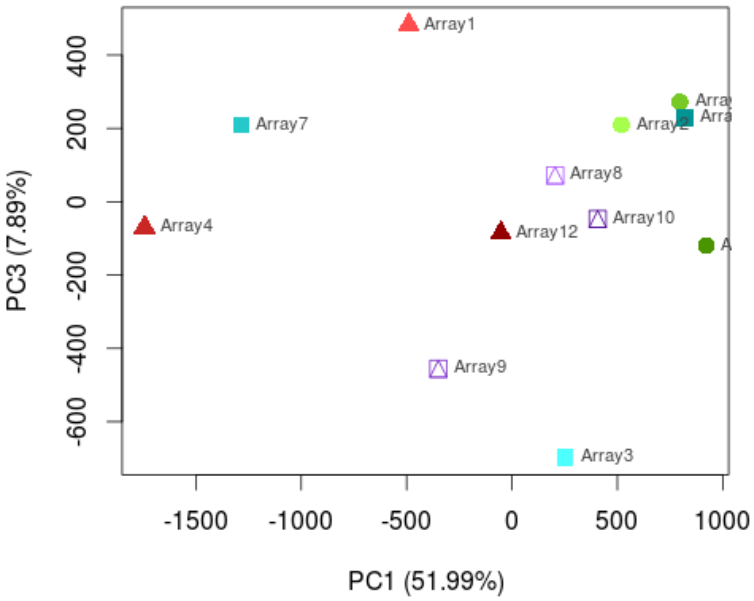


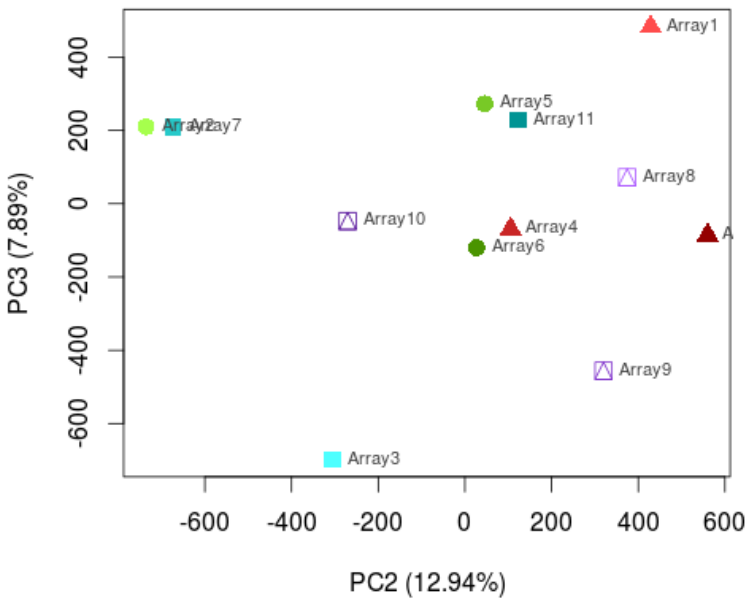

Relative Log Expression (RLE)
RLE distributions should be centered around 0
Group1
Group2
Group3
Group4
4
2
0
-2
-4
-6
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10

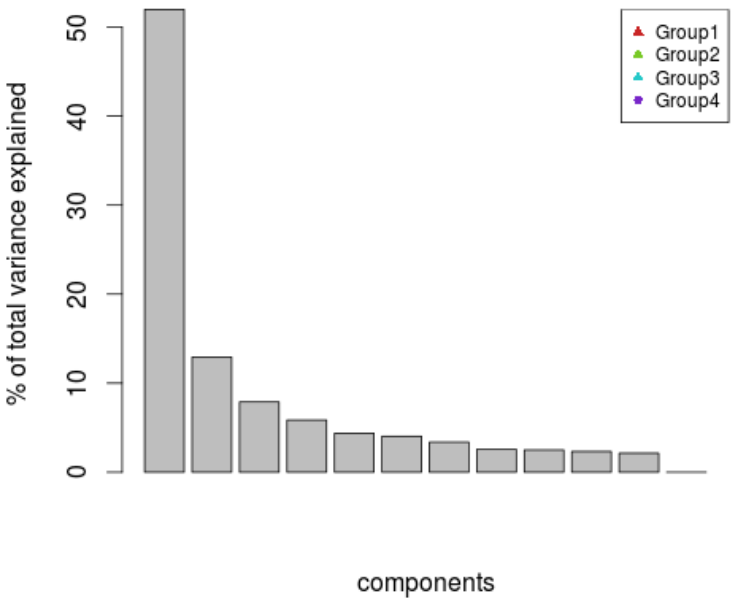

Color Key
and Density Plot
Density
0
5
15
25
0.88
0.92
0.96
1
Value
Raw data correlation plot
correlation method: pearson
cluster method: ward
Array5
Array8
Array4
Array1
Array11
Array10
Array3
Array2
Array7
Array9
Array6
Array12
Array12
Array6
Array9
Array7
Array2
Array3
Array10
Array11
Array1
Array4
Array8
Array5

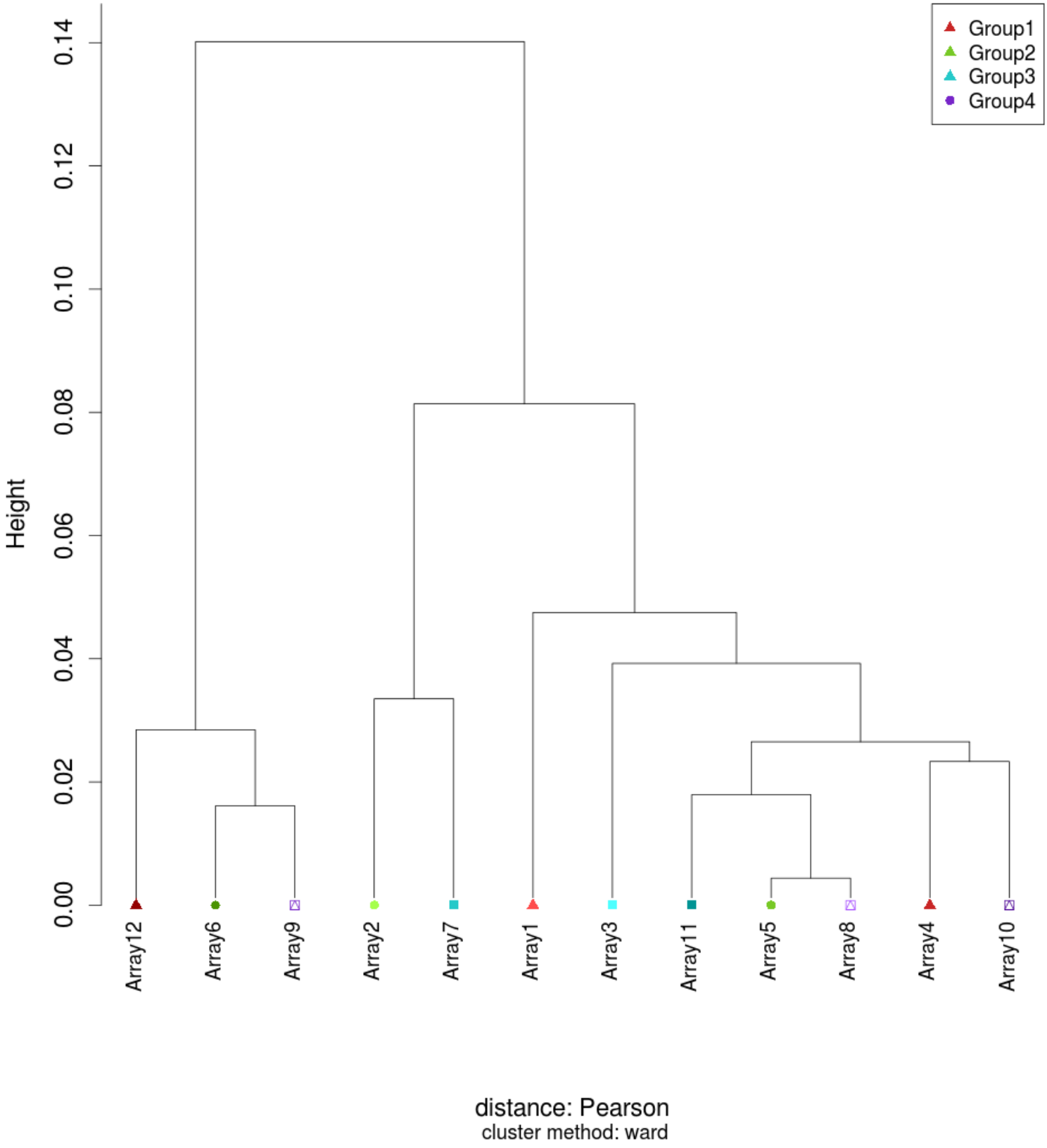
PCA analysis of Raw data





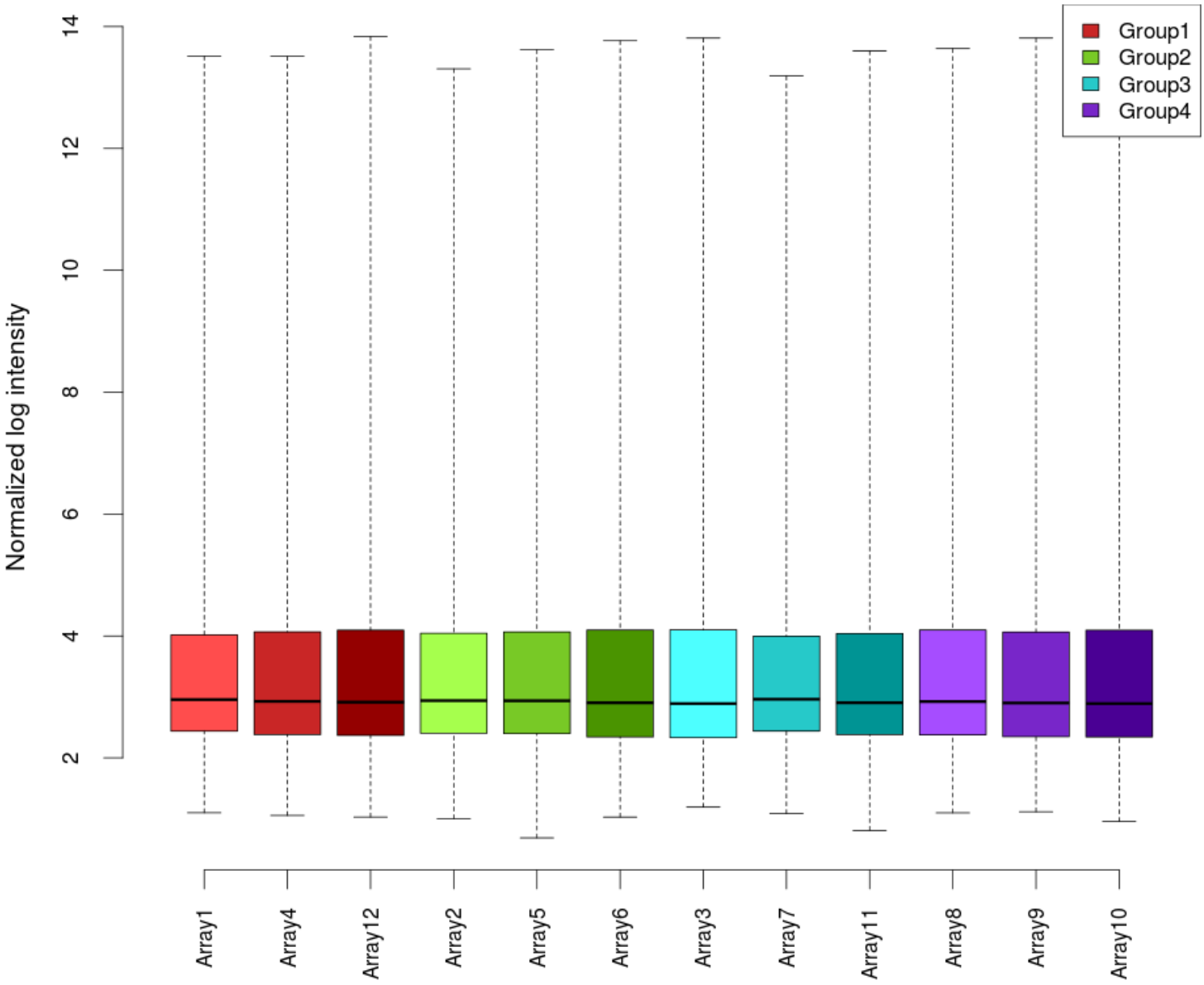


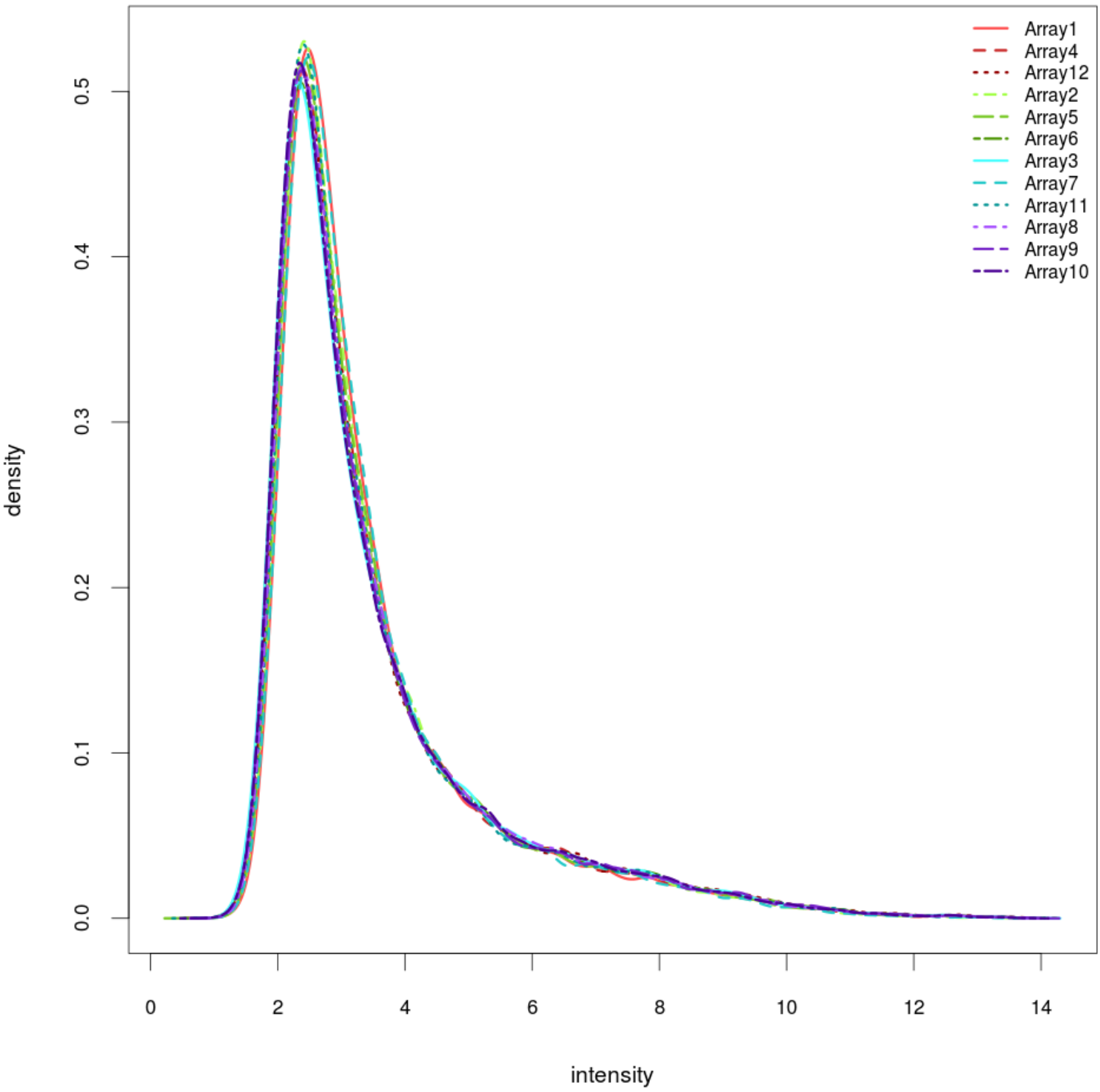



Cluster dendrogram of raw data
Group1
Group2
Group3
Group4
Height
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Array12
Array6
Array9
Array2
Array7
Array1
Array3
Array11
Array5
Array8
Array4
Array10
distance: Pearson
cluster method: ward

Pre-processing of Raw Data

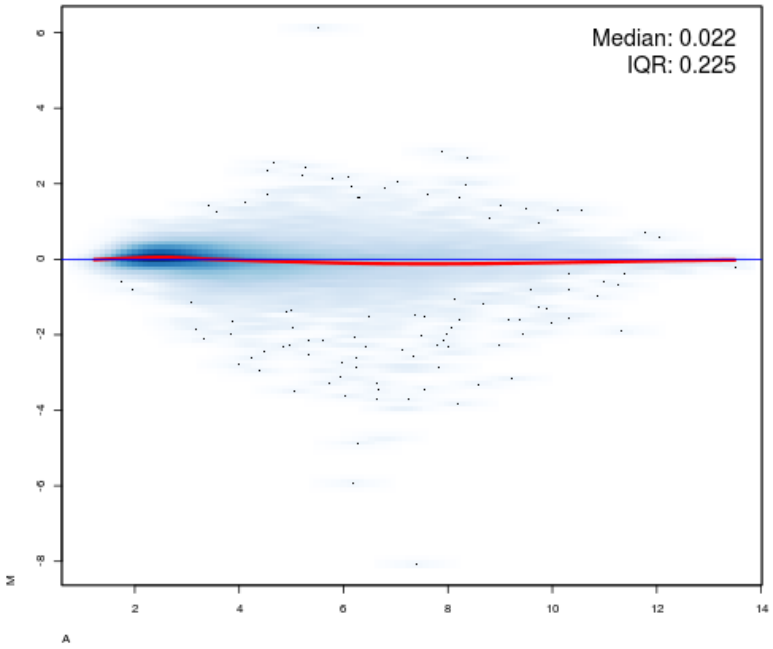
Method: GCRMA
Annotation: hgu133plus2_Hs_ENSG


Boxplot after GCRMA
Distributions should be comparable between arrays
Normalized log intensity
2
4
6
8
10
12
14
Group1
Group2
Group3
Group4
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10

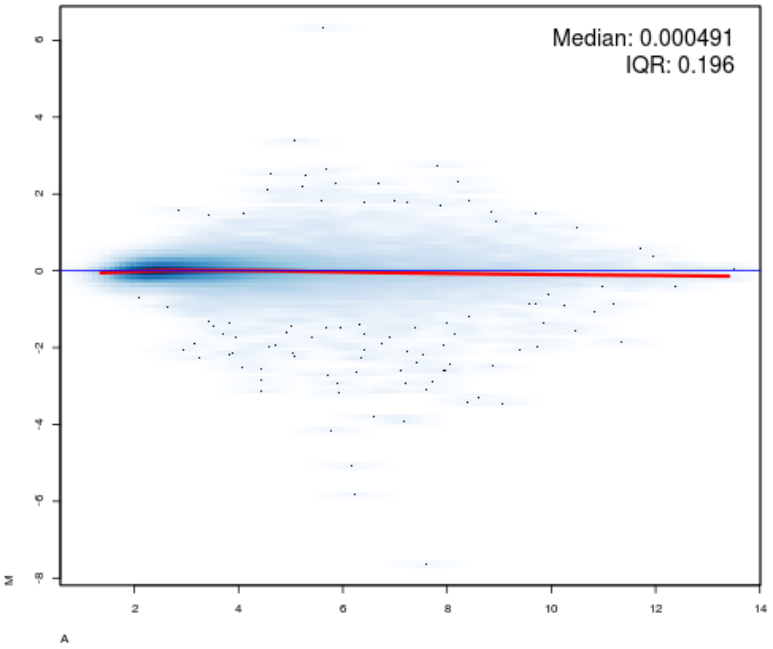

Density histogram after GCRMA
Curves should be comparable between arrays
Array1
Array4
Array12
Array2
Array5
Array6
Array3
Array7
Array11
Array8
Array9
Array10
density
0.0
0.1
0.2
0.3
0.4
0.5
0
2
4
6
8
10
12
14
intensity

MA plots afterGCRMAnormalization 1 / 2

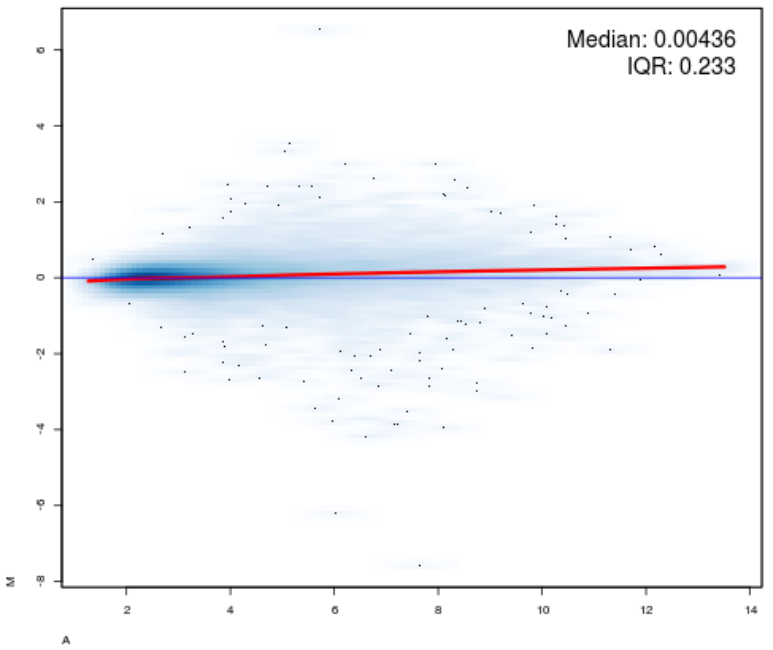
Array1 vs pseudo-median reference chip



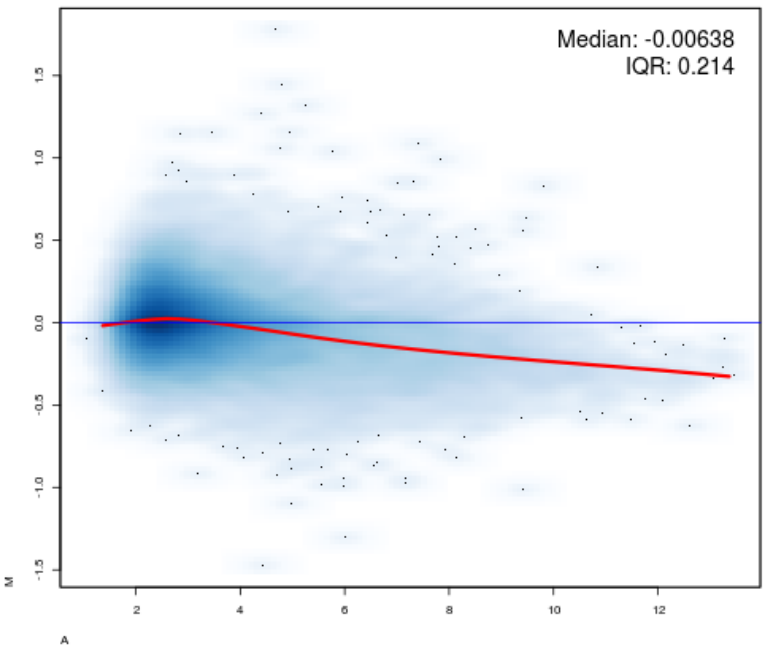
Array4 vs pseudo-median reference chip



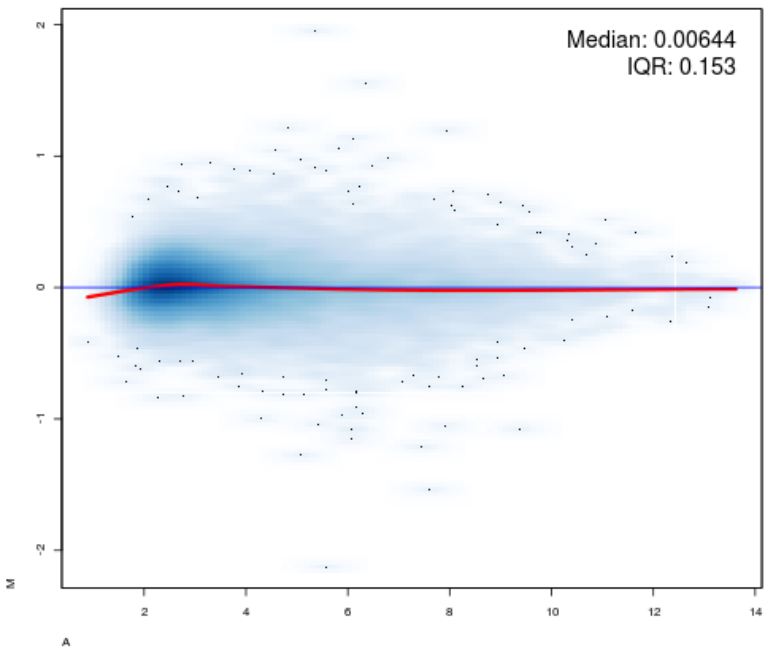
Array12 vs pseudo-median reference chip



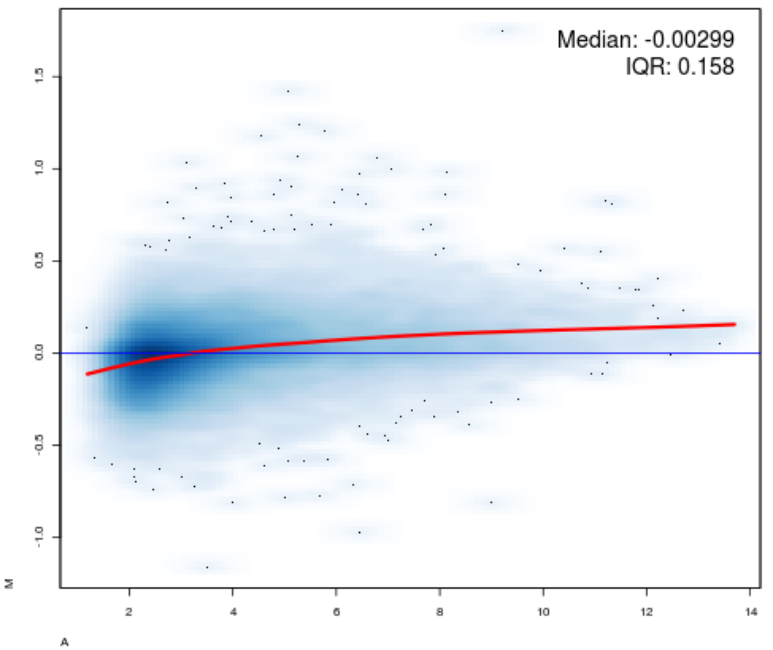
Array2 vs pseudo-median reference chip



Array5 vs pseudo-median reference chip

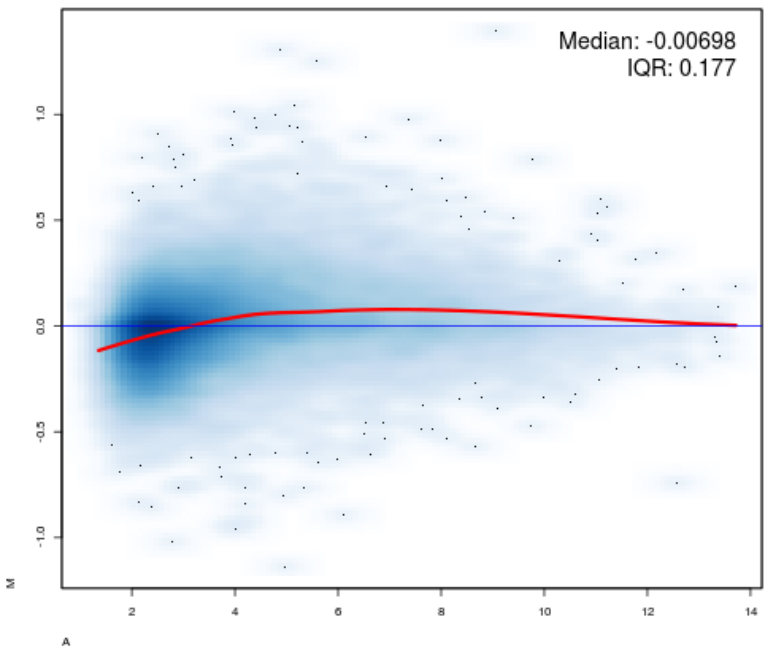


Array6 vs pseudo-median reference chip

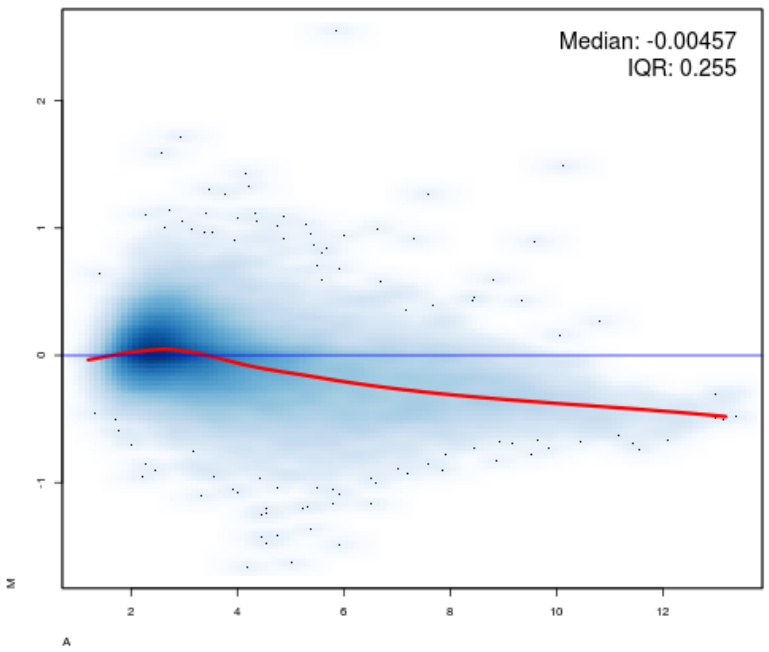


MA plots afterGCRMAnormalization 2 / 2

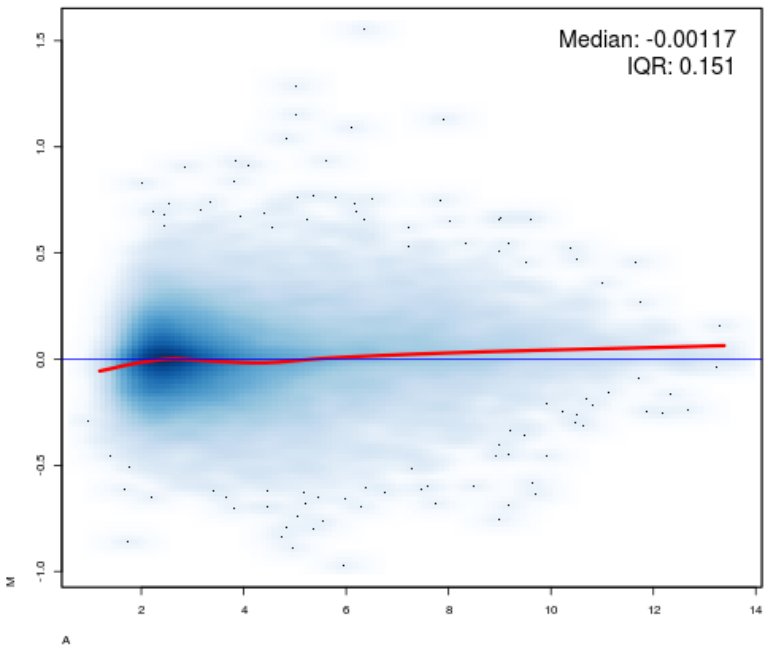
Array3 vs pseudo-median reference chip



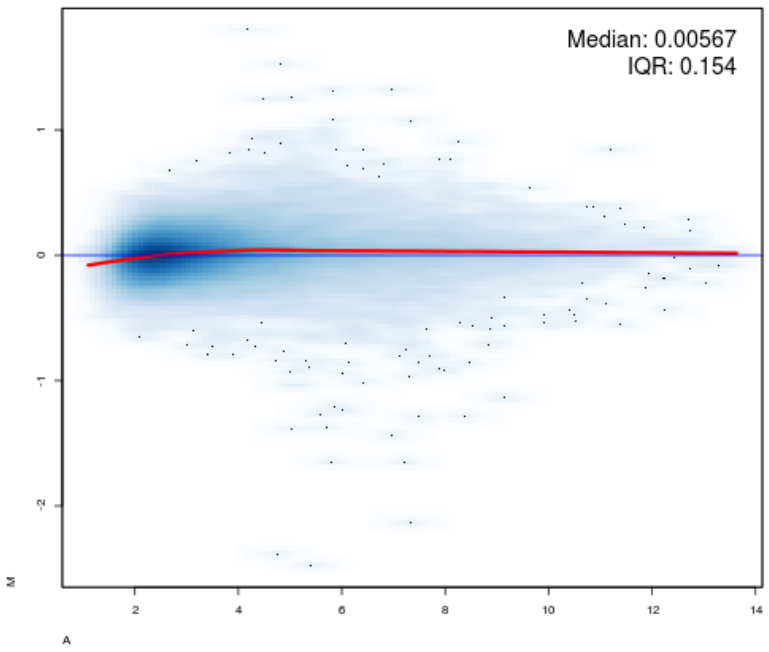
Array7 vs pseudo-median reference chip



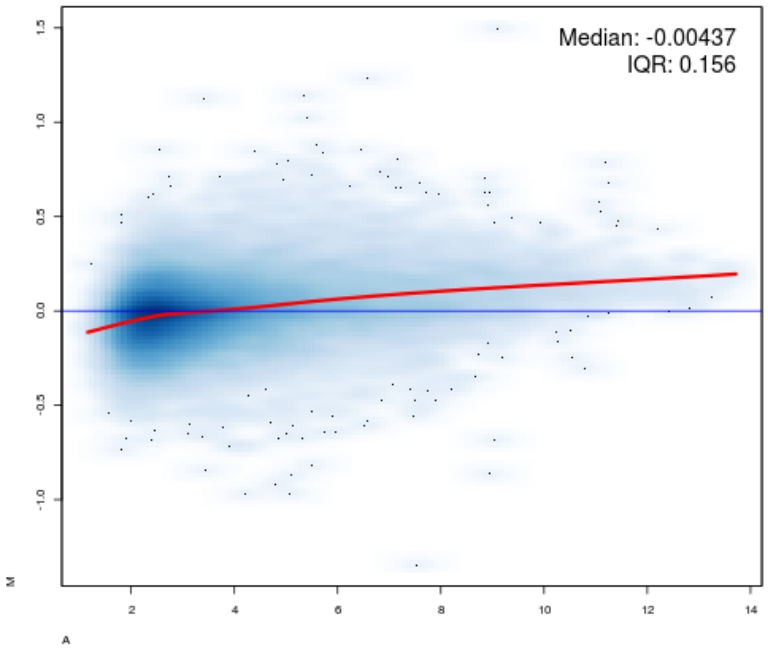
Array11 vs pseudo-median reference chip



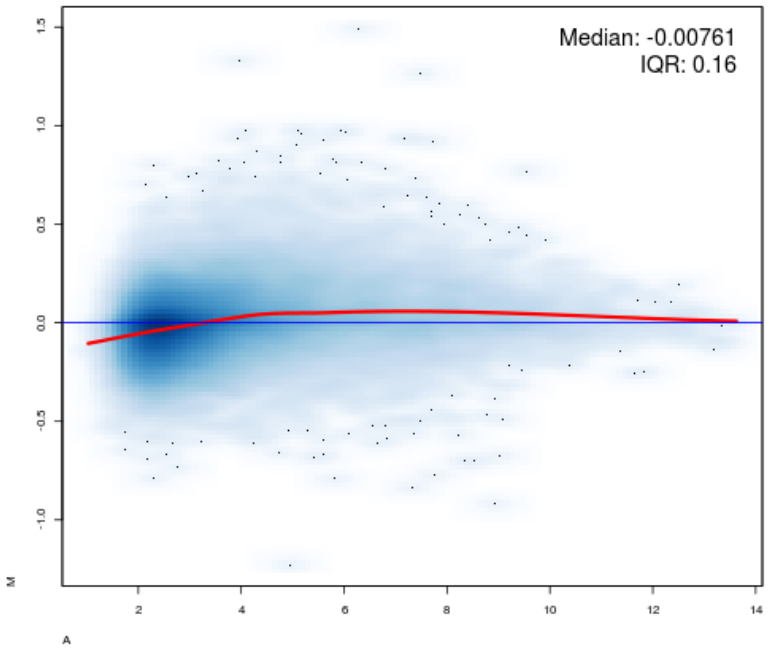
Array8 vs pseudo-median reference chip

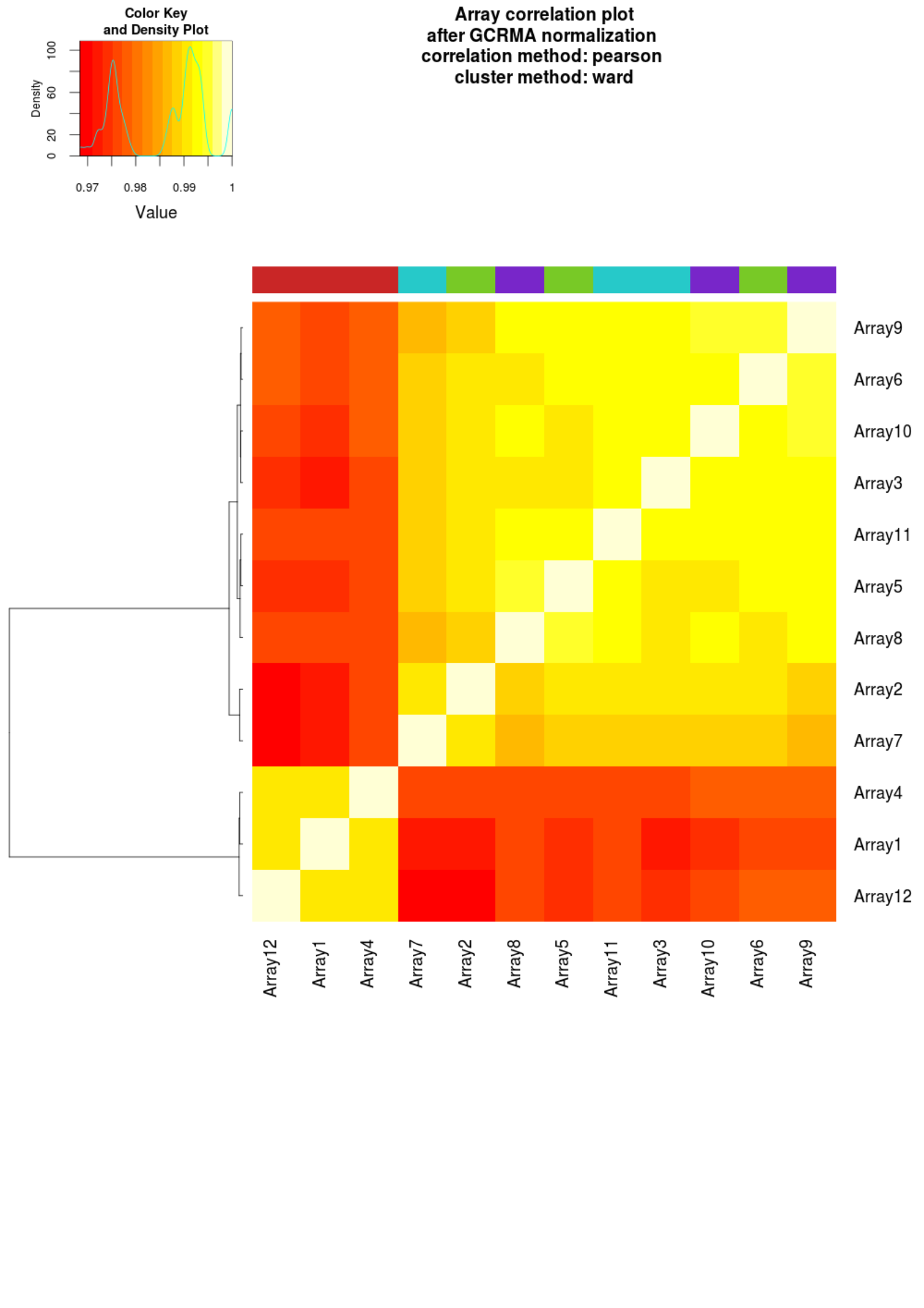


Array9 vs pseudo-median reference chip

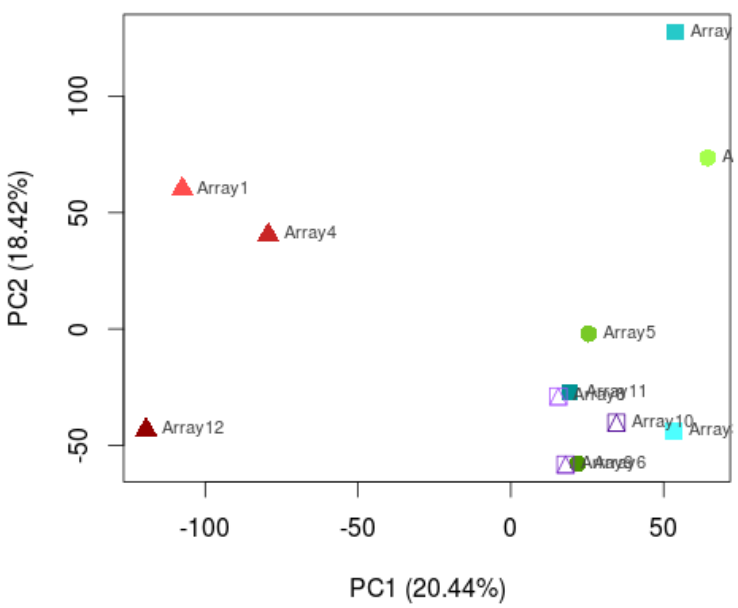


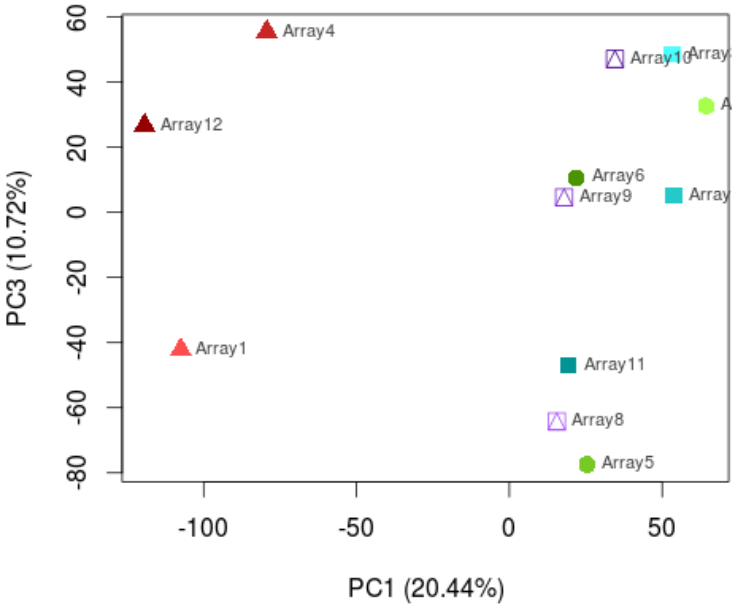
Array10 vs pseudo-median reference chip

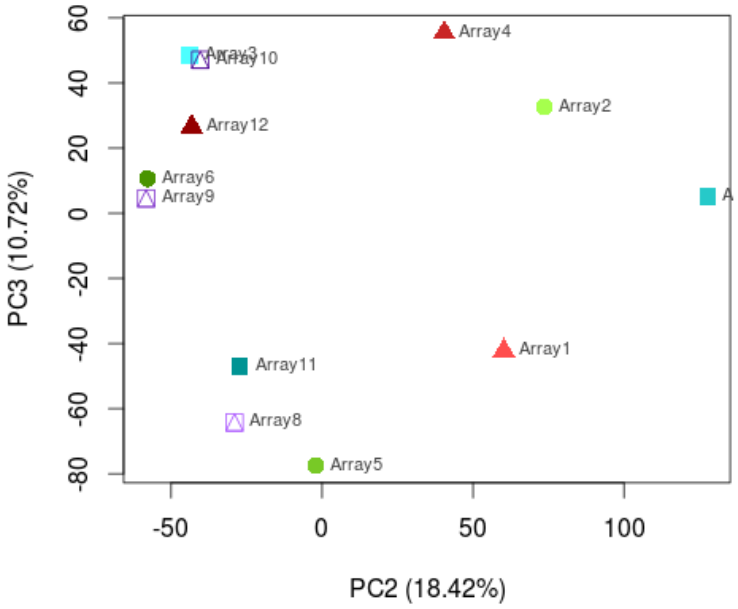


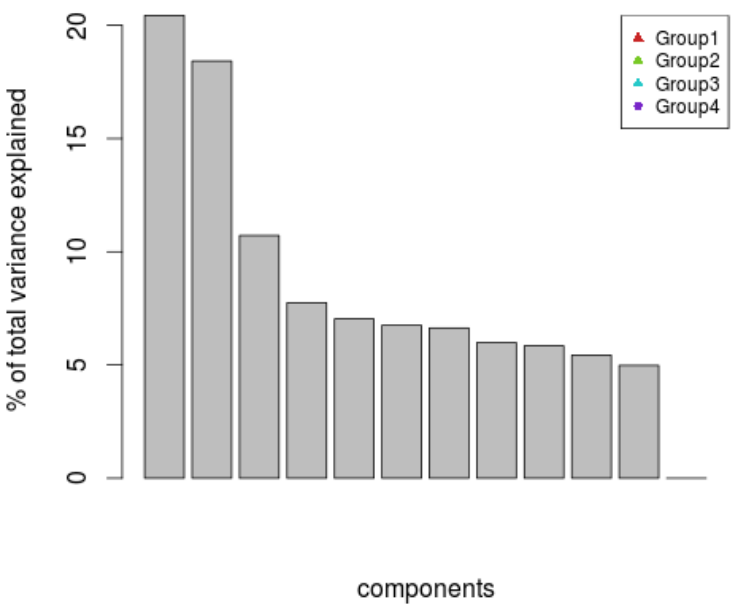

Color Key
and Density Plot
Density
0
20
60
100
0.97
0.98
0.99
1
Value
Array correlation plot
after GCRMA normalization
correlation method: pearson
cluster method: ward
Array9
Array6
Array10
Array3
Array11
Array5
Array8
Array2
Array7
Array4
Array1
Array12
Array12
Array1
Array4
Array7
Array2
Array8
Array5
Array11
Array3
Array10
Array6
Array9

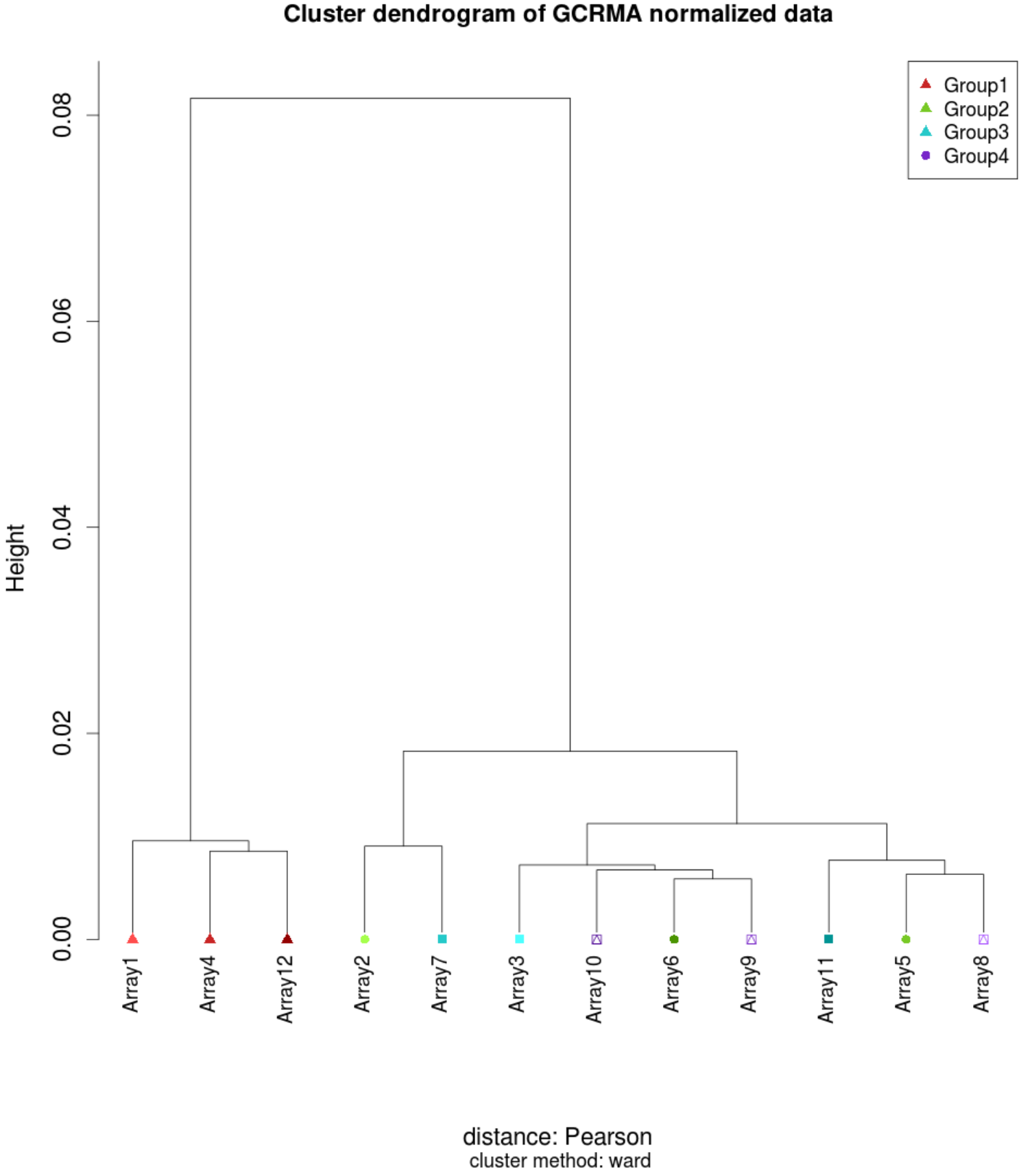
PCA analysis after GCRMA normalization










Cluster dendrogram of GCRMA normalized data
Group1
Group2
Group3
Group4
Height
0.00
0.02
0.04
0.06
0.08
Array1
Array4
Array12
Array2
Array7
Array3
Array10
Array6
Array9
Array11
Array5
Array8
distance: Pearson
cluster method: ward