Quality Control & Pre-processing Evaluation

of

E-MEXP-1956.raw.1_0

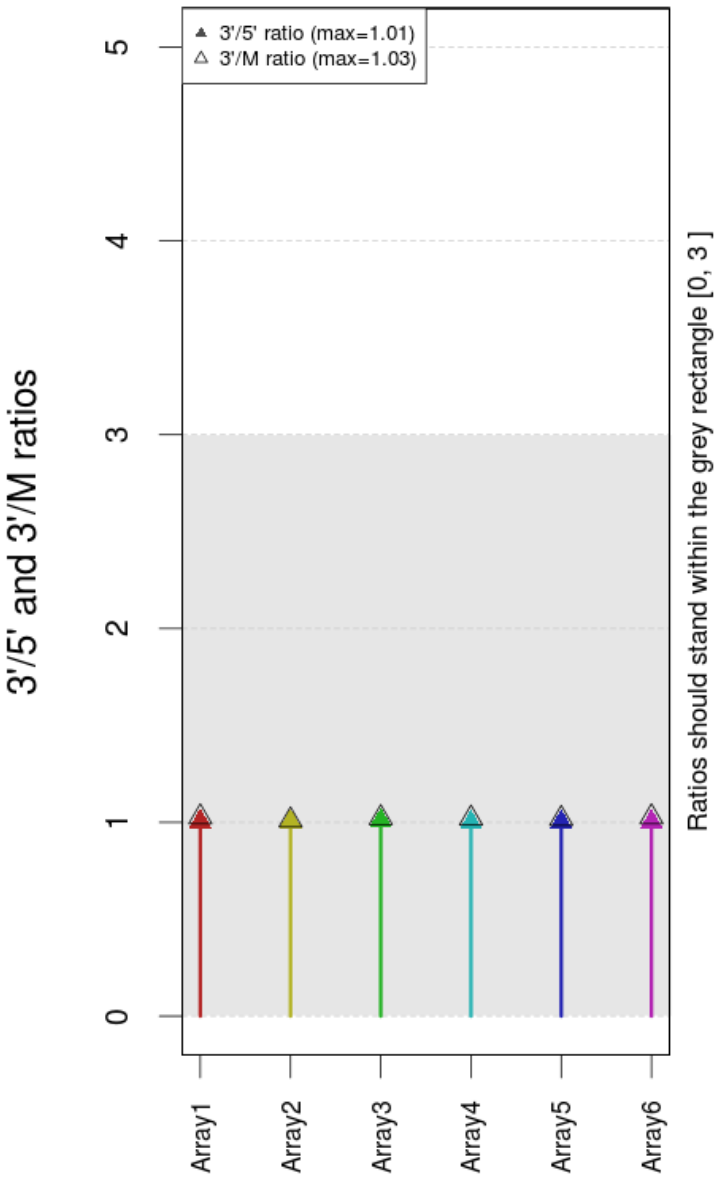
REPORT

Array names and grouping

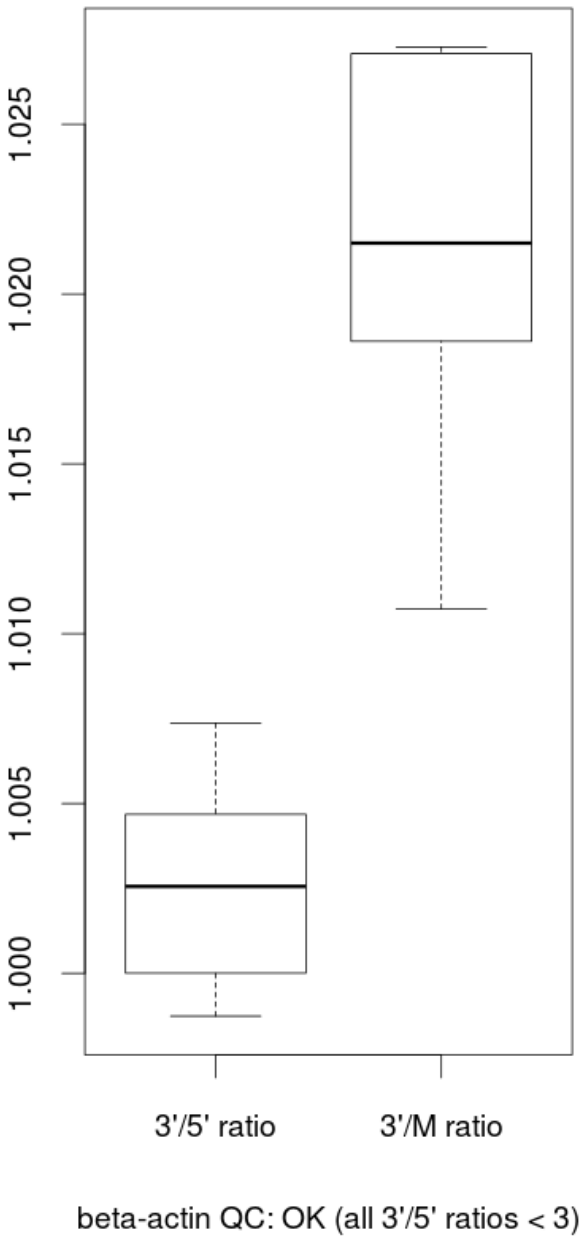
ArrayDataFile	SourceName	FactorValue
08-F2C14.CEL	Array1	Group1
12-F3C18.CEL	Array2	Group1
11-F3C15.CEL	Array3	Group1
07-F2C03.CEL	Array4	Group1
10-F6C04.CEL	Array5	Group1
09-F6C01.CEL	Array6	Group1

Summary of raw data quality indicators

RNA degradation of beta-actin

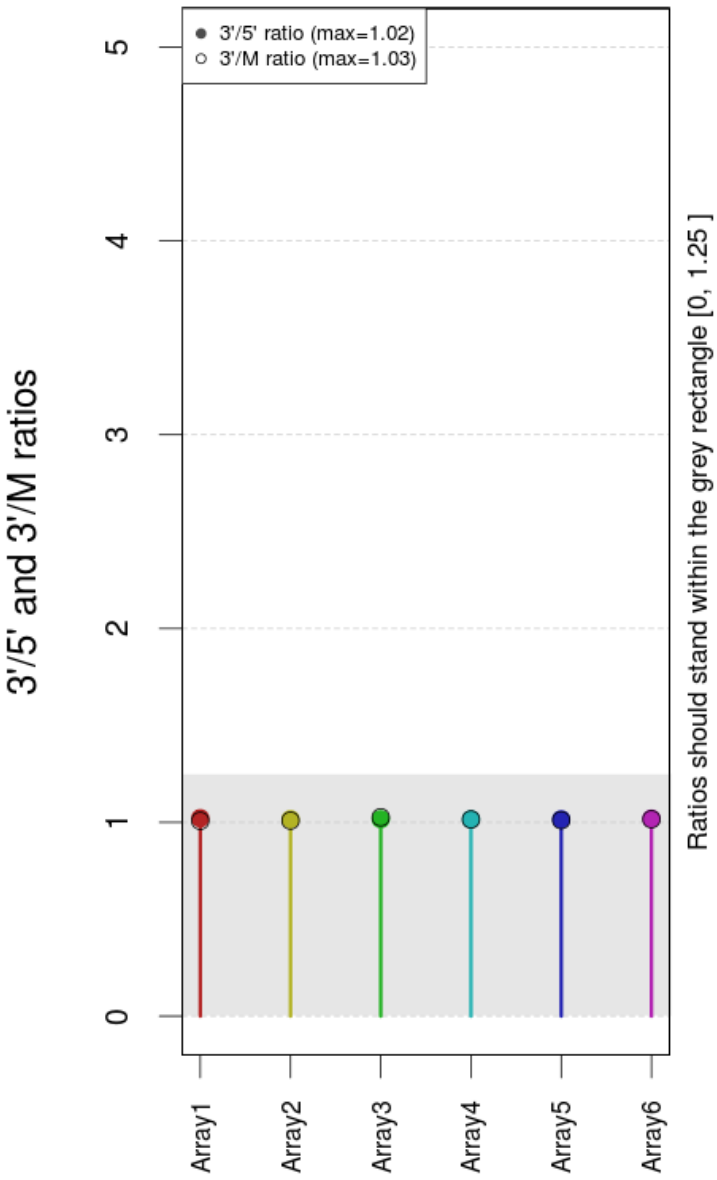


Boxplot of beta-actin ratios

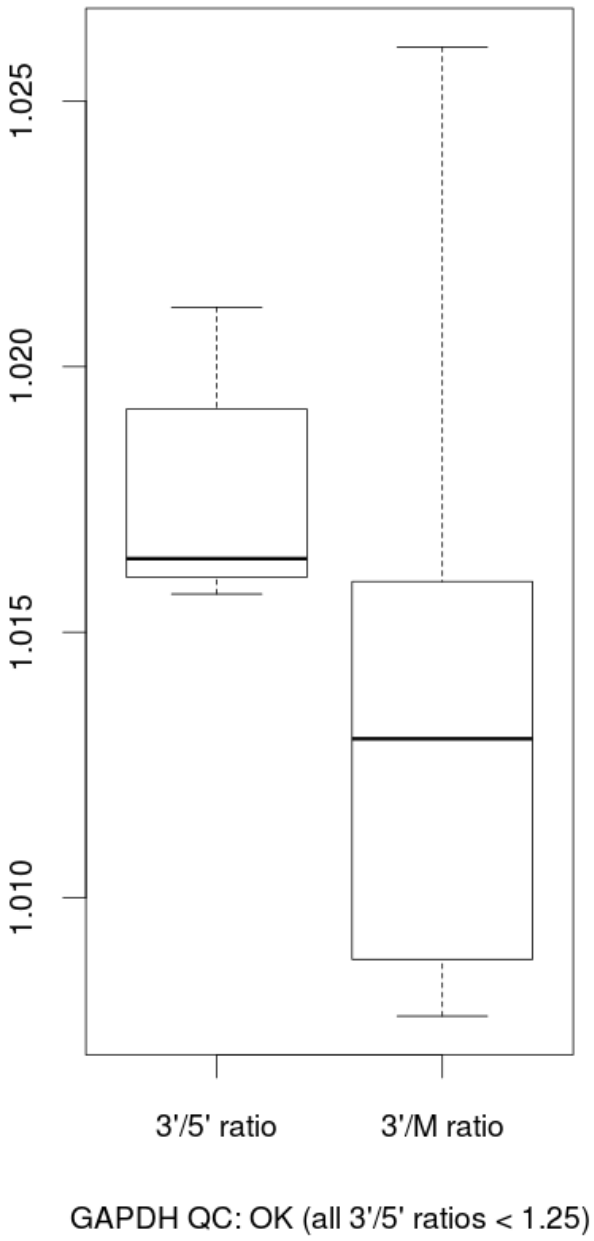


beta-actin QC: OK (all 3'/5' ratios < 3)

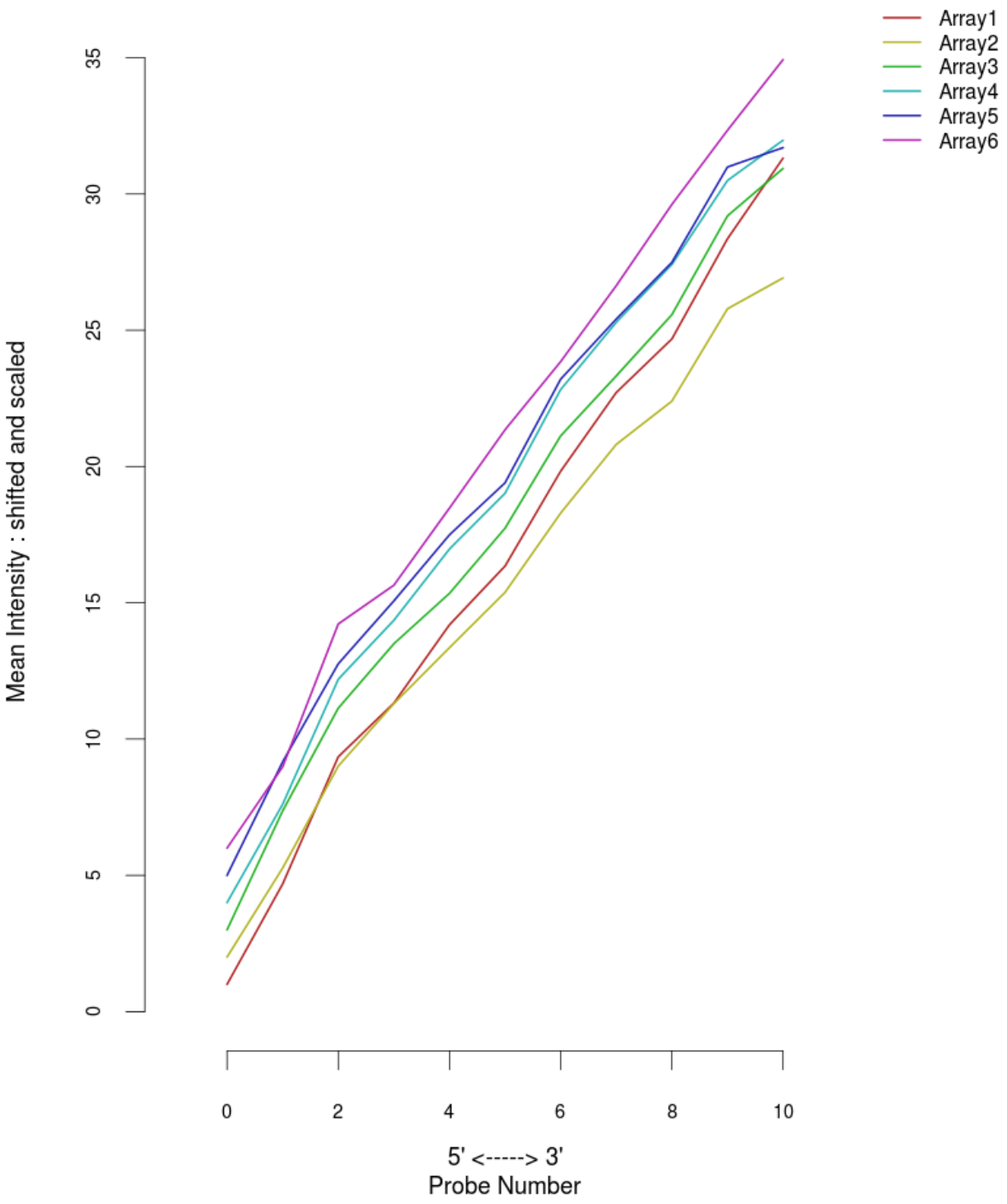
RNA degradation of GAPDH



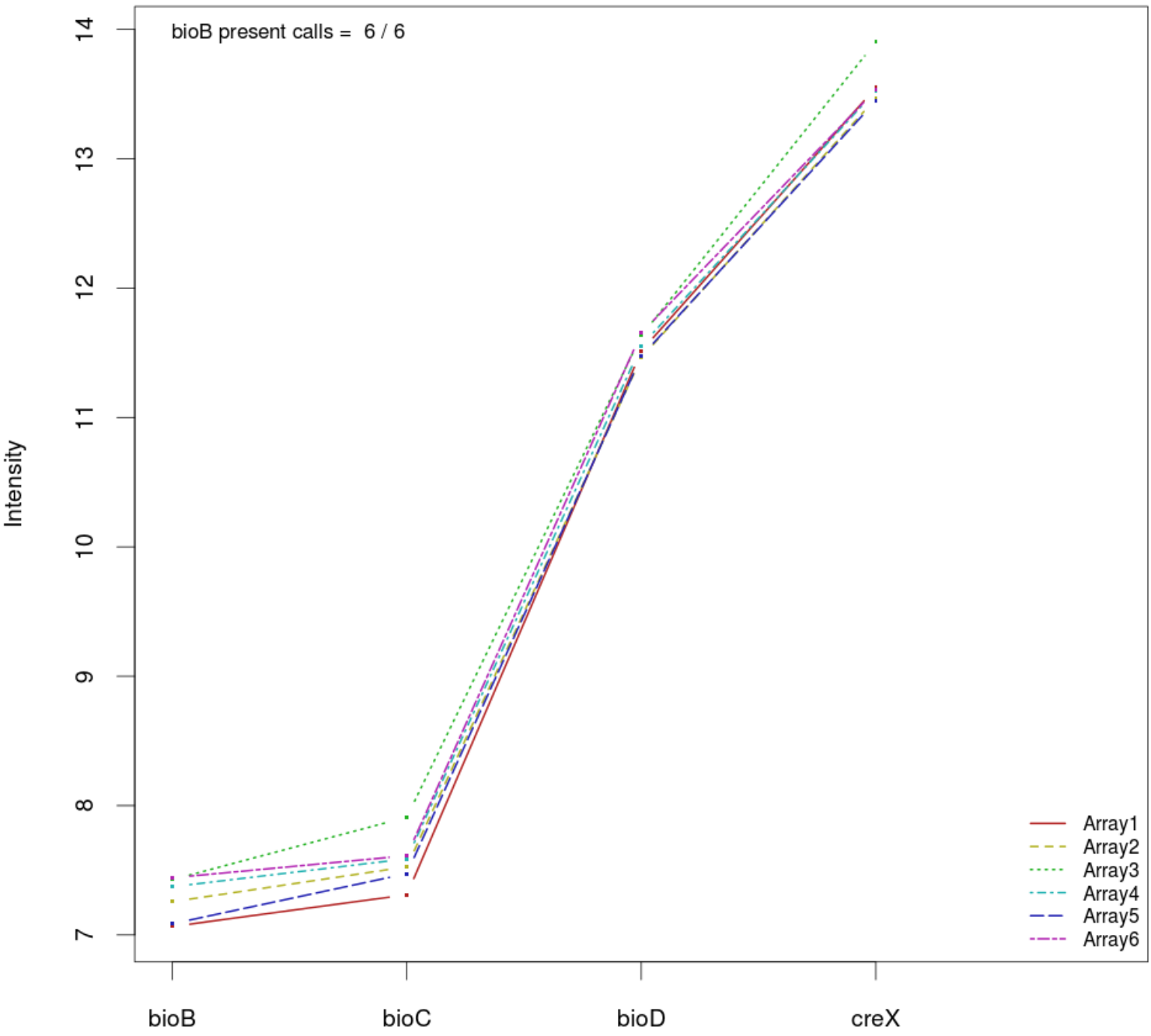
Boxplot of GAPDH ratios



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


RNA degradation plot
Array1
Array2
Array3
Array4
Array5
Array6
Mean Intensity : shifted and scaled
0
5
10
15
20
25
30
35
0
2
4
6
8
10
5' <-----> 3'
Probe Number

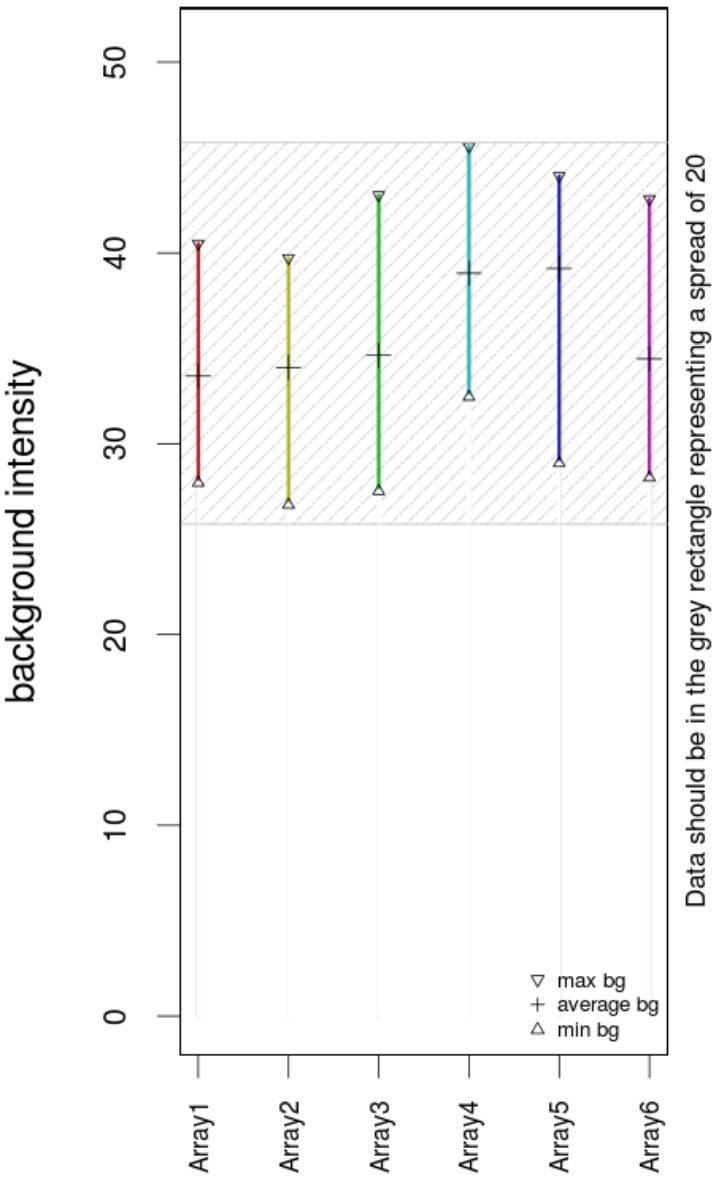
Spike-in Hybridization controls intensities and calls



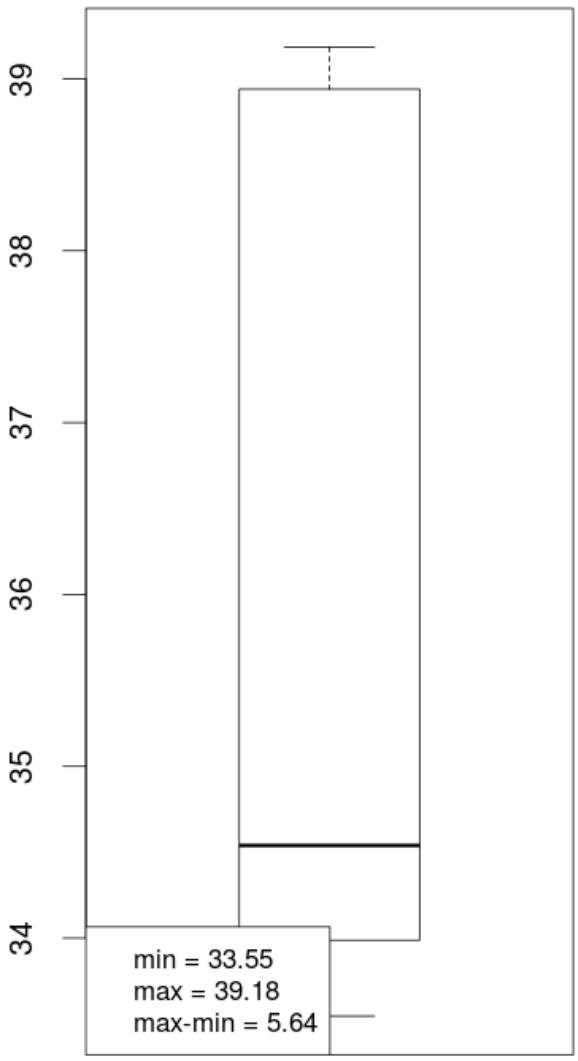
Intensities: OK (bioB < bioC < bioD < creX for all arrays)

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

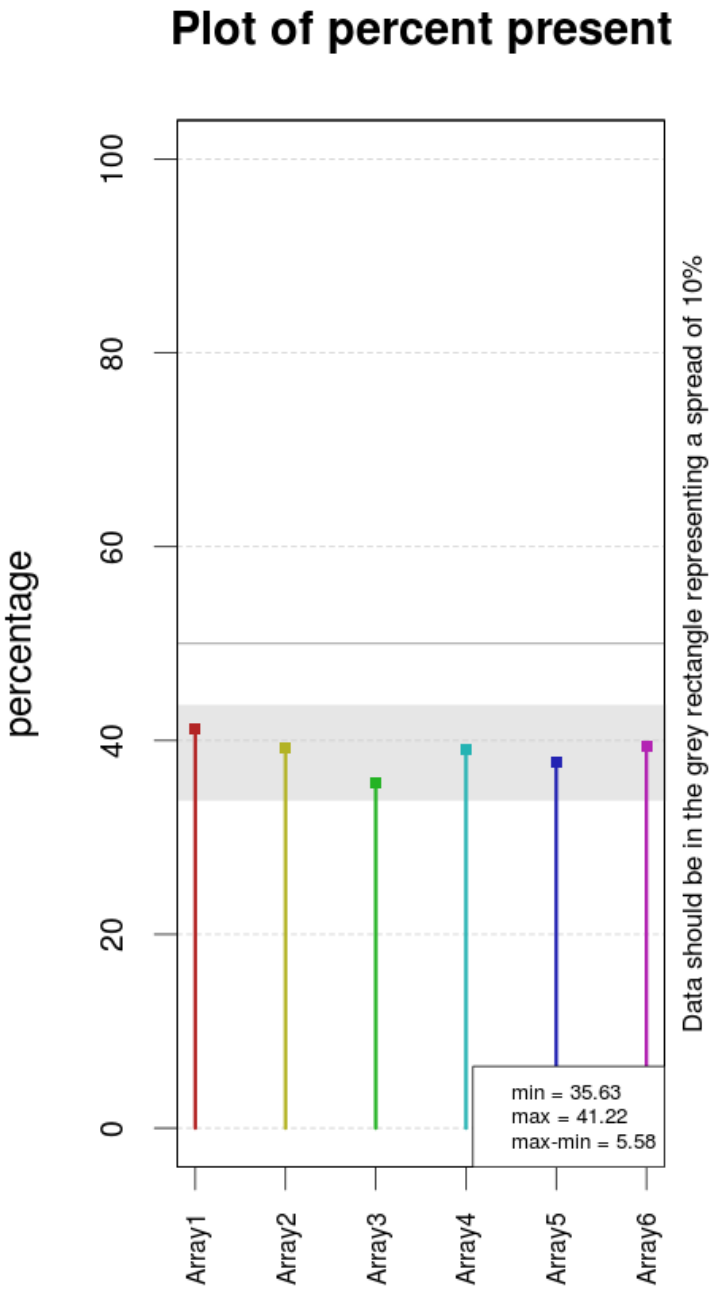
Plot of background intensity

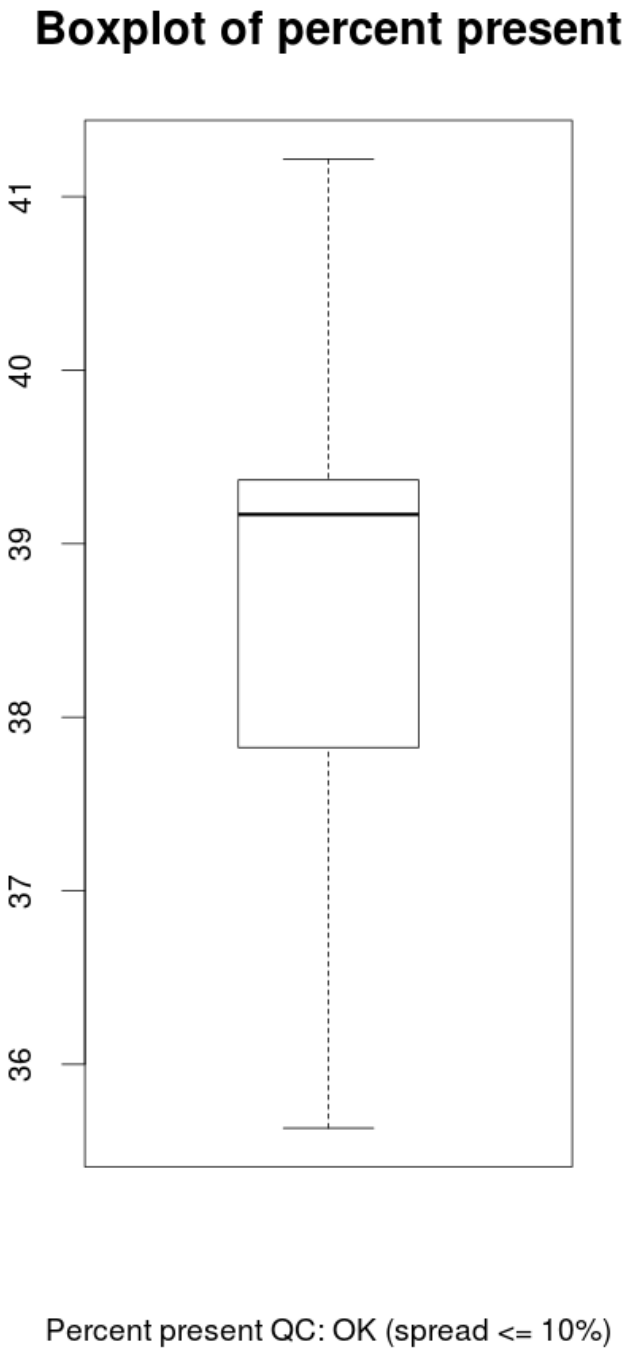


Average background intensity

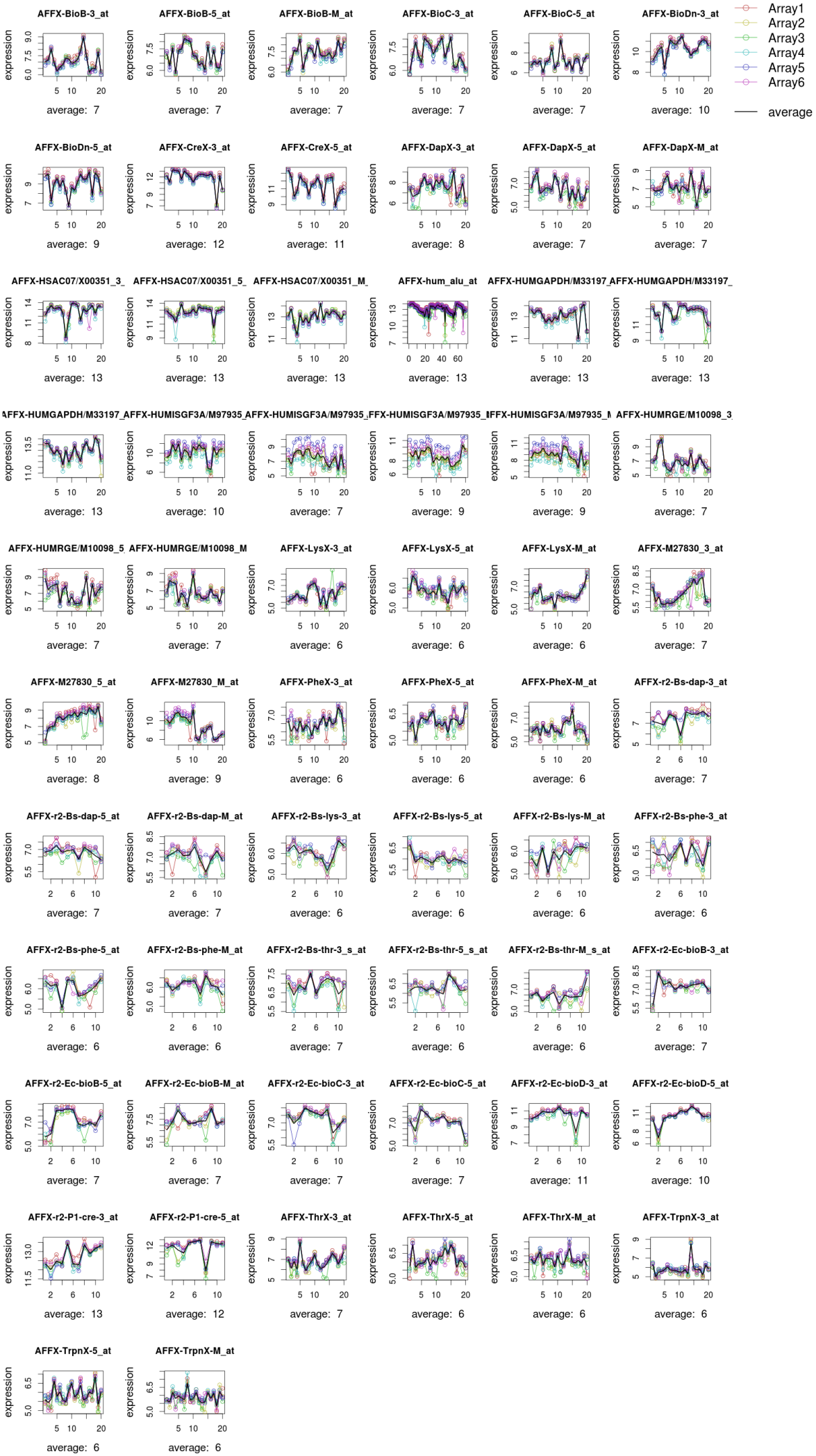


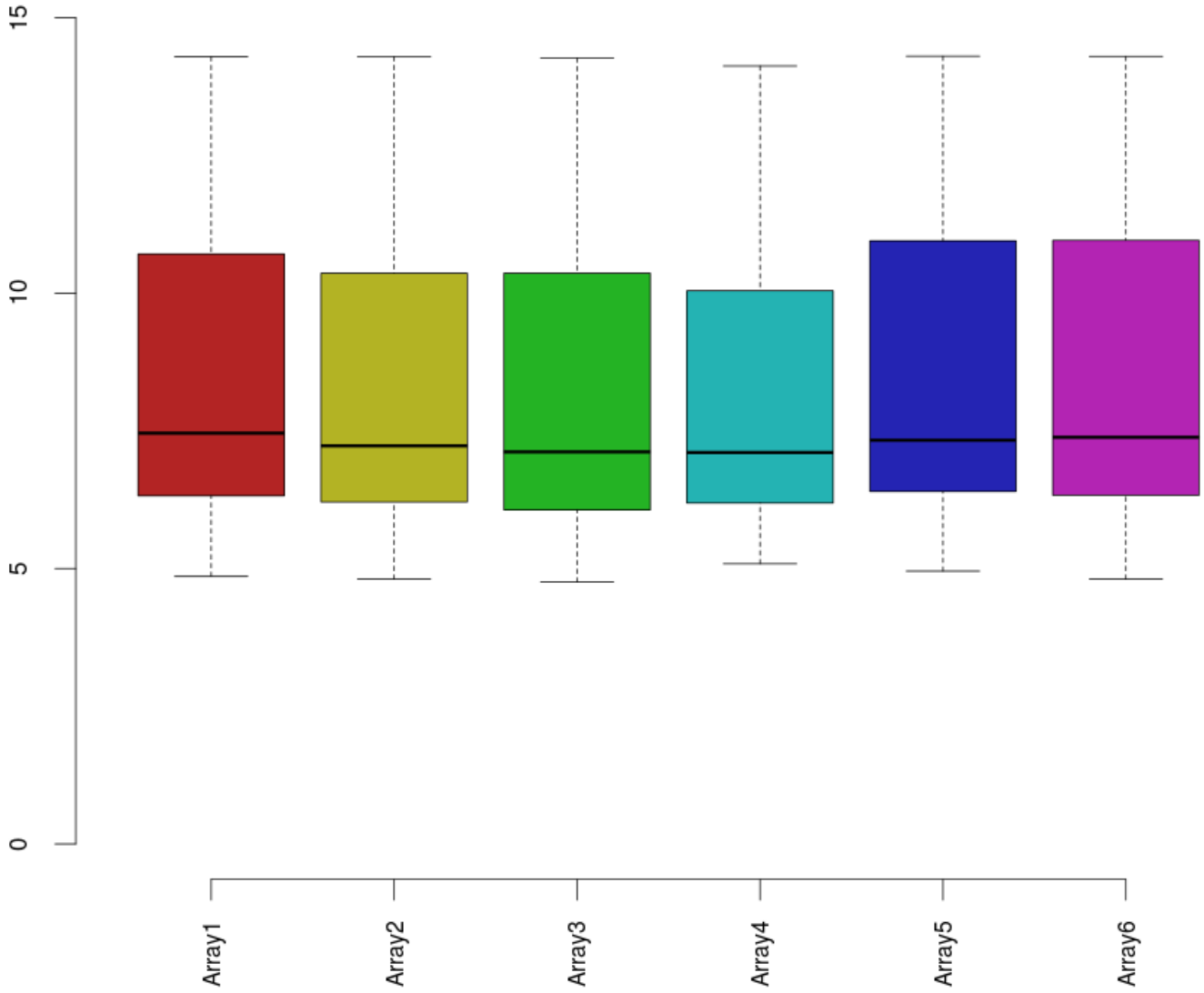
Background QC: OK (spread <= 20)

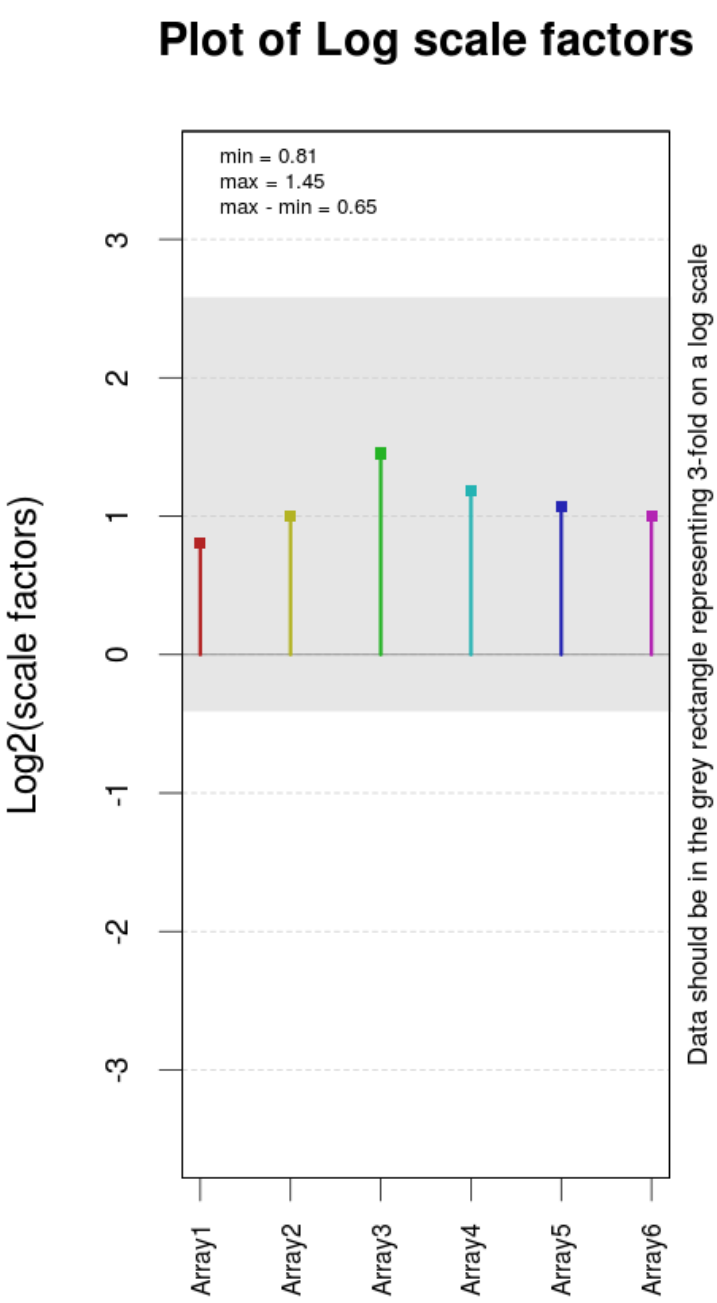

Plot of percent present
percentage
0
20
40
60
80
100
Array1
Array2
Array3
Array4
Array5
Array6
min = 35.63
max = 41.22
max-min = 5.58
Data should be in the grey rectangle representing a spread of 10%

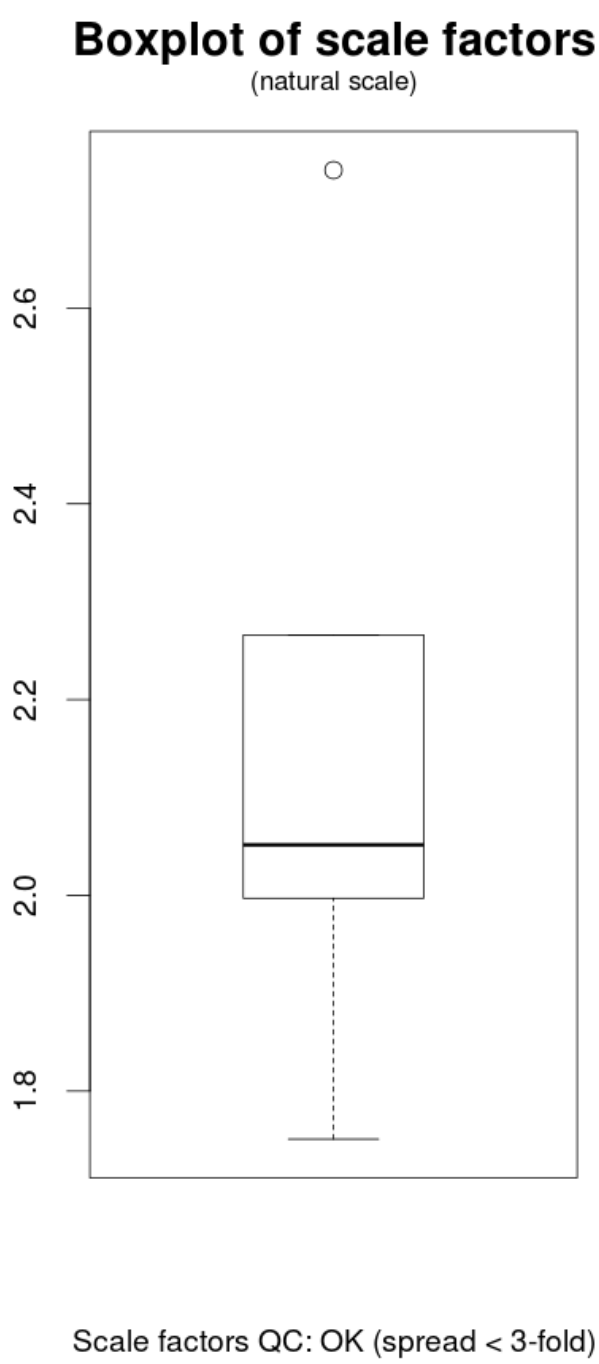

Boxplot of percent present
36
37
38
39
40
41
Percent present QC: OK (spread <= 10%)

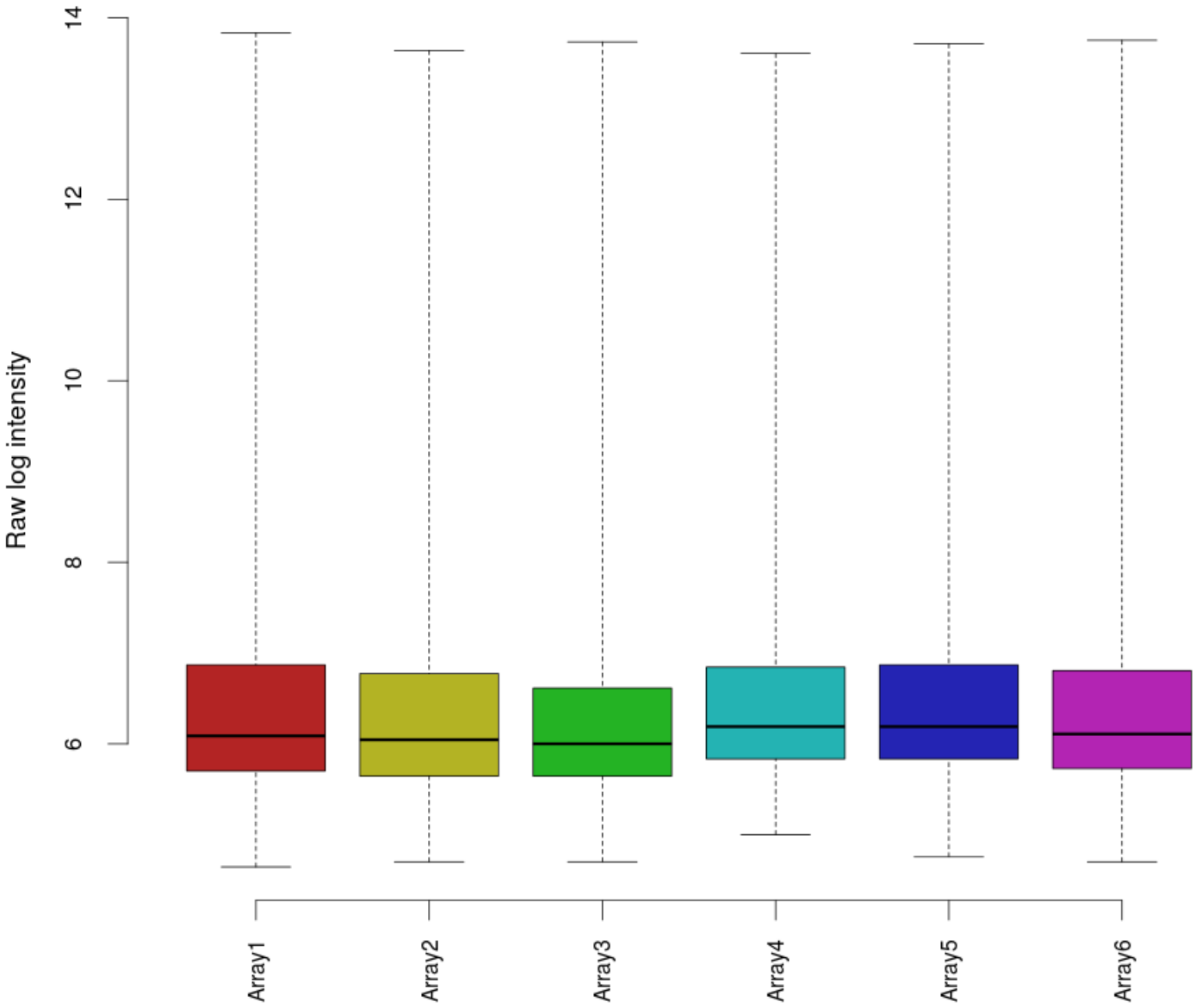
affx control profiles

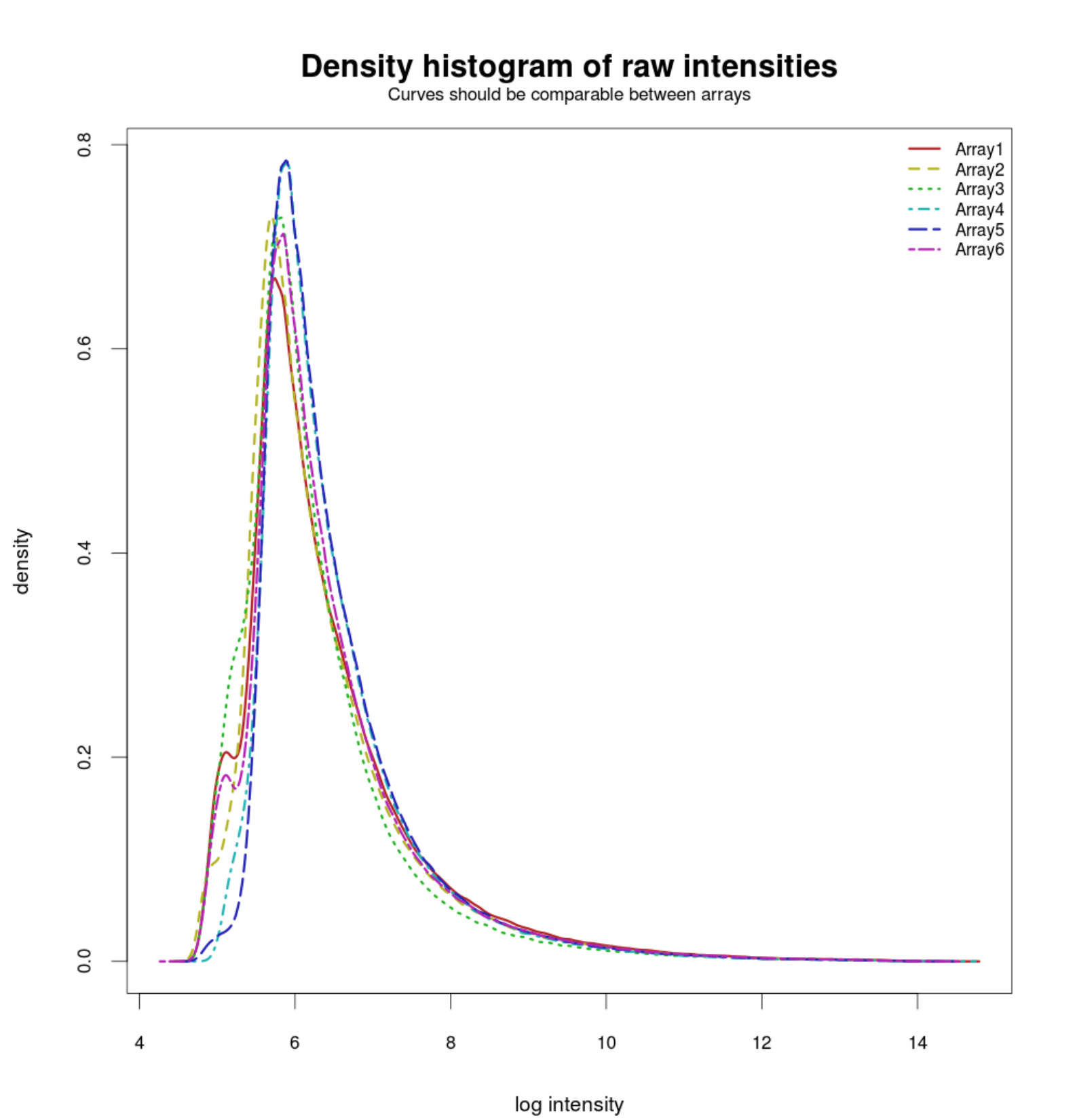



affx controls
15
10
5
0
Array1
Array2
Array3
Array4
Array5
Array6


Plot of Log scale factors
min = 0.81
max = 1.45
max - min = 0.65
Log2(scale factors)
3
2
1
0
-1
-2
-3
Array1
Array2
Array3
Array4
Array5
Array6
Data should be in the grey rectangle representing 3-fold on a log scale

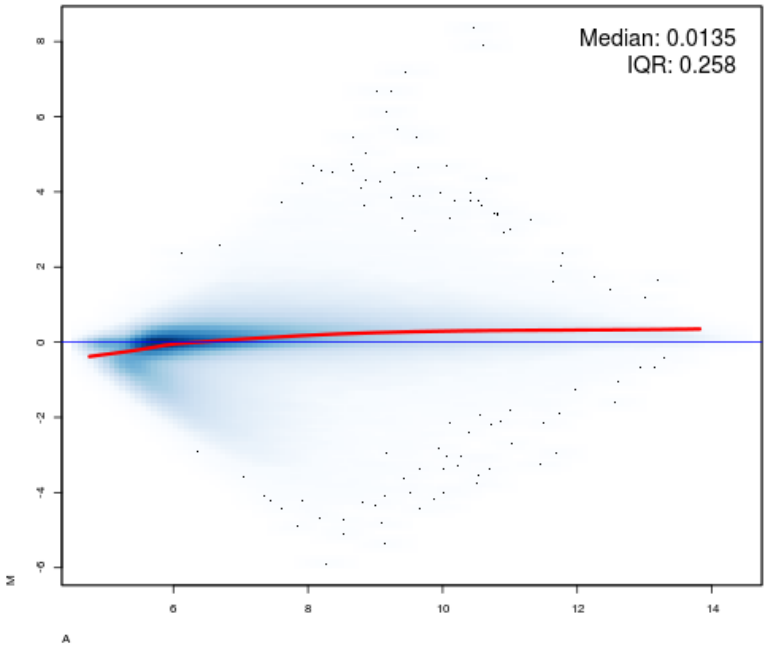

Boxplot of scale factors
(natural scale)
2.6
2.4
2.2
2.0
1.8
Scale factors QC: OK (spread < 3-fold)


Boxplot of raw intensities
Distributions should be comparable between arrays
Raw log intensity
14
12
10
8
6
Array1
Array2
Array3
Array4
Array5
Array6

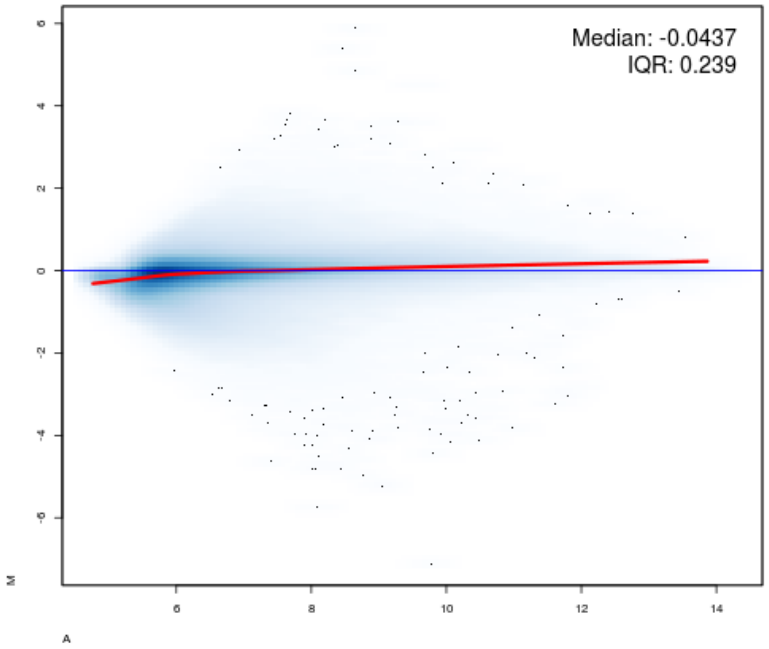

Density histogram of raw intensities
Curves should be comparable between arrays
Array1
Array2
Array3
Array4
Array5
Array6
density
0.0
0.2
0.4
0.6
0.8
4
6
8
10
12
14
log intensity

MA plots of raw data

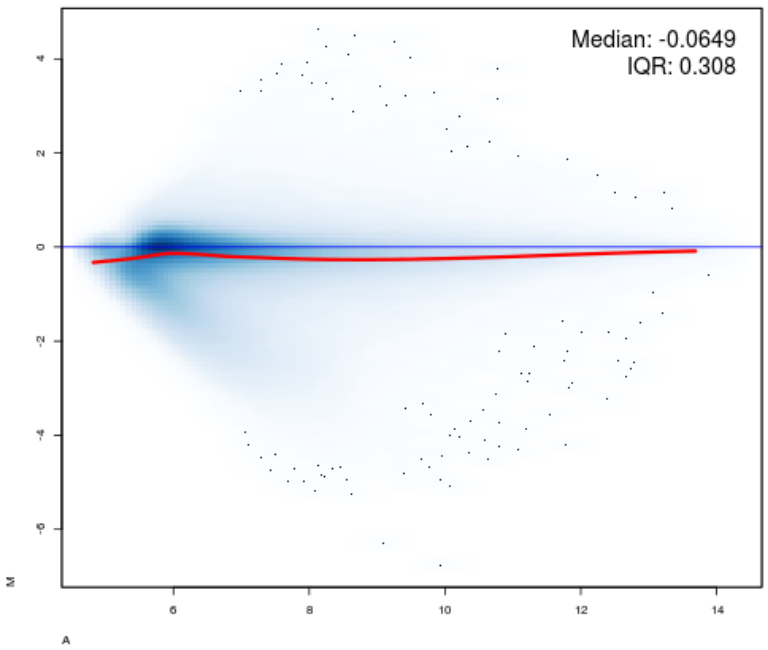
Array1 vs pseudo-median reference chip



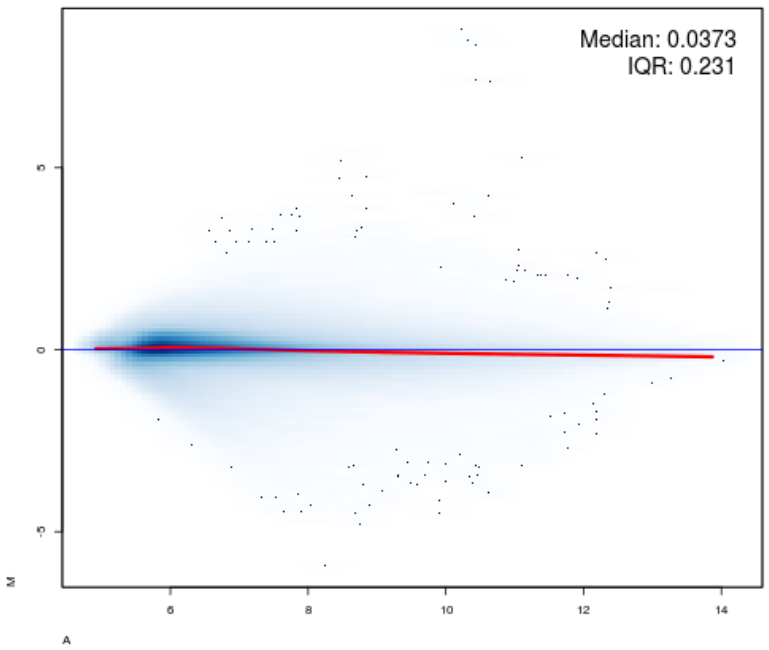
Array2 vs pseudo-median reference chip



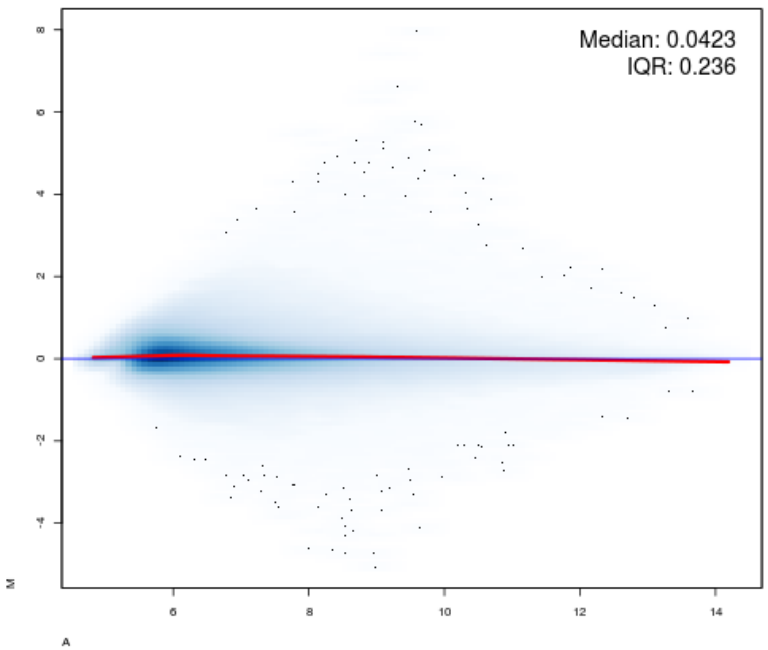
Array3 vs pseudo-median reference chip



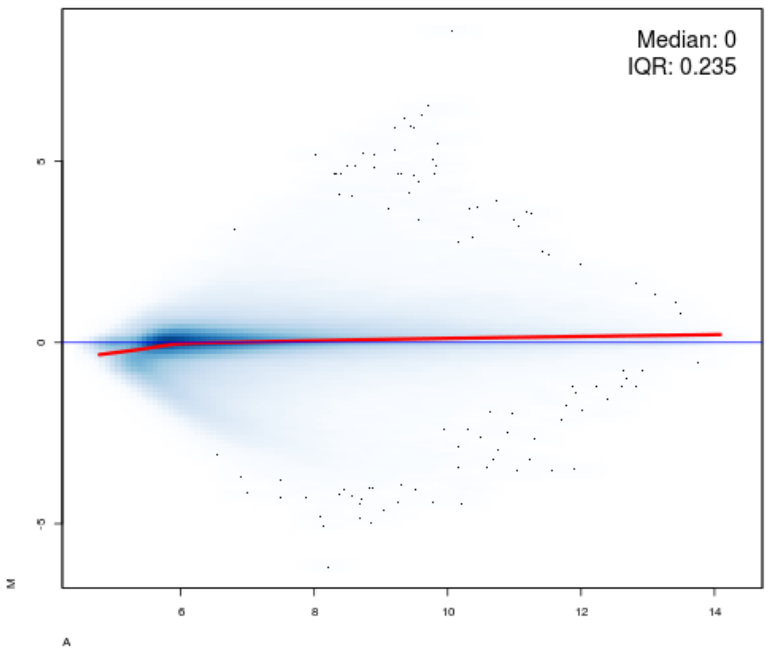
Array4 vs pseudo-median reference chip



Array5 vs pseudo-median reference chip

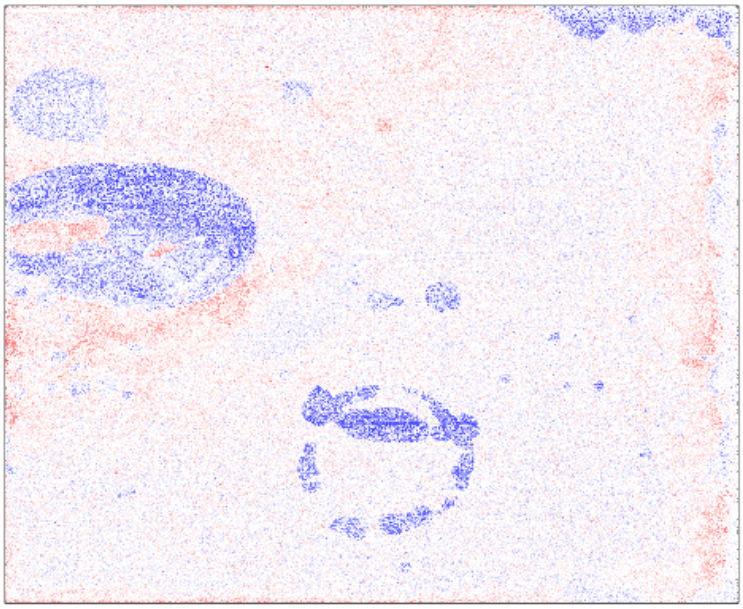


Array6 vs pseudo-median reference chip

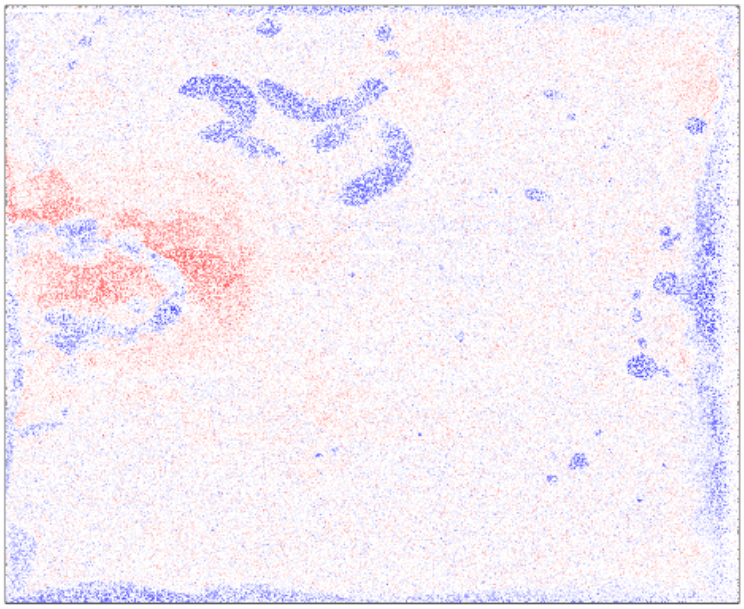


2D virtual PLM image for model characteristic: resids

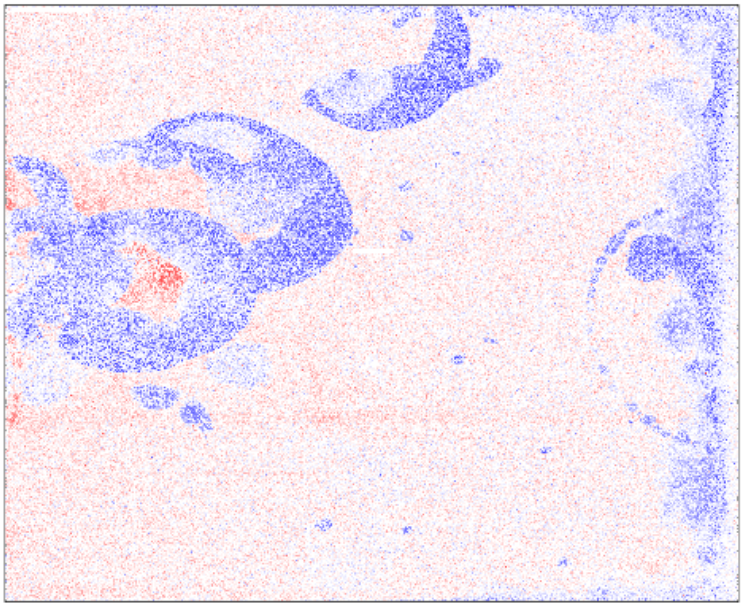
Array1



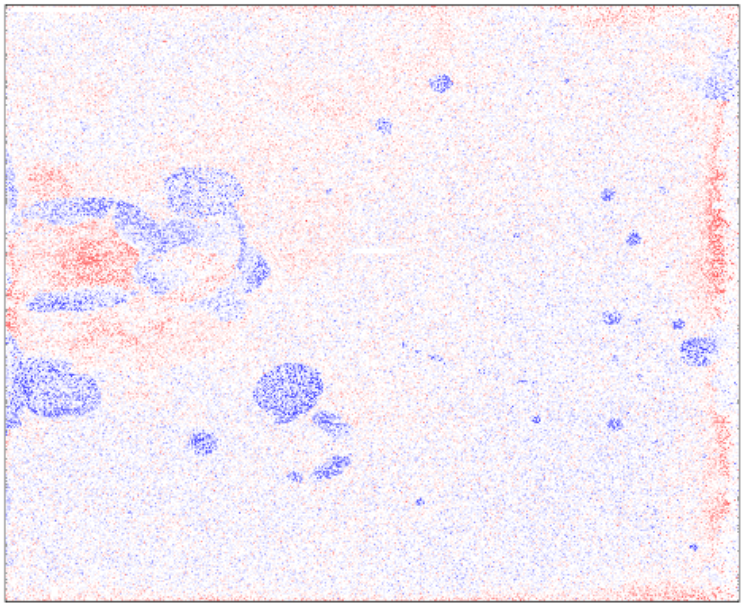
Array2



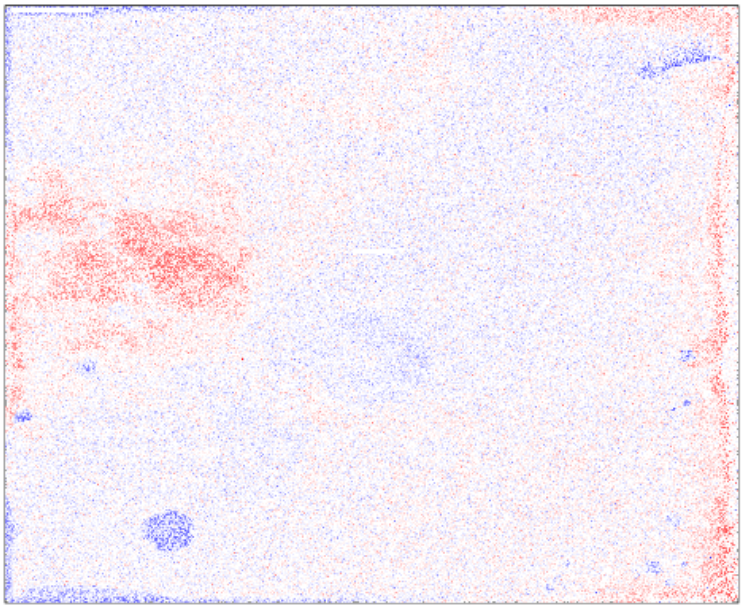
Array3



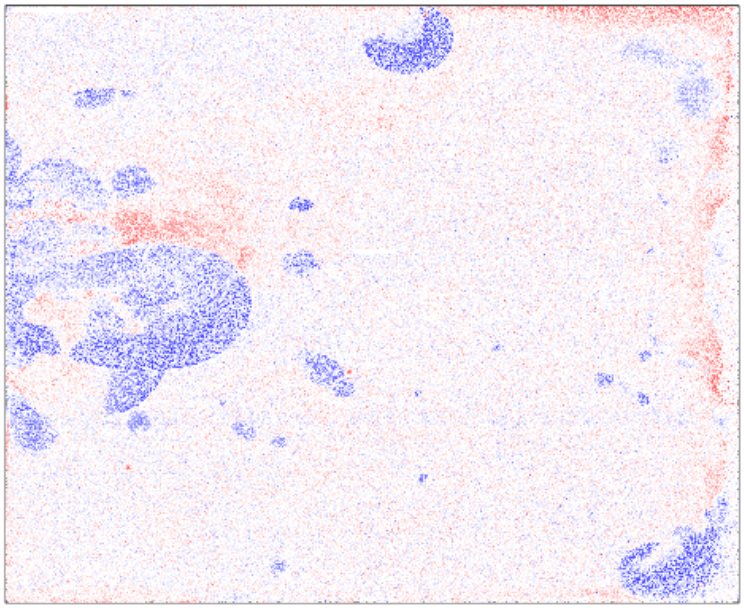
Array4

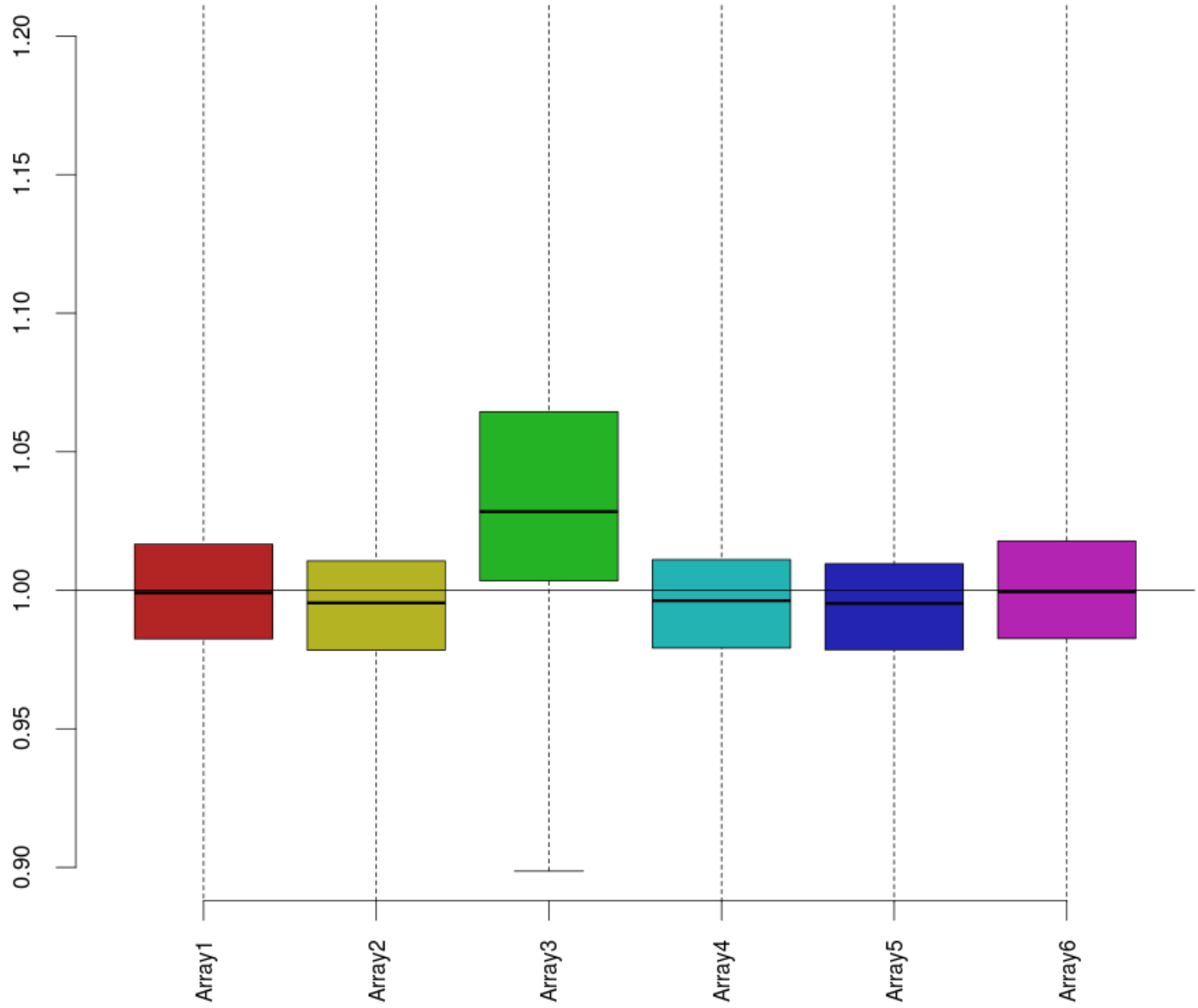


Array5



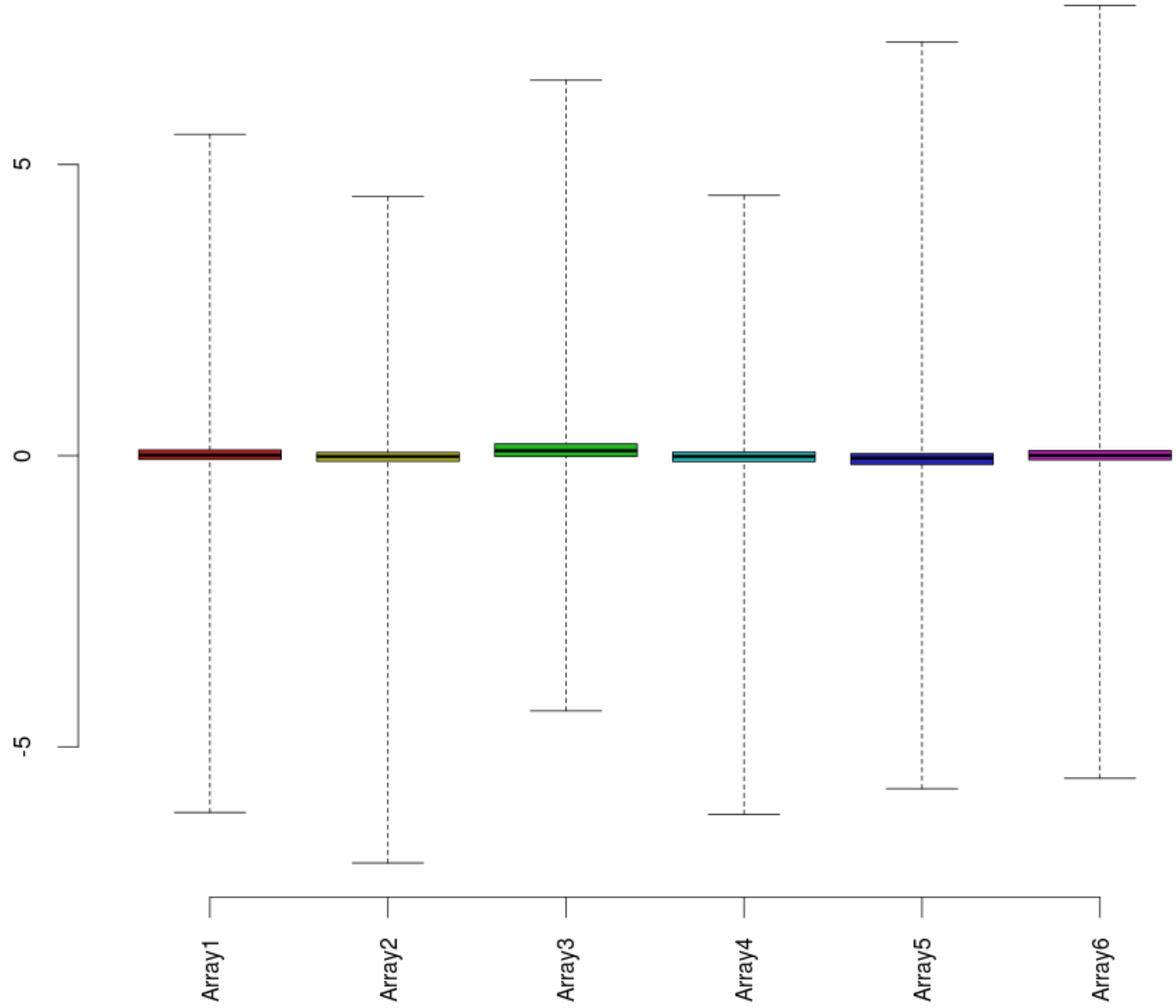
Array6

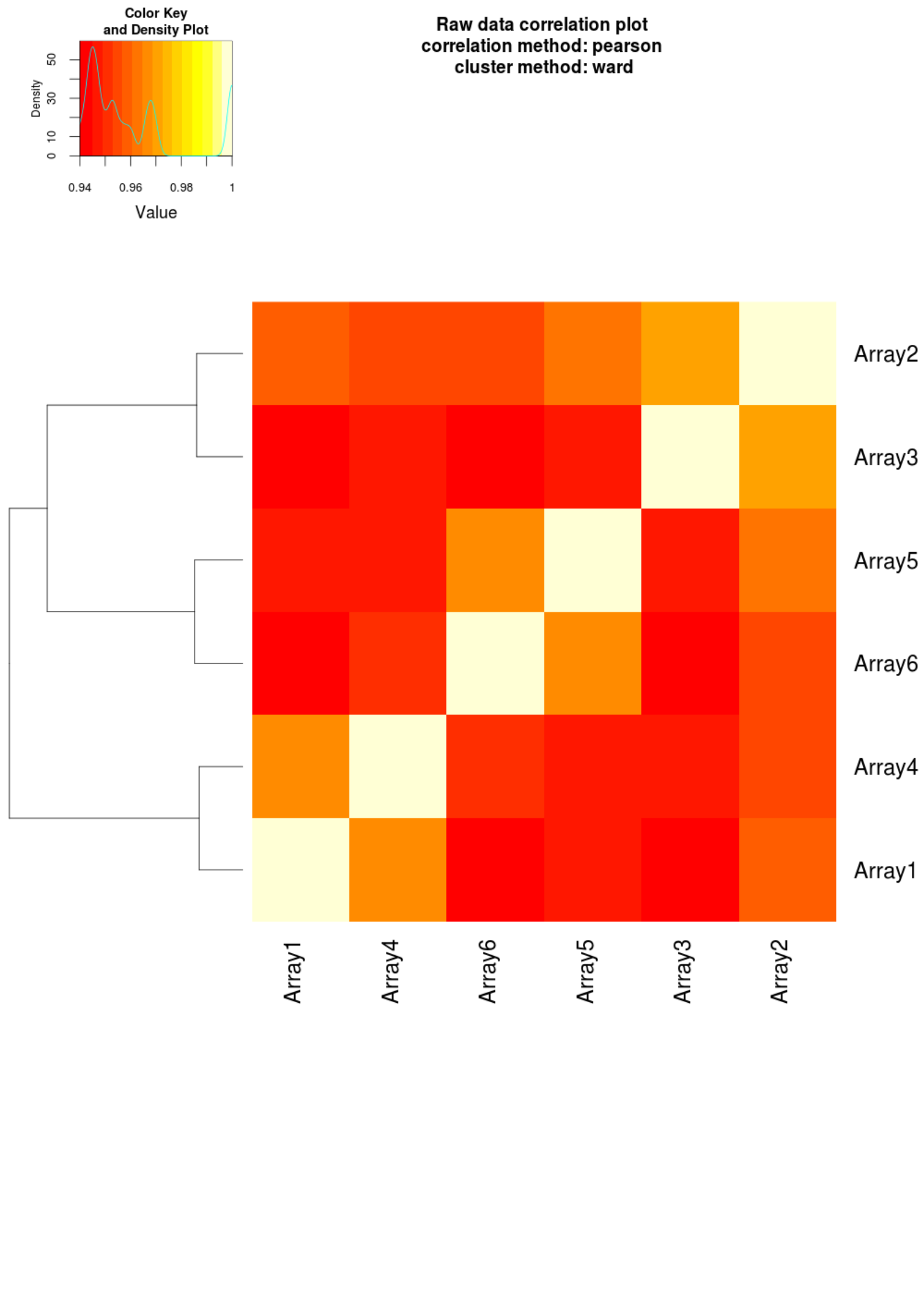



Normalized Unscaled Standard Errors (NUSE)
NUSE median value should be < 1.1
1.20
1.15
1.10
1.05
1.00
0.95
0.90
Array1
Array2
Array3
Array4
Array5
Array6

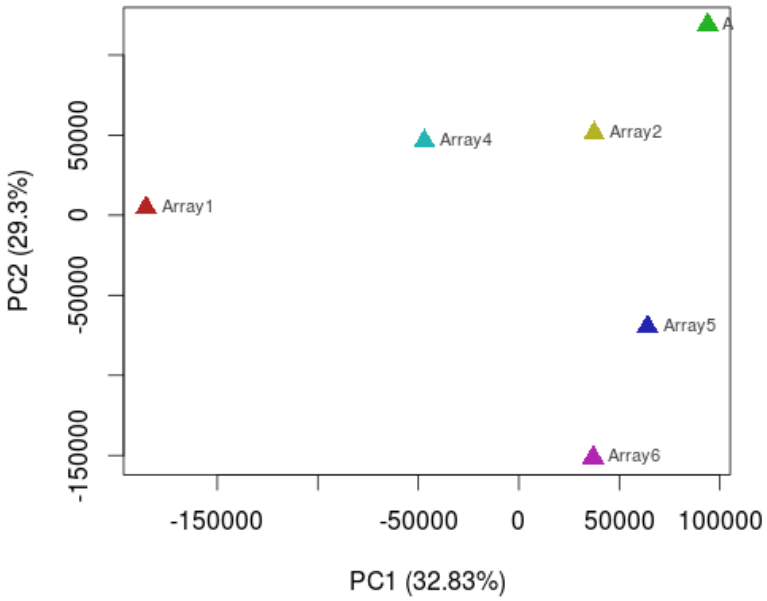
Relative Log Expression (RLE)

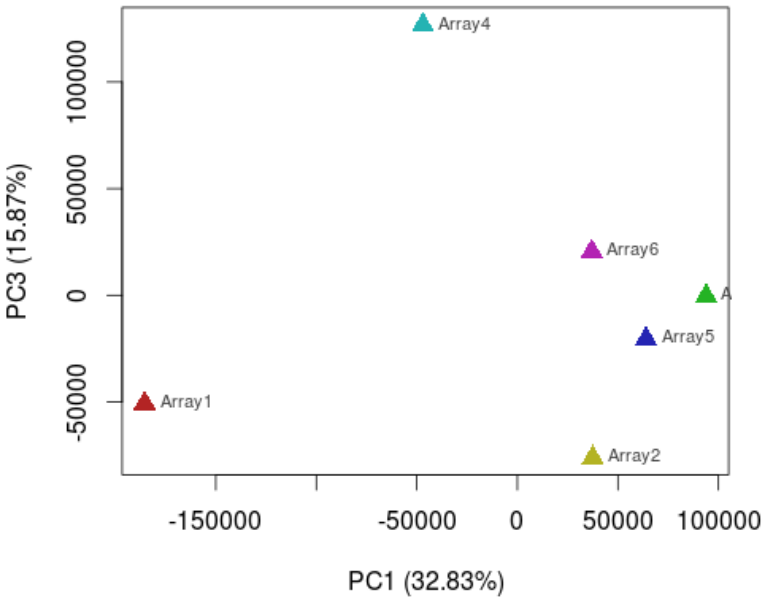
RLE distributions should be centered around 0

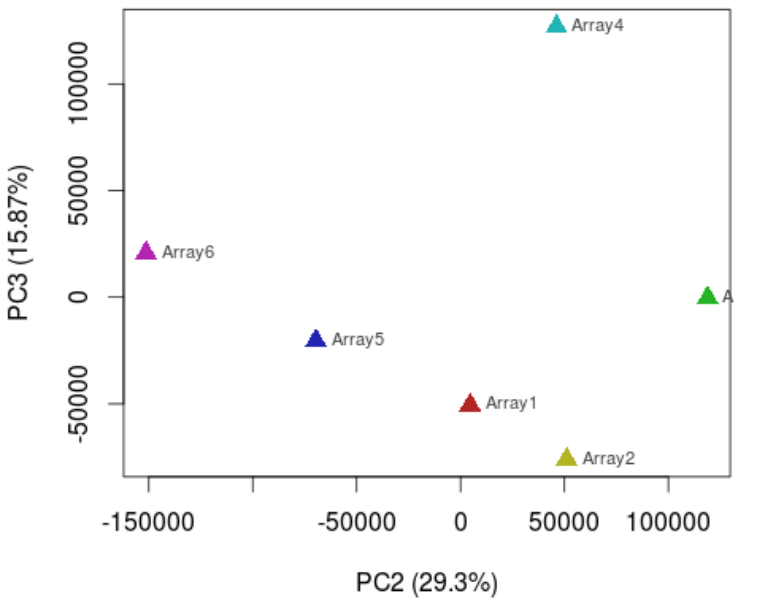


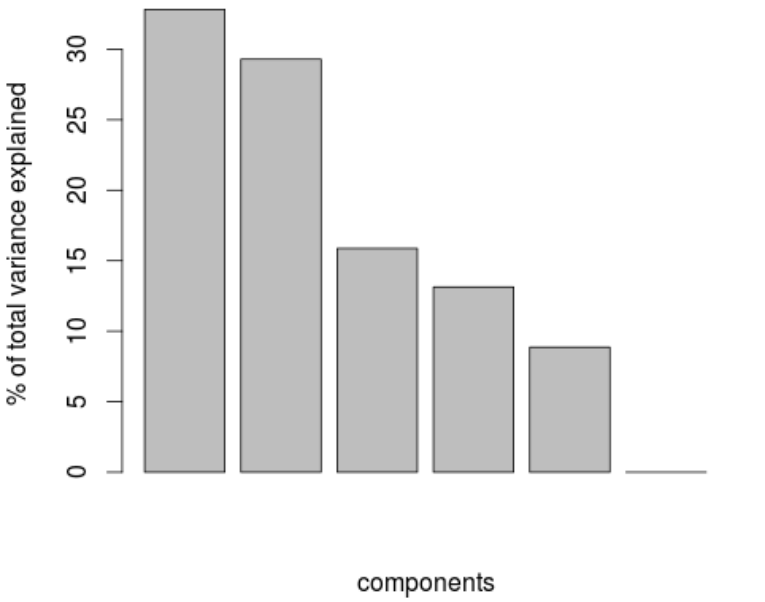

Color Key
and Density Plot
Density
0
10
30
50
0.94
0.96
0.98
1
Value
Raw data correlation plot
correlation method: pearson
cluster method: ward
Array2
Array3
Array5
Array6
Array4
Array1
Array1
Array4
Array6
Array5
Array3
Array2

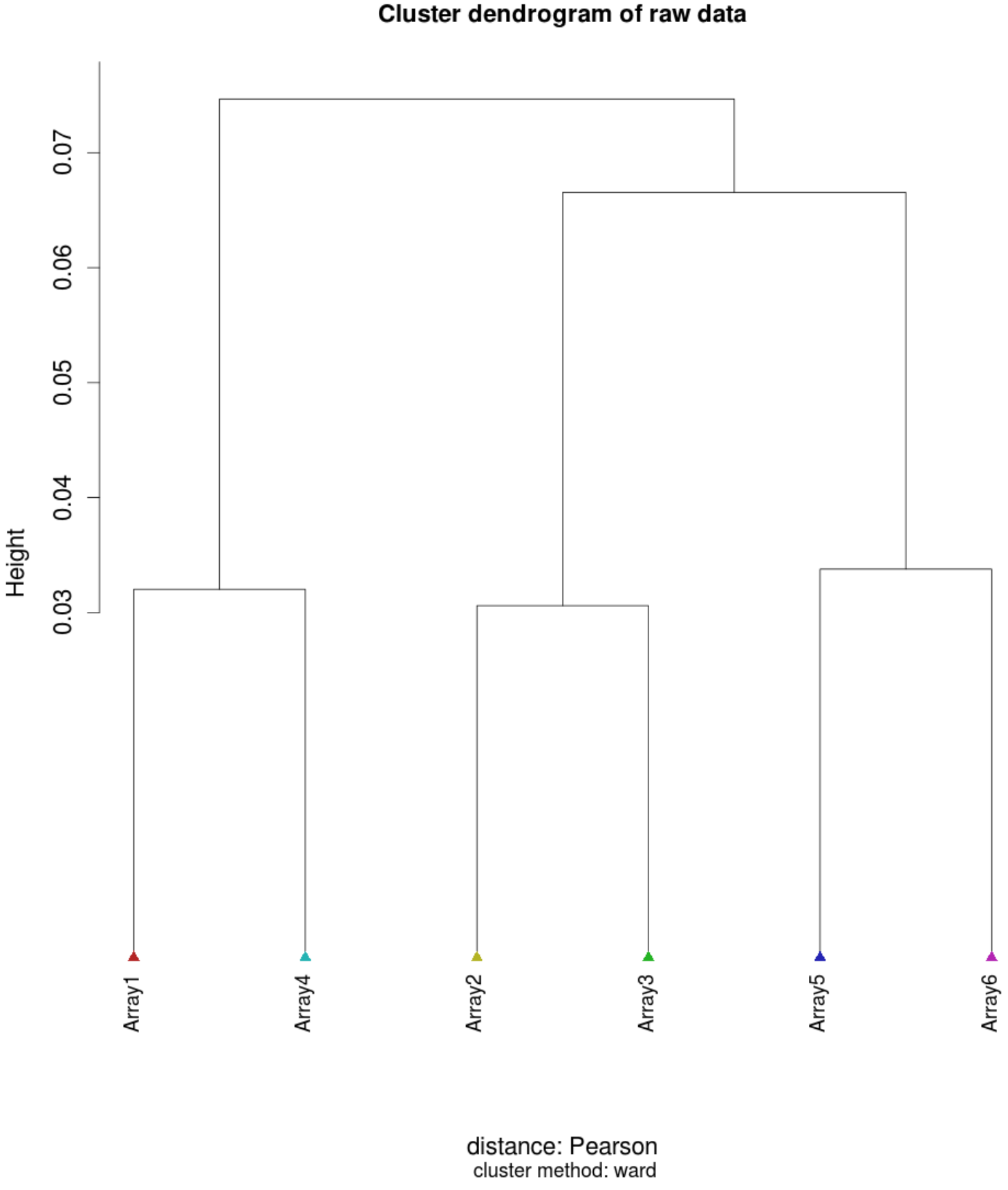
PCA analysis of Raw data





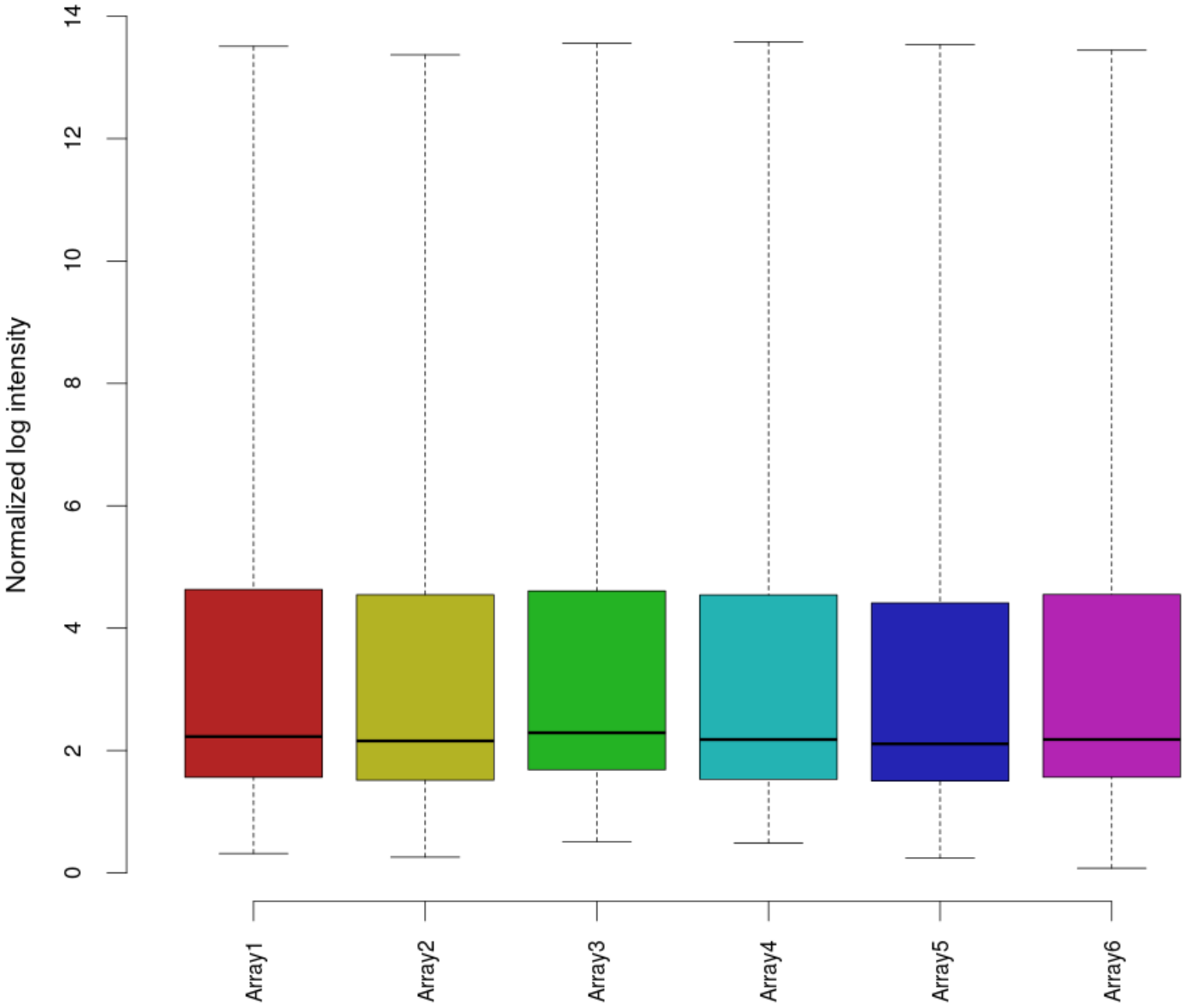


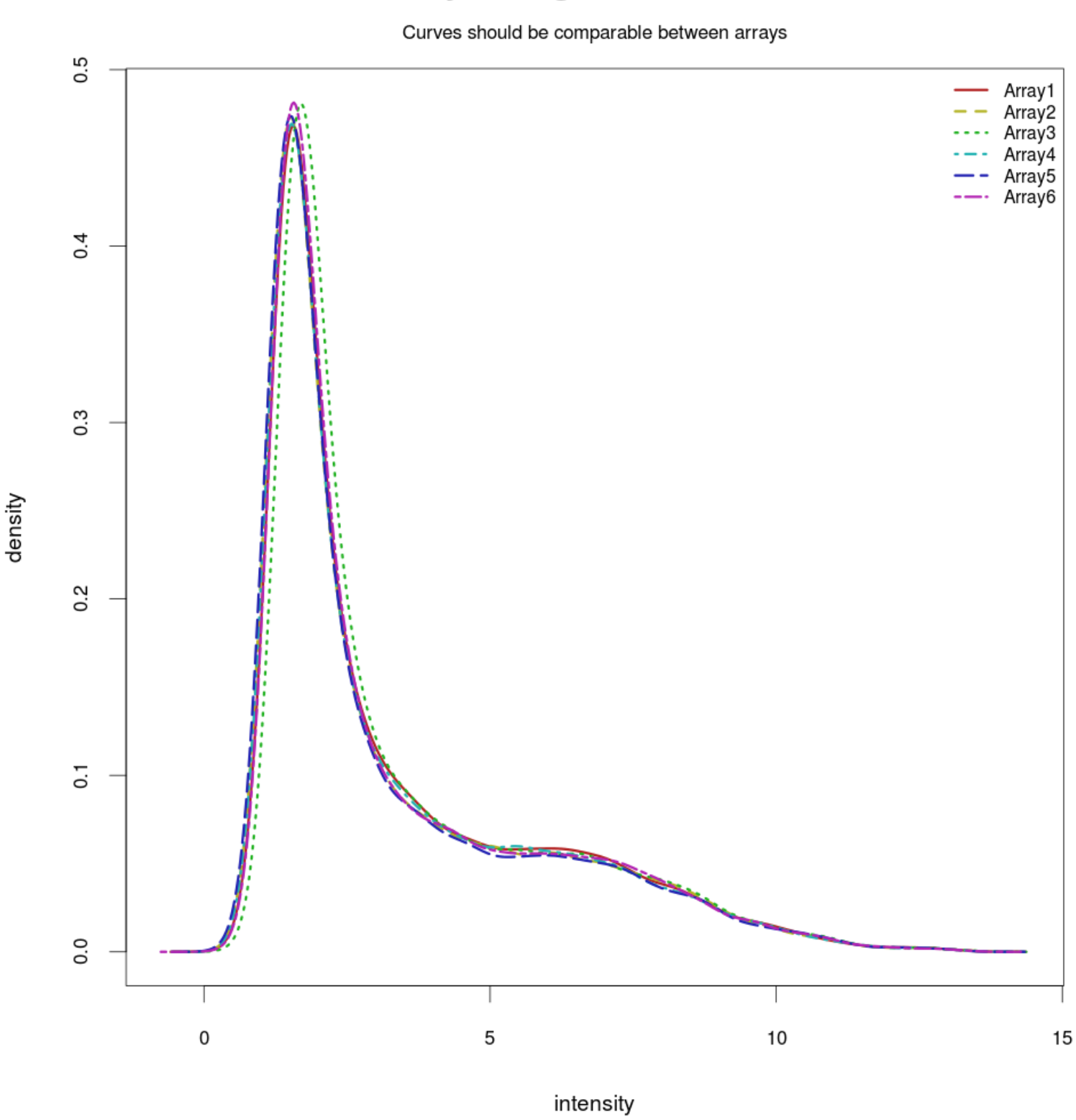



Cluster dendrogram of raw data
Height
0.03
0.04
0.05
0.06
0.07
Array1
Array4
Array2
Array3
Array5
Array6
distance: Pearson
cluster method: ward

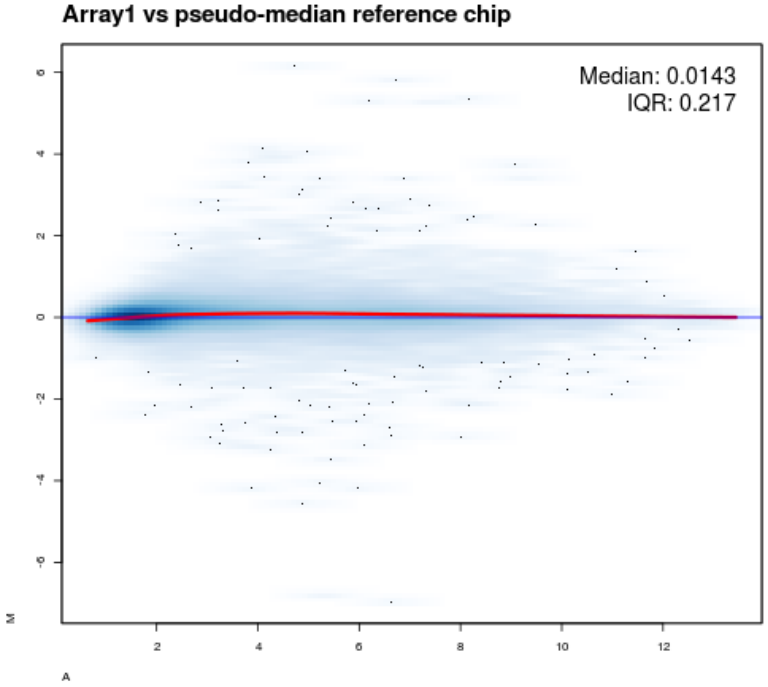
Pre-processing of Raw Data

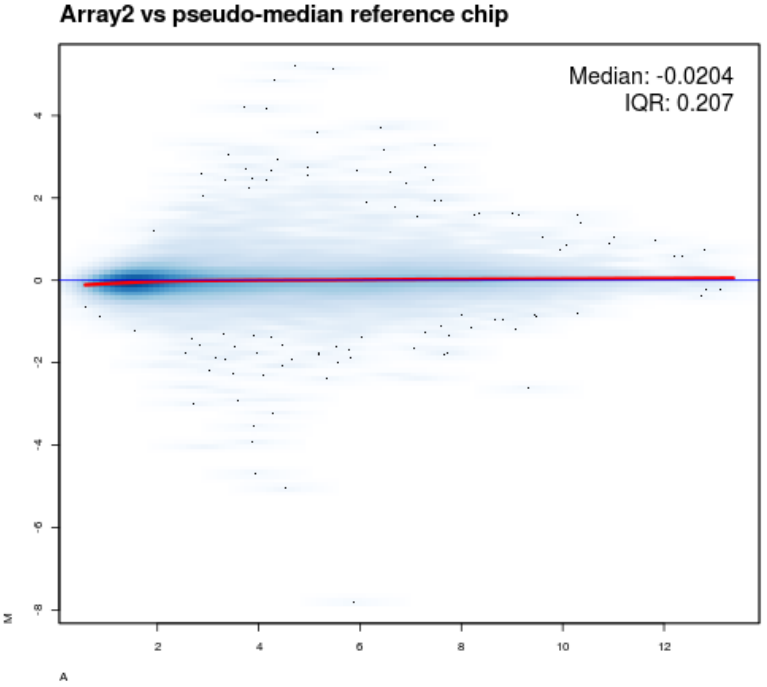
Method: GCRMA
Annotation: hgu133plus2_Hs_ENSG

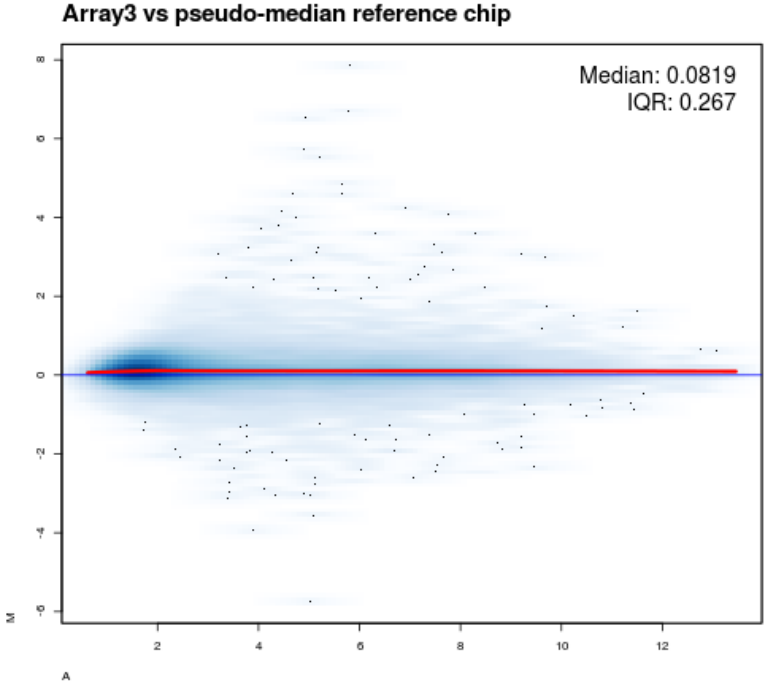

Boxplot after GCRMA
Distributions should be comparable between arrays
Normalized log intensity
0
2
4
6
8
10
12
14
Array1
Array2
Array3
Array4
Array5
Array6

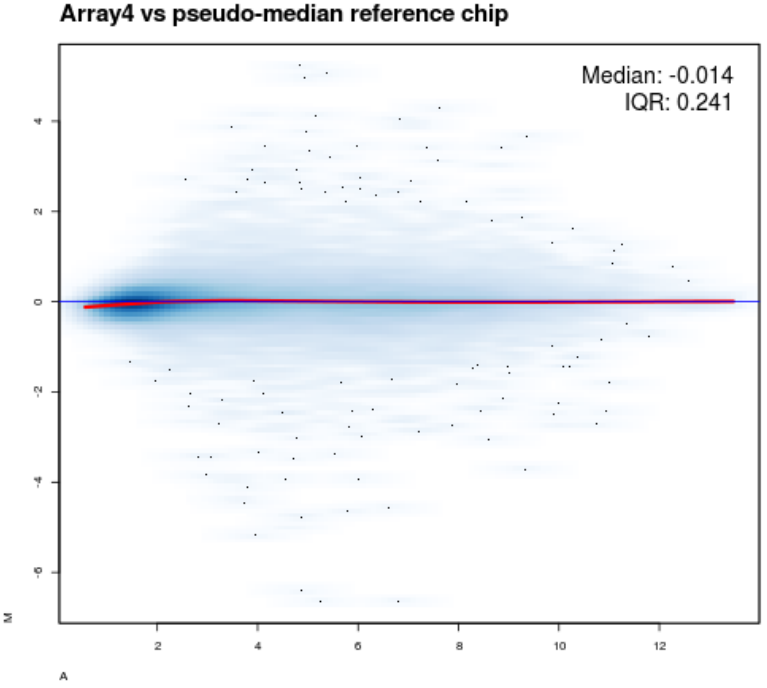

Density histogram after GCRMA
Curves should be comparable between arrays
Array1
Array2
Array3
Array4
Array5
Array6
density
0.0
0.1
0.2
0.3
0.4
0.5
0
5
10
15
intensity

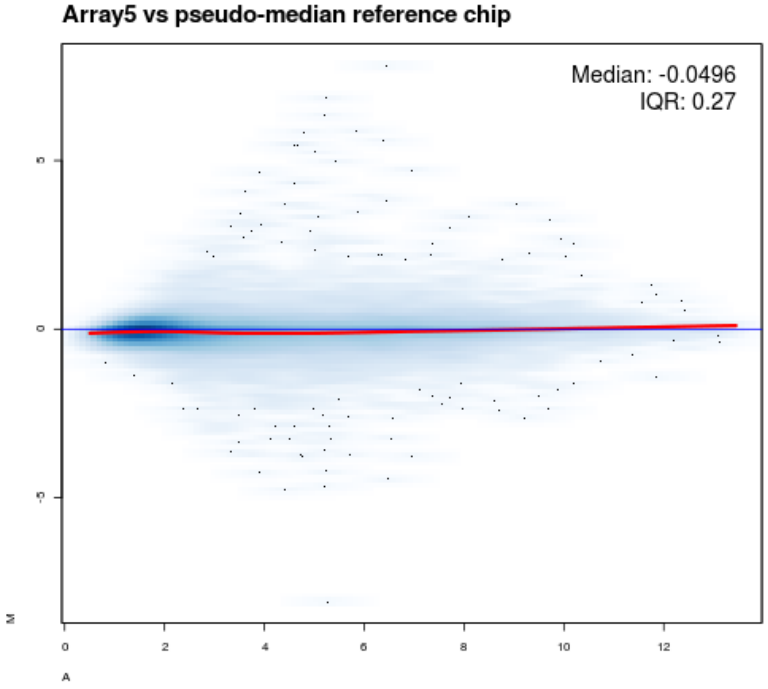
MA plots afterGCRMAnormalization

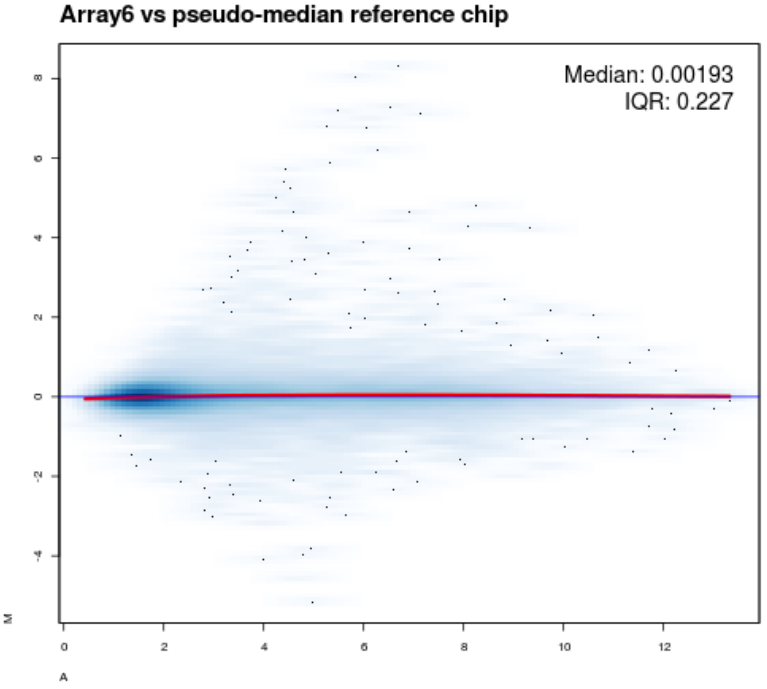


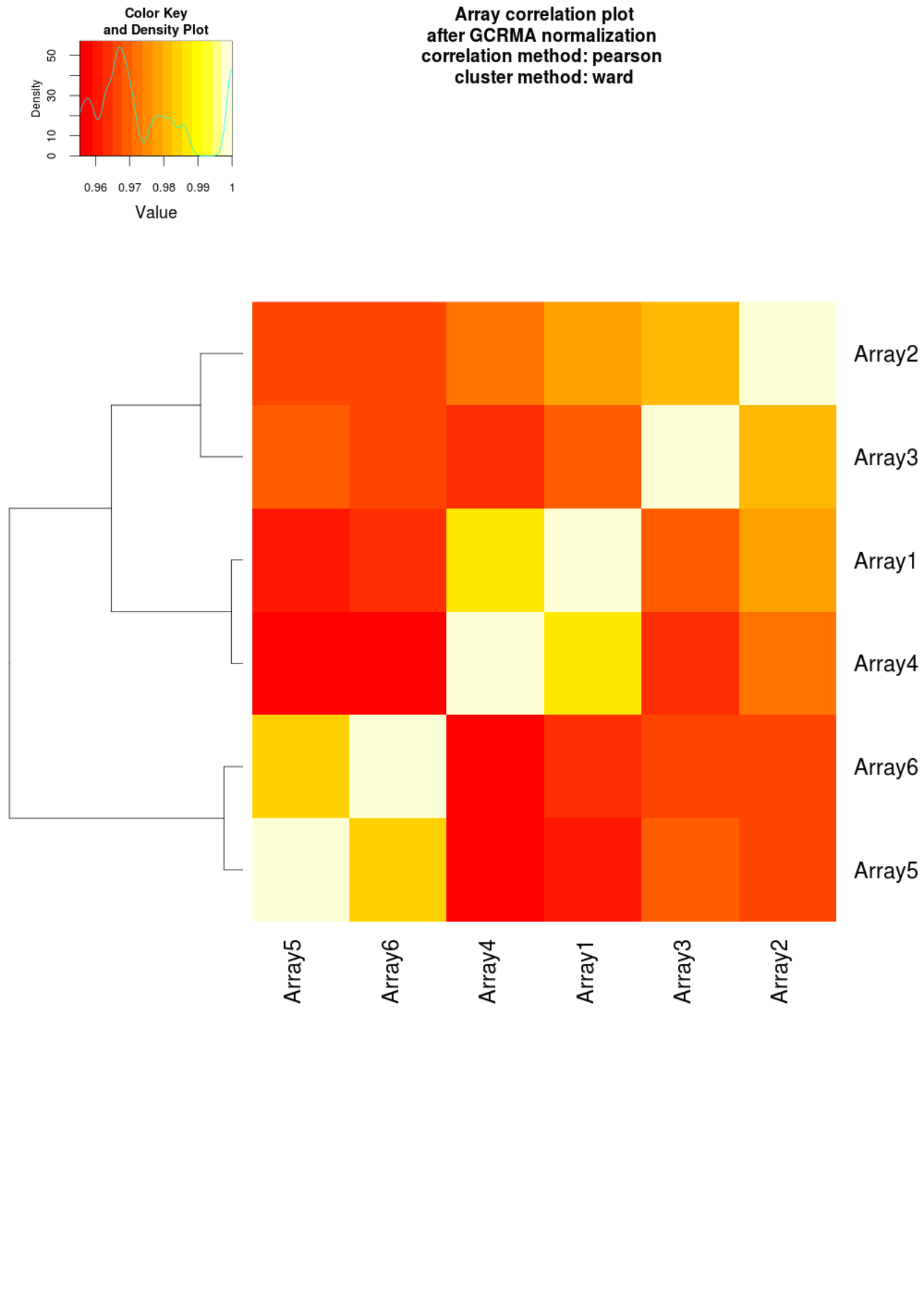




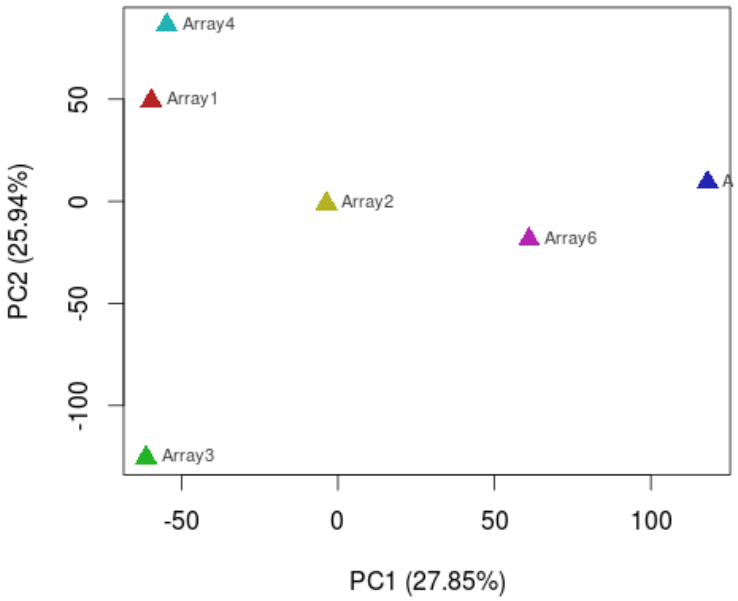


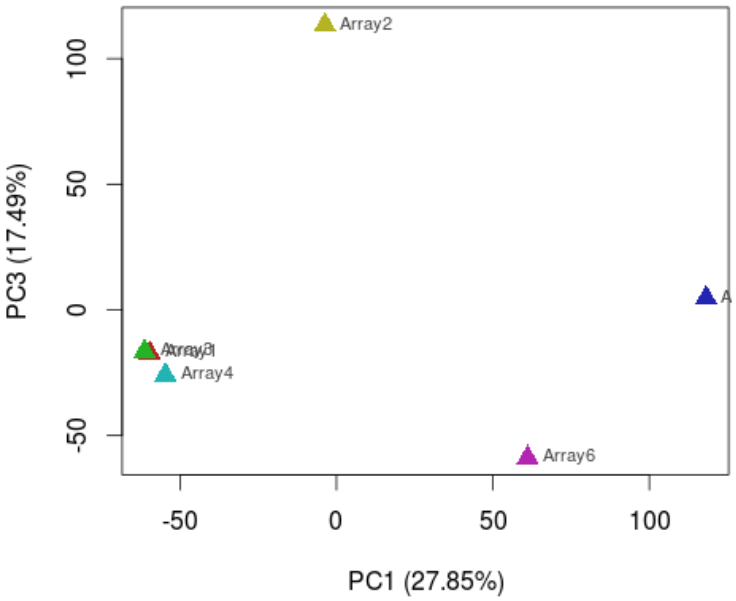


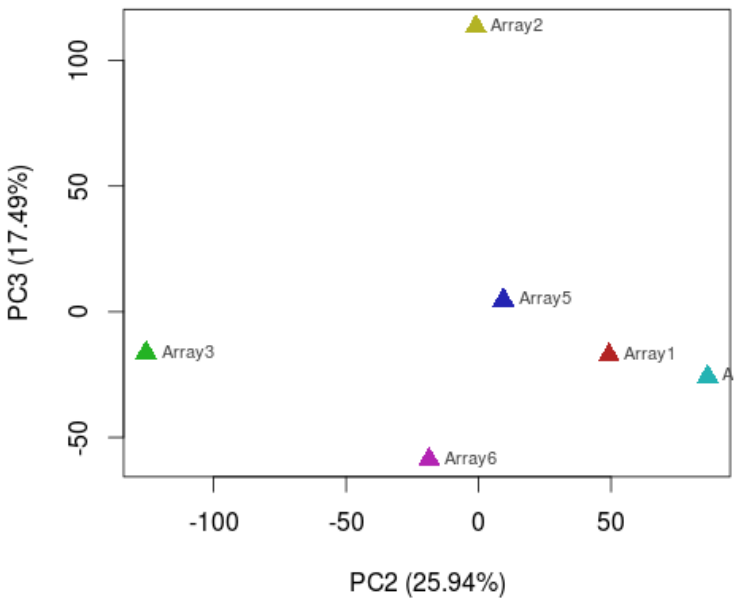


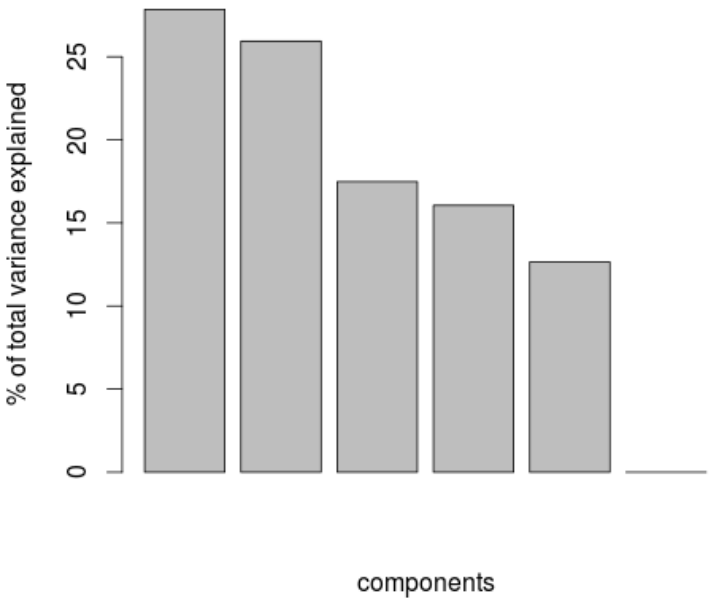

Color Key
and Density Plot
Density
0
10
30
50
0.96
0.97
0.98
0.99
1
Value
Array correlation plot
after GCRMA normalization
correlation method: pearson
cluster method: ward
Array2
Array3
Array1
Array4
Array6
Array5
Array5
Array6
Array4
Array1
Array3
Array2

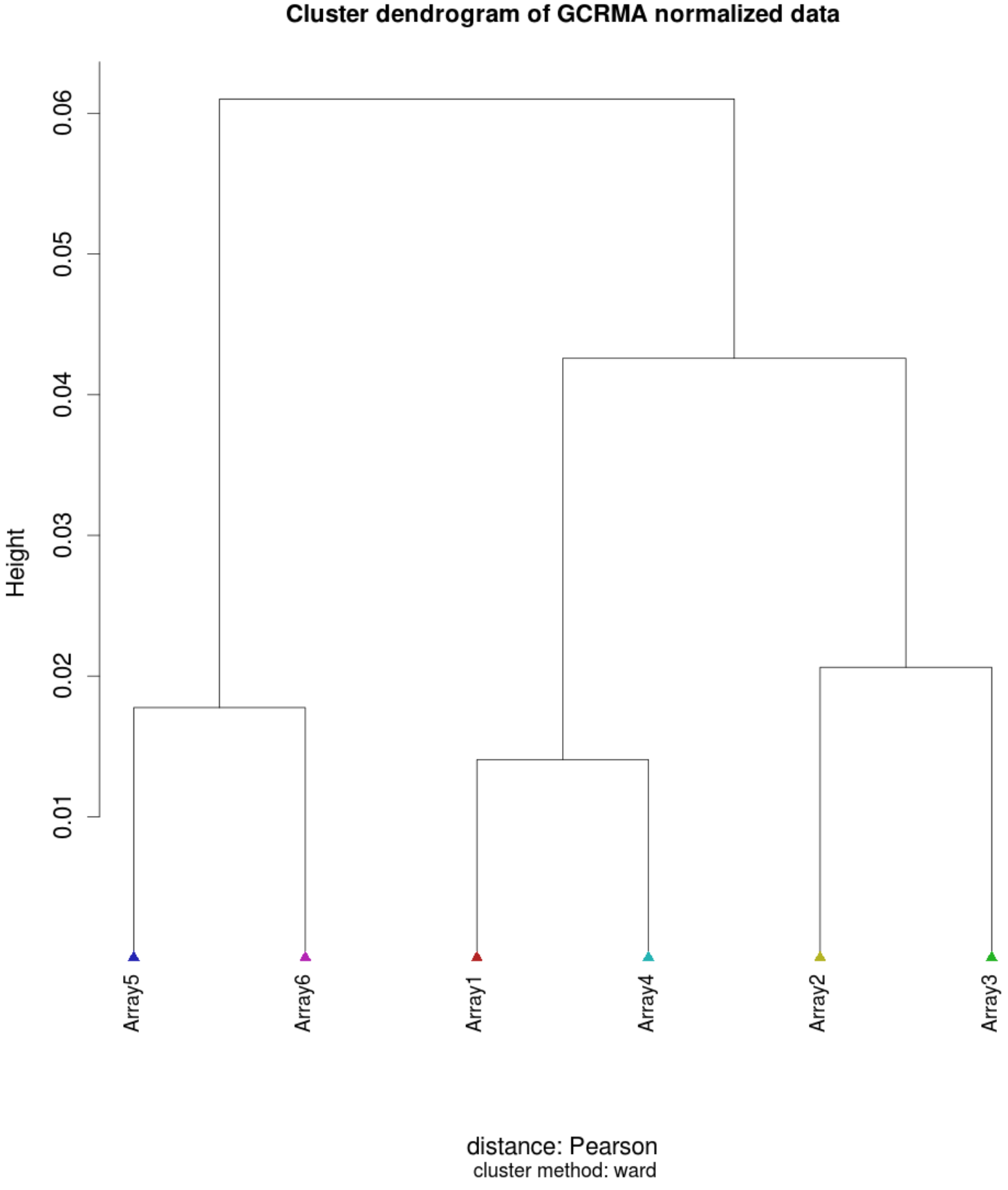
PCA analysis after GCRMA normalization










Cluster dendrogram of GCRMA normalized data
Height
0.06
0.05
0.04
0.03
0.02
0.01
Array5
Array6
Array1
Array4
Array2
Array3
distance: Pearson
cluster method: ward