Bioinformatika és genomanalízis az orvostudományban

Nagy áteresztésű technikák

Cserző Miklós

2020

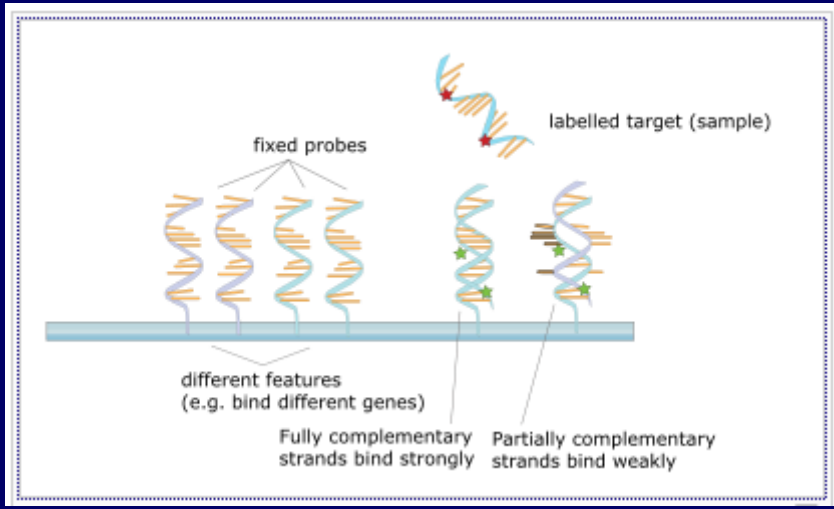
https://semmelweis.zoom.us/j/96102872458?pwd=Rk1PL2tqS21sdlUwc3B4eDFCZkNKQT09

A mai előadás

- A gén-chip működési elve
- Gén-chipek típusai
- Minőségellenőrzés
- Az adatok statisztikai előkészítése
- Gén-chipes adatok elemzése
- Génexpressziós adatbázisok

Hogyan működik a gén-chip?

- Rövid, egyszálú DNS darabok
- Kovalensen kötve egy hordozó felülethez
- A mintában jelzett, feldarabolt DNS van
- A komplementer kötődik a hordozón levő párjához



A gyártás

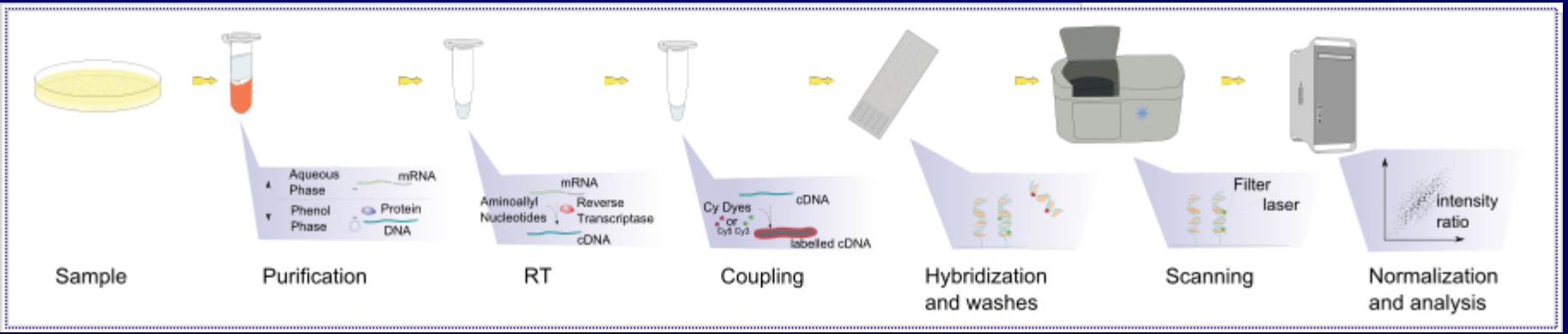
- Robotok készítik nyomdai technikával
- A próbák négyzethálósan vannak elrendezve
- A pontok száma néhány 10 és néhány millió között
- A próbák mérete: általában 20 és 60 bázis
- A próbák lehetnek egy adott készletből
- Vagy igény szerint tetszőlegesen

Kik gyártják?

- Néhány nagy gyártó dominál: Affymetrix, Illumina, Agilent, Nimblegen, stb..
- A termékek nem kompatibilisek egymással
  - Sem fizikai méretükben
  - Sem tartalmuk alapján
- A tipikus ár: 200 – 1000 $
- Ez csak a chip, a készülék külön jön
- Szolgáltatást is lehet rendelni

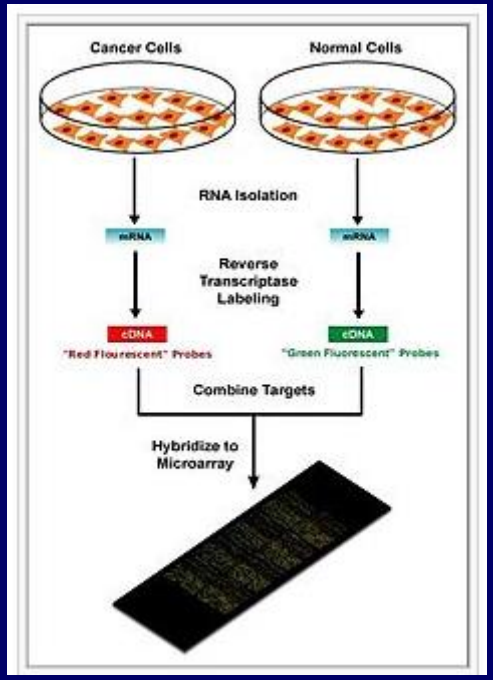
A mérés menete

- Tisztítás
- Reverz transzkripció
- Festés
- Hibridizálás
- Leolvasás
- Kiértékelés



A minta festése

- Kétszeres festés
  - A mintát és a referenciát két különböző festékkel jelölik – zöld és piros
  - Összekeverik és a két mintából származó DNS versenyez a próbákért
  - Relatív skála



Egyszeres festés

- Egy festéket alkalmaznak
- Csak egy mintát visznek fel a chip-re
- A referenciát külön viszik fel egy másik chip-re
- A két minta nem zavarja egymást
- Kétszer annyiba kerül

Mit lehet vizsgálni vele?

- "gene expression profiling"
  - A gének aktivitását mérjük a minta mRNS tartalmán keresztül
  - Gyakran mérési sorozatok formájában – expresszió változás stimulus hatására
- "comparative genomic hybridization"
  - Közeli rokon fajok összehasonlítása
- "geneID"
  - Genetikai anyag alapú azonosítás

Folytatás…

- "ChiP-on-chip" – cromatin-immuno-prcipitálás
  - A DNS-t kovalensen kötjük a fehérjéhez
  - A DNS-hez kötő fehérjét antitesttel kicsapjuk
  - Ez viszi magával a DNS-darabot is
  - Ezt keressük meg a gén-chipe-en
- "SNP-chip" – "single nucloetide polymorphism"
  - Pontmutációkat keresünk a vizsgált genomban

Folytatás…

- "alternative splicing"
  - Egy gén változatait keressük
  - A próbákat az exon/intron határokra tervezzük
- "tiling array"
  - Egy nagyobb kromoszómális régió teljes térképe
  - Rövid, átfedő szakaszok formájában

A gén-chipek bioinformatikája

A legfontosabb kihívások:

- A kisérlet megtervezése
- A kisérleti bizonytalanság kezelése
- Az adatok statisztikai elemzése
- Szabványok kidolgozása
- Az adatok kezelése

Kisérlettervezés

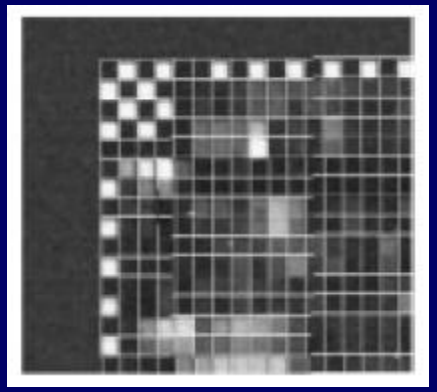
- Kisérleti kérdésfeltevés
  - Csak biológiailag megismételhető kérdést érdemes vizsgálni
- A kisérleti protokol megismételhetősége
  - Ugynabból a mintából két párhuzamos koherens eredményt ad?
- A hiridizálás megismételhetősége
- Milyen próbák legyenek a chip-en?
  - Egyértelmű legyen a hibridizálás
  - A próba valóban azt jelezze, amire kiváncsi vagyok

Kisérleti bizonytalanság

- Ha RNS expressziós szintet mérünk, az nem feltétlenül arányos a fehérje szinttel
- És a fehérje aktivitásával sem
- Egy gén adhat keresztreakciót egy másik génnek szánt próbával
- A próba lehet téves – nincs is olyan DNS-darab a genomban

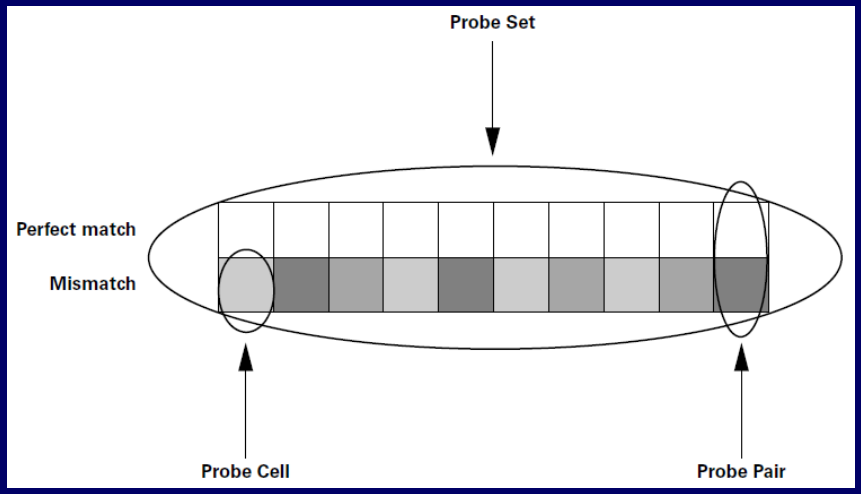
Statisztikus feldolgozás

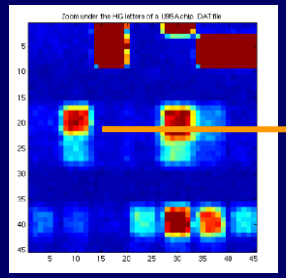
- A gén-chip kísérlet eredménye egy kép
- Az első feladat: beolvasni a képet, értelmezni és alkalmassá tenni a további feldolgozásra
  - Képfelismerés az iránypontok alapján

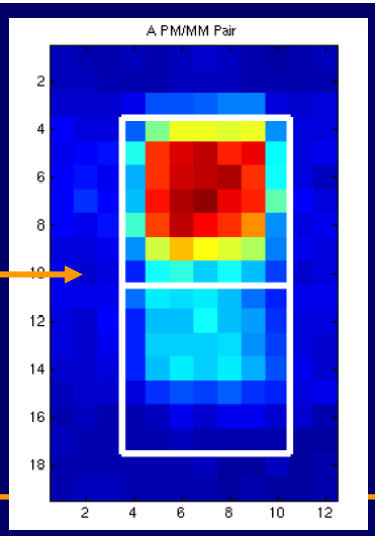


Belső ellenőrző pontok

- Egy próba a chip több pontján is jelen van
- Egy gént több próbával is elérünk
- Minden próbának van saját kontrolja
- "Perfect Match" – "MisMatch" párok
- Ismert mennyiségű, idegen DNS

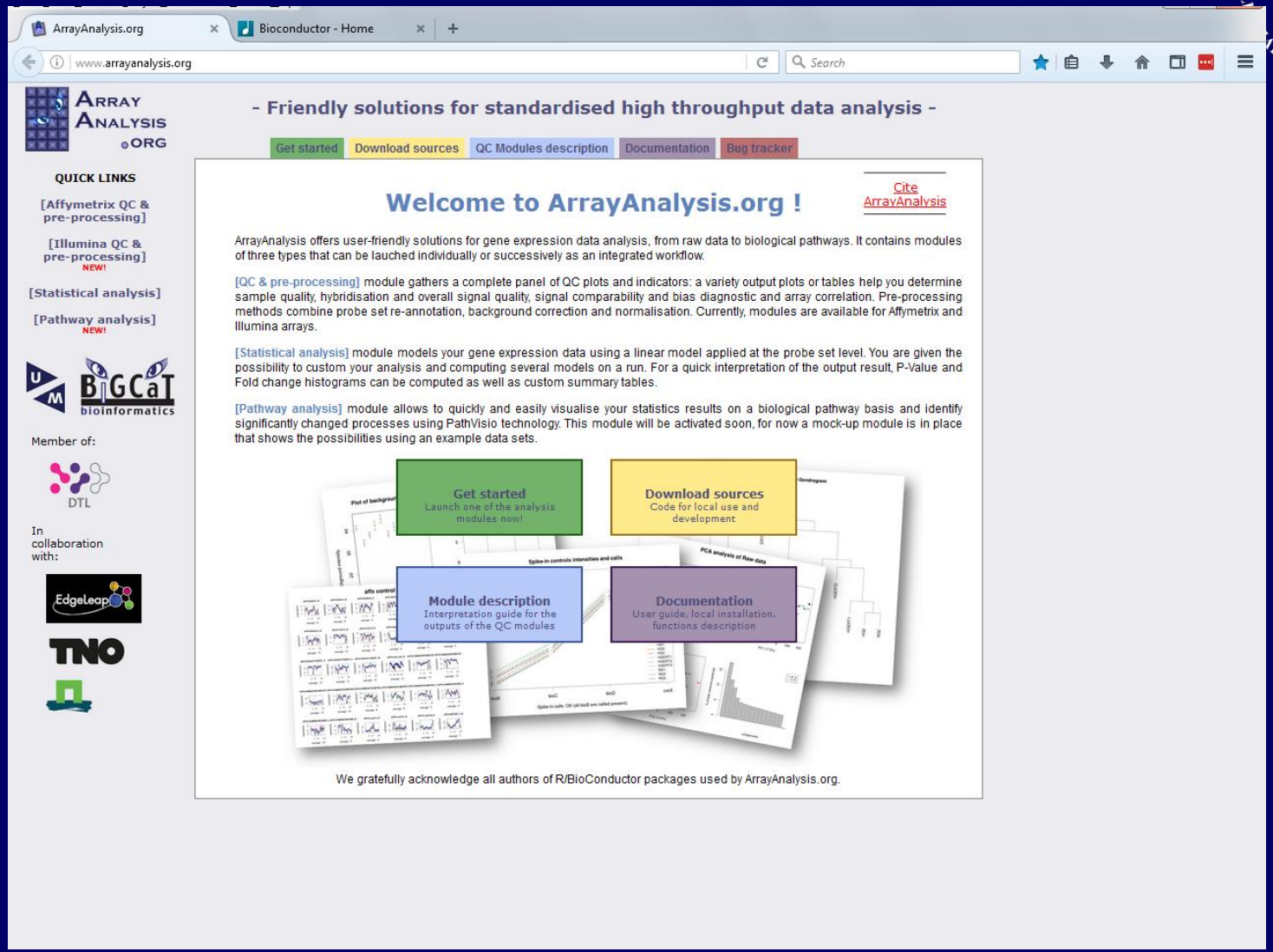


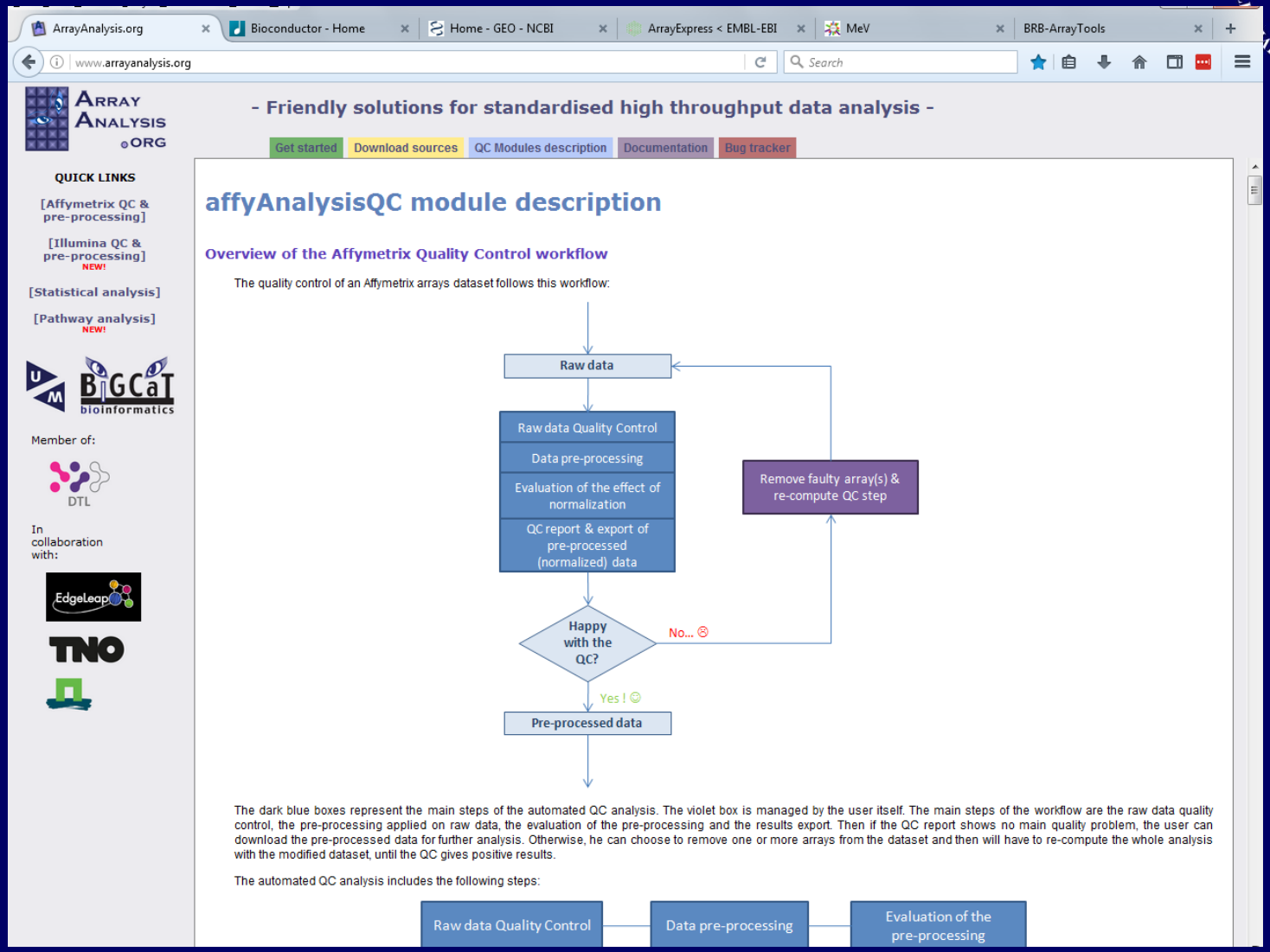


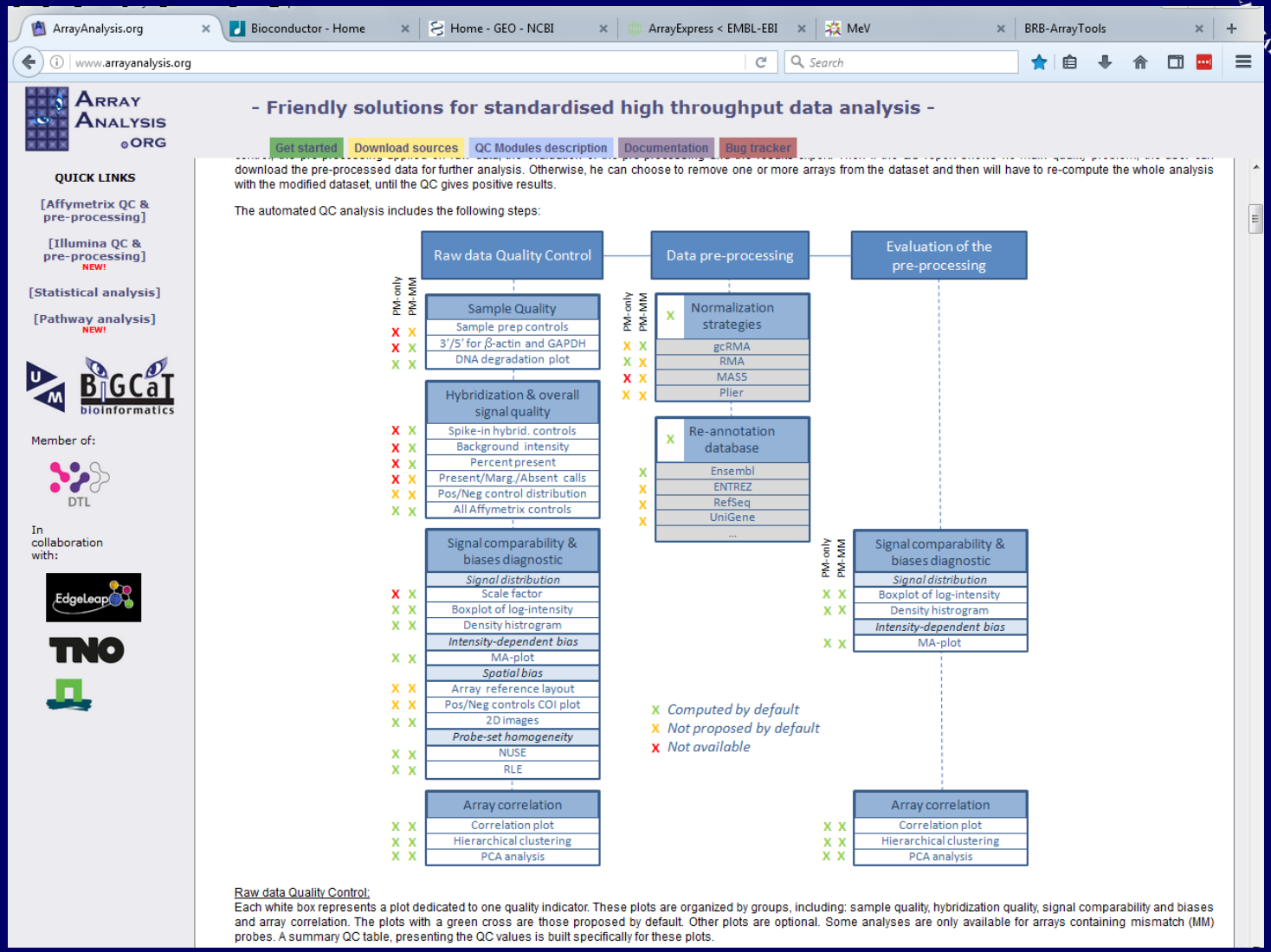


Minőségellenőrzés

- Affymetrix chip-ekre ingyenesen elérhető on-line ellenőrző rendszer
- Web: http://www.arrayanalysis.org/
- Az alap: “Bioconductor”
- Parancssor alapú rendszer – közösségi fejlesztés
- Rugalmas, sokrétű, de barátságtalan
- Web: http://www.bioconductor.org/


ArrayAnalysis.org
Bioconductor - Home
www.arrayanalysis.org
Search
ARRAY ANALYSIS .ORG
- Friendly solutions for standardised high throughput data analysis -
Get started
Download sources
QC Modules description
Documentation
Bug tracker
QUICK LINKS
[Affymetrix QC & pre-processing]
[Illumina QC & pre-processing]
NEW!
[Statistical analysis]
[Pathway analysis]
NEW!
BiGCaT bioinformatics
Member of:
DTL
In collaboration with:
EdgeLeap
TNO
Welcome to ArrayAnalysis.org !
Cite ArrayAnalysis
ArrayAnalysis offers user-friendly solutions for gene expression data analysis, from raw data to biological pathways. It contains modules of three types that can be lauched individually or successively as an integrated workflow.
[QC & pre-processing] module gathers a complete panel of QC plots and indicators: a variety output plots or tables help you determine sample quality, hybridisation and overall signal quality, signal comparability and bias diagnostic and array correlation. Pre-processing methods combine probe set re-annotation, background correction and normalisation. Currently, modules are available for Affymetrix and Illumina arrays.
[Statistical analysis] module models your gene expression data using a linear model applied at the probe set level. You are given the possibility to custom your analysis and computing several models on a run. For a quick interpretation of the output result, P-Value and Fold change histograms can be computed as well as custom summary tables.
[Pathway analysis] module allows to quickly and easily visualise your statistics results on a biological pathway basis and identify significantly changed processes using PathVisio technology. This module will be activated soon, for now a mock-up module is in place that shows the possibilities using an example data sets.
Get started
Launch one of the analysis modules now!
Download sources
Code for local use and development
Module description
Interpretation guide for the outputs of the QC modules
Documentation
User guide, local installation, functions description
We gratefully acknowledge all authors of R/BioConductor packages used by ArrayAnalysis.org.


ArrayAnalysis.org
Bioconductor - Home
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.arrayanalysis.org
Search
ARRAY ANALYSIS .ORG
- Friendly solutions for standardised high throughput data analysis -
Get started
Download sources
QC Modules description
Documentation
Bug tracker
QUICK LINKS
[Affymetrix QC & pre-processing]
[Illumina QC & pre-processing]
NEW!
[Statistical analysis]
[Pathway analysis]
NEW!
BiGCaT bioinformatics
Member of:
DTL
In collaboration with:
EdgeLeap
TNO
affyAnalysisQC module description
Overview of the Affymetrix Quality Control workflow
The quality control of an Affymetrix arrays dataset follows this workflow:
Raw data
Raw data Quality Control
Data pre-processing
Evaluation of the effect of normalization
QC report & export of pre-processed (normalized) data
Remove faulty array(s) & re-compute QC step
Happy with the QC?
No... ☹
Yes ! ☺
Pre-processed data
The dark blue boxes represent the main steps of the automated QC analysis. The violet box is managed by the user itself. The main steps of the workflow are the raw data quality control, the pre-processing applied on raw data, the evaluation of the pre-processing and the results export. Then if the QC report shows no main quality problem, the user can download the pre-processed data for further analysis. Otherwise, he can choose to remove one or more arrays from the dataset and then will have to re-compute the whole analysis with the modified dataset, until the QC gives positive results.
The automated QC analysis includes the following steps:
Raw data Quality Control
Data pre-processing
Evaluation of the pre-processing


ArrayAnalysis.org
Bioconductor - Home
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.arrayanalysis.org
Search
ARRAY ANALYSIS .ORG
- Friendly solutions for standardised high throughput data analysis -
Get started
Download sources
QC Modules description
Documentation
Bug tracker
QUICK LINKS
[Affymetrix QC & pre-processing]
[Illumina QC & pre-processing] NEW!
[Statistical analysis]
[Pathway analysis] NEW!
BiGCaT bioinformatics
Member of:
DTL
In collaboration with:
EdgeLeap
TNO
download the pre-processed data for further analysis. Otherwise, he can choose to remove one or more arrays from the dataset and then will have to re-compute the whole analysis with the modified dataset, until the QC gives positive results.
The automated QC analysis includes the following steps:
Raw data Quality Control
Data pre-processing
Evaluation of the pre-processing
PM-only
PM-MM
Sample Quality
Sample prep controls
3'/5' for β-actin and GAPDH
DNA degradation plot
Hybridization & overall signal quality
Spike-in hybrid. controls
Background intensity
Percent present
Present/Marg./Absent calls
Pos/Neg control distribution
All Affymetrix controls
Signal comparability & biases diagnostic
Signal distribution
Scale factor
Boxplot of log-intensity
Density histrogram
Intensity-dependent bias
MA-plot
Spatial bias
Array reference layout
Pos/Neg controls COI plot
2D images
Probe-set homogeneity
NUSE
RLE
Array correlation
Correlation plot
Hierarchical clustering
PCA analysis
Normalization strategies
gcRMA
RMA
MAS5
Plier
Re-annotation database
Ensembl
ENTREZ
RefSeq
UniGene
...
Signal comparability & biases diagnostic
Signal distribution
Boxplot of log-intensity
Density histrogram
Intensity-dependent bias
MA-plot
Array correlation
Correlation plot
Hierarchical clustering
PCA analysis
X Computed by default
X Not proposed by default
X Not available
Raw data Quality Control:
Each white box represents a plot dedicated to one quality indicator. These plots are organized by groups, including: sample quality, hybridization quality, signal comparability and biases and array correlation. The plots with a green cross are those proposed by default. Other plots are optional. Some analyses are only available for arrays containing mismatch (MM) probes. A summary QC table, presenting the QC values is built specifically for these plots.

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

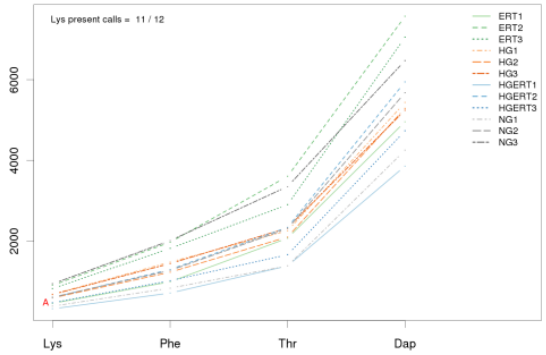
Array correlation
- Correlation plot
- PCA analysis
- Clustering

BiGCaT bioinformatics

Sample prep controls: Dap, Thr, Phe and Lys -PolyA unlabeled spikes

These probe sets are designed from several B. subtilis genes (dap, thr, phe and lys). They are spiked at the beginning of the chip processing and used to assess the overall success of the target prep steps.
The ploy-A controls Dap, Thr, Phe and Lys should be called present at a decreasing intensity, to verify that there was no bias during the retro-transcription between highly expressed genes and low expressed genes. The linearity for lys, phe and thr (dap is present at a much higher concentration) is affected by a double amplification.

Spike-in Sample Prep controls intensities and calls



Intensity: OK (Lys < Phe < Thr < Dap for all arrays)
Present calls: 1 Lys not called present

In this example, the intensity is increasing from Lys to Dap for all arrays and the lowest expected intensity (Lys) are called present for 11 arrays among the 12. ERT1 array's Lys probeset is called Absent (red 'A' on the graph).

[Technical documentation of the function]

RNA quality control: 3'/5' ratio for beta-actin and GAPDH

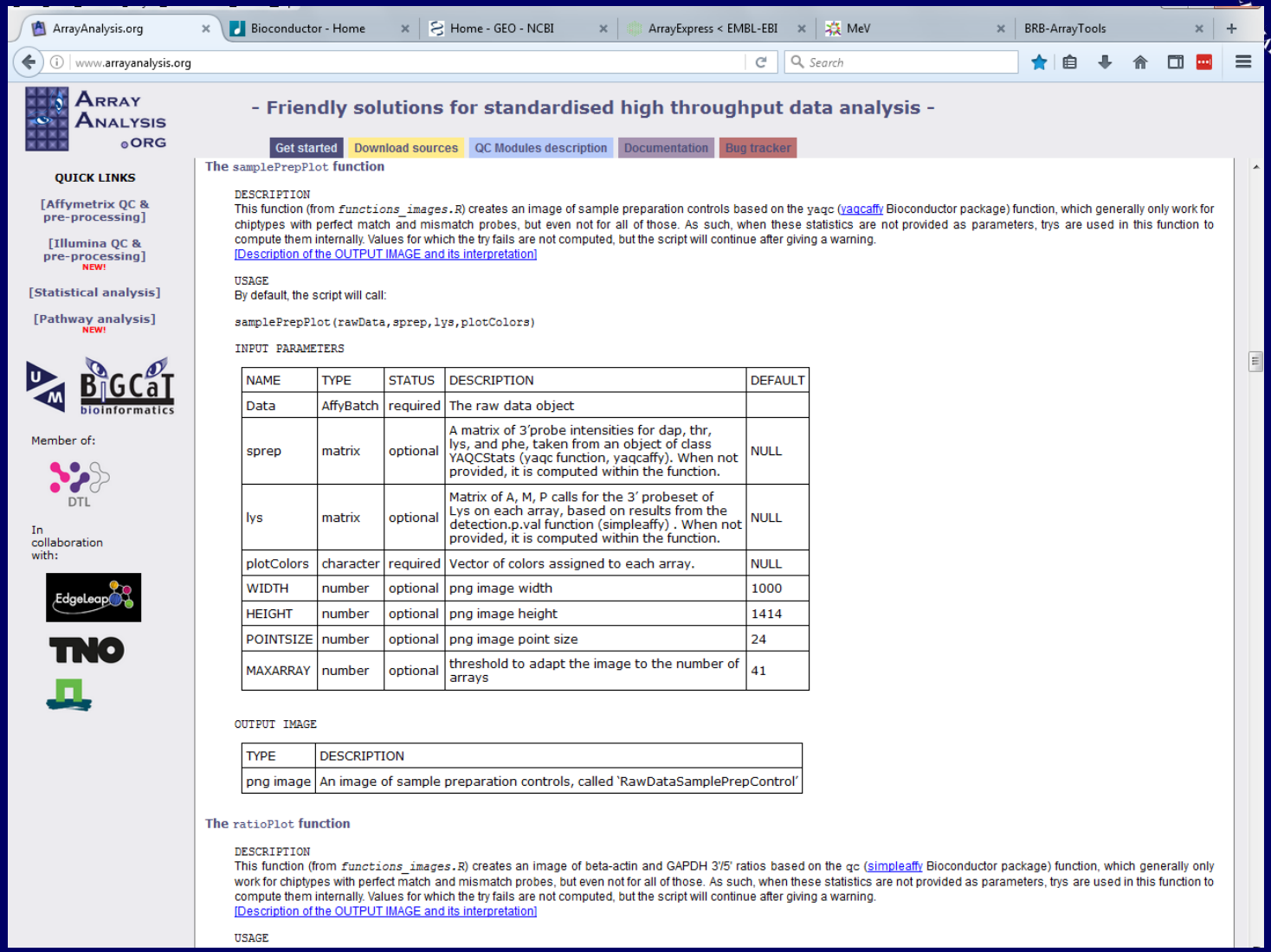
Because beta-actin and GAPDH are expressed is most cell types and are relatively long genes, Affymetrix chips use them as controls of the RNA quality. Three probe-sets are designed on 3 regions of these genes (5', mid (called M) and 3' extremities). Similar intensities for their 3 regions indicate that the transcripts were not truncated and labeled equally along the sequence. By the way, since RNA degradation starts from the 5' end of the molecule, it is common that the probeset intensity at that end is slightly lower.

For an array of good quality, Affymetrix recommends that the 3'/5' ratio should not exceed :
- 3 for beta-actin
- 1.25 for GAPDH

These values were set for human samples so the ratio may be slightly higher for other species.

The QC pipeline proposes the following graphs to evaluate these controls:

www.arrayanalysis.org/main.html#SP



The `samplePrepPlot` function

DESCRIPTION

This function (from *functions_images.R*) creates an image of sample preparation controls based on the `yaqc` (yaqcaffy Bioconductor package) function, which generally only work for chiptypes with perfect match and mismatch probes, but even not for all of those. As such, when these statistics are not provided as parameters, trys are used in this function to compute them internally. Values for which the try fails are not computed, but the script will continue after giving a warning.

[Description of the OUTPUT IMAGE and its interpretation]

USAGE

By default, the script will call:

```
samplePrepPlot(rawData,sprep,lys,plotColors)
```

INPUT PARAMETERS

NAME	TYPE	STATUS	DESCRIPTION	DEFAULT
Data	AffyBatch	required	The raw data object	
sprep	matrix	optional	A matrix of 3'probe intensities for dap, thr, lys, and phe, taken from an object of class YAQCStats (yaqc function, yaqcaffy). When not provided, it is computed within the function.	NULL
lys	matrix	optional	Matrix of A, M, P calls for the 3' probeset of Lys on each array, based on results from the detection.p.val function (simpleaffy) . When not provided, it is computed within the function.	NULL
plotColors	character	required	Vector of colors assigned to each array.	NULL
WIDTH	number	optional	png image width	1000
HEIGHT	number	optional	png image height	1414
POINTSIZE	number	optional	png image point size	24
MAXARRAY	number	optional	threshold to adapt the image to the number of arrays	41

OUTPUT IMAGE

TYPE	DESCRIPTION
png image	An image of sample preparation controls, called 'RawDataSamplePrepControl'

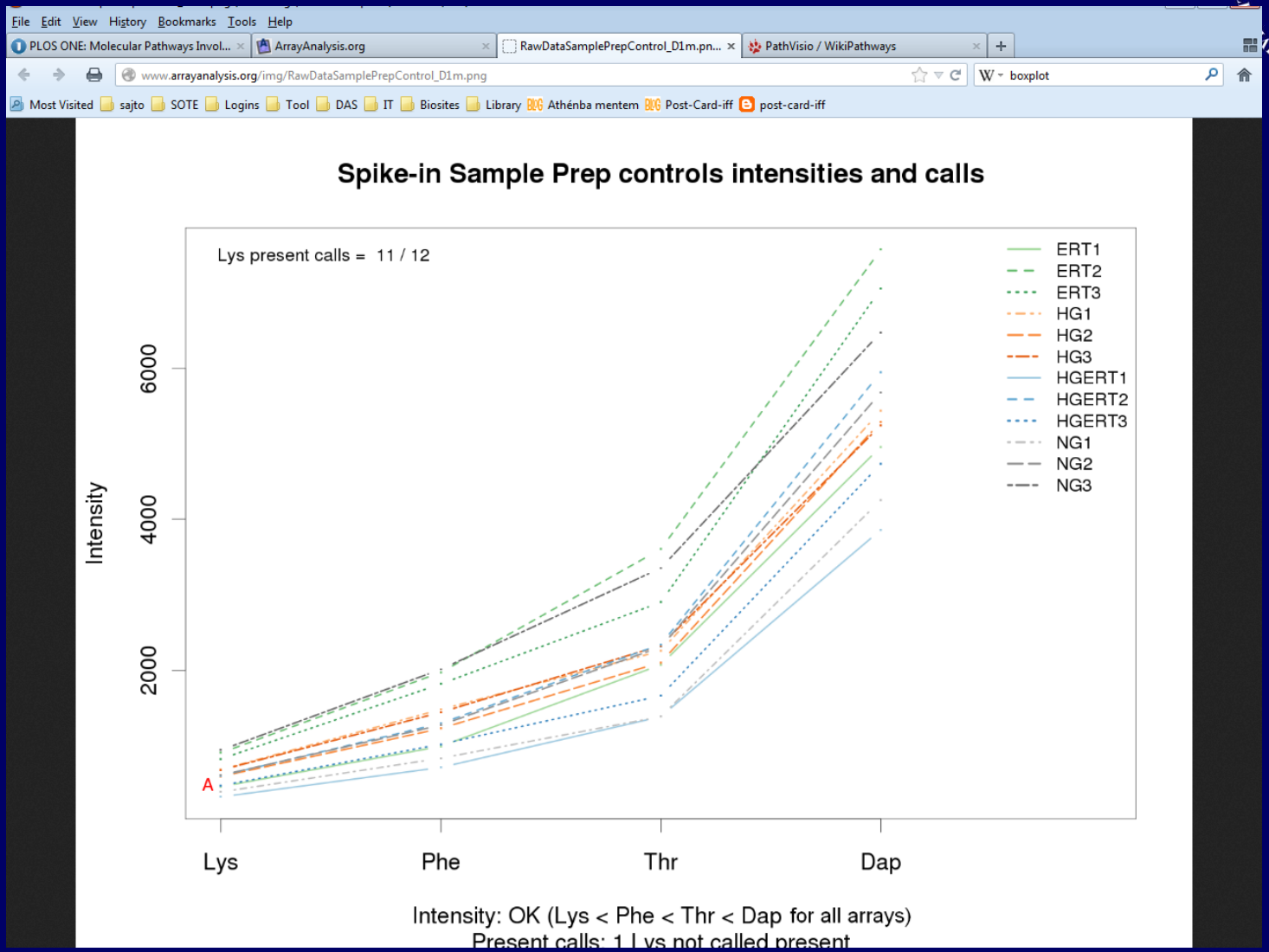
The `ratioPlot` function

DESCRIPTION

This function (from *functions_images.R*) creates an image of beta-actin and GAPDH 3'/5' ratios based on the `qc` (simpleaffy Bioconductor package) function, which generally only work for chiptypes with perfect match and mismatch probes, but even not for all of those. As such, when these statistics are not provided as parameters, trys are used in this function to compute them internally. Values for which the try fails are not computed, but the script will continue after giving a warning.

[Description of the OUTPUT IMAGE and its interpretation]

USAGE


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataSamplePrepControl_D1m.pn...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataSamplePrepControl_D1m.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Spike-in Sample Prep controls intensities and calls
Lys present calls = 11 / 12
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
Intensity
2000
4000
6000
A
Lys
Phe
Thr
Dap
Intensity: OK (Lys < Phe < Thr < Dap for all arrays)
Present calls: 1 Lys not called present

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

Array correlation
- Correlation plot
- PCA analysis
- Clustering

BiGCaT bioinformatics

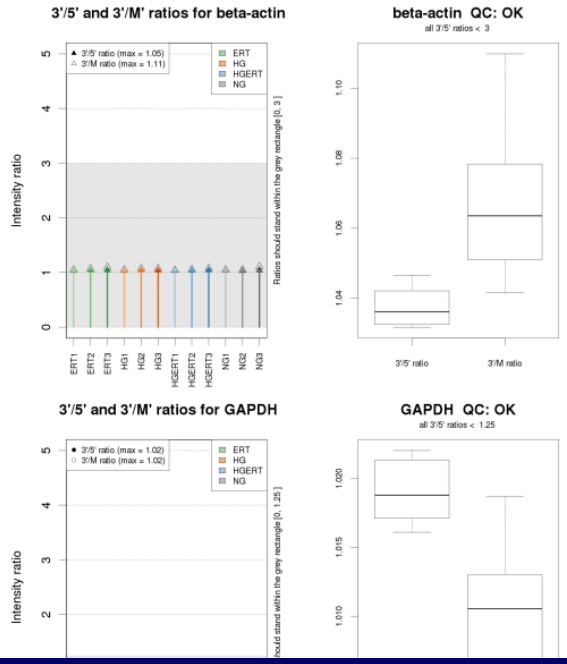
RNA quality control: 3'/5' ratio for beta-actin and GAPDH

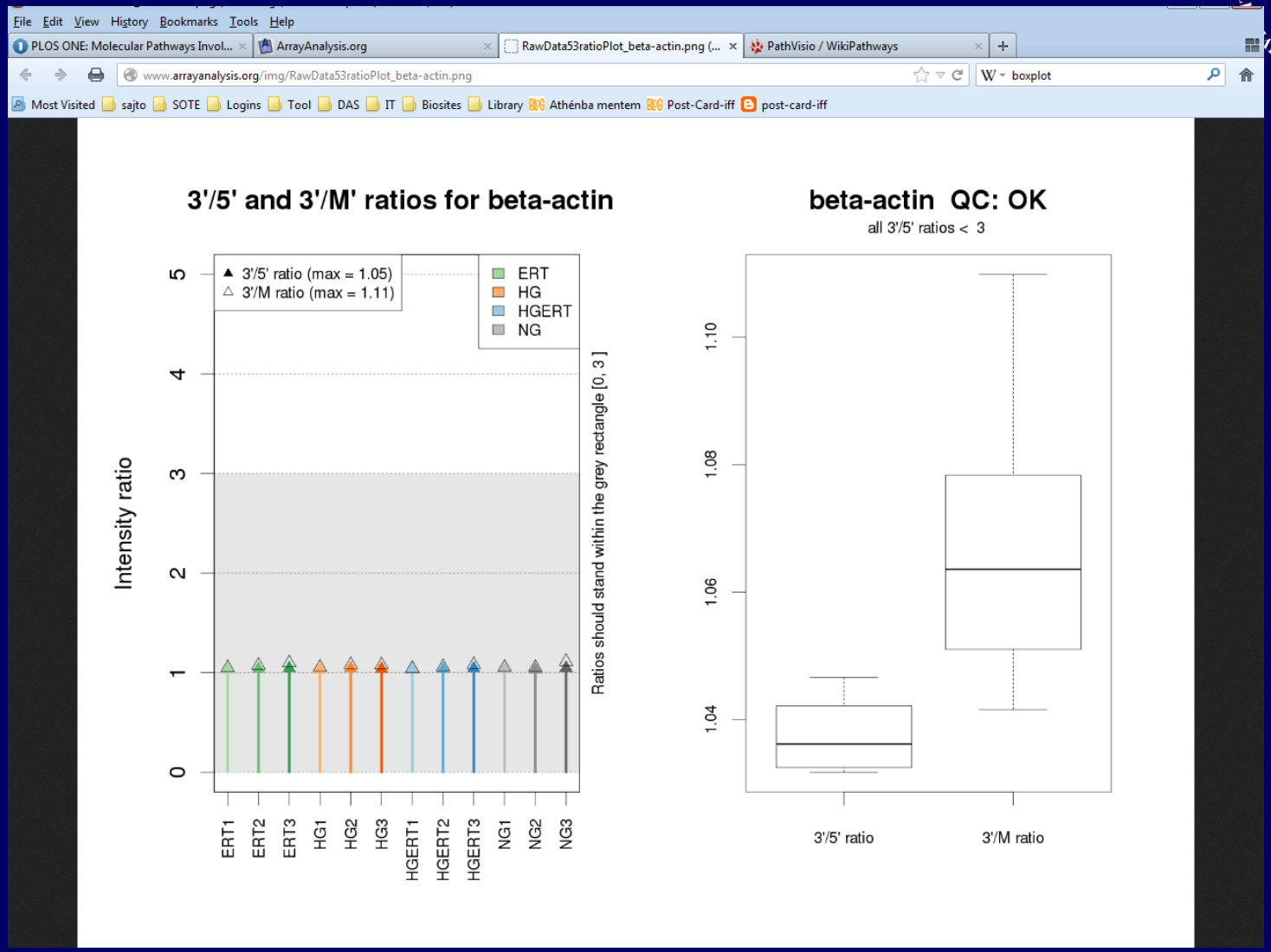
Because beta-actin and GAPDH are expressed is most cell types and are relatively long genes, Affymetrix chips use them as controls of the RNA quality. Three probe-sets are designed on 3 regions of these genes (5', mid (called M) and 3' extremities). Similar intensities for their 3 regions indicate that the transcripts were not truncated and labeled equally along the sequence. By the way, since RNA degradation starts from the 5' end of the molecule, it is common that the probeset intensity at that end is slightly lower.

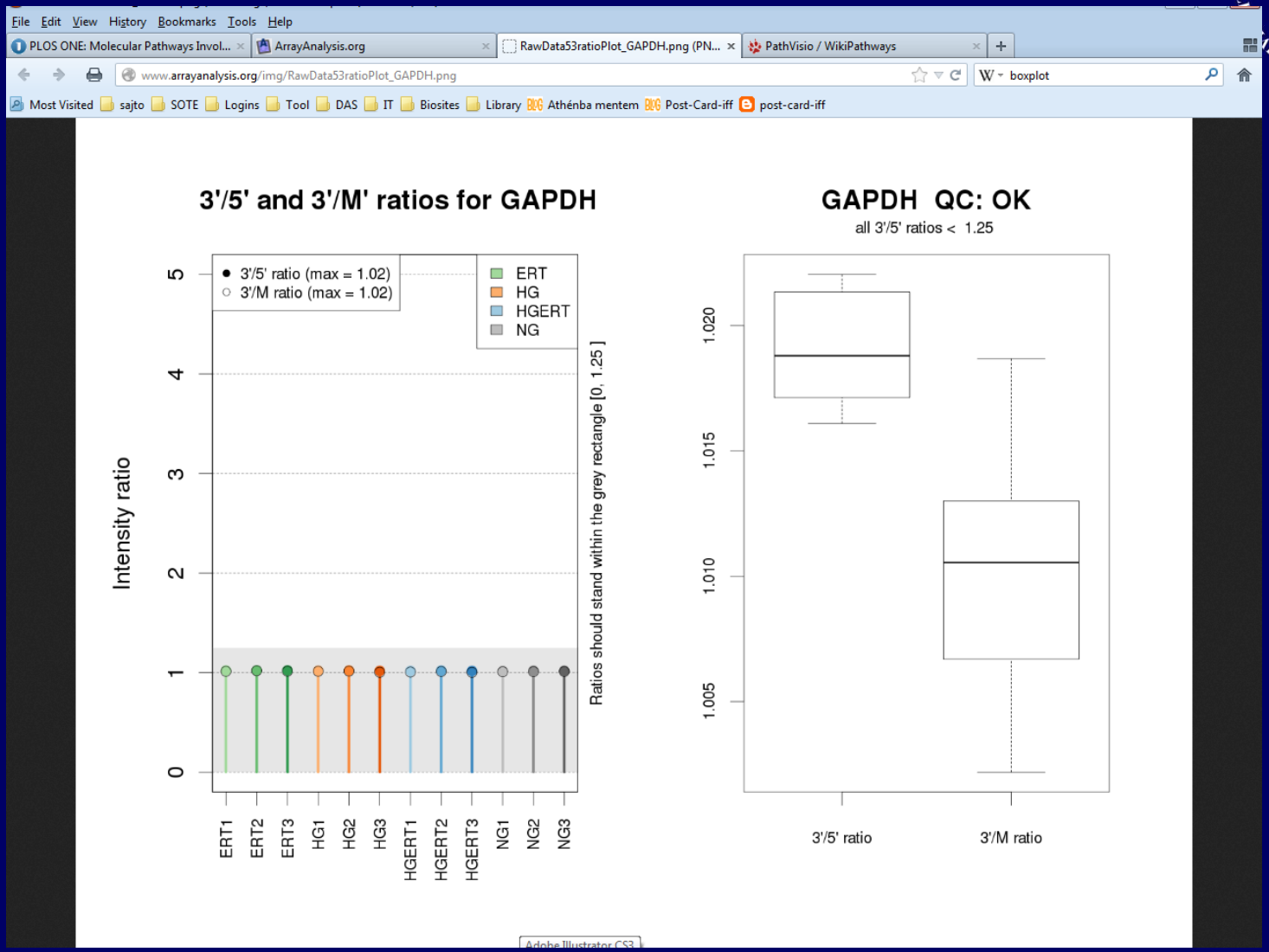
For an array of good quality, Affymetrix recommends that the 3'/5' ratio should not exceed :
- 3 for beta-actin
- 1.25 for GAPDH

These values were set for human samples so the ratio may be slightly higher for other species.

The QC pipeline proposes the following graphs to evaluate these controls:




File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawData53ratioPlot_beta-actin.png (...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawData53ratioPlot_beta-actin.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
3'/5' and 3'/M' ratios for beta-actin
3'/5' ratio (max = 1.05)
3'/M ratio (max = 1.11)
ERT
HG
HGERT
NG
Intensity ratio
Ratios should stand within the grey rectangle [0, 3]
ERT1 ERT2 ERT3 HG1 HG2 HG3 HGERT1 HGERT2 HGERT3 NG1 NG2 NG3
beta-actin QC: OK
all 3'/5' ratios < 3
3'/5' ratio
3'/M ratio


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawData53ratioPlot_GAPDH.png (PN...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawData53ratioPlot_GAPDH.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
3'/5' and 3'/M' ratios for GAPDH
3'/5' ratio (max = 1.02)
3'/M ratio (max = 1.02)
ERT
HG
HGERT
NG
Intensity ratio
Ratios should stand within the grey rectangle [0, 1.25]
ERT1 ERT2 ERT3 HG1 HG2 HG3 HGERT1 HGERT2 HGERT3 NG1 NG2 NG3
GAPDH QC: OK
all 3'/5' ratios < 1.25
1.020
1.015
1.010
1.005
3'/5' ratio
3'/M ratio

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

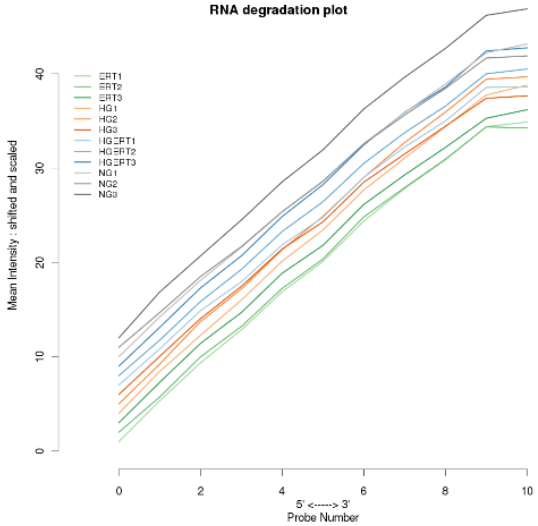
[Pipeline Description]

- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

BiGCaT bioinformatics

Overall RNA quality control: RNA degradation plot

In Affymetrix arrays, a probe-set is dedicated to each target. A probe-set is composed by several probes (classically 11) all targeting the probe-set target sequence. The RNA degradation plot proposes to plot the average intensity of each probes across all probe-sets, ordered from the 5' to the 3' end. Indeed since RNA degradation starts from the 5' end of the molecule, we would expect probe intensities to be globally lowered at that end of a probe set when compared to the 3' end. The RNA degradation plot aims at visualizing this trend. RNA which is too degraded will have a very high slope from 5' to 3'. The standardized slope of the curves is thus used as a quantitative indicator of the RNA degradation. An array with unexpected degradation is identified because it has a bigger slope and should stand out.



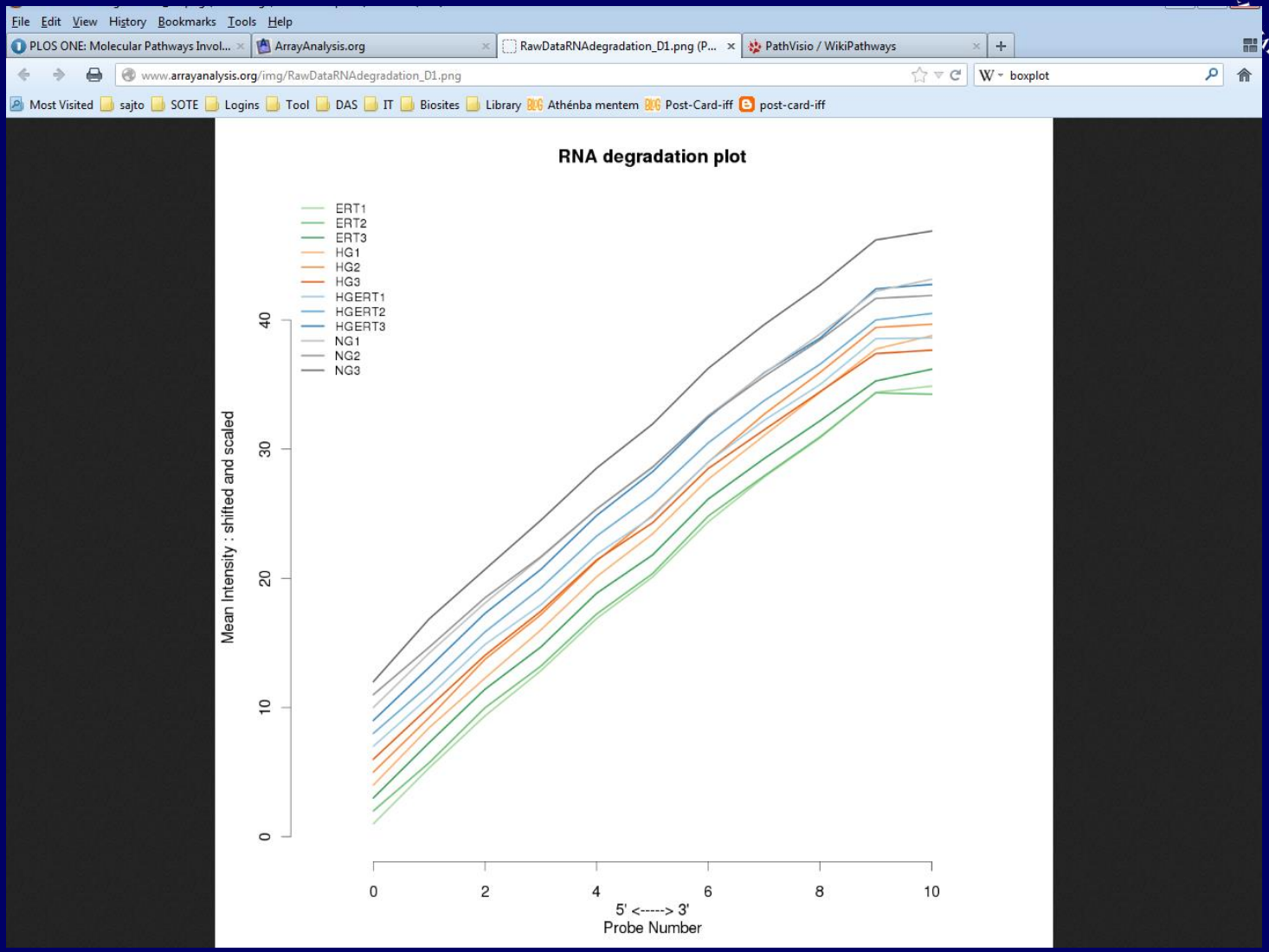
Each line corresponds to an array. In this example, all arrays give curve with a regular slope, exept between the 10th and the 11th probes.

[Technical documentation of the function]

[Top]

Hybridization and overall signal quality

Hybridization spike-in controls: BioB, BioC, BioD and CreX


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataRNAdegradation_D1.png (P...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataRNAdegradation_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
RNA degradation plot
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
Mean Intensity : shifted and scaled
0
10
20
30
40
0
2
4
6
8
10
5' <-----> 3'
Probe Number

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

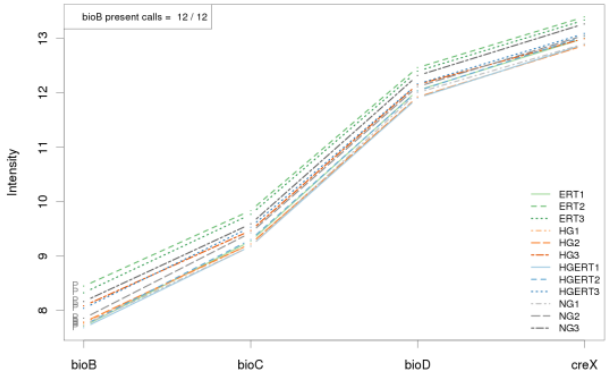
- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

BiGCaT bioinformatics

Hybridization spike-in controls: BioB, BioC, BioD and CreX

Affymetrix arrays include spike-in hybridization controls: 4 targets are spiked before the labeling step with 4 different concentrations; from the lowest to the highest: BioB, BioC, BioD and CreX. bioB, bioC and bioD are genes in the biotin synthesis pathway of E. coli, and cre is the recombinase gene from P1 bacteriophage and are not expected to cross-hybrid with non-bacterial and non-viral samples. Intensity pattern for these 4 controls should show the increase in target concentration. Other patterns would be a sign of bad hybridization.
If BioB is not flagged "present", this would also be a sign of bad quality, indicating that the sensitivity may not be sufficient for the array.

Spike-in controls intensities and calls



Spike-in calls: OK (all bioB are called present)

Each line corresponds to an array. The legend at the top left corner indicated the number of present calls among all BioB probe-sets. The text below the graph concludes about these calls. In this example, all arrays have a BioB probe-set called present.

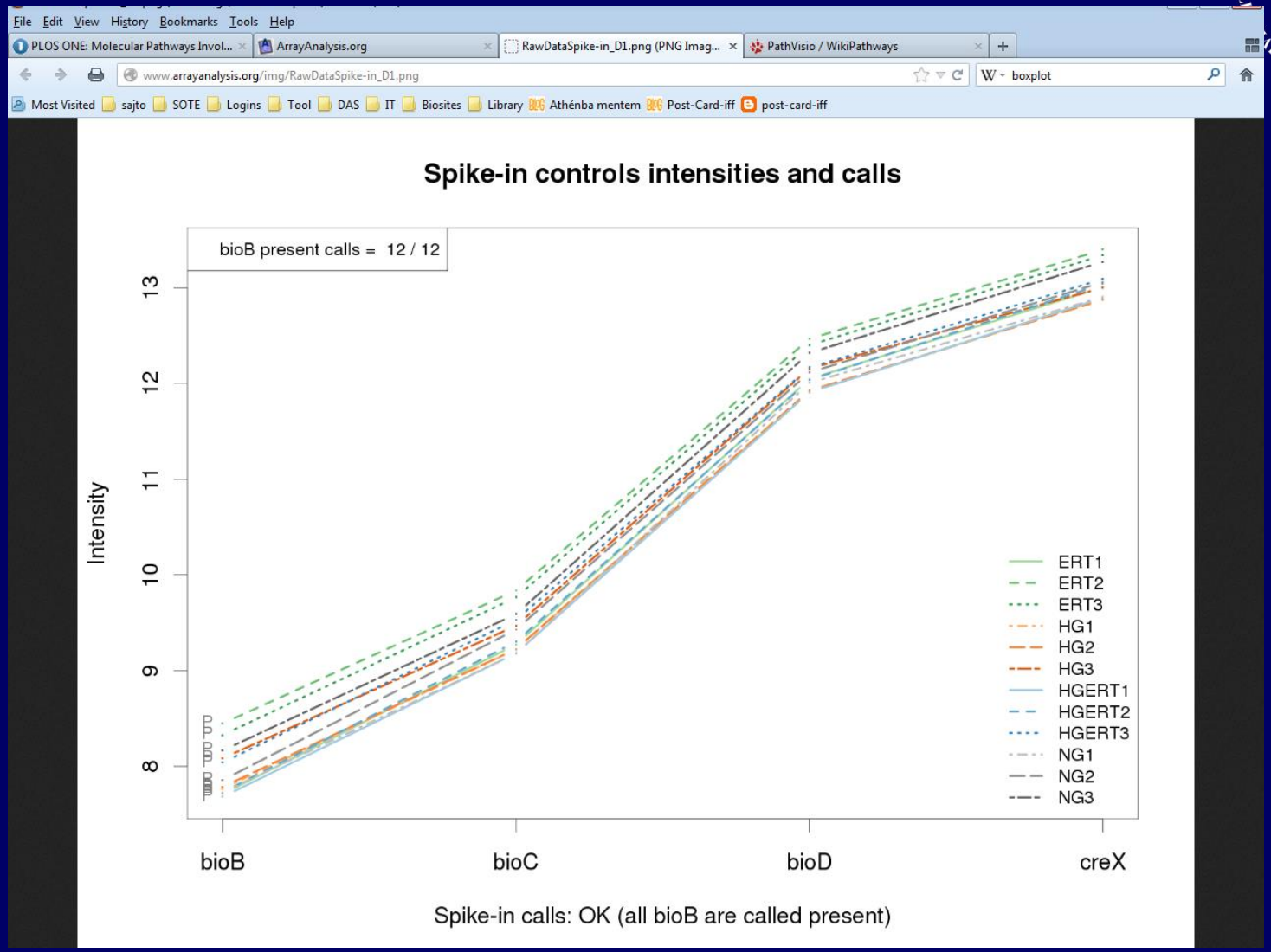
[Technical documentation of the function]

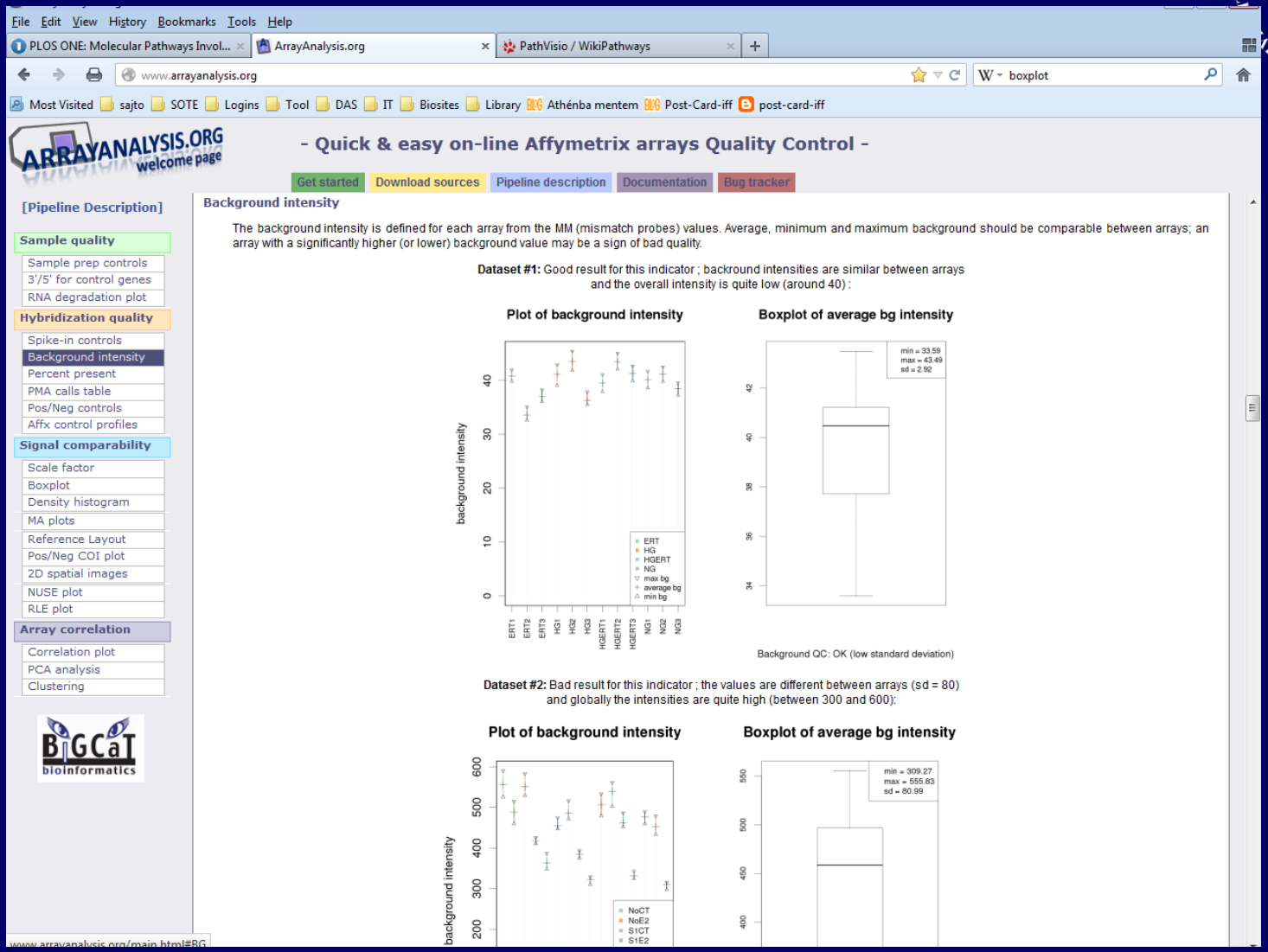
Background intensity

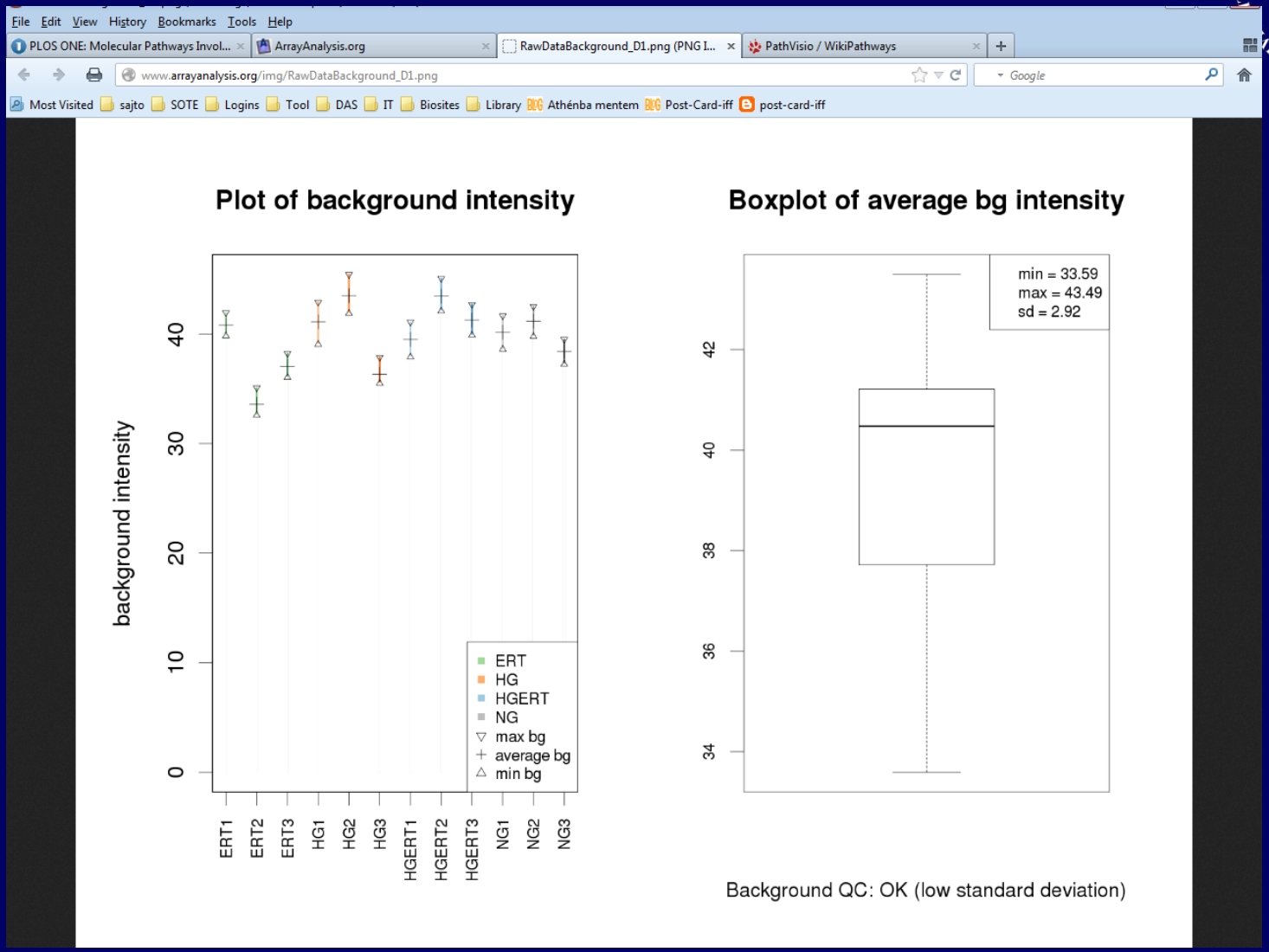
The background intensity is defined for each array from the MM (mismatch probes) values. Average, minimum and maximum background should be comparable between arrays; an array with a significantly higher (or lower) background value may be a sign of bad quality.

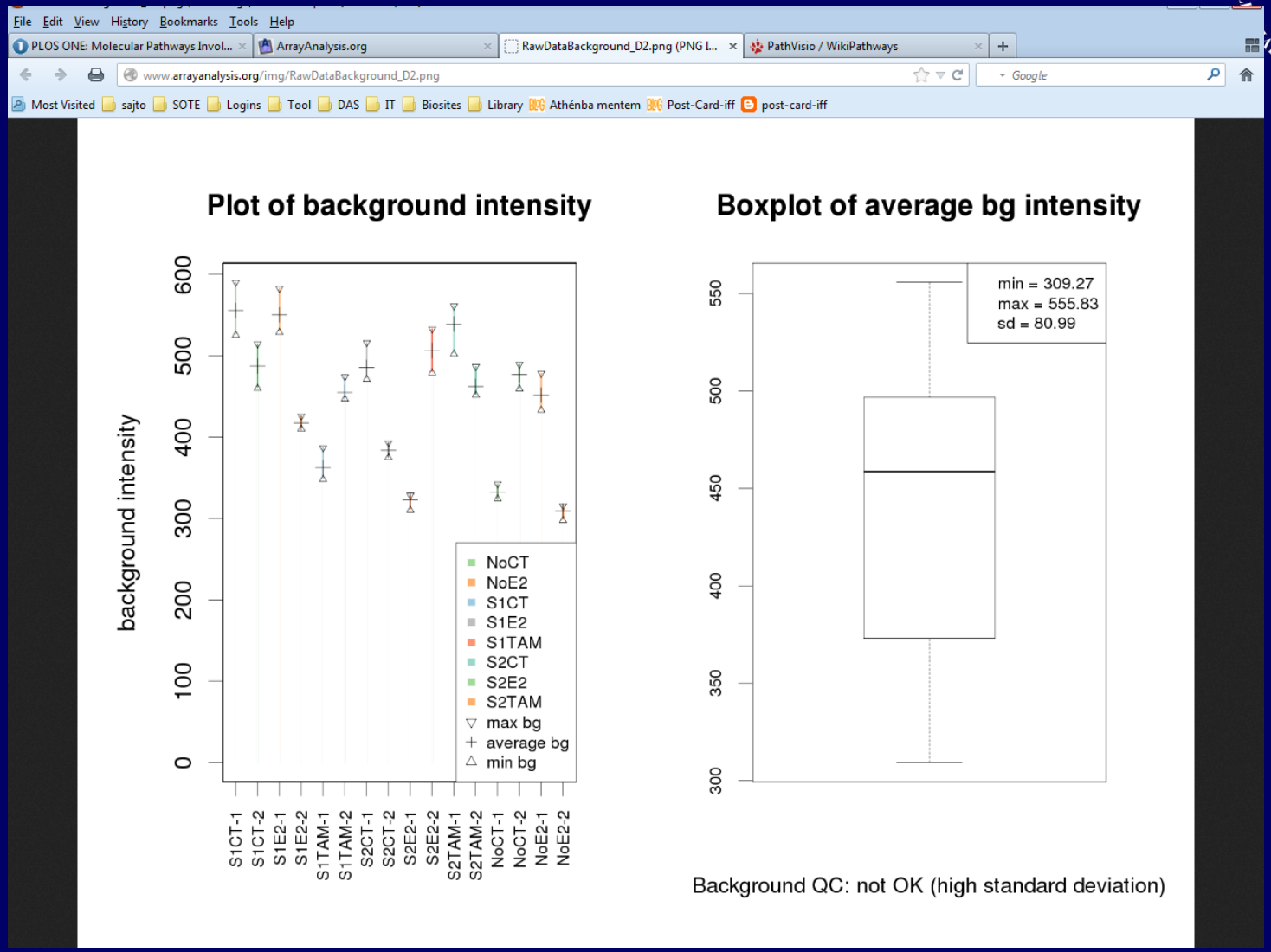
Dataset #1: Good result for this indicator ; backround intensities are similar between arrays and the overall intensity is quite low (around 40) :

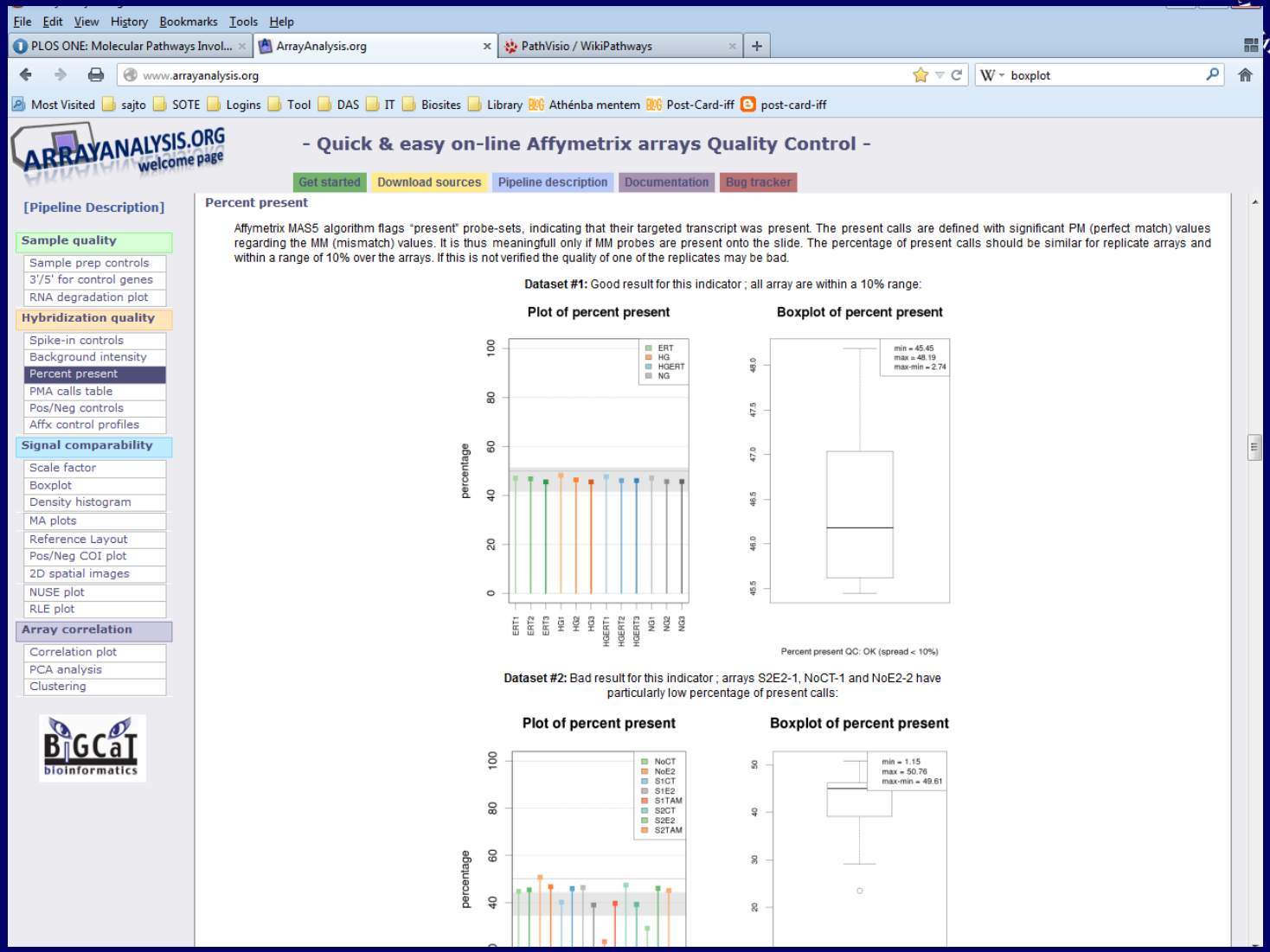
Plot of background intensity | **Boxplot of average bg intensity**

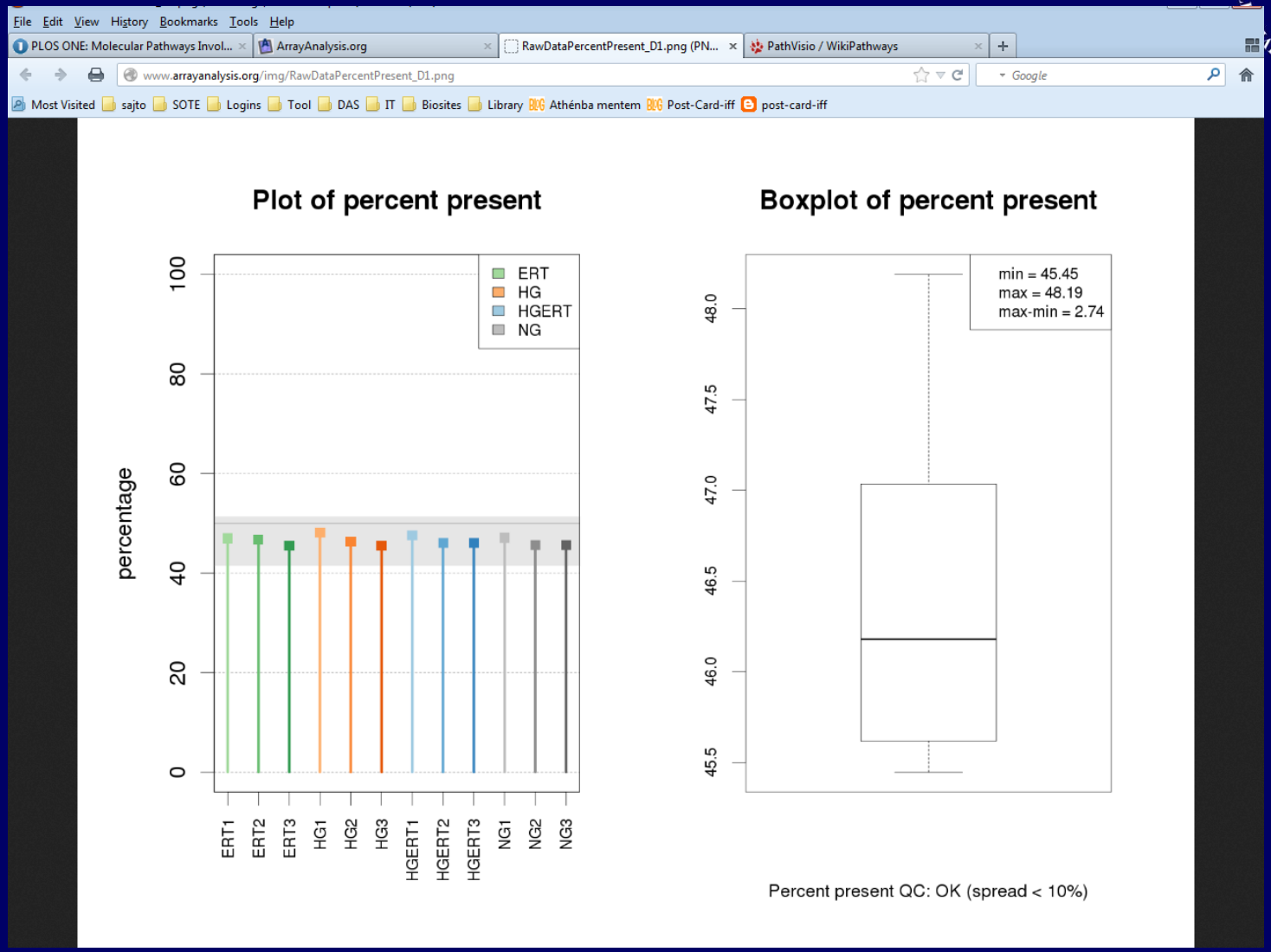

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataSpike-in_D1.png (PNG Imag...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataSpike-in_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Spike-in controls intensities and calls
bioB present calls = 12 / 12
Intensity
8
9
10
11
12
13
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
bioB
bioC
bioD
creX
Spike-in calls: OK (all bioB are called present)

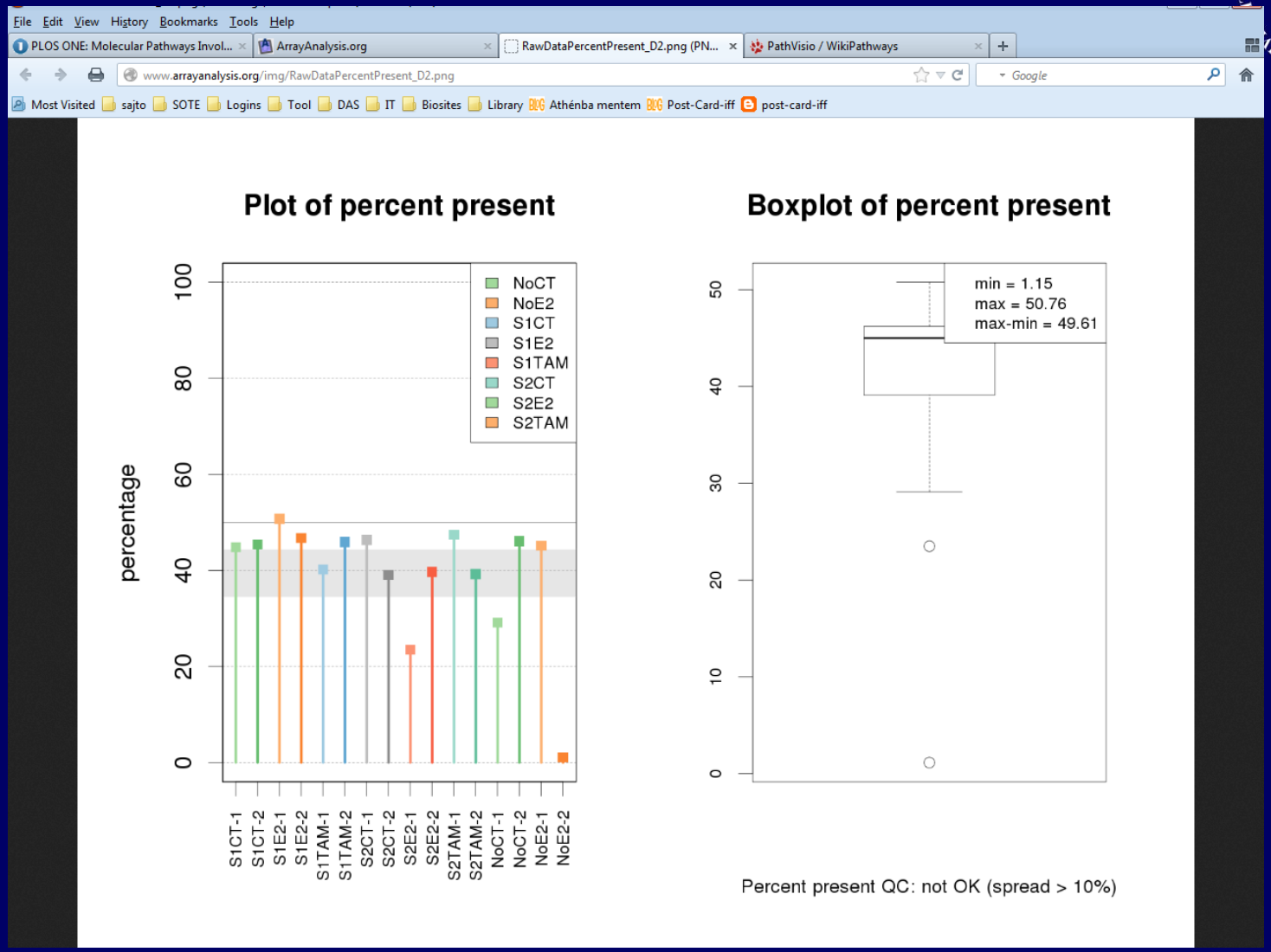

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
PathVisio / WikiPathways
www.arrayanalysis.org
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
ARRAYANALYSIS.ORG welcome page
- Quick & easy on-line Affymetrix arrays Quality Control -
Get started
Download sources
Pipeline description
Documentation
Bug tracker
[Pipeline Description]
Sample quality
Sample prep controls
3'/5' for control genes
RNA degradation plot
Hybridization quality
Spike-in controls
Background intensity
Percent present
PMA calls table
Pos/Neg controls
Affx control profiles
Signal comparability
Scale factor
Boxplot
Density histogram
MA plots
Reference Layout
Pos/Neg COI plot
2D spatial images
NUSE plot
RLE plot
Array correlation
Correlation plot
PCA analysis
Clustering
BiGCaT bioinformatics
Background intensity
The background intensity is defined for each array from the MM (mismatch probes) values. Average, minimum and maximum background should be comparable between arrays; an array with a significantly higher (or lower) background value may be a sign of bad quality.
Dataset #1: Good result for this indicator ; backround intensities are similar between arrays and the overall intensity is quite low (around 40) :
Plot of background intensity
Boxplot of average bg intensity
background intensity
ERT
HG
HGERT
NG
max bg
average bg
min bg
min = 33.59
max = 43.49
sd = 2.92
Background QC: OK (low standard deviation)
Dataset #2: Bad result for this indicator ; the values are different between arrays (sd = 80) and globally the intensities are quite high (between 300 and 600):
Plot of background intensity
Boxplot of average bg intensity
NoCT
NoE2
S1CT
S1E2
min = 309.27
max = 555.83
sd = 80.99


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataBackground_D1.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataBackground_D1.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Plot of background intensity
background intensity
0
10
20
30
40
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
ERT
HG
HGERT
NG
max bg
average bg
min bg
Boxplot of average bg intensity
min = 33.59
max = 43.49
sd = 2.92
34
36
38
40
42
Background QC: OK (low standard deviation)


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataBackground_D2.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataBackground_D2.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Plot of background intensity
background intensity
0
100
200
300
400
500
600
S1CT-1
S1CT-2
S1E2-1
S1E2-2
S1TAM-1
S1TAM-2
S2CT-1
S2CT-2
S2E2-1
S2E2-2
S2TAM-1
S2TAM-2
NoCT-1
NoCT-2
NoE2-1
NoE2-2
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
max bg
average bg
min bg
Boxplot of average bg intensity
min = 309.27
max = 555.83
sd = 80.99
300
350
400
450
500
550
Background QC: not OK (high standard deviation)


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
PathVisio / WikiPathways
www.arrayanalysis.org
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
ARRAYANALYSIS.ORG welcome page
- Quick & easy on-line Affymetrix arrays Quality Control -
Get started
Download sources
Pipeline description
Documentation
Bug tracker
[Pipeline Description]
Sample quality
Sample prep controls
3'/5' for control genes
RNA degradation plot
Hybridization quality
Spike-in controls
Background intensity
Percent present
PMA calls table
Pos/Neg controls
Affx control profiles
Signal comparability
Scale factor
Boxplot
Density histogram
MA plots
Reference Layout
Pos/Neg COI plot
2D spatial images
NUSE plot
RLE plot
Array correlation
Correlation plot
PCA analysis
Clustering
BiGCaT bioinformatics
Percent present
Affymetrix MAS5 algorithm flags "present" probe-sets, indicating that their targeted transcript was present. The present calls are defined with significant PM (perfect match) values regarding the MM (mismatch) values. It is thus meaningfull only if MM probes are present onto the slide. The percentage of present calls should be similar for replicate arrays and within a range of 10% over the arrays. If this is not verified the quality of one of the replicates may be bad.
Dataset #1: Good result for this indicator ; all array are within a 10% range:
Plot of percent present
Boxplot of percent present
ERT
HG
HGERT
NG
percentage
min = 45.45
max = 48.19
max-min = 2.74
Percent present QC: OK (spread < 10%)
Dataset #2: Bad result for this indicator ; arrays S2E2-1, NoCT-1 and NoE2-2 have particularly low percentage of present calls:
Plot of percent present
Boxplot of percent present
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
min = 1.15
max = 50.76
max-min = 49.61


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPercentPresent_D1.png (PN...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPercentPresent_D1.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Plot of percent present
ERT
HG
HGERT
NG
percentage
0
20
40
60
80
100
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
Boxplot of percent present
min = 45.45
max = 48.19
max-min = 2.74
45.5
46.0
46.5
47.0
47.5
48.0
Percent present QC: OK (spread < 10%)


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPercentPresent_D2.png (PN...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPercentPresent_D2.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Plot of percent present
percentage
0
20
40
60
80
100
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
S1CT-1
S1CT-2
S1E2-1
S1E2-2
S1TAM-1
S1TAM-2
S2CT-1
S2CT-2
S2E2-1
S2E2-2
S2TAM-1
S2TAM-2
NoCT-1
NoCT-2
NoE2-1
NoE2-2
Boxplot of percent present
min = 1.15
max = 50.76
max-min = 49.61
0
10
20
30
40
50
Percent present QC: not OK (spread > 10%)

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

BiGCaT bioinformatics

Present/Margin/Absent (PMA) table

Optionaly, and for arrays containing perfect match (PM) and mismatch (MM) probes, we propose to create the table of the probeset present calls: "P" for Present, "M" for Margin and "A" for absent. See description of Percent Present plot for explanations about the compuation of these values. The output table is a text file (.txt), each column is tab-separated so it is easily open with any spredsheet editor like Excel:

ENSG_ID	Array1.present	Array2.present	Array3.present	Array4.present	Array5.present	Array6.present
ENSG00000000005	A	A	A	A	A	A
ENSG00000000419	P	P	P	P	P	P
ENSG00000000460	P	P	P	P	P	P
ENSG00000000938	A	A	A	A	A	A
ENSG00000000971	A	A	A	A	A	A
ENSG00000001036	P	P	P	P	P	P
ENSG00000001084	P	P	P	P	P	P
ENSG00000001167	P	P	P	P	P	P
ENSG00000001460	A	A	A	A	P	A
ENSG00000001461	P	P	P	P	P	P
ENSG00000001497	P	P	P	P	A	P
ENSG00000001561	P	P	P	P	P	P
ENSG00000001617	P	P	P	P	P	P
ENSG00000001626	A	A	A	A	A	A
ENSG00000001629	P	P	P	P	P	P
ENSG00000001631	A	P	A	A	A	A
ENSG00000002016	M	P	A	A	A	P
ENSG00000002330	P	P	P	P	P	P
ENSG00000002549	P	P	P	P	P	P
ENSG00000002586	P	P	P	P	P	P
ENSG00000002587	A	A	A	A	A	A
ENSG00000002726	A	M	A	A	P	A
ENSG00000002745	A	A	A	A	A	A
ENSG00000002746	P	A	A	A	P	A
ENSG00000002822	P	P	P	P	P	P
ENSG00000002834	P	P	P	P	P	P
ENSG00000002919	P	P	P	P	P	P
ENSG00000002933	A	A	A	A	A	A
ENSG00000003056	P	P	P	P	P	P
ENSG00000003096	P	P	P	P	P	P
ENSG00000003137	A	A	A	A	A	A
ENSG00000003147	P	P	P	P	P	P
ENSG00000003249	P	P	P	P	P	P
ENSG00000003393	P	P	P	P	P	P
ENSG00000003400	A	A	A	A	A	A

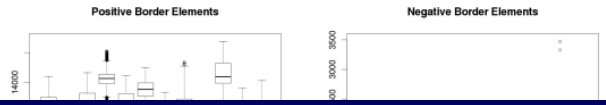
This table may be useful for further investigation on particular genes.

[Technical documentation of the function]

Positive and Negative control distribution

Affymetrix arrays contain border elements, positioned on the outer edge of the array. They serve for the automatic setting of the grids but also as controls for signal intensity. For each array the intensities for all border elements are collected. Elements with an intensity greater the 1.2 times the mean for that group are assumed to be positive controls. Elements with a signal less that 0.8 of the mean are assumed to be negative controls. Elements falling in between these cut offs are not used in further calculations. This graph presents boxplots of the positive and negative distribution. The means and spread of positive elements should be comparable between arrays. Dissimilarity can arise either from non-uniform hybridization or gridding problems. The negative elements represent spots with no hybridization signal, so they are expected to be close to 0. Boxplots that are strongly elevated relative to all the others reflect a higher background level.

Dataset #1: Good result for this indicator ; positive and negative controls are in separate intensity ranges. Negative controls have similar distributions between arrays:



File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Involl... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

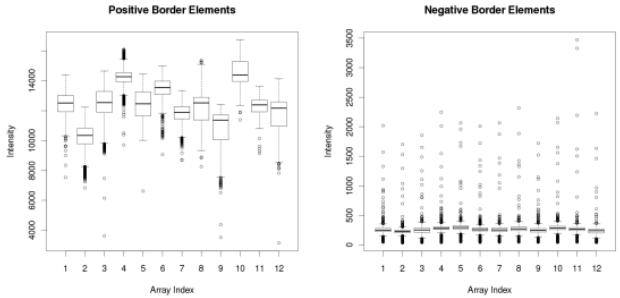
Array correlation
- Correlation plot
- PCA analysis
- Clustering

BiGCaT bioinformatics

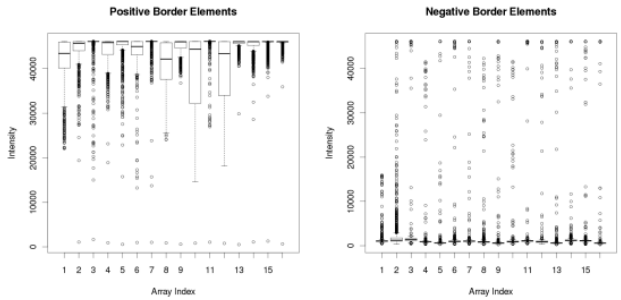
Positive and Negative control distribution

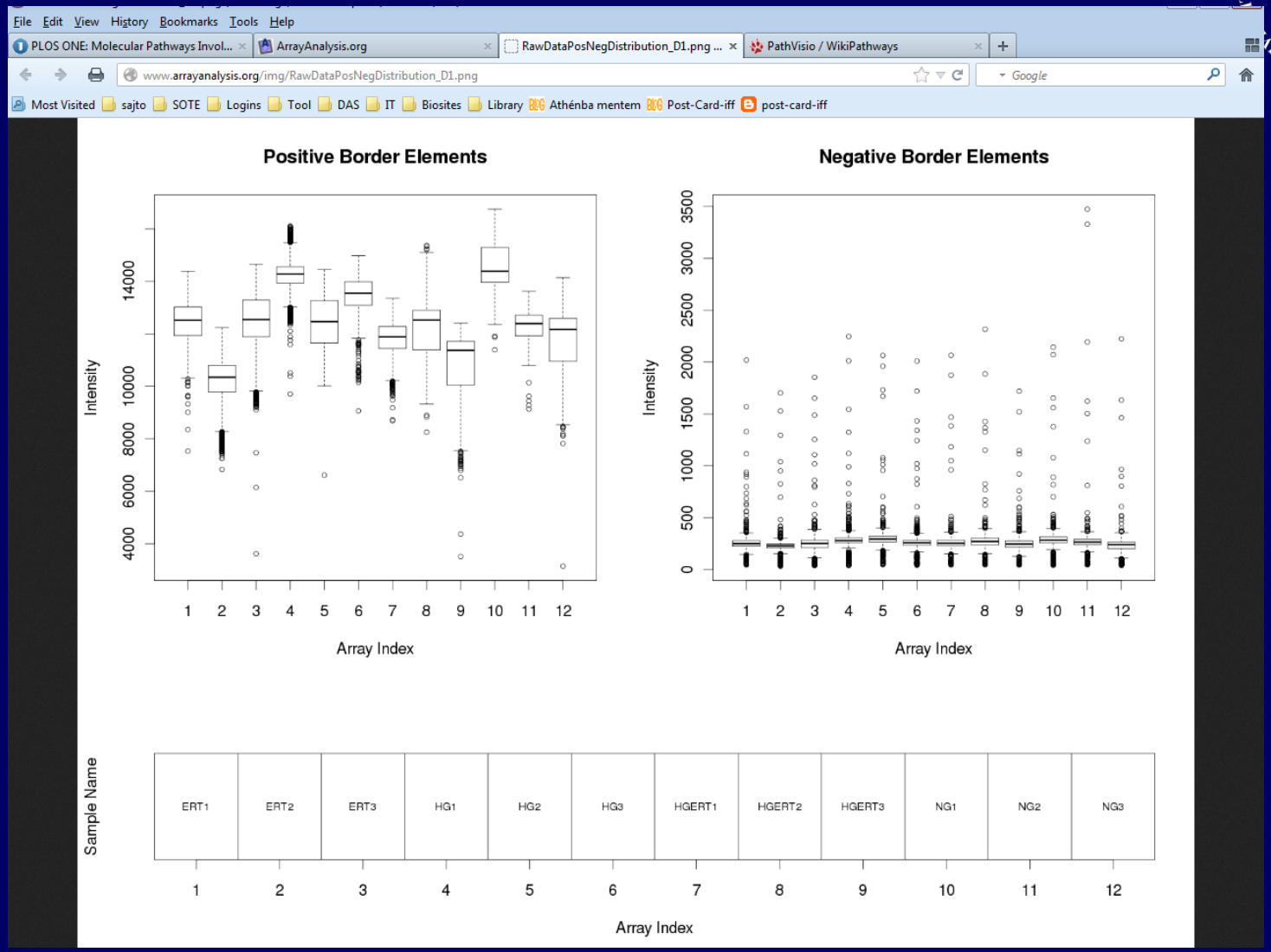
Affymetrix arrays contain border elements, positioned on the outer edge of the array. They serve for the automatic setting of the grids but also as controls for signal intensity. For each array the intensities for all border elements are collected. Elements with an intensity greater the 1.2 times the mean for that group are assumed to be positive controls. Elements with a signal less that 0.8 of the mean are assumed to be negative controls. Elements falling in between these cut offs are not used in further calculations. This graph presents boxplots of the positive and negative distribution. The means and spread of positive elements should be comparable between arrays. Dissimilarity can arise either from non-uniform hybridization or gridding problems. The negative elements represent spots with no hybridization signal, so they are expected to be close to 0. Boxplots that are strongly elevated relative to all the others reflect a higher background level.

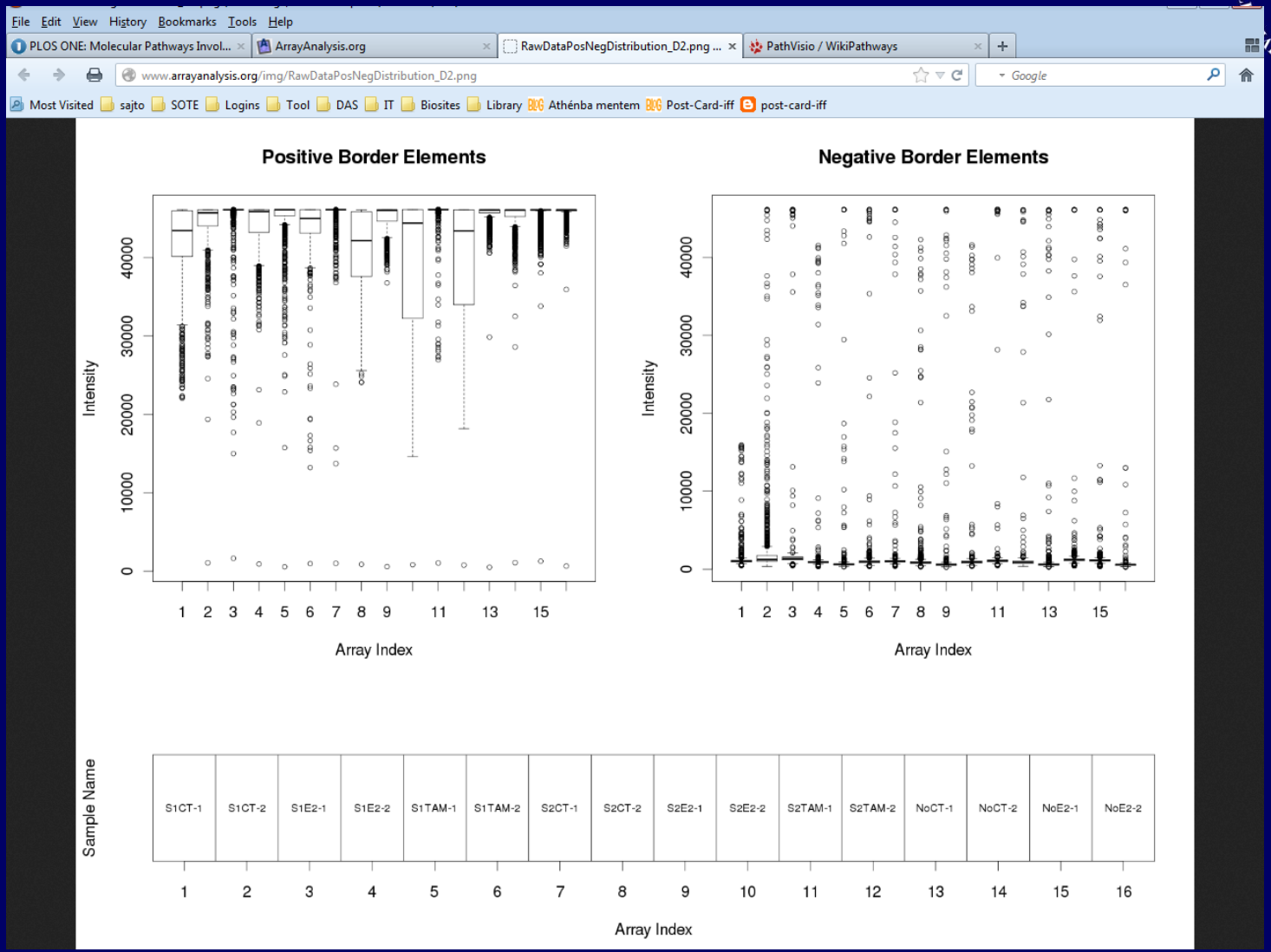
Dataset #1: Good result for this indicator ; positive and negative controls are in separate intensity ranges. Negative controls have similar distributions between arrays:



Dataset #2: Bad result for this indicator ; positive controls reach the saturation level (50,000) and negative controls are spread between 0 and 50,000. Array 2 has significatively higher negative controls.




File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPosNegDistribution_D1.png ...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPosNegDistribution_D1.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Positive Border Elements
Intensity
4000 6000 8000 10000 14000
1 2 3 4 5 6 7 8 9 10 11 12
Array Index
Negative Border Elements
Intensity
0 500 1000 1500 2000 2500 3000 3500
1 2 3 4 5 6 7 8 9 10 11 12
Array Index
Sample Name
ERT1 ERT2 ERT3 HG1 HG2 HG3 HGERT1 HGERT2 HGERT3 NG1 NG2 NG3
1 2 3 4 5 6 7 8 9 10 11 12
Array Index


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPosNegDistribution_D2.png ...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPosNegDistribution_D2.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Positive Border Elements
Negative Border Elements
Intensity
0 10000 20000 30000 40000
1 2 3 4 5 6 7 8 9 11 13 15
Array Index
Sample Name
S1CT-1 S1CT-2 S1E2-1 S1E2-2 S1TAM-1 S1TAM-2 S2CT-1 S2CT-2 S2E2-1 S2E2-2 S2TAM-1 S2TAM-2 NoCT-1 NoCT-2 NoE2-1 NoE2-2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Array Index

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

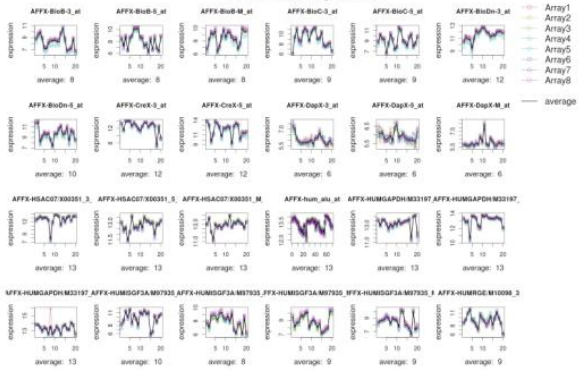
Array correlation
- Correlation plot
- PCA analysis
- Clustering

BiGCaT bioinformatics

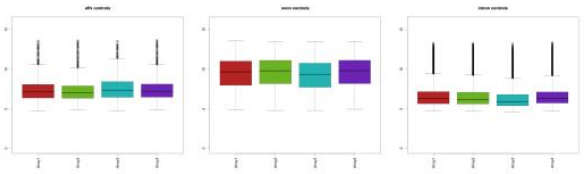
Profiles and boxplots of all controls (AFFX, INTRON, EXON)

Affymetrix arrays contain several control probesets, most of them annotated with the "AFFX" prefix. This is mainly the case for the spike-in sample prep and hybridization controls, GAPDH and beta-actin 3'/5' controls, but other control probesets may be present, depending of your array type. The first image contains a list of graphs, one per "AFFX"-annotated probeset, plotting log-intensity profiles of all probes for all arrays. Oulier arrays may have different intensity profiles compared to other arrays. The general tendance is represented by an average curve, plotted in black. Some control probesets contain too many probes so the profile plot cannot be read properly and used for QC.

affx control profiles



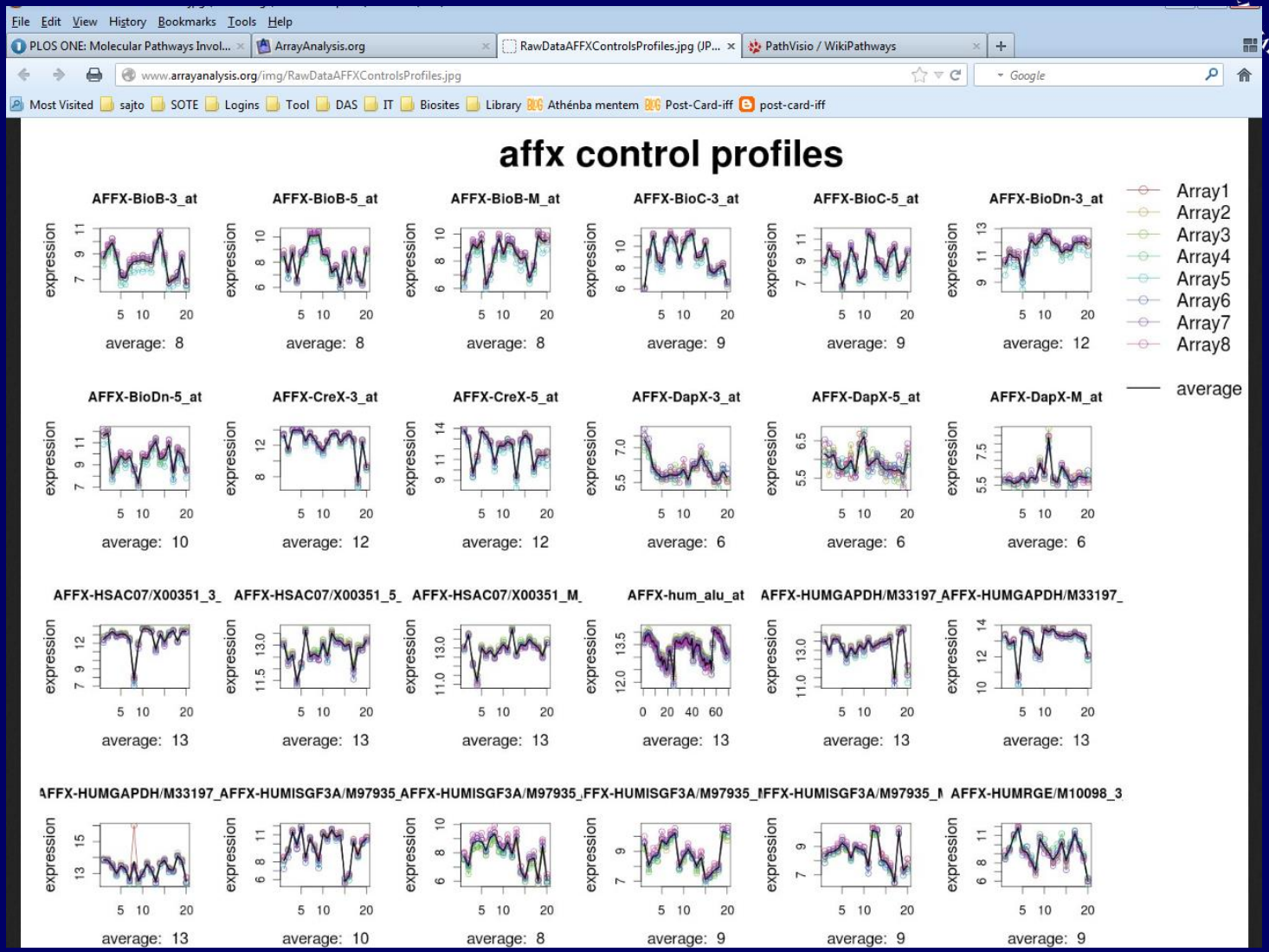
The second image is a boxplot summarizing all AFFX control log-intensities for each array. Other boxplots may be generated, depending on your array type, to present EXON and INTRON controls.

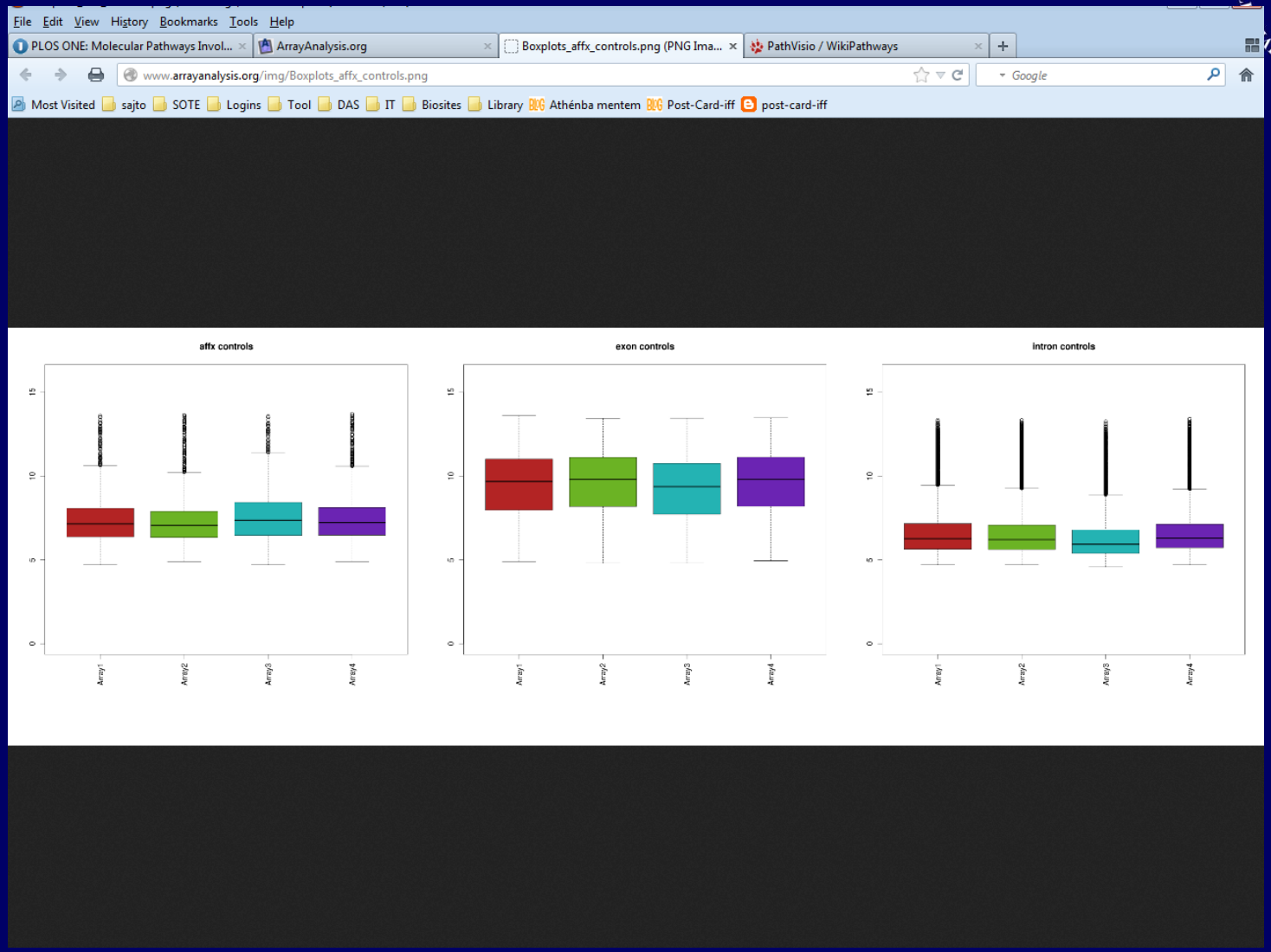


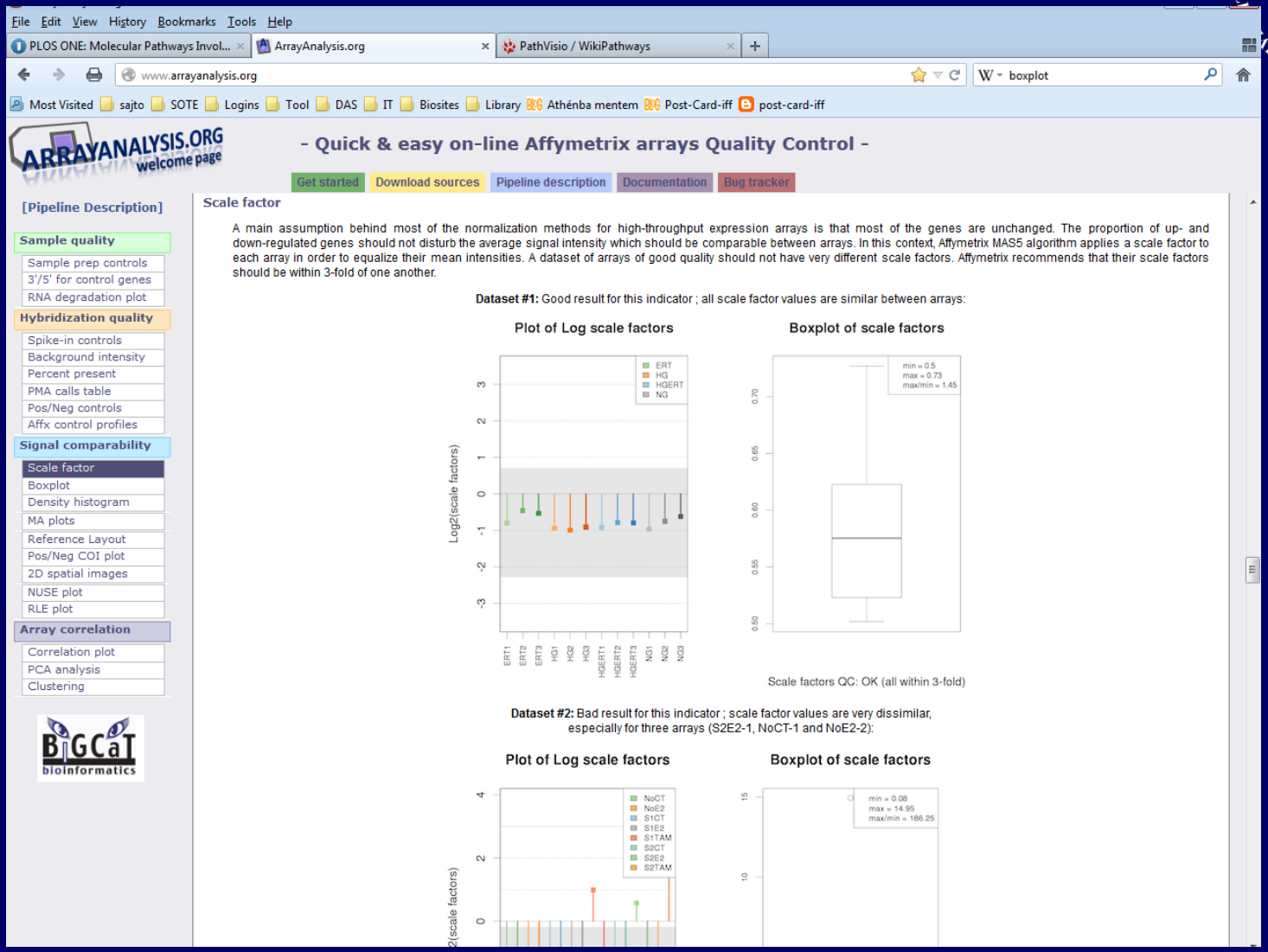
[Technical documentation of the function]

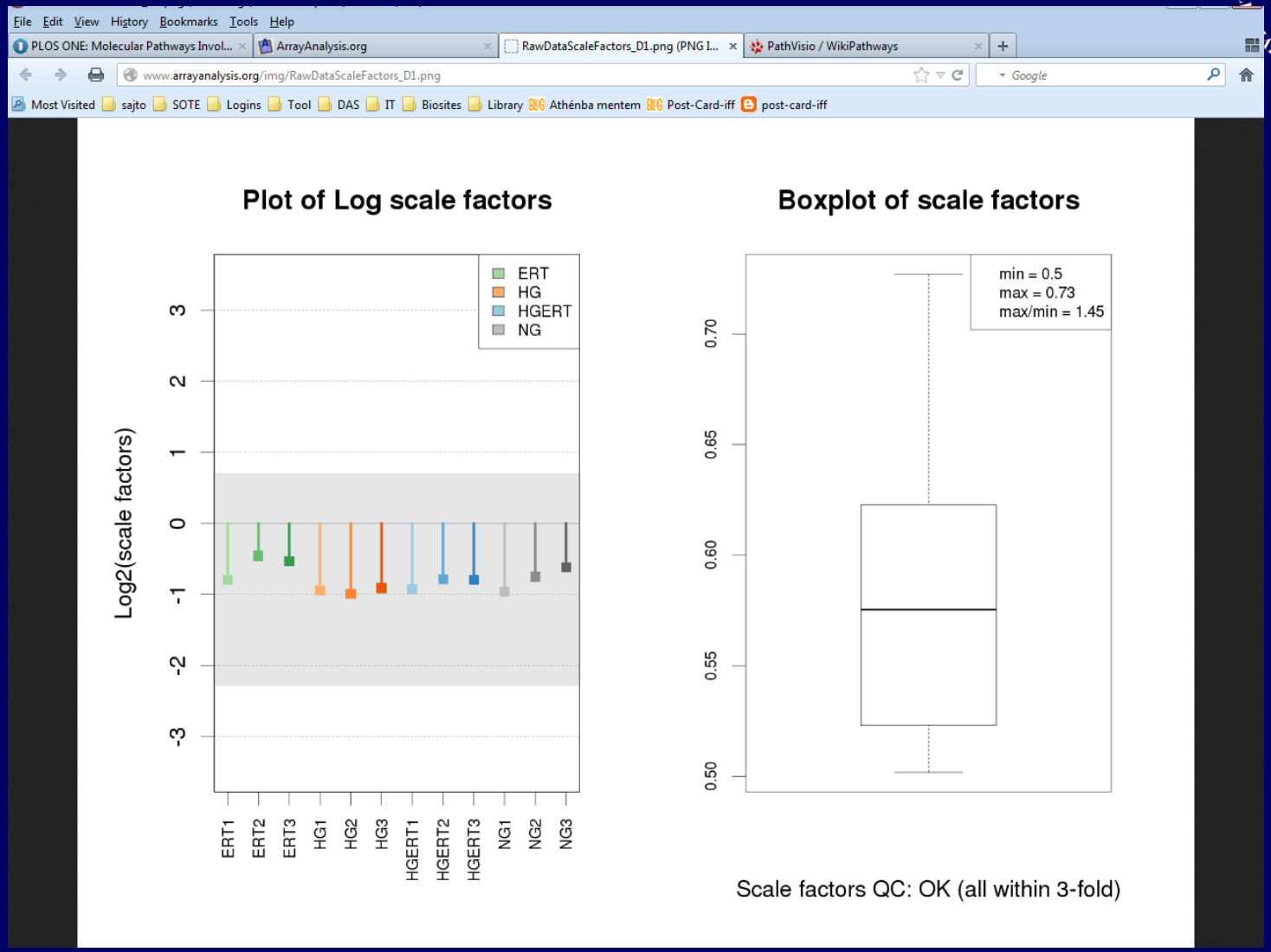
[Top]

Signal comparability and biases diagnostic


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataAFFXControlsProfiles.jpg (JP...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataAFFXControlsProfiles.jpg
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
affx control profiles
AFFX-BioB-3_at
AFFX-BioB-5_at
AFFX-BioB-M_at
AFFX-BioC-3_at
AFFX-BioC-5_at
AFFX-BioDn-3_at
AFFX-BioDn-5_at
AFFX-CreX-3_at
AFFX-CreX-5_at
AFFX-DapX-3_at
AFFX-DapX-5_at
AFFX-DapX-M_at
AFFX-HSAC07/X00351_3_
AFFX-HSAC07/X00351_5_
AFFX-HSAC07/X00351_M_
AFFX-hum_alu_at
AFFX-HUMGAPDH/M33197_
AFFX-HUMGAPDH/M33197_
AFFX-HUMGAPDH/M33197_
AFFX-HUMISGF3A/M97935_
AFFX-HUMISGF3A/M97935_
AFFX-HUMISGF3A/M97935_
AFFX-HUMISGF3A/M97935_
AFFX-HUMRGE/M10098_3
expression
average: 8
average: 8
average: 8
average: 9
average: 9
average: 12
average: 10
average: 12
average: 12
average: 6
average: 6
average: 6
average: 13
average: 13
average: 13
average: 13
average: 13
average: 13
average: 13
average: 10
average: 8
average: 9
average: 9
average: 9
Array1
Array2
Array3
Array4
Array5
Array6
Array7
Array8
average


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
Boxplots_affx_controls.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/Boxplots_affx_controls.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
affx controls
exon controls
intron controls
Array1
Array2
Array3
Array4


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
PathVisio / WikiPathways
www.arrayanalysis.org
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
ARRAYANALYSIS.ORG welcome page
- Quick & easy on-line Affymetrix arrays Quality Control -
Get started
Download sources
Pipeline description
Documentation
Bug tracker
[Pipeline Description]
Sample quality
Sample prep controls
3'/5' for control genes
RNA degradation plot
Hybridization quality
Spike-in controls
Background intensity
Percent present
PMA calls table
Pos/Neg controls
Affx control profiles
Signal comparability
Scale factor
Boxplot
Density histogram
MA plots
Reference Layout
Pos/Neg COI plot
2D spatial images
NUSE plot
RLE plot
Array correlation
Correlation plot
PCA analysis
Clustering
BiGCaT bioinformatics
Scale factor
A main assumption behind most of the normalization methods for high-throughput expression arrays is that most of the genes are unchanged. The proportion of up- and down-regulated genes should not disturb the average signal intensity which should be comparable between arrays. In this context, Affymetrix MAS5 algorithm applies a scale factor to each array in order to equalize their mean intensities. A dataset of arrays of good quality should not have very different scale factors. Affymetrix recommends that their scale factors should be within 3-fold of one another.
Dataset #1: Good result for this indicator ; all scale factor values are similar between arrays:
Plot of Log scale factors
Boxplot of scale factors
Log2(scale factors)
ERT
HG
HGERT
NG
min = 0.5
max = 0.73
max/min = 1.45
Scale factors QC: OK (all within 3-fold)
Dataset #2: Bad result for this indicator ; scale factor values are very dissimilar, especially for three arrays (S2E2-1, NoCT-1 and NoE2-2):
Plot of Log scale factors
Boxplot of scale factors
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
min = 0.08
max = 14.95
max/min = 186.25


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataScaleFactors_D1.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataScaleFactors_D1.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Plot of Log scale factors
ERT
HG
HGERT
NG
Log2(scale factors)
3
2
1
0
-1
-2
-3
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
Boxplot of scale factors
min = 0.5
max = 0.73
max/min = 1.45
0.70
0.65
0.60
0.55
0.50
Scale factors QC: OK (all within 3-fold)


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataScaleFactors_D2.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataScaleFactors_D2.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Plot of Log scale factors
Log2(scale factors)
4
2
0
-2
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
S1CT-1
S1CT-2
S1E2-1
S1E2-2
S1TAM-1
S1TAM-2
S2CT-1
S2CT-2
S2E2-1
S2E2-2
S2TAM-1
S2TAM-2
NoCT-1
NoCT-2
NoE2-1
NoE2-2
Boxplot of scale factors
15
10
5
0
min = 0.08
max = 14.95
max/min = 186.25
Scale factors QC: not OK (more than 3-fold)

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

Array correlation
- Correlation plot
- PCA analysis
- Clustering

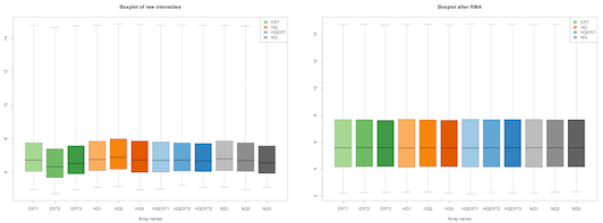
BiGCaT bioinformatics

Boxplots of log-intensities

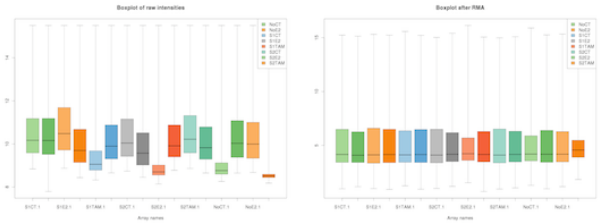
Boxplots of log-intensity distribution are plotted for between-array comparison. The distributions of raw PM (perfect match probes) log-intensities are not expected to be identical but still not totally different while the distributions of normalized (and summarized) probe-set log-intensities are expected to be more comparable if not identical (some normalization methods make the distributions even). Drawing these boxplots before and after normalization allows also checking the normalization step.

Left : Raw data // Right : Normalized data.

Dataset #1: Raw signal distributions are similar even before normalization:



Dataset #2: Four arrays have particularly low intensities (S1TAM-1, S2E2-1, NoCT-1 and NoE2-2). The normalization step does not manage to correct totally these differences:

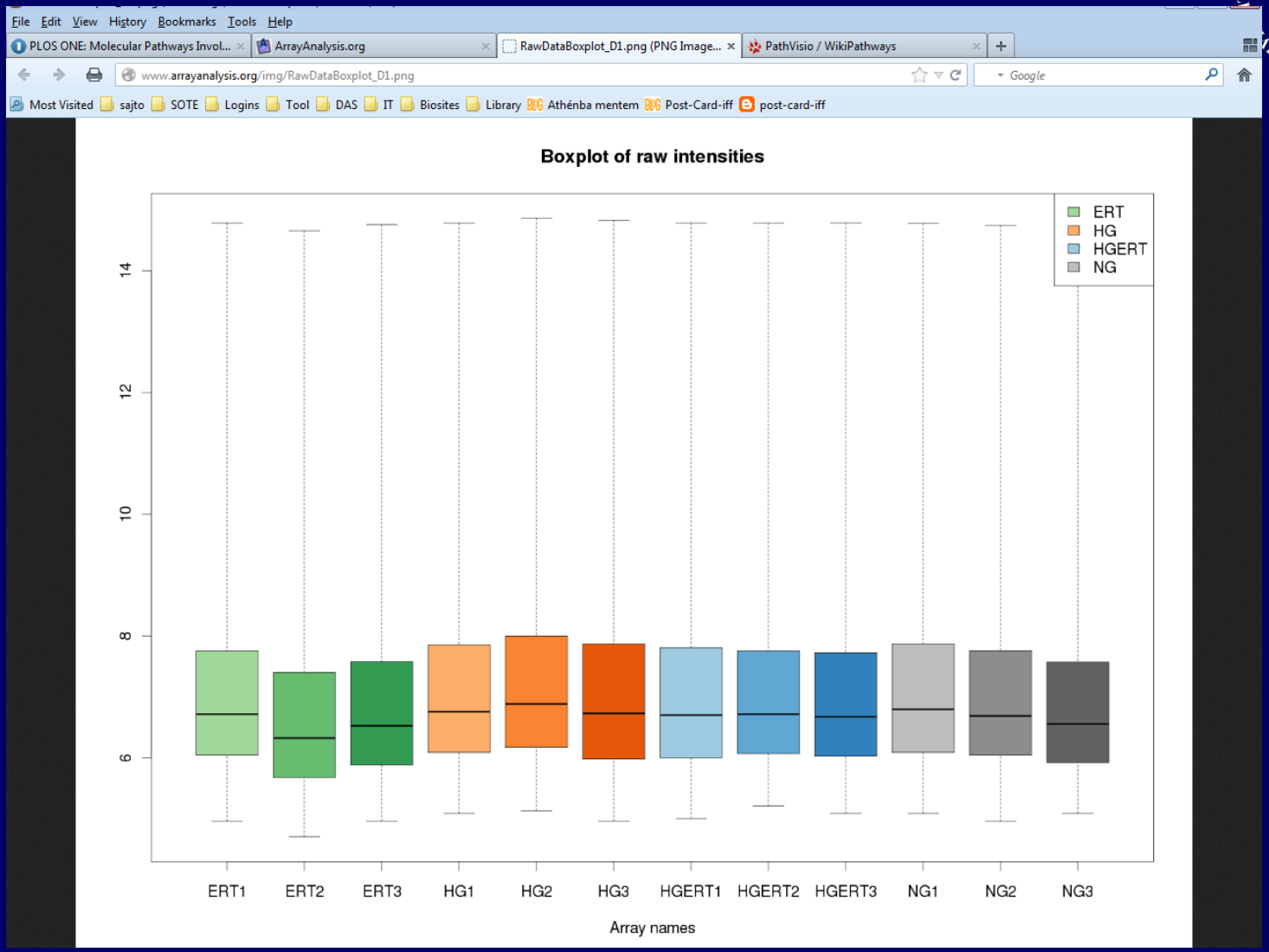


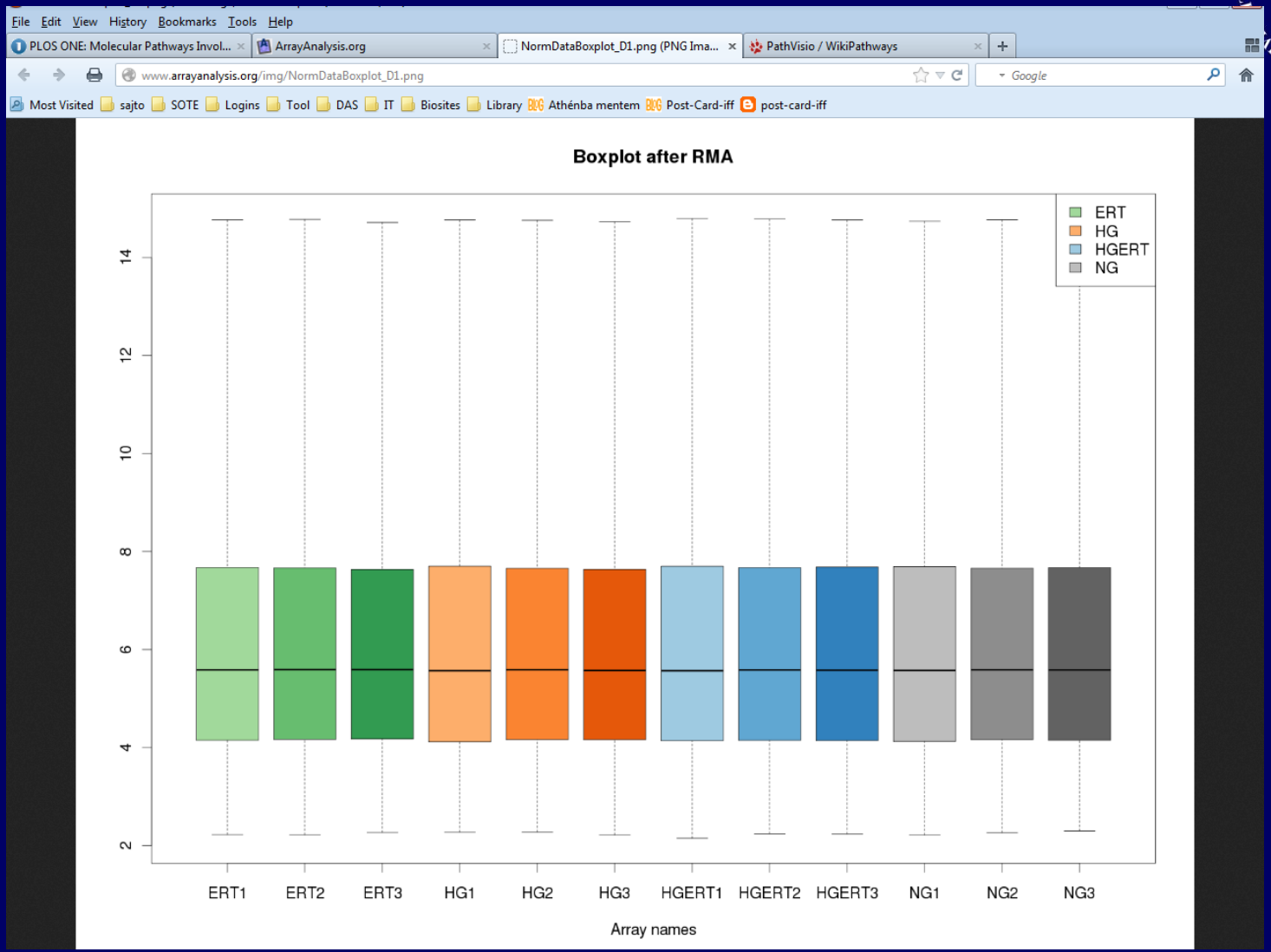
[Technical documentation of the function]

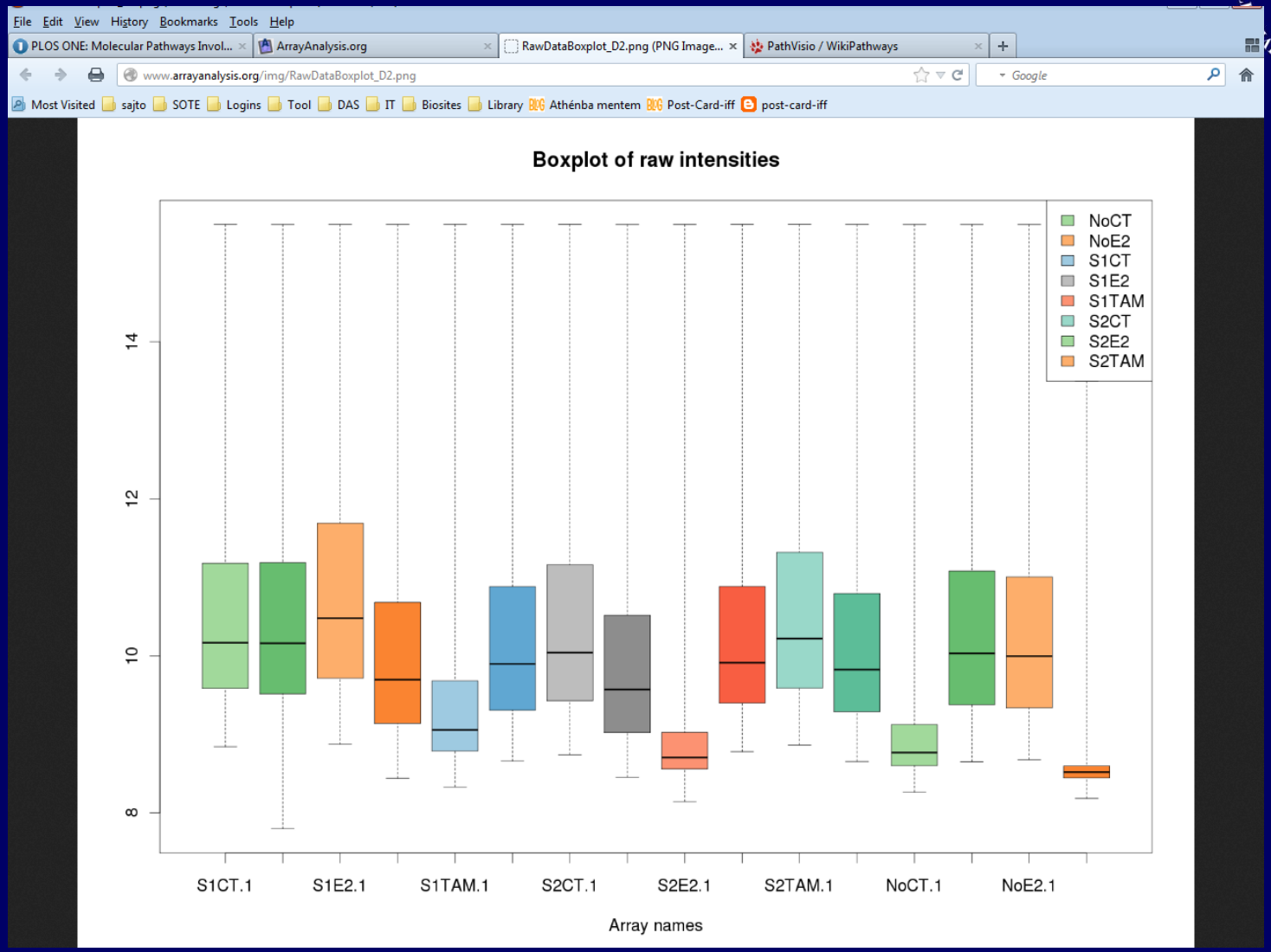
Density histogram of log-intensities

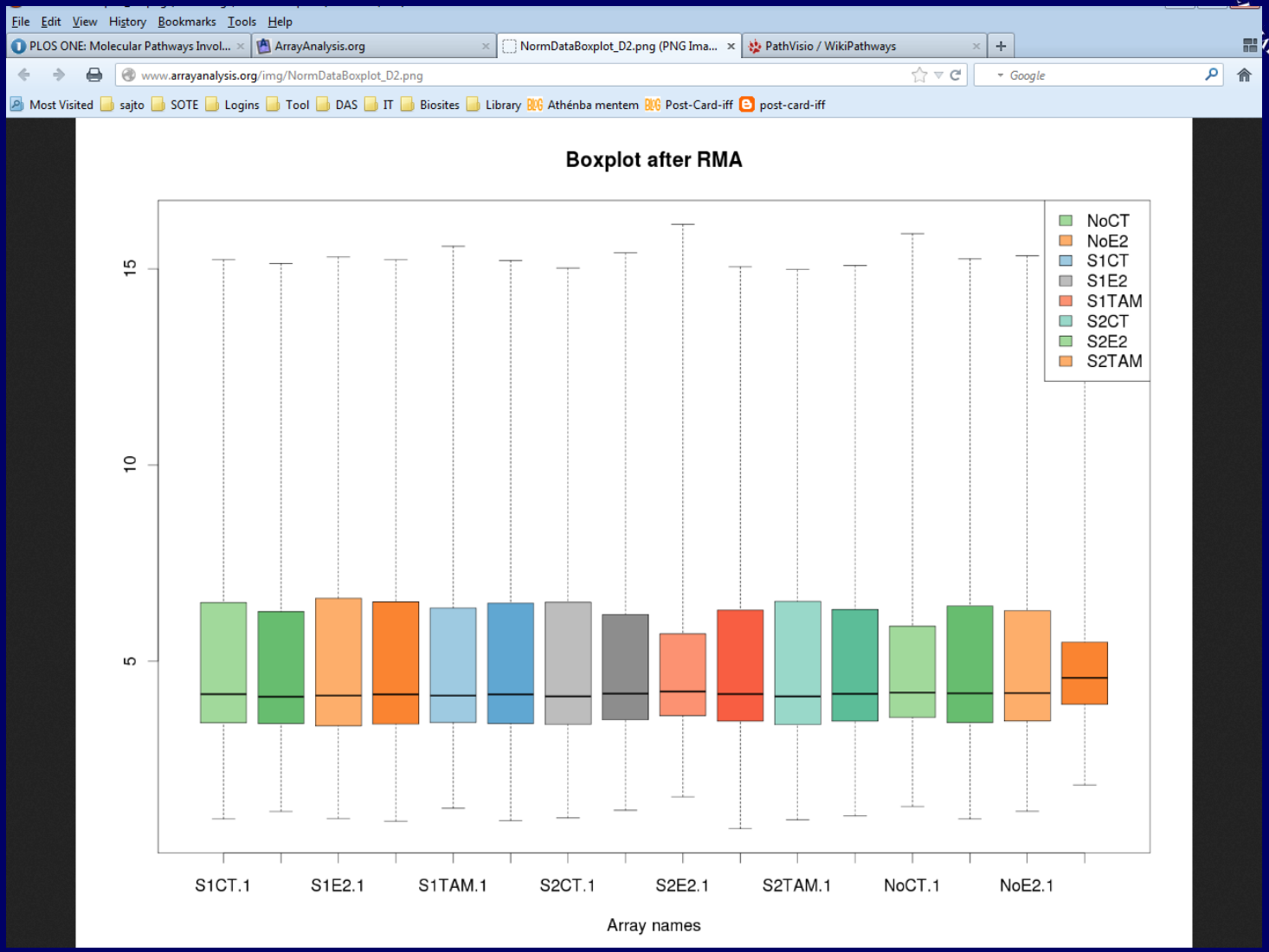
Density plots of log-intensity distribution of each array are superposed on a single graph for a better comparison between arrays and for an identification of arrays with weird distribution. As for the boxplots, the density distributions of raw PM (perfect match probes) log-intensities are not expected to be identical but still not totally different while the distributions of normalized probe-set log-intensities are expected to be more. Drawing these plots before and after normalization allows also checking the normalization step.

Left : Raw data // Right : Normalized data.


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataBoxplot_D1.png (PNG Image...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataBoxplot_D1.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Boxplot of raw intensities
ERT
HG
HGERT
NG
14
12
10
8
6
ERT1 ERT2 ERT3 HG1 HG2 HG3 HGERT1 HGERT2 HGERT3 NG1 NG2 NG3
Array names


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataBoxplot_D1.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataBoxplot_D1.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Boxplot after RMA
ERT
HG
HGERT
NG
2
4
6
8
10
12
14
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
Array names


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataBoxplot_D2.png (PNG Image...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataBoxplot_D2.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Boxplot of raw intensities
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
8
10
12
14
S1CT.1
S1E2.1
S1TAM.1
S2CT.1
S2E2.1
S2TAM.1
NoCT.1
NoE2.1
Array names


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataBoxplot_D2.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataBoxplot_D2.png
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Boxplot after RMA
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM
15
10
5
S1CT.1
S1E2.1
S1TAM.1
S2CT.1
S2E2.1
S2TAM.1
NoCT.1
NoE2.1
Array names

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

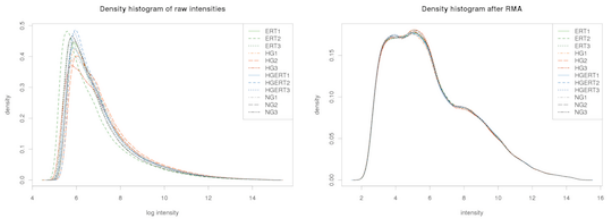
BiGCaT bioinformatics

Density histogram of log-intensities

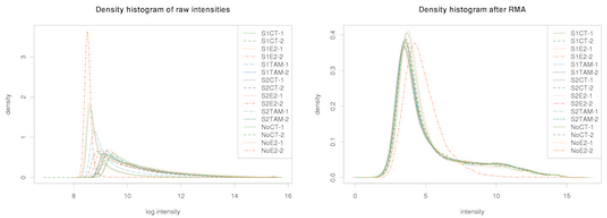
Density plots of log-intensity distribution of each array are superposed on a single graph for a better comparison between arrays and for an identification of arrays with weird distribution. As for the boxplots, the density distributions of raw PM (perfect match probes) log-intensities are not expected to be identical but still not totally different while the distributions of normalized probe-set log-intensities are expected to be more. Drawing these plots before and after normalization allows also checking the normalization step.

Left : Raw data // Right : Normalized data.

Dataset #1: Raw signal distributions are similar even before normalization:



Dataset #2: Raw distributions are different when compared between arrays.
After the normalization, array NoE2-2 still presents a different distribution:



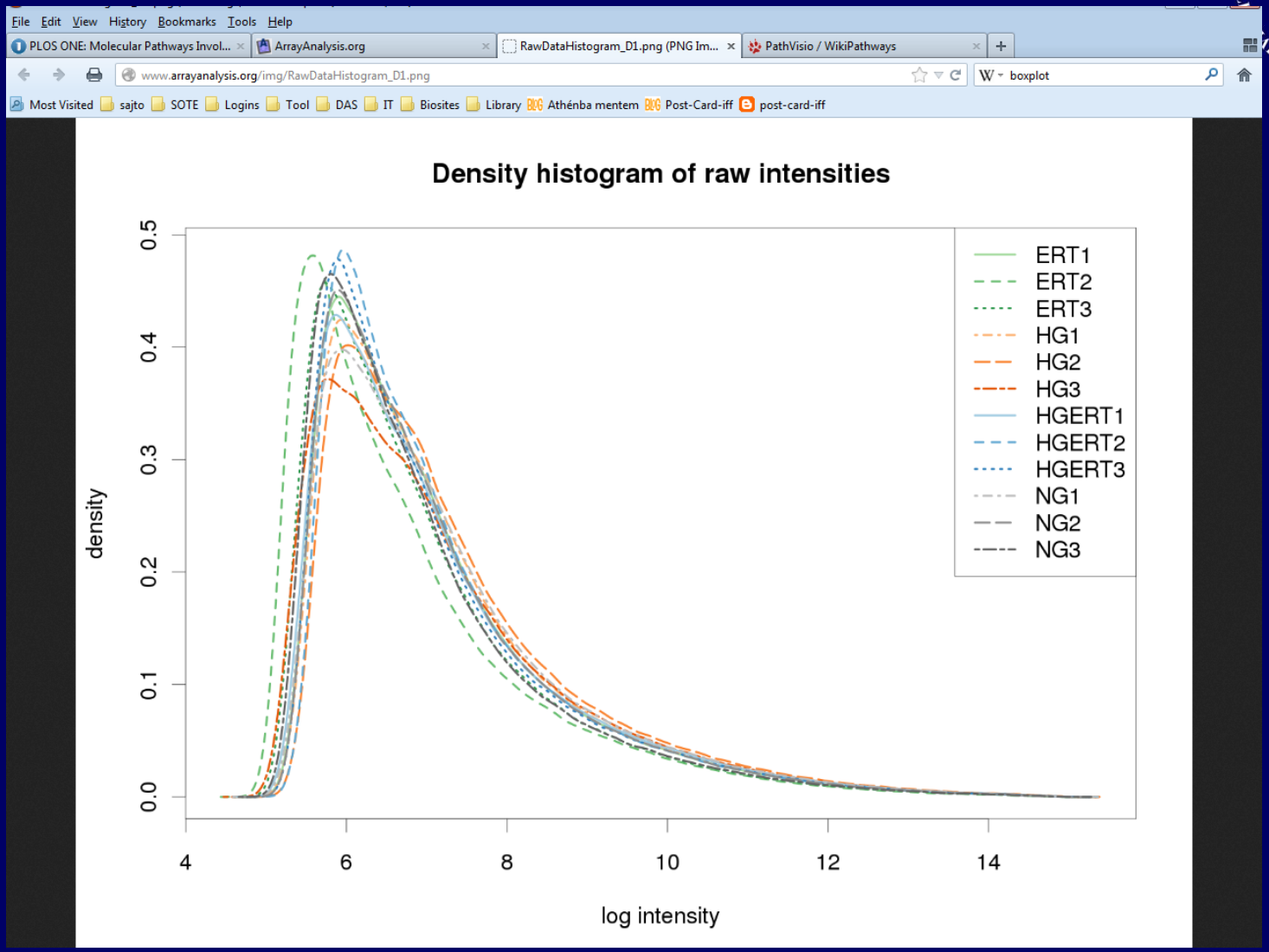
[Technical documentation of the function]

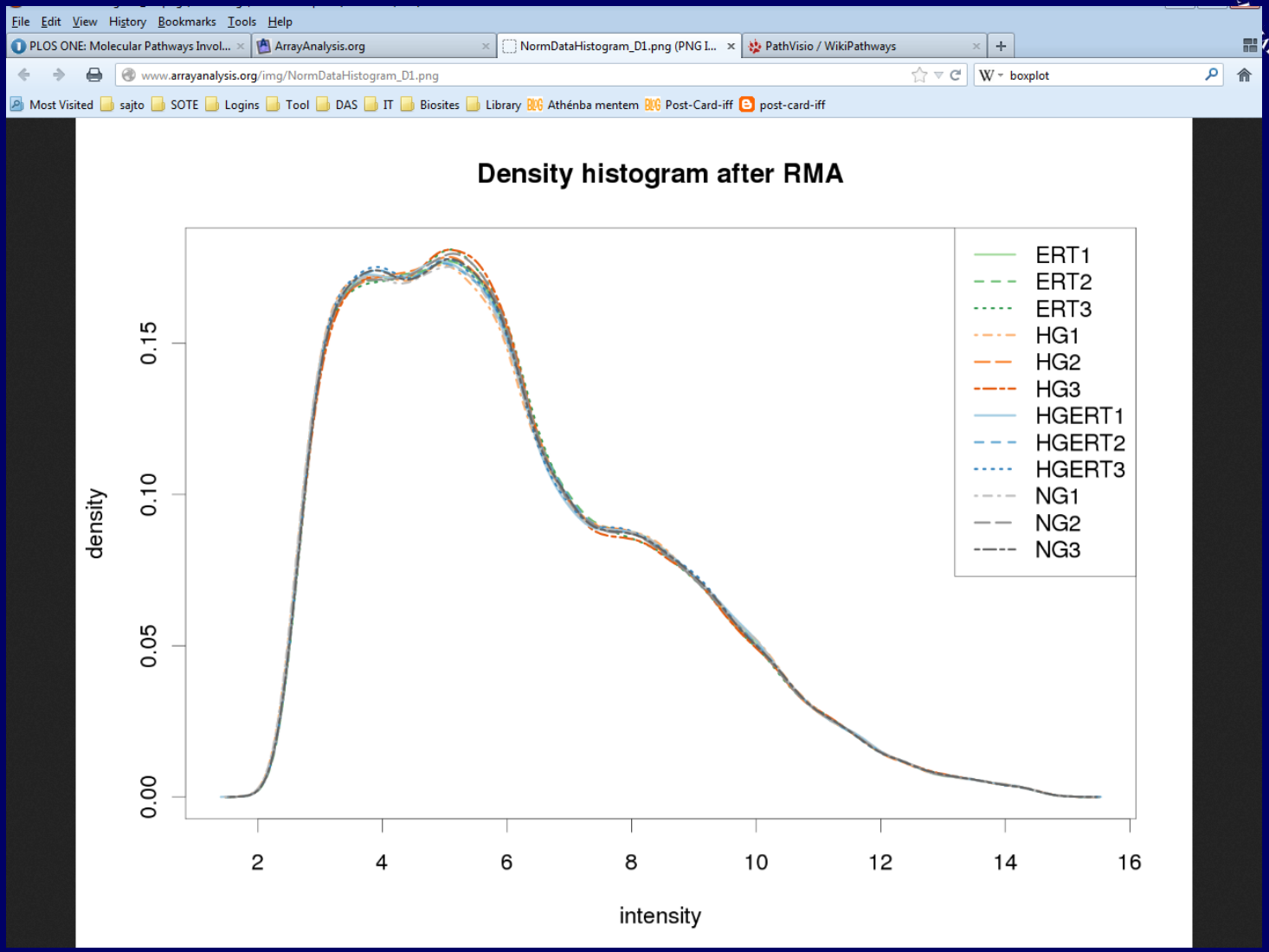
Unsmoothed density histogram

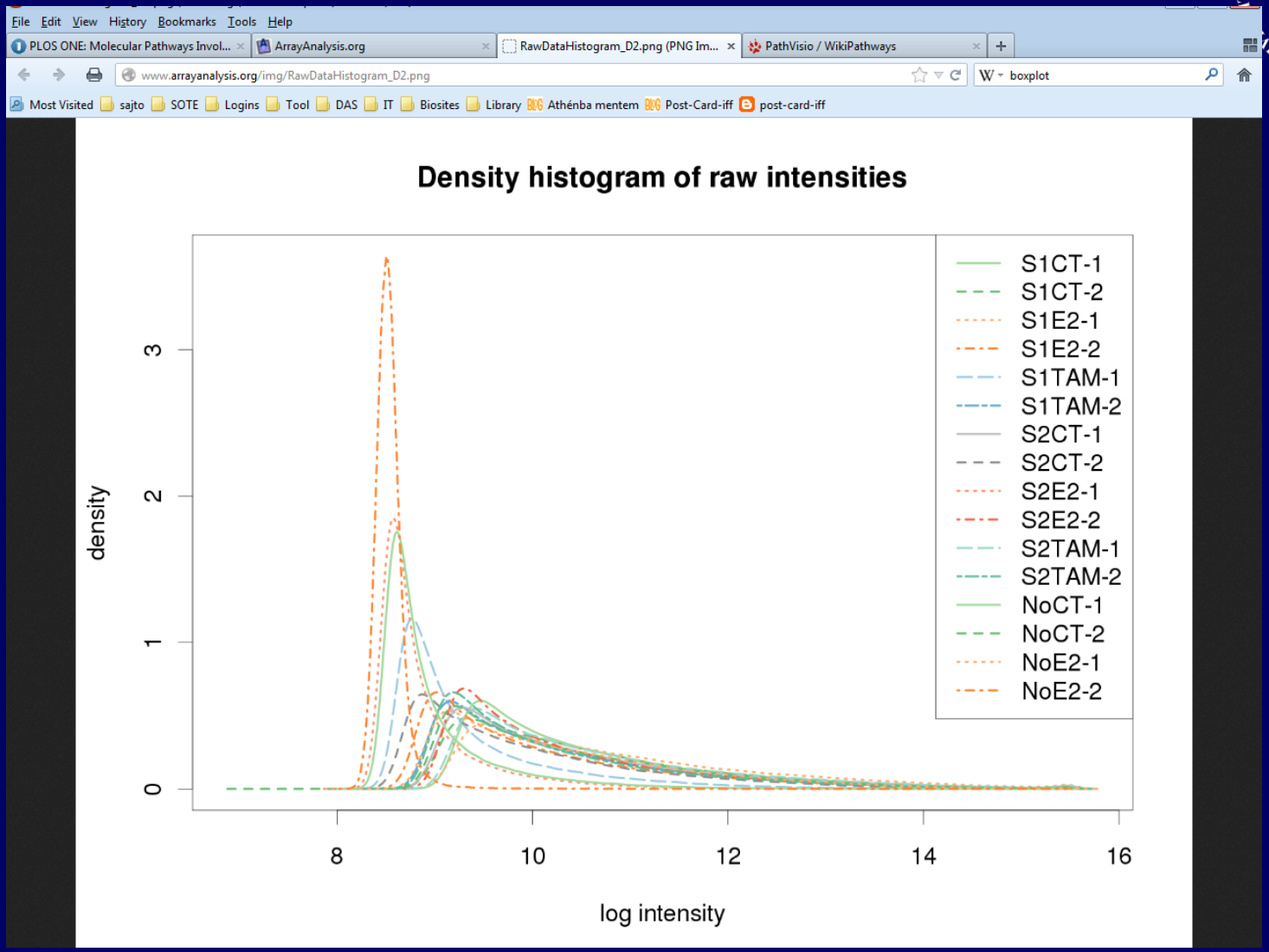
We propose an alternative function to draw the density histogram, using an unsmoothed density curve of the intensities for all arrays in the raw or normalized dataset. This function is not called by default but is implemented. You may modify the function calls to be able to use this particular version.

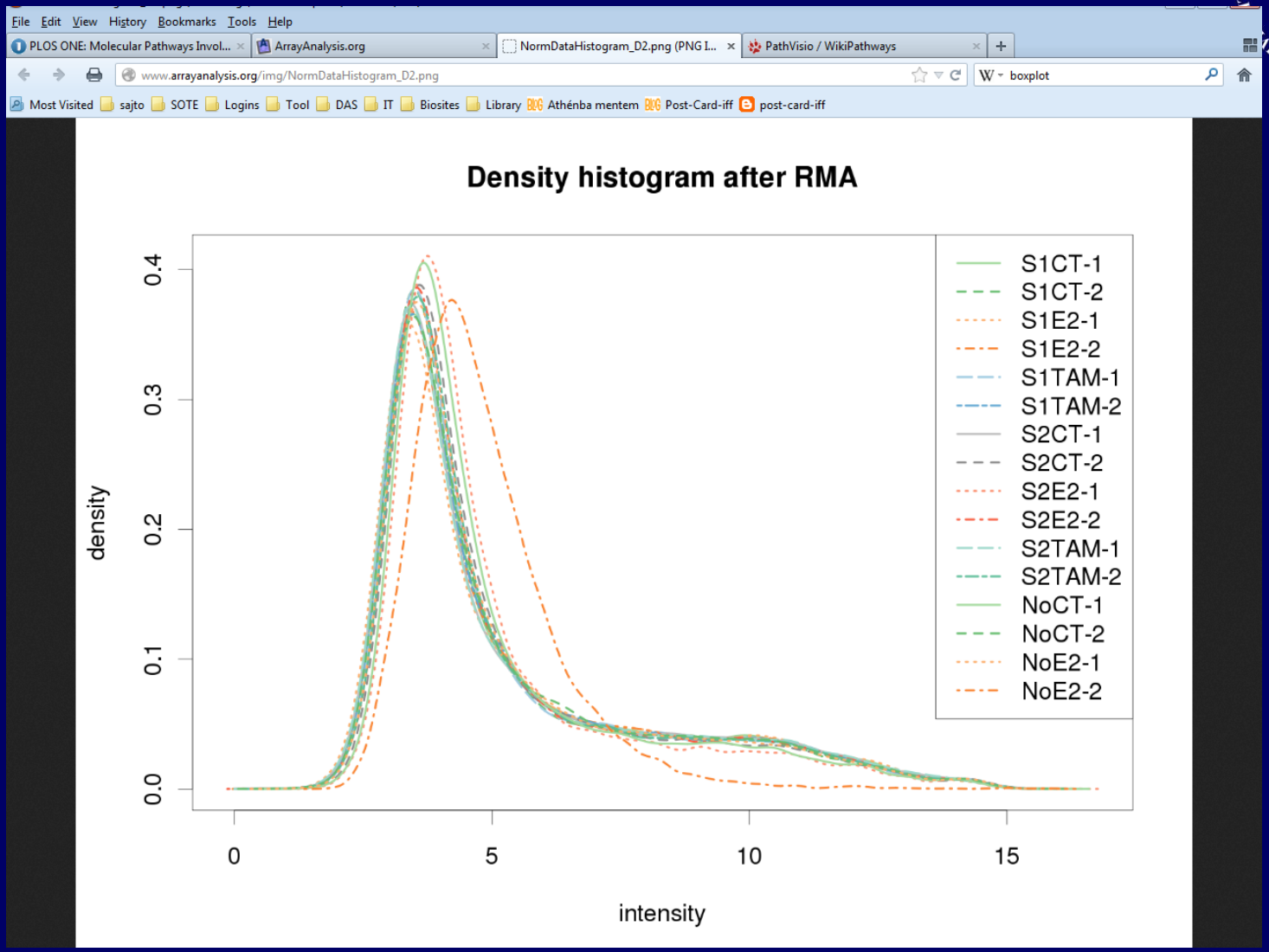
Raw signal distributions shown on an unsmoothed density histogram

Density histogram of raw intensities


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataHistogram_D1.png (PNG Im...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataHistogram_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Density histogram of raw intensities
density
0.0
0.1
0.2
0.3
0.4
0.5
4
6
8
10
12
14
log intensity
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataHistogram_D1.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataHistogram_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Density histogram after RMA
density
0.00
0.05
0.10
0.15
2
4
6
8
10
12
14
16
intensity
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataHistogram_D2.png (PNG Im...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataHistogram_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Density histogram of raw intensities
density
log intensity
0
1
2
3
8
10
12
14
16
S1CT-1
S1CT-2
S1E2-1
S1E2-2
S1TAM-1
S1TAM-2
S2CT-1
S2CT-2
S2E2-1
S2E2-2
S2TAM-1
S2TAM-2
NoCT-1
NoCT-2
NoE2-1
NoE2-2


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataHistogram_D2.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataHistogram_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Density histogram after RMA
density
0.0
0.1
0.2
0.3
0.4
0
5
10
15
intensity
S1CT-1
S1CT-2
S1E2-1
S1E2-2
S1TAM-1
S1TAM-2
S2CT-1
S2CT-2
S2E2-1
S2E2-2
S2TAM-1
S2TAM-2
NoCT-1
NoCT-2
NoE2-1
NoE2-2

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

boxplot

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

Array correlation
- Correlation plot
- PCA analysis
- Clustering

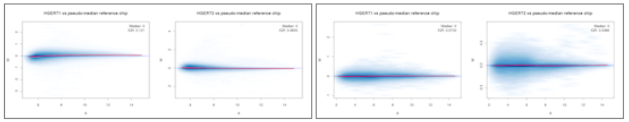
BiGCaT bioinformatics

MA plot

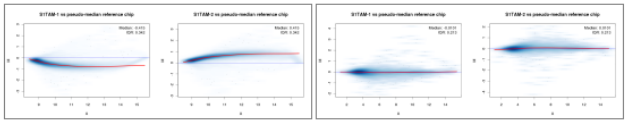
The MA plots allow pairewise comparison of log-intensity of each array to a reference array and identification of intensity-dependent biases. The Y axis of the plot contains the log-ratio intensity of one array to the reference median array, which is called 'M' while the X axis contains the average log-intensity of both arrays - called 'A'. Within a group of replicates, the probe levels are not likely to differ a lot so we expect a MA plot centered on the Y=0 axis from low to high intensities. When the MA plots is computed for each replicate group separately, the references array is the median array of each group. A smooth Loess regression curve is plotted to facilitated the comparison to the Y=0 axis. The normalization is expected to correct for intensity-dependent biases: these graphs plotted before and after normalization allow checking the efficiency of this correction.

Left : Raw data // Right : Normalized data.

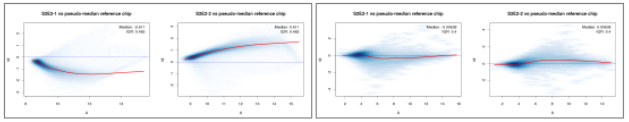
Dataset #1: Raw signal distributions does not show any intensity-dependent bias:



Dataset #2.1: There is an intensity-bias in the raw data that is corrected after normalization



Dataset #2.2: The intensity-bias in the raw data is so strong that the normalization cannot correct it totally



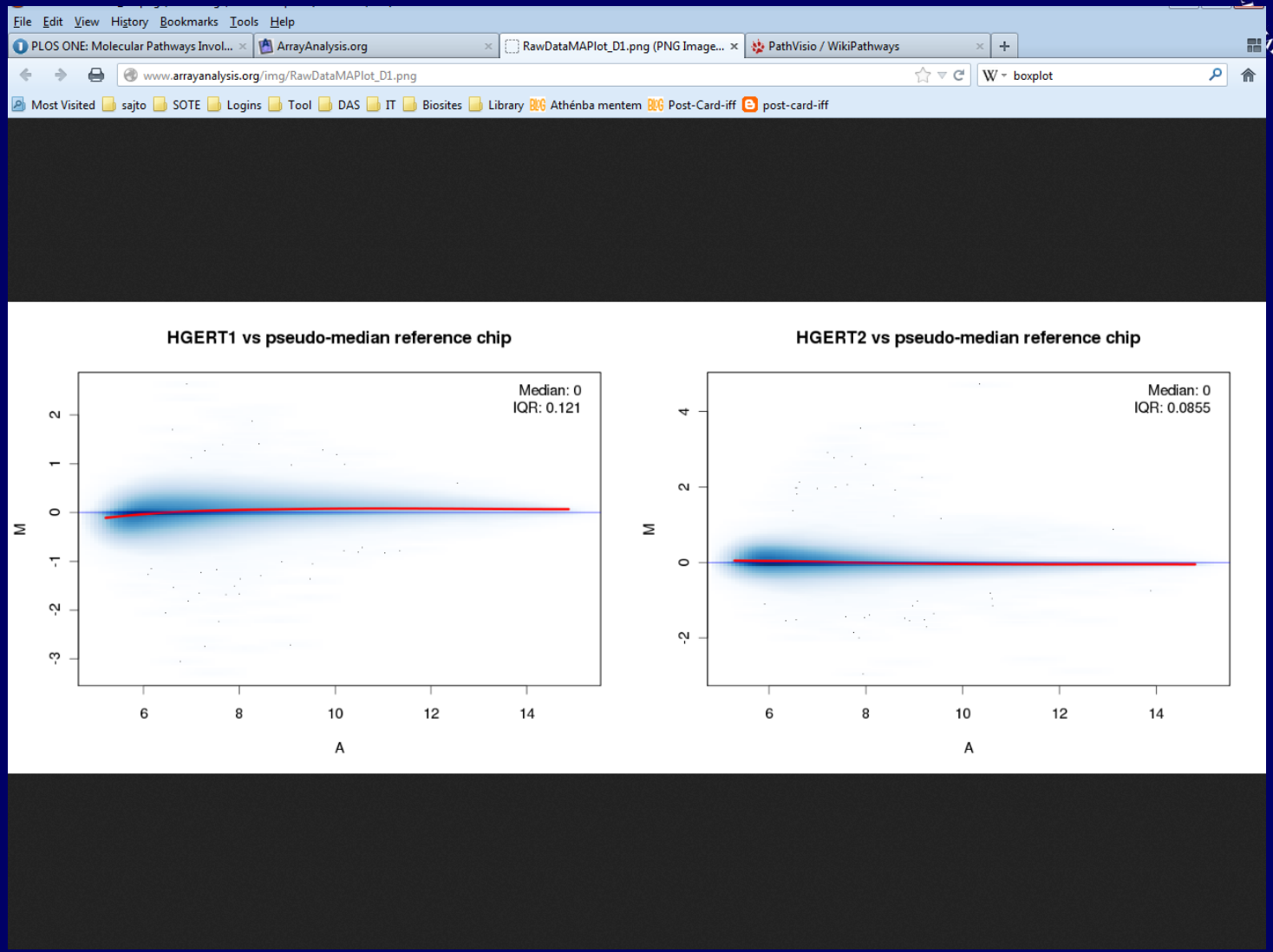
[Technical documentation of the function]

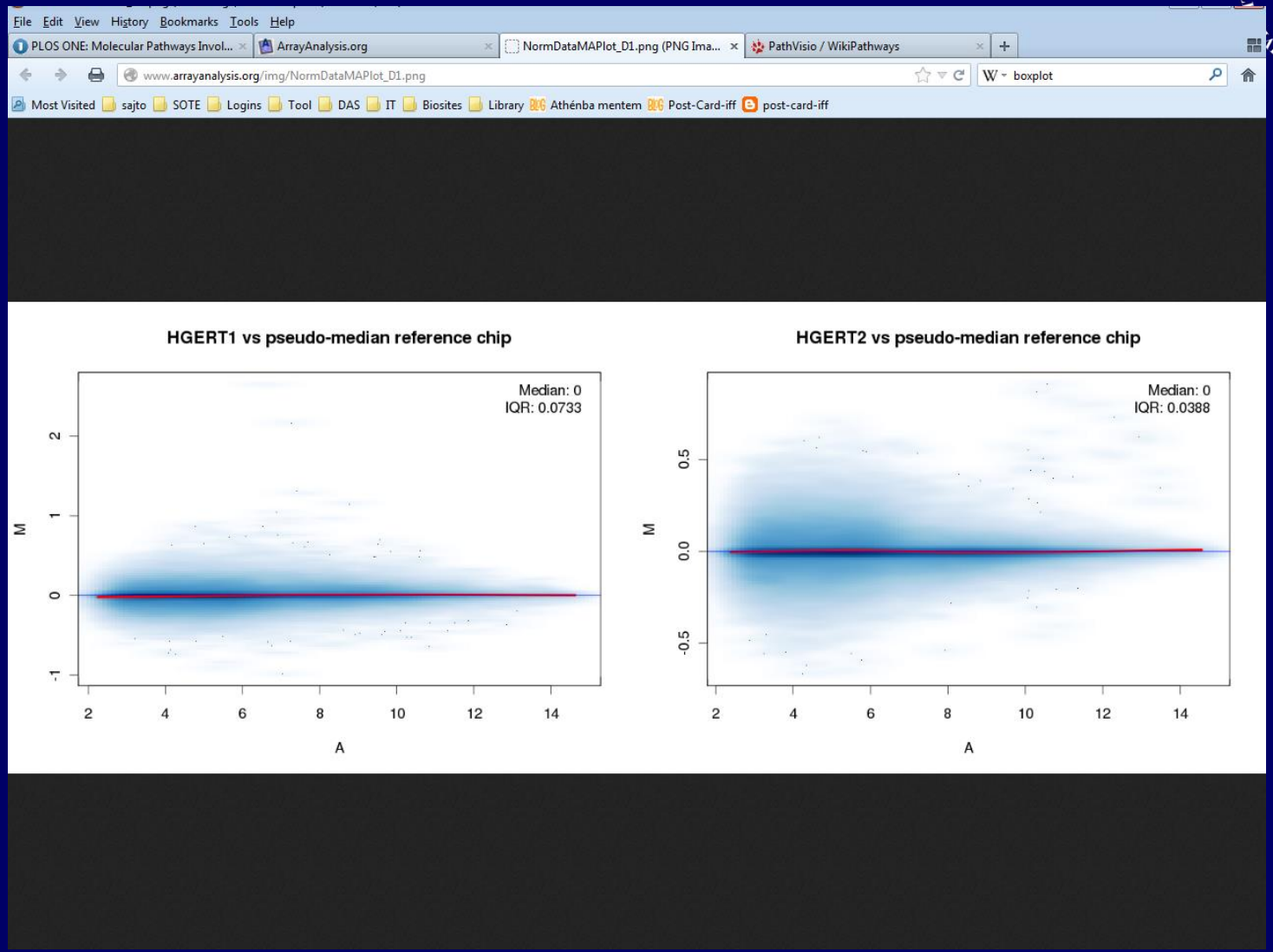
[Top]

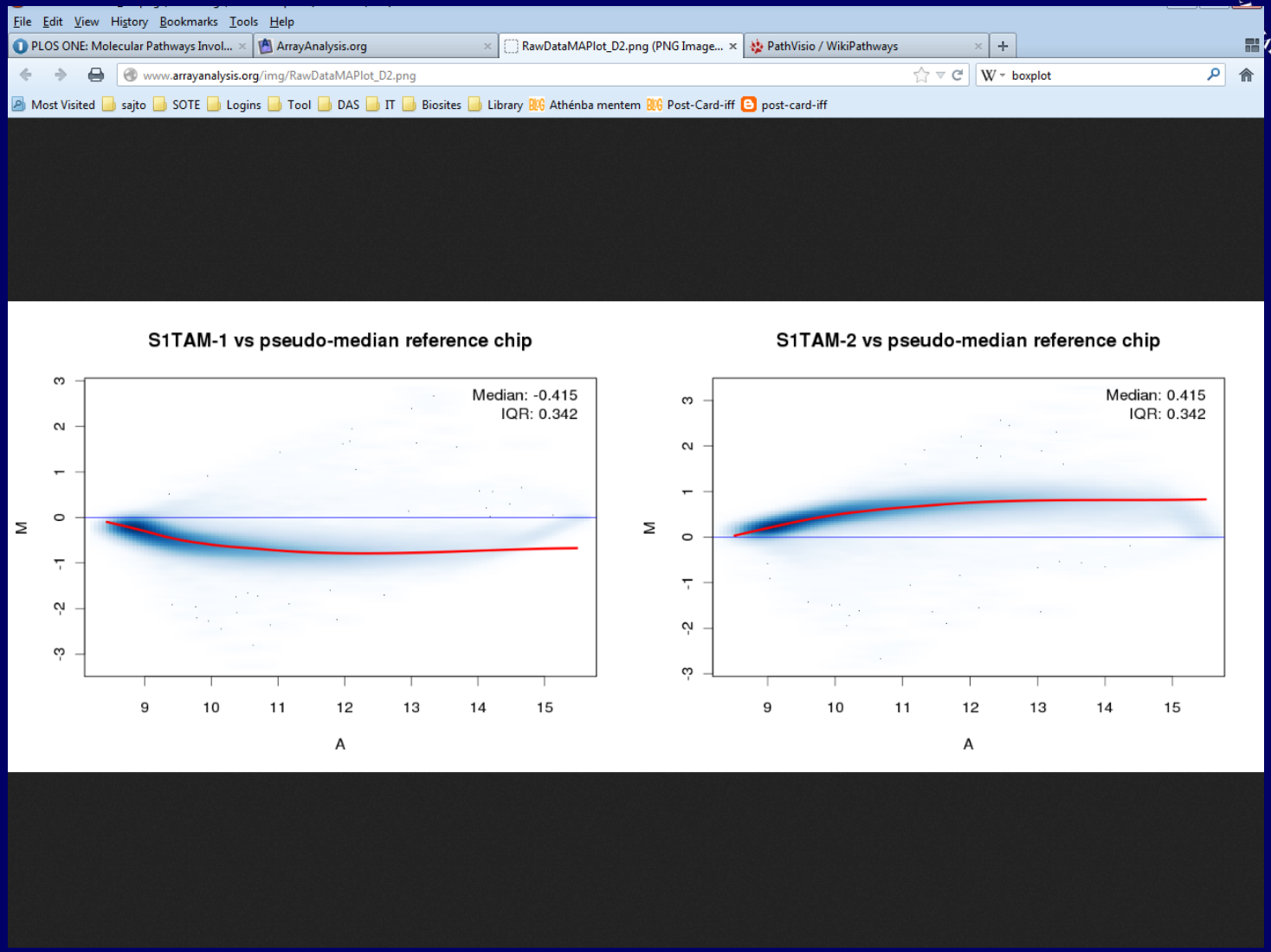
Spatial biases

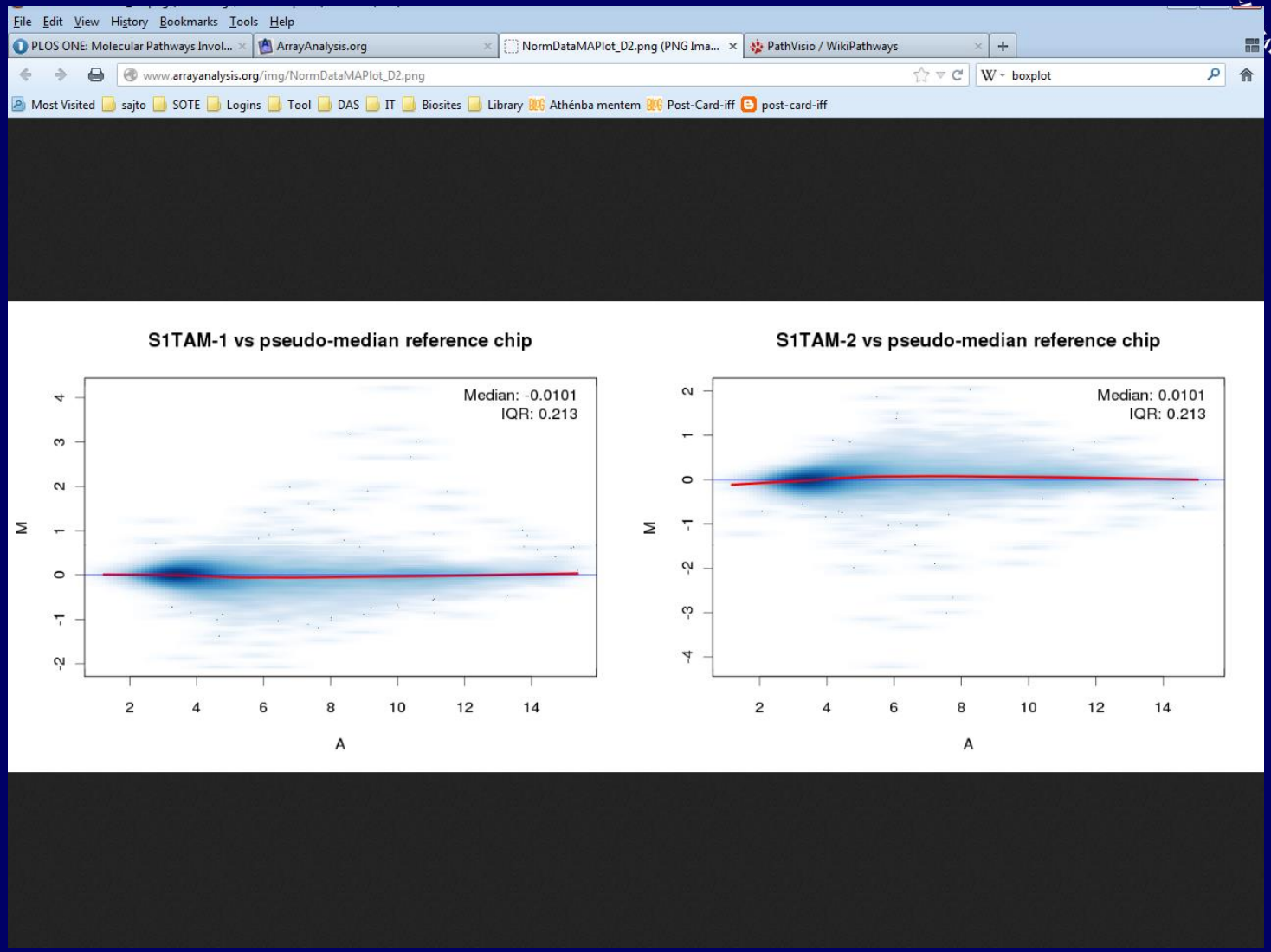
Array reference layout

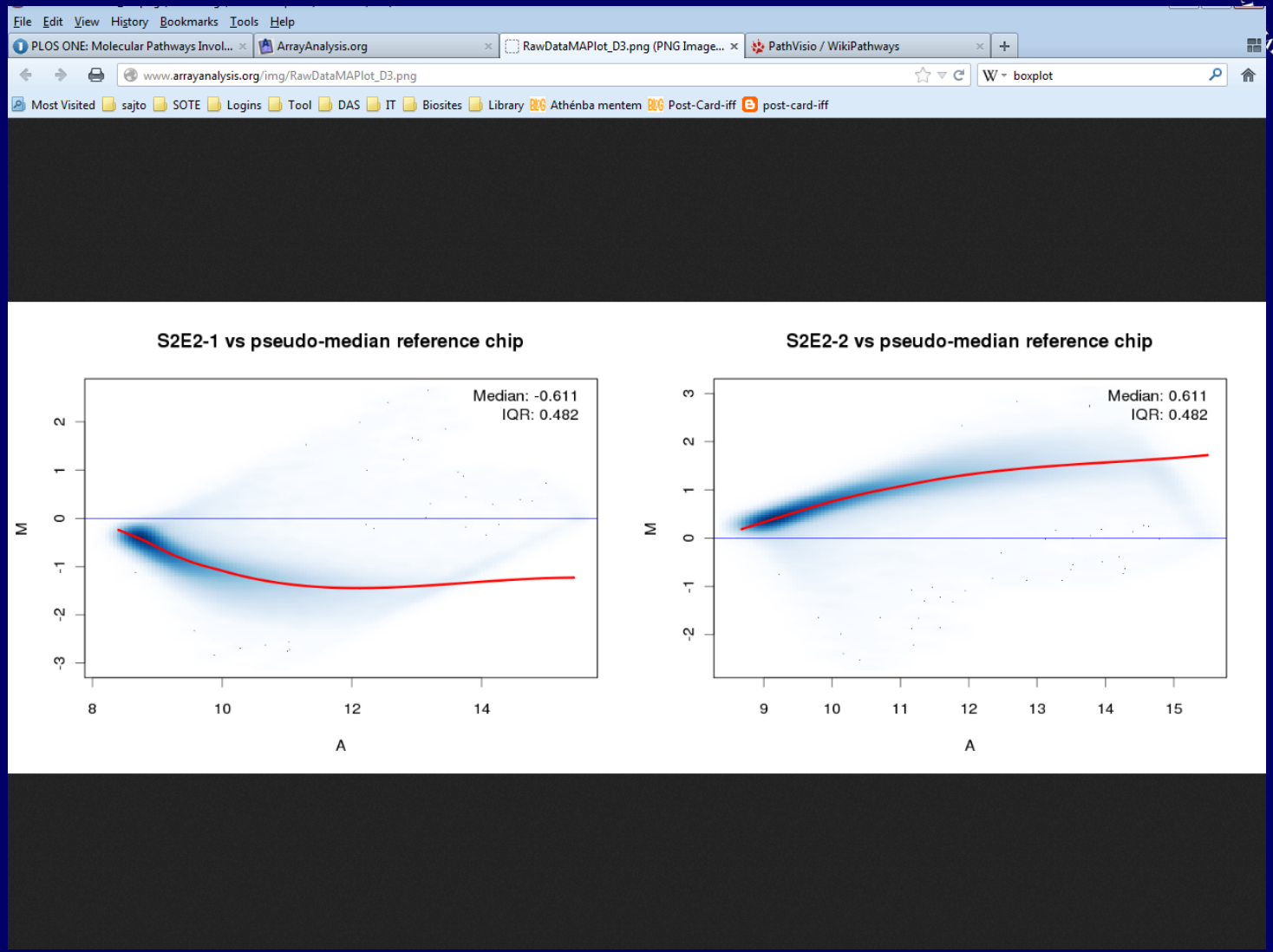
This graph presents the layout of the grid by color-coding a 2D image of one array according to the probe type (mainly regular probes, control probes and control regions), using available annotation libraries. Thus no data are plotted; the plot only shows the position of control and regular probes on the array. If applicable, a distinction is made between perfect match (PM) and mismatch (MM) probes.

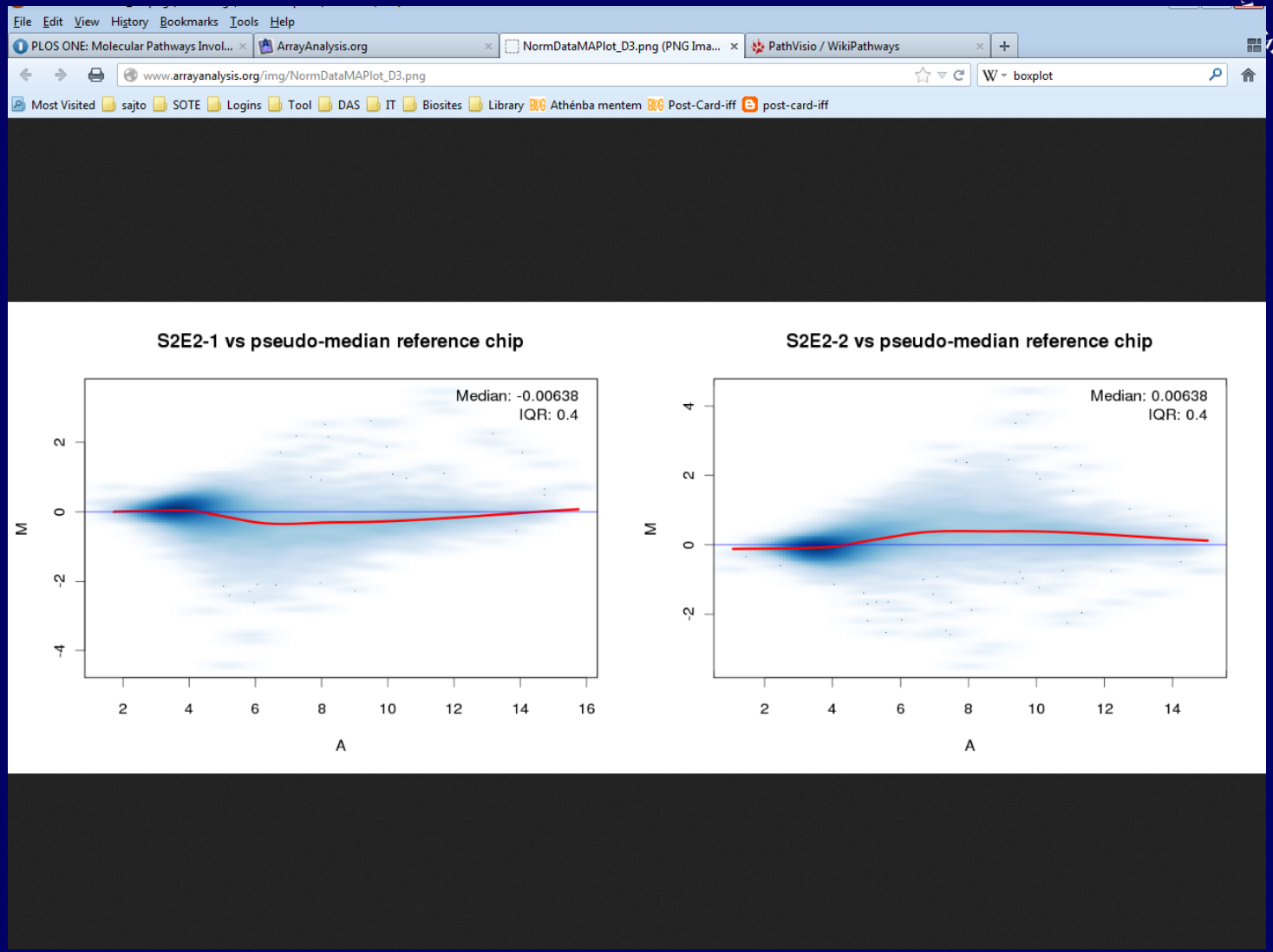

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataMAPlot_D1.png (PNG Image...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataMAPlot_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
HGERT1 vs pseudo-median reference chip
Median: 0
IQR: 0.121
M
A
HGERT2 vs pseudo-median reference chip
Median: 0
IQR: 0.0855
M
A

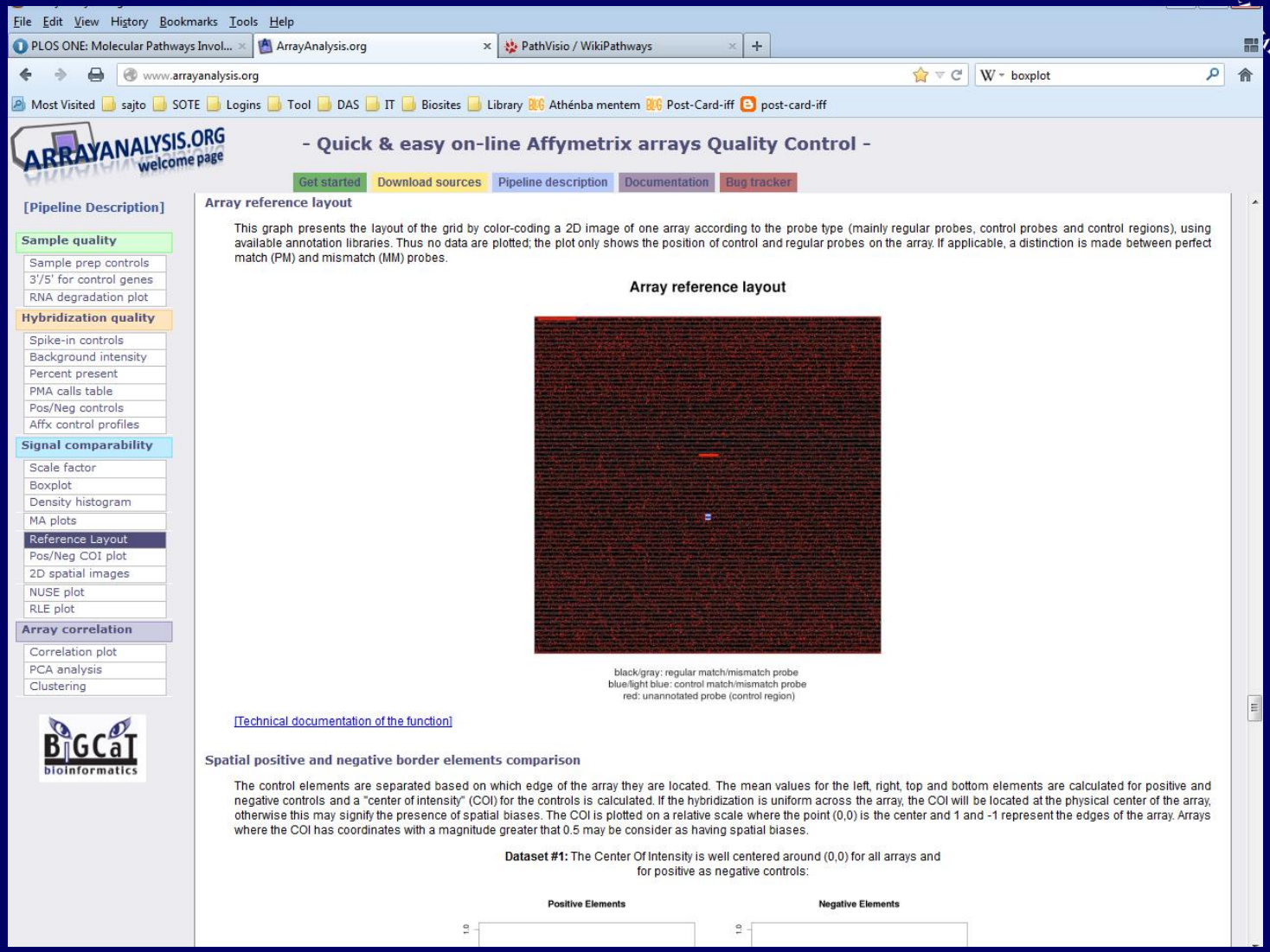

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataMAPlot_D1.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataMAPlot_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
HGERT1 vs pseudo-median reference chip
Median: 0
IQR: 0.0733
M
A
HGERT2 vs pseudo-median reference chip
Median: 0
IQR: 0.0388
M
A

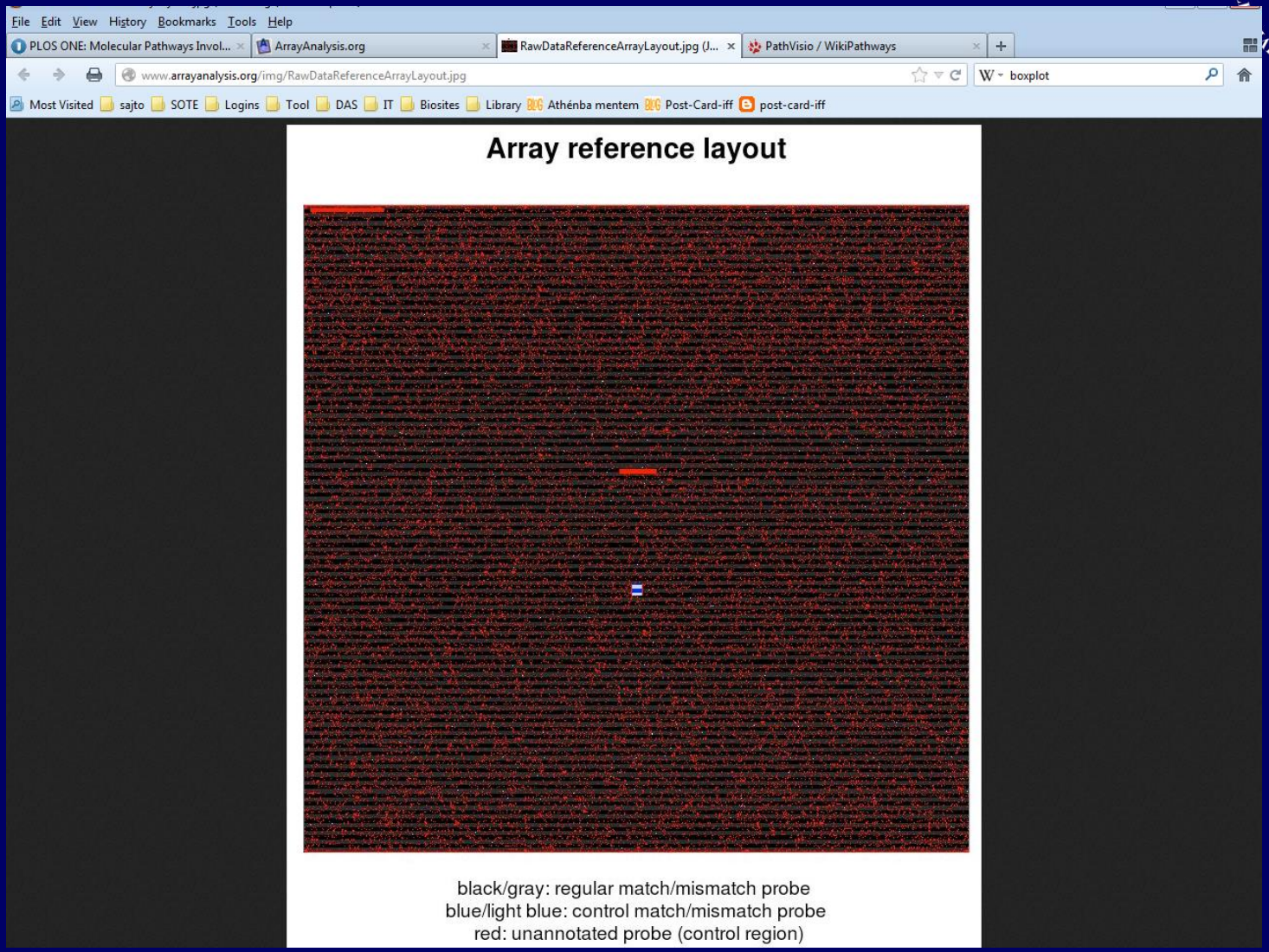

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataMAPlot_D2.png (PNG Image...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataMAPlot_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
S1TAM-1 vs pseudo-median reference chip
Median: -0.415
IQR: 0.342
M
A
S1TAM-2 vs pseudo-median reference chip
Median: 0.415
IQR: 0.342
M
A


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataMAPlot_D2.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataMAPlot_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
S1TAM-1 vs pseudo-median reference chip
Median: -0.0101
IQR: 0.213
M
A
S1TAM-2 vs pseudo-median reference chip
Median: 0.0101
IQR: 0.213
M
A


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataMAPlot_D3.png (PNG Image...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataMAPlot_D3.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
S2E2-1 vs pseudo-median reference chip
Median: -0.611
IQR: 0.482
M
A
S2E2-2 vs pseudo-median reference chip
Median: 0.611
IQR: 0.482
M
A


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataMAPlot_D3.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataMAPlot_D3.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
S2E2-1 vs pseudo-median reference chip
Median: -0.00638
IQR: 0.4
M
A
S2E2-2 vs pseudo-median reference chip
Median: 0.00638
IQR: 0.4
M
A


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
PathVisio / WikiPathways
www.arrayanalysis.org
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
ARRAYANALYSIS.ORG welcome page
- Quick & easy on-line Affymetrix arrays Quality Control -
Get started
Download sources
Pipeline description
Documentation
Bug tracker
[Pipeline Description]
Sample quality
Sample prep controls
3'/5' for control genes
RNA degradation plot
Hybridization quality
Spike-in controls
Background intensity
Percent present
PMA calls table
Pos/Neg controls
Affx control profiles
Signal comparability
Scale factor
Boxplot
Density histogram
MA plots
Reference Layout
Pos/Neg COI plot
2D spatial images
NUSE plot
RLE plot
Array correlation
Correlation plot
PCA analysis
Clustering
BiGCaT bioinformatics
Array reference layout
This graph presents the layout of the grid by color-coding a 2D image of one array according to the probe type (mainly regular probes, control probes and control regions), using available annotation libraries. Thus no data are plotted; the plot only shows the position of control and regular probes on the array. If applicable, a distinction is made between perfect match (PM) and mismatch (MM) probes.
Array reference layout
black/gray: regular match/mismatch probe
blue/light blue: control match/mismatch probe
red: unannotated probe (control region)
[Technical documentation of the function]
Spatial positive and negative border elements comparison
The control elements are separated based on which edge of the array they are located. The mean values for the left, right, top and bottom elements are calculated for positive and negative controls and a "center of intensity" (COI) for the controls is calculated. If the hybridization is uniform across the array, the COI will be located at the physical center of the array, otherwise this may signify the presence of spatial biases. The COI is plotted on a relative scale where the point (0,0) is the center and 1 and -1 represent the edges of the array. Arrays where the COI has coordinates with a magnitude greater that 0.5 may be consider as having spatial biases.
Dataset #1: The Center Of Intensity is well centered around (0,0) for all arrays and for positive as negative controls:
Positive Elements
Negative Elements


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataReferenceArrayLayout.jpg (J...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataReferenceArrayLayout.jpg
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Array reference layout
black/gray: regular match/mismatch probe
blue/light blue: control match/mismatch probe
red: unannotated probe (control region)

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

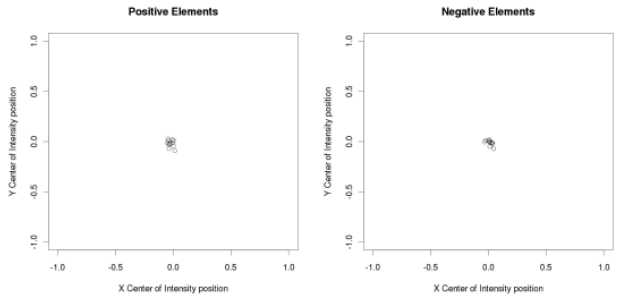
- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

BiGCaT bioinformatics

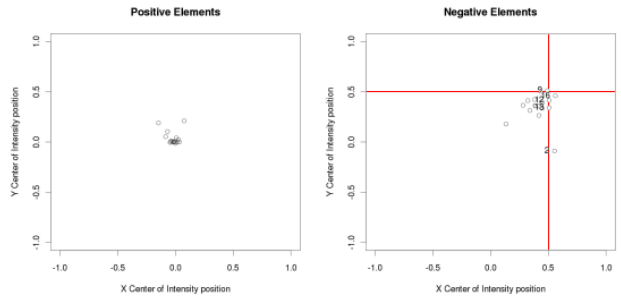
Spatial positive and negative border elements comparison

The control elements are separated based on which edge of the array they are located. The mean values for the left, right, top and bottom elements are calculated for positive and negative controls and a "center of intensity" (COI) for the controls is calculated. If the hybridization is uniform across the array, the COI will be located at the physical center of the array, otherwise this may signify the presence of spatial biases. The COI is plotted on a relative scale where the point (0,0) is the center and 1 and -1 represent the edges of the array. Arrays where the COI has coordinates with a magnitude greater that 0.5 may be consider as having spatial biases.

Dataset #1: The Center Of Intensity is well centered around (0,0) for all arrays and for positive as negative controls:

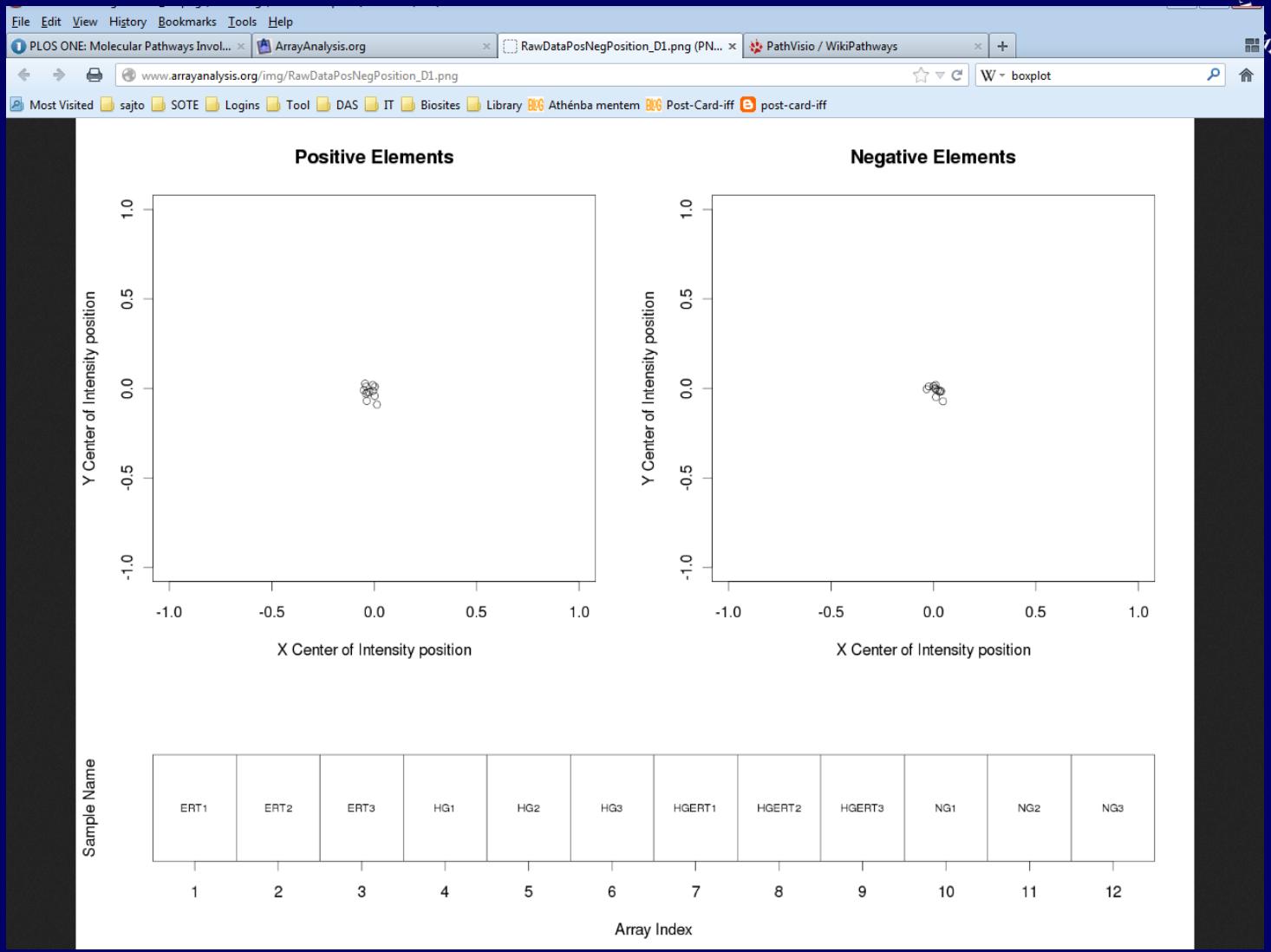


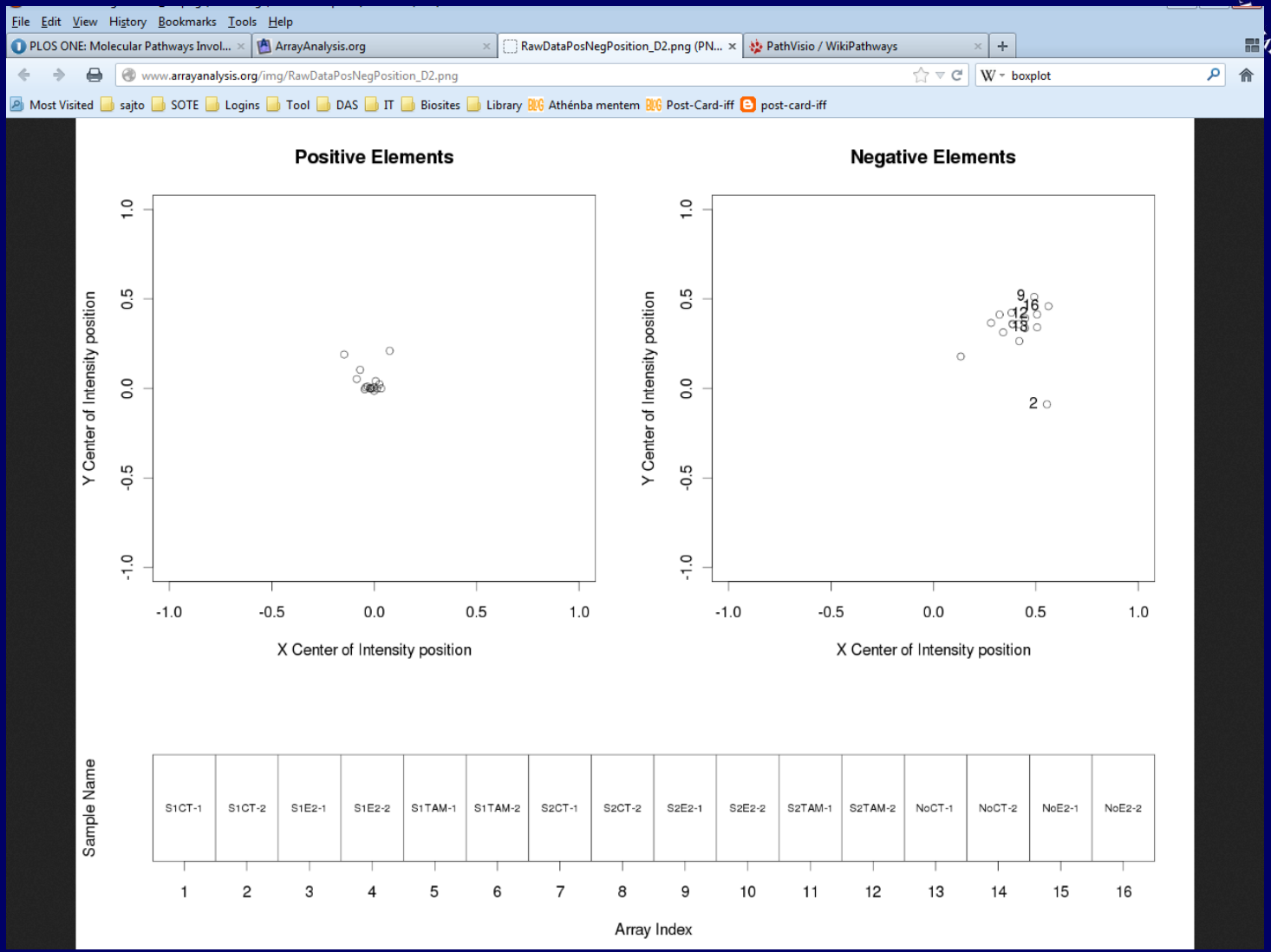
Dataset #2: Arrays with a COI out of the recommended range are tagged on the plot. In this dataset, arrays 2, 9, 12, 13 and 16 have negative controls COI out of range:

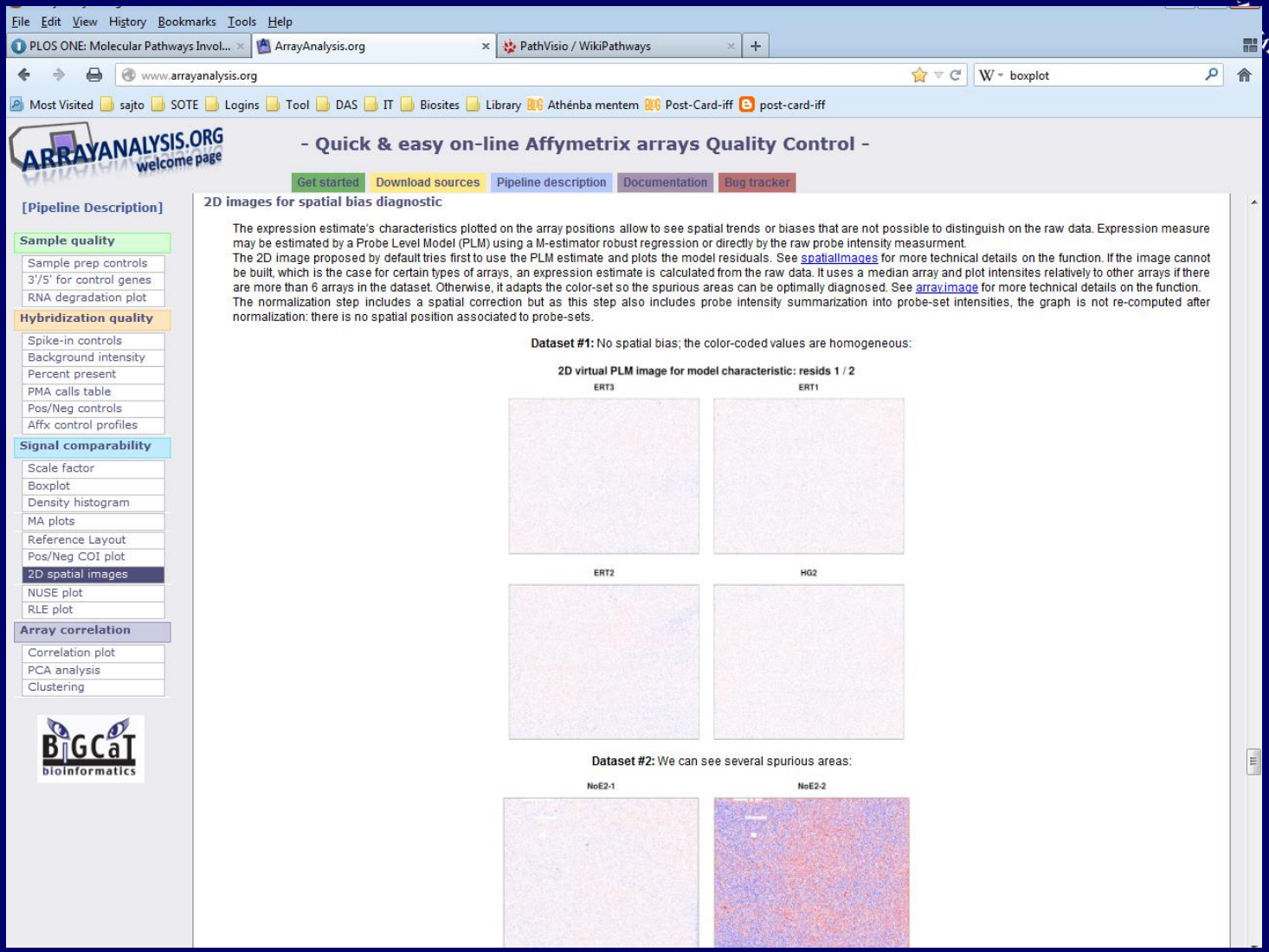


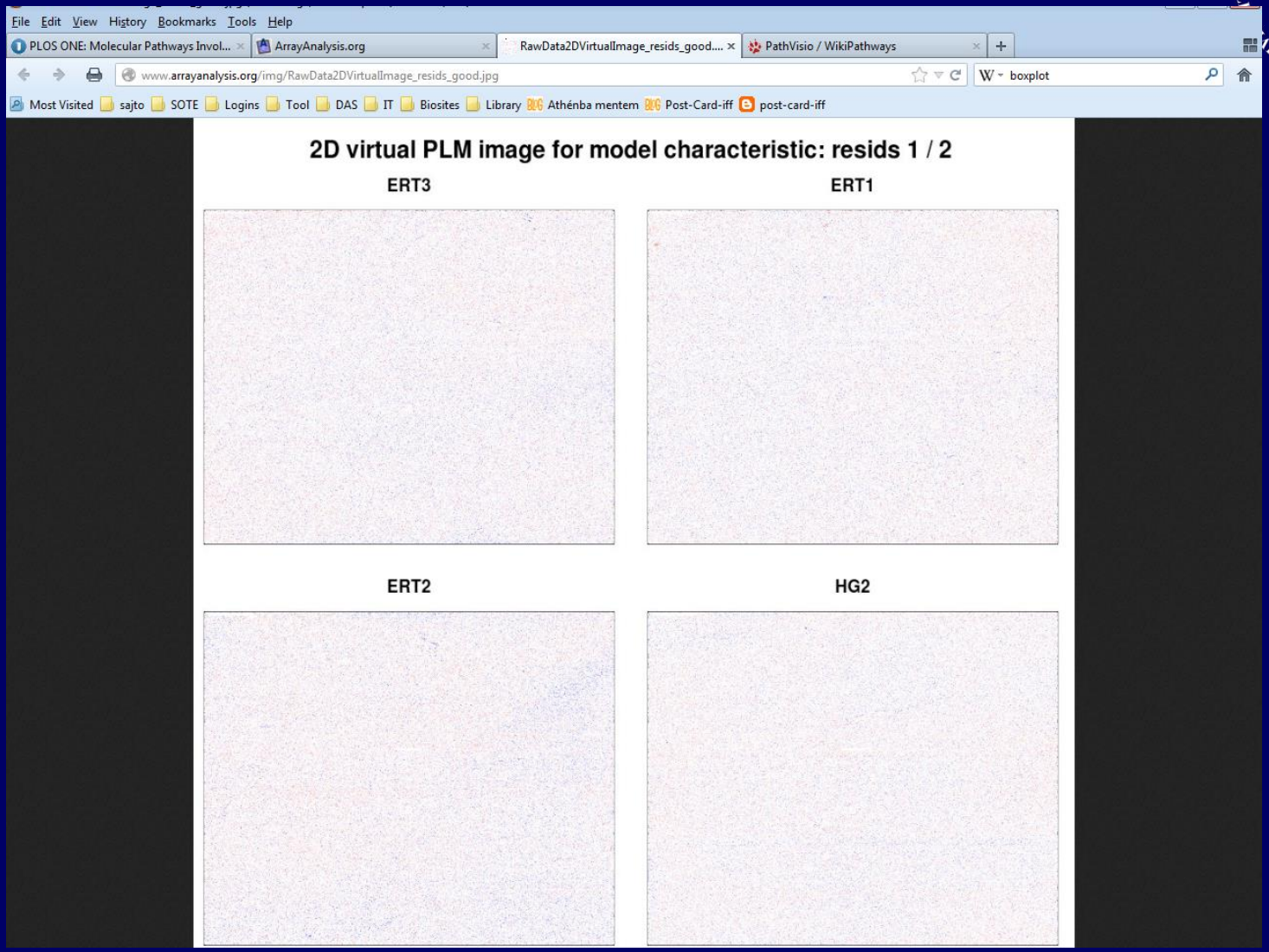
[Technical documentation of the function]

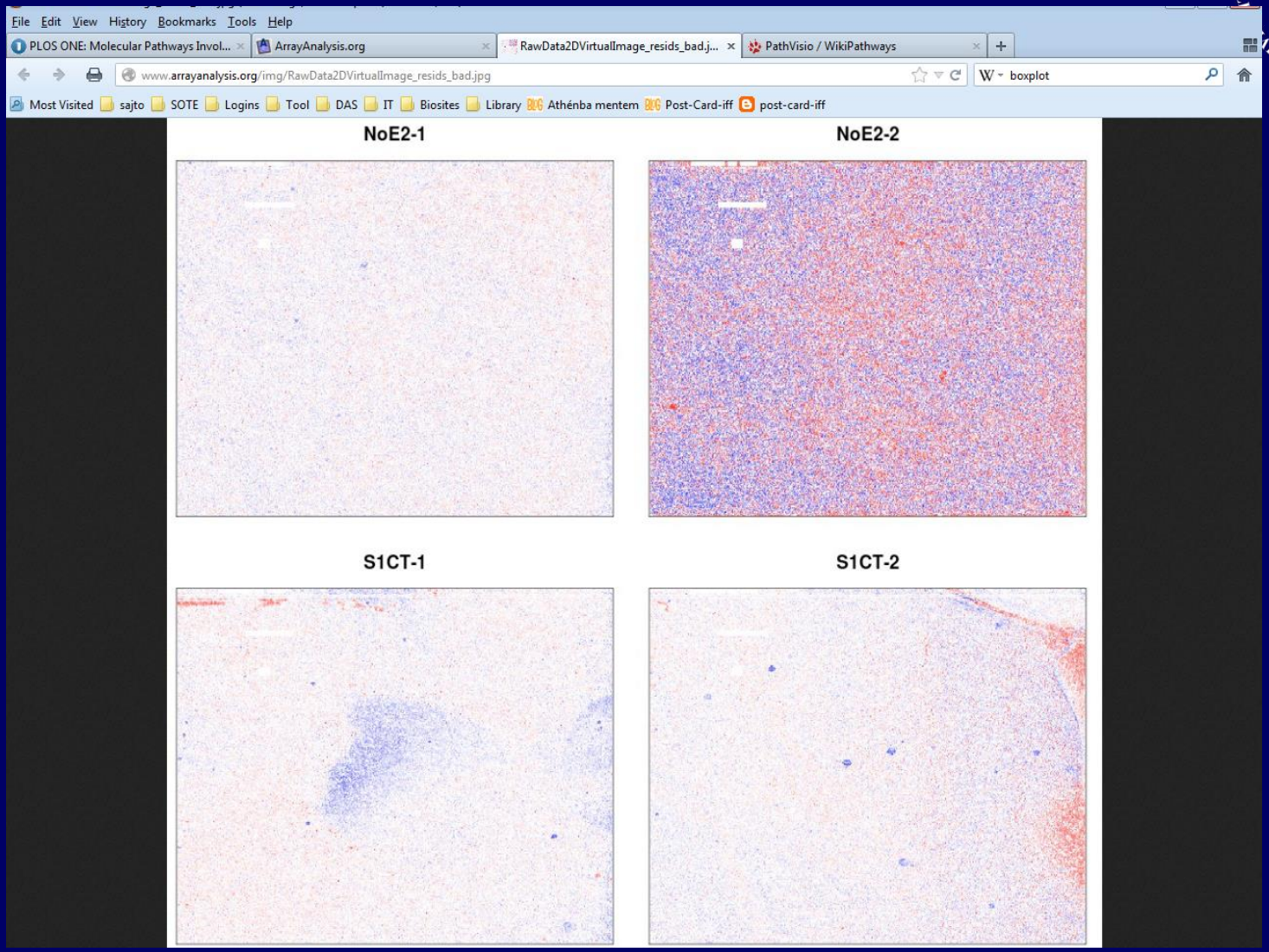
www.arrayanalysis.org/main.html#PS

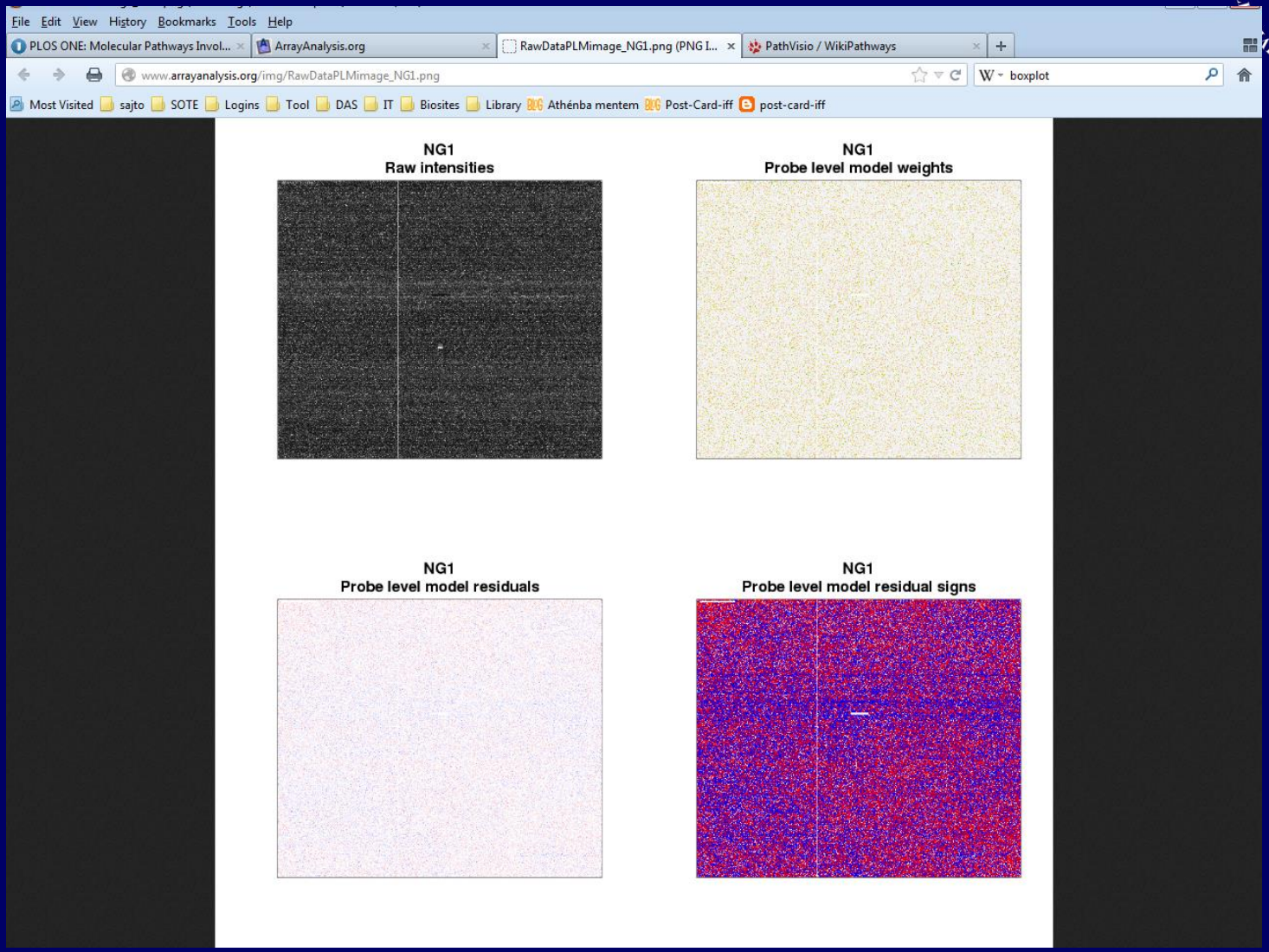

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPosNegPosition_D1.png (PN...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPosNegPosition_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Positive Elements
Negative Elements
Y Center of Intensity position
X Center of Intensity position
1.0
0.5
0.0
-0.5
-1.0
Sample Name
ERT1
ERT2
ERT3
HG1
HG2
HG3
HGERT1
HGERT2
HGERT3
NG1
NG2
NG3
1
2
3
4
5
6
7
8
9
10
11
12
Array Index


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPosNegPosition_D2.png (PN...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPosNegPosition_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Positive Elements
Negative Elements
Y Center of Intensity position
X Center of Intensity position
-1.0
-0.5
0.0
0.5
1.0
Sample Name
S1CT-1
S1CT-2
S1E2-1
S1E2-2
S1TAM-1
S1TAM-2
S2CT-1
S2CT-2
S2E2-1
S2E2-2
S2TAM-1
S2TAM-2
NoCT-1
NoCT-2
NoE2-1
NoE2-2
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
Array Index


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
PathVisio / WikiPathways
www.arrayanalysis.org
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
ARRAYANALYSIS.ORG welcome page
- Quick & easy on-line Affymetrix arrays Quality Control -
Get started
Download sources
Pipeline description
Documentation
Bug tracker
[Pipeline Description]
Sample quality
Sample prep controls
3'/5' for control genes
RNA degradation plot
Hybridization quality
Spike-in controls
Background intensity
Percent present
PMA calls table
Pos/Neg controls
Affx control profiles
Signal comparability
Scale factor
Boxplot
Density histogram
MA plots
Reference Layout
Pos/Neg COI plot
2D spatial images
NUSE plot
RLE plot
Array correlation
Correlation plot
PCA analysis
Clustering
BiGCaT bioinformatics
2D images for spatial bias diagnostic
The expression estimate's characteristics plotted on the array positions allow to see spatial trends or biases that are not possible to distinguish on the raw data. Expression measure may be estimated by a Probe Level Model (PLM) using a M-estimator robust regression or directly by the raw probe intensity measurment.
The 2D image proposed by default tries first to use the PLM estimate and plots the model residuals. See spatialImages for more technical details on the function. If the image cannot be built, which is the case for certain types of arrays, an expression estimate is calculated from the raw data. It uses a median array and plot intensites relatively to other arrays if there are more than 6 arrays in the dataset. Otherwise, it adapts the color-set so the spurious areas can be optimally diagnosed. See arrayimage for more technical details on the function.
The normalization step includes a spatial correction but as this step also includes probe intensity summarization into probe-set intensities, the graph is not re-computed after normalization: there is no spatial position associated to probe-sets.
Dataset #1: No spatial bias; the color-coded values are homogeneous:
2D virtual PLM image for model characteristic: resids 1 / 2
ERT3
ERT1
ERT2
HG2
Dataset #2: We can see several spurious areas:
NoE2-1
NoE2-2


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawData2DVirtualImage_resids_good....
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawData2DVirtualImage_resids_good.jpg
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
2D virtual PLM image for model characteristic: resids 1 / 2
ERT3
ERT1
ERT2
HG2


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawData2DVirtualImage_resids_bad.j...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawData2DVirtualImage_resids_bad.jpg
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
NoE2-1
NoE2-2
S1CT-1
S1CT-2


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataPLMimage_NG1.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataPLMimage_NG1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
NG1
Raw intensities
NG1
Probe level model weights
NG1
Probe level model residuals
NG1
Probe level model residual signs

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

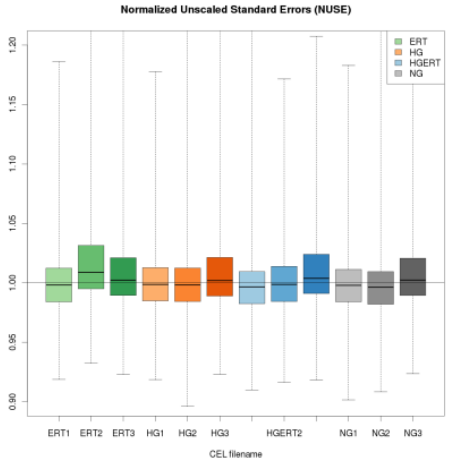
[Pipeline Description]

- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

BiGCaT bioinformatics

NUSE plot

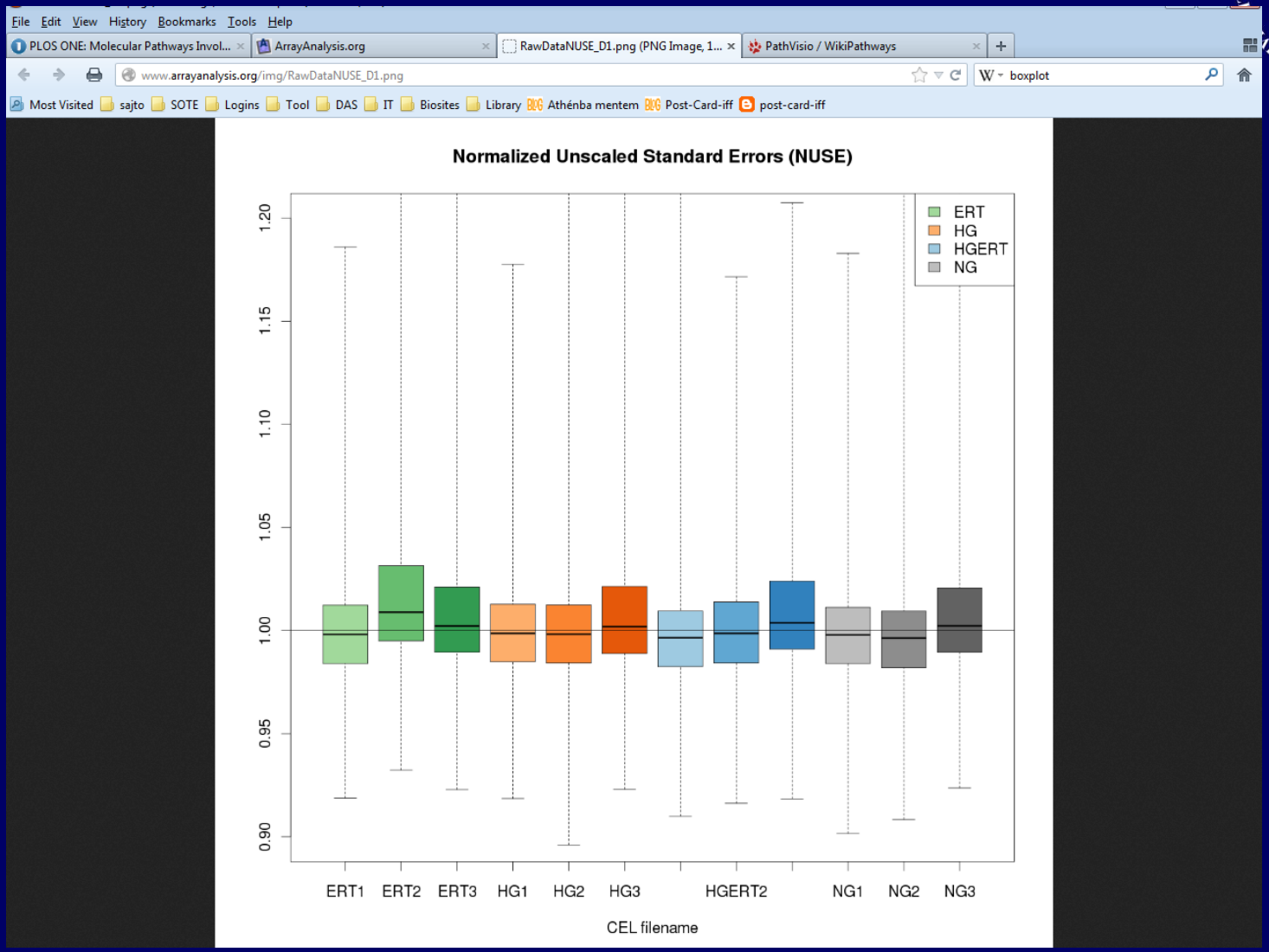
The Normalized Unscaled Standard Error (NUSE) is the individual probe error fitting the Probe-Level Model (the PLM models expression measures using a M-estimator robust regression). The NUSE values are standardized at the probe-set level across the arrays: median values for each probe-set are set to 1. The boxplots allow checking (1) if all distributions are centered near 1 – typically an array with a boxplot centered around 1.1 shows bad quality and (2) if one array has globally higher spread of NUSE distribution than others, which may also be a sign of low quality.



[Technical documentation of the function]

RLE plot

The Relative Log Expression (RLE) values are computed by calculating for each probe-set the ratio between the expression of a probe-set and the median expression of this probe-set across all arrays of the experiment. It is assumed that most probe-sets are not changed across the arrays, so it is expected that these ratios are around 0 on a log scale. The boxplots presenting the distribution of these log-ratios should then be centered near 0 and have similar spread. Other behavior would be a sign of low quality.


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataNUSE_D1.png (PNG Image, 1...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataNUSE_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Normalized Unscaled Standard Errors (NUSE)
ERT
HG
HGERT
NG
1.20
1.15
1.10
1.05
1.00
0.95
0.90
ERT1 ERT2 ERT3 HG1 HG2 HG3 HGERT2 NG1 NG2 NG3
CEL filename

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

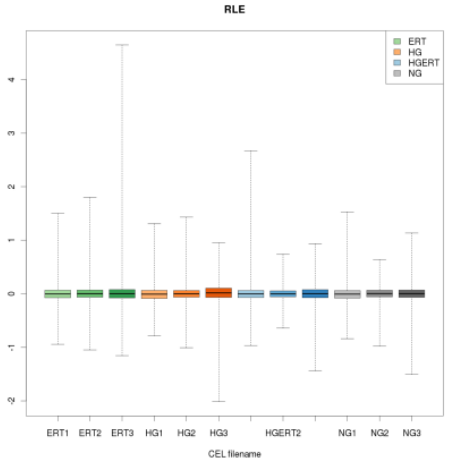
[Pipeline Description]

- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

BiGCaT bioinformatics

RLE plot

The Relative Log Expression (RLE) values are computed by calculating for each probe-set the ratio between the expression of a probe-set and the median expression of this probe-set across all arrays of the experiment. It is assumed that most probe-sets are not changed across the arrays, so it is expected that these ratios are around 0 on a log scale. The boxplots presenting the distribution of these log-ratios should then be centered near 0 and have similar spread. Other behavior would be a sign of low quality.



[Technical documentation of the function]

[Top]

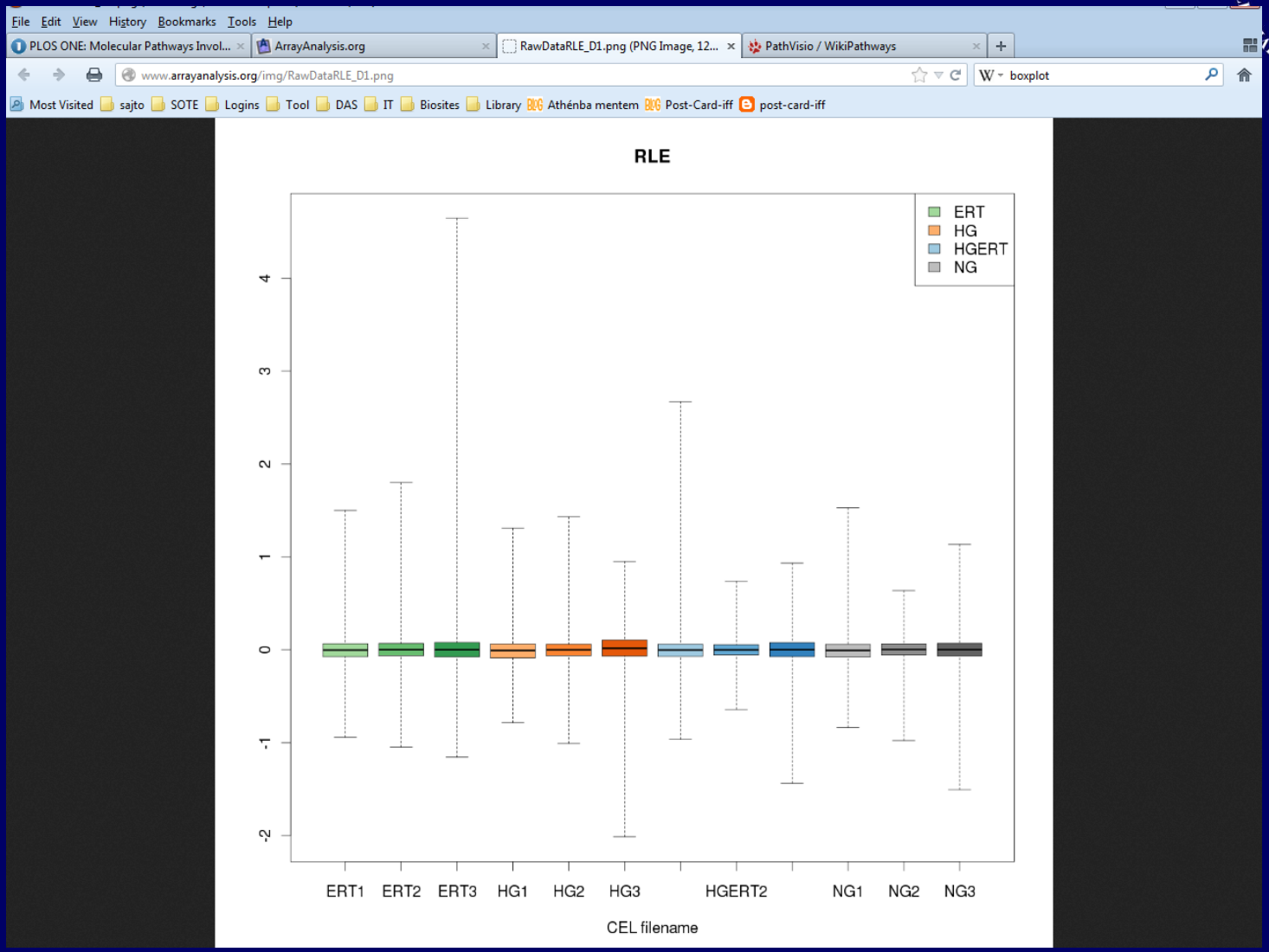
Correlation between arrays

Correlation plot

A correlation coefficient is computed for each pair of arrays in the dataset and is presented qualitatively on a colored matrix. The minimal value of this coefficient (given on the legend) gives a good idea of the dataset homogeneity: low coefficients indicate important differences between array intensities. We suggest to plot it before and after normalization: as the normalization makes the arrays more comparable, the correlation should be higher after this step.

Left : Raw data // Right : Normalized data.

Dataset #1: Signals are highly correlated, even before normalization. The lower correlation coefficient is 0.955 before and 0.99 after normalizatrion:


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataRLE_D1.png (PNG Image, 12...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataRLE_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
RLE
ERT
HG
HGERT
NG
4
3
2
1
0
-1
-2
ERT1 ERT2 ERT3 HG1 HG2 HG3 HGERT2 NG1 NG2 NG3
CEL filename

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

Array correlation
- Correlation plot
- PCA analysis
- Clustering

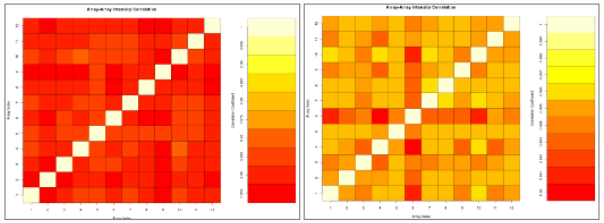
BiGCaT bioinformatics

Correlation plot

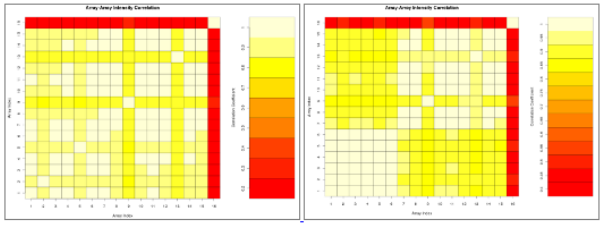
A correlation coefficient is computed for each pair of arrays in the dataset and is presented qualitatively on a colored matrix. The minimal value of this coefficient (given on the legend) gives a good idea of the dataset homogeneity: low coefficients indicate important differences between array intensities. We suggest to plot it before and after normalization: as the normalization makes the arrays more comparable, the correlation should be higher after this step.

Left : Raw data // Right : Normalized data.

Dataset #1: Signals are highly correlated, even before normalization. The lower correlation coefficient is 0.955 before and 0.99 after normalizatrion:



Dataset #2: Array 16 has a correlation coefficient <50% with all other arrays, even after normalization. Arrays 9 and 13 are correlated with a coefficient <90%.



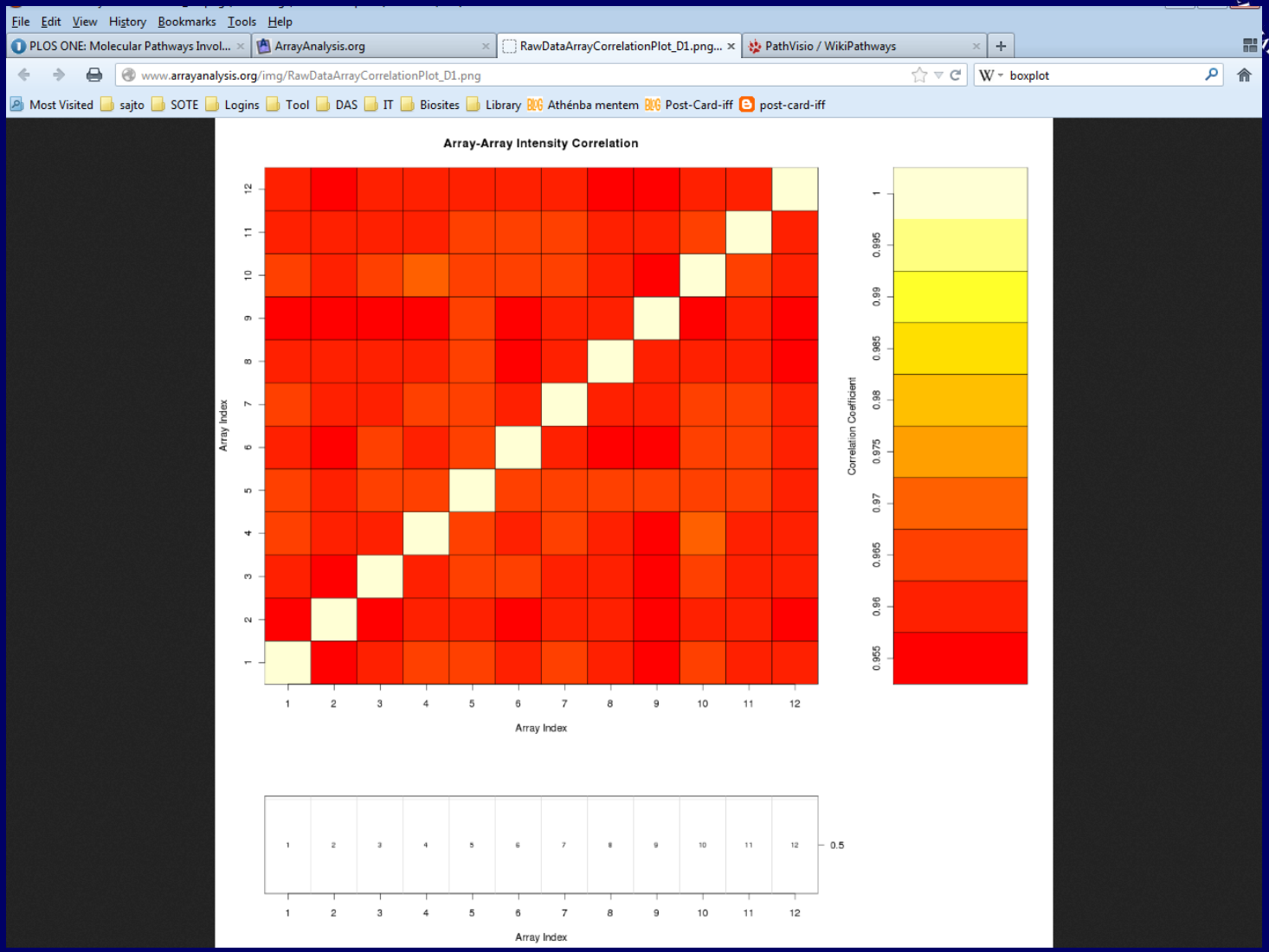
[Technical documentation of the function]

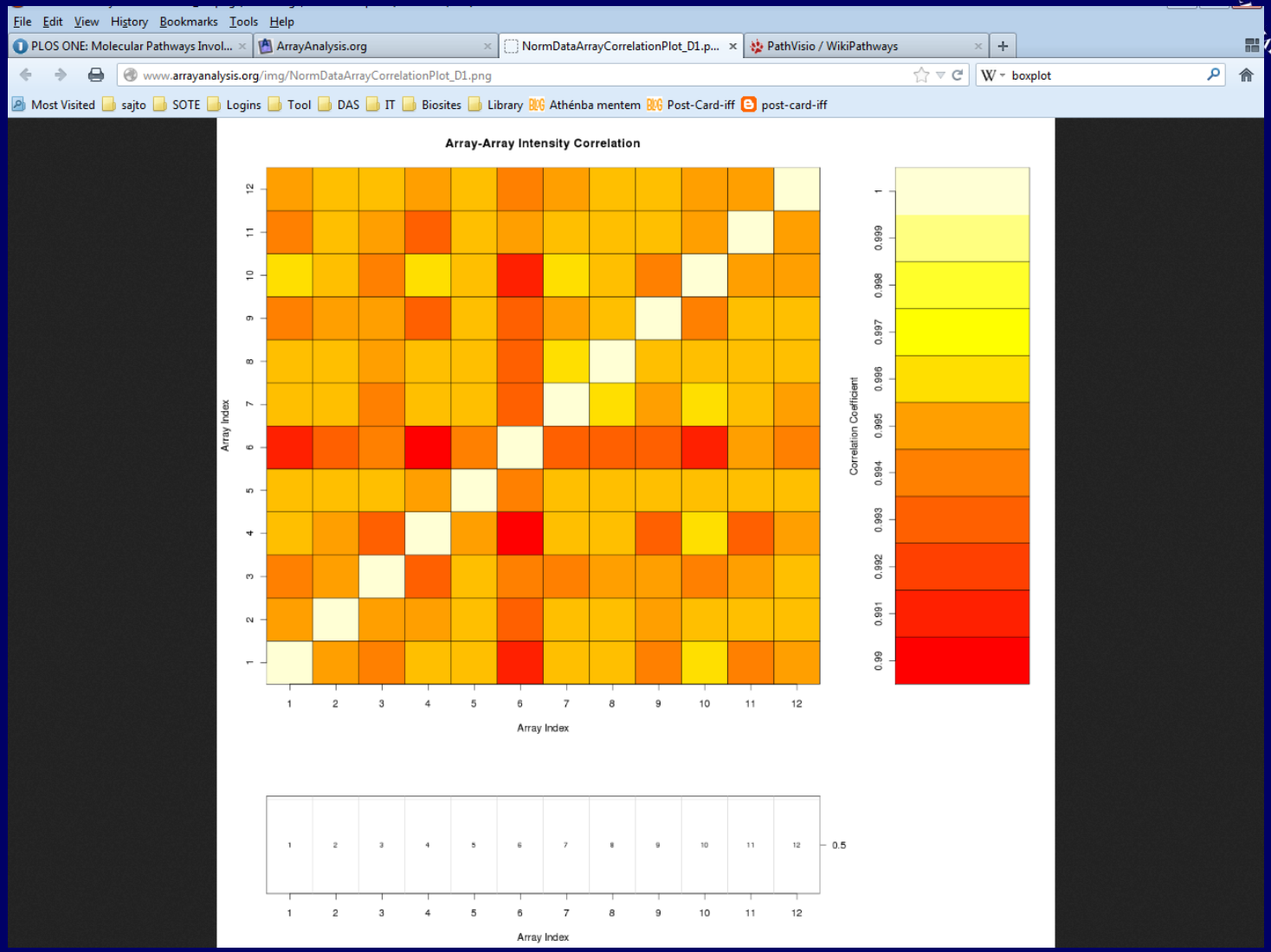
PCA analysis

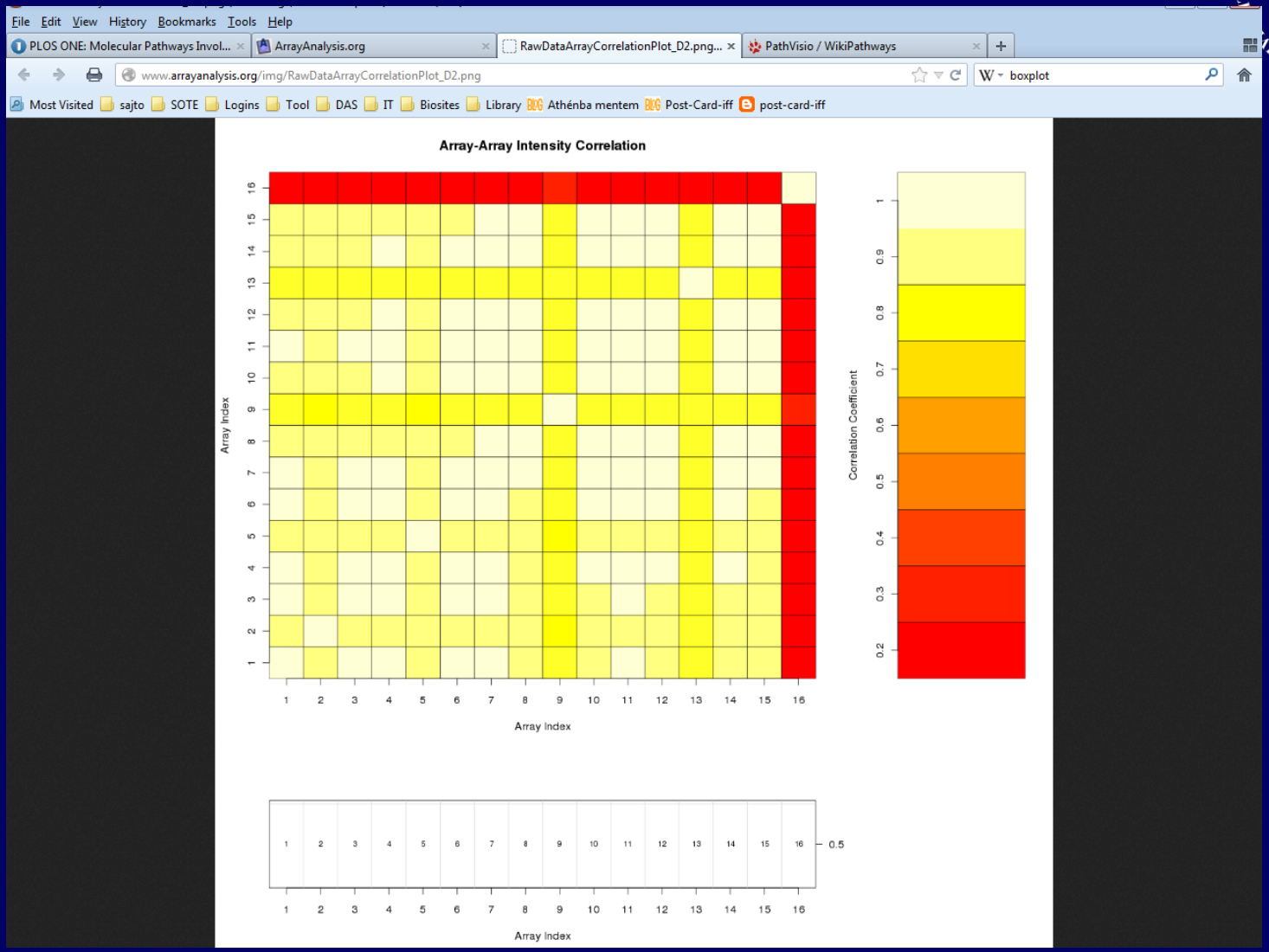
The PCA (Principal Component Analysis) gives an other view of the correlations of expression between arrays: the data are projected on several axes (or components), ordered by decreasing significativity; the first principal component (PC1) explains most of the variations of expression. (in the following example, PC1 explains almost 47% of the variance). Clusters of genes on a PCA plot present a strong correlation of expression signals. This analysis is proposed before and after normalization.

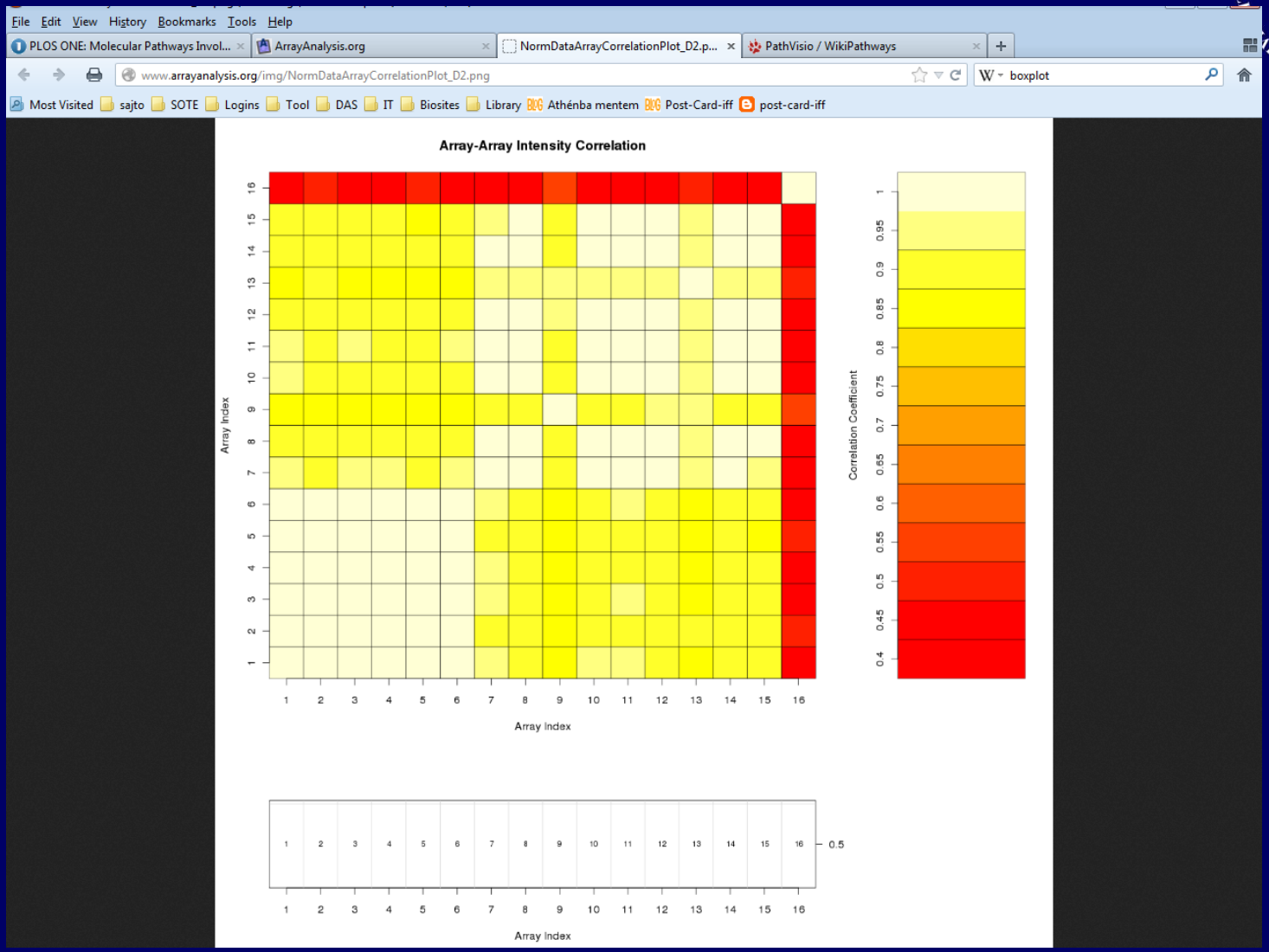
PCA analysis after RMA normalization

www.arrayanalysis.org/main.html#CP


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataArrayCorrelationPlot_D1.png...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataArrayCorrelationPlot_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Array-Array Intensity Correlation
Array Index
Array Index
Correlation Coefficient
1 0.995 0.99 0.985 0.98 0.975 0.97 0.965 0.96 0.955
0.5
Array Index


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataArrayCorrelationPlot_D1.p...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataArrayCorrelationPlot_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Array-Array Intensity Correlation
Array Index
Array Index
Correlation Coefficient
1
0.999
0.998
0.997
0.996
0.995
0.994
0.993
0.992
0.991
0.99
0.5
Array Index


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataArrayCorrelationPlot_D2.png...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataArrayCorrelationPlot_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Array-Array Intensity Correlation
Array Index
Array Index
Correlation Coefficient
0.5
Array Index


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataArrayCorrelationPlot_D2.p...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataArrayCorrelationPlot_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Array-Array Intensity Correlation
Array Index
Array Index
Correlation Coefficient
0.5
Array Index

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

Sample quality
- Sample prep controls
- 3'/5' for control genes
- RNA degradation plot

Hybridization quality
- Spike-in controls
- Background intensity
- Percent present
- PMA calls table
- Pos/Neg controls
- Affx control profiles

Signal comparability
- Scale factor
- Boxplot
- Density histogram
- MA plots
- Reference Layout
- Pos/Neg COI plot
- 2D spatial images
- NUSE plot
- RLE plot

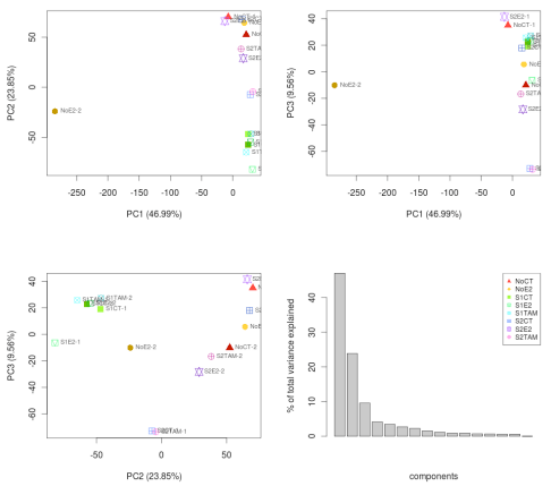
Array correlation
- Correlation plot
- PCA analysis
- Clustering

BiGCaT bioinformatics

PCA analysis

The PCA (Principal Component Analysis) gives an other view of the correlations of expression between arrays: the data are projected on several axes (or components), ordered by decreasing significativity; the first principal component (PC1) explains most of the variations of expression. (in the following example, PC1 explains almost 47% of the variance). Clusters of genes on a PCA plot present a strong correlation of expression signals. This analysis is proposed before and after normalization.

PCA analysis after RMA normalization



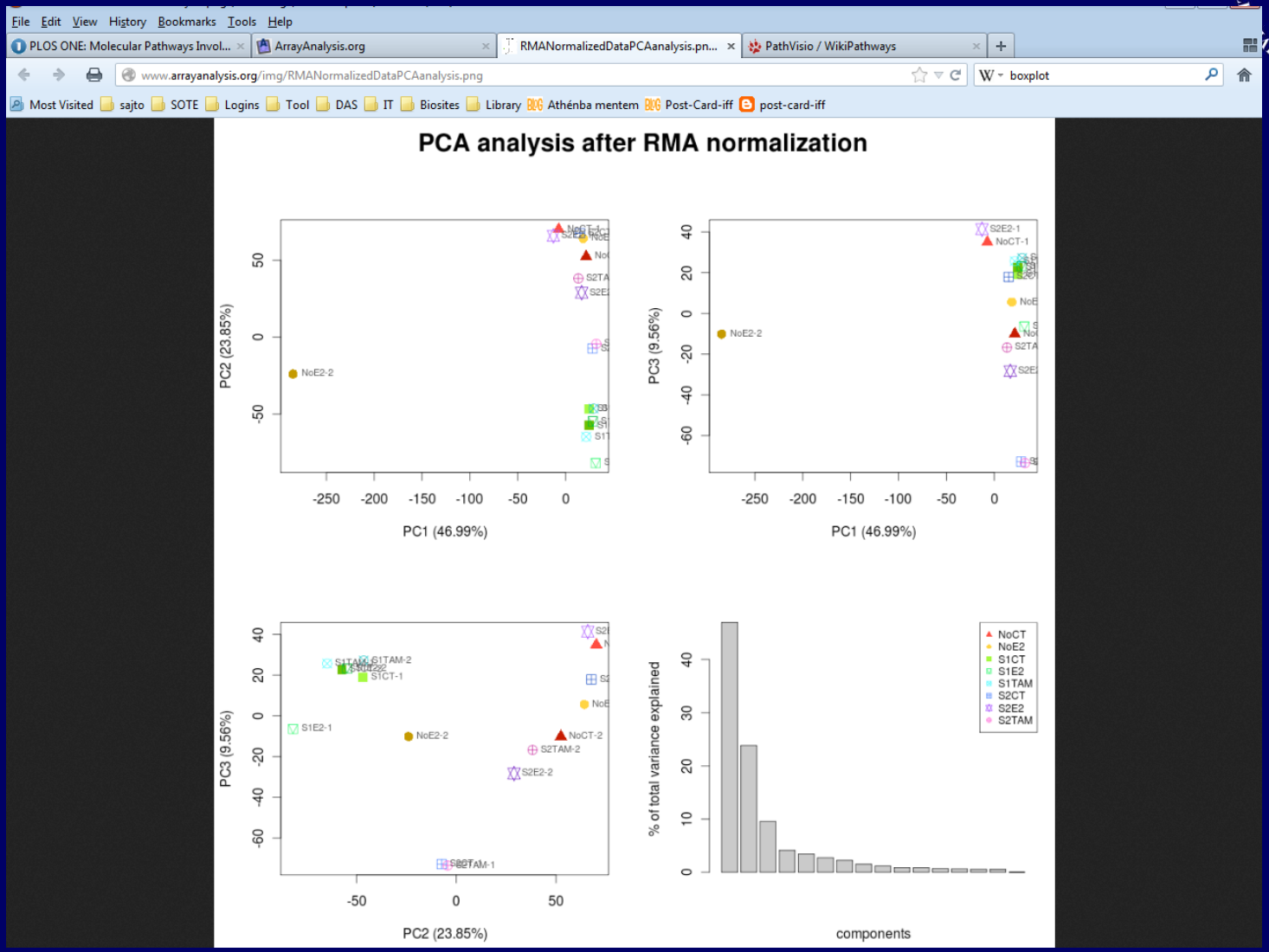
The PCA graph presents 4 plots: the array data are projected respectively on PC1 versus PC2, PC1 versus PC3 and PC2 versus PC3. The fourth plot is an histogram of the percentage of variance explained per each component, by decreasing order of significance: PC1, PC2, PC3...
We clearly see on this example that PC1 projects outlier array NoE2-2 very far from other arrays, because this variance is the most important in this dataset. By the way, PC2 and PC3 gives interesting results that are not spoiled by the outlier array: arrays from groups S1CT and S1TAM are all grouped together in the PC2 versus PC3 plot.

[Technical documentation of the function]

Hierarchical clustering

The Hierarchical Clustering plot is computed in two steps: first it computes an expression measure distance between all pairs of arrays and then it creates the tree from these distances. The distance absolute values are of interest as well as the groups of arrays that emerged from this analysis.

Left : Raw data // Right : Normalized data.


File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RMANormalizedDataPCAanalysis.pn...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RMANormalizedDataPCAanalysis.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
PCA analysis after RMA normalization
PC2 (23.85%)
PC1 (46.99%)
NoE2-2
PC3 (9.56%)
PC1 (46.99%)
S2E2-1
NoCT-1
NoE2-2
PC3 (9.56%)
PC2 (23.85%)
S1E2-1
NoE2-2
NoCT-2
S2TAM-2
S2E2-2
% of total variance explained
components
NoCT
NoE2
S1CT
S1E2
S1TAM
S2CT
S2E2
S2TAM

File Edit View History Bookmarks Tools Help

PLOS ONE: Molecular Pathways Invol... | ArrayAnalysis.org | PathVisio / WikiPathways

www.arrayanalysis.org

Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff

ARRAYANALYSIS.ORG welcome page

- Quick & easy on-line Affymetrix arrays Quality Control -

Get started | Download sources | Pipeline description | Documentation | Bug tracker

[Pipeline Description]

- **Sample quality**
  - Sample prep controls
  - 3'/5' for control genes
  - RNA degradation plot
- **Hybridization quality**
  - Spike-in controls
  - Background intensity
  - Percent present
  - PMA calls table
  - Pos/Neg controls
  - Affx control profiles
- **Signal comparability**
  - Scale factor
  - Boxplot
  - Density histogram
  - MA plots
  - Reference Layout
  - Pos/Neg COI plot
  - 2D spatial images
  - NUSE plot
  - RLE plot
- **Array correlation**
  - Correlation plot
  - PCA analysis
  - Clustering

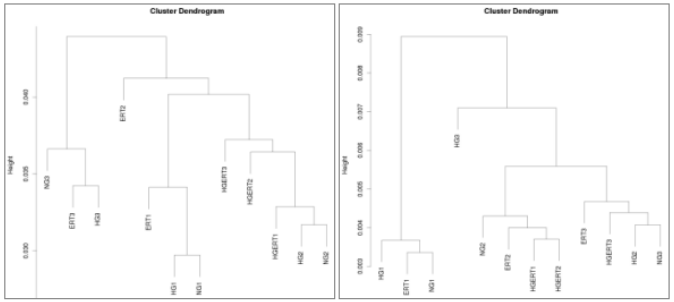
BiGCaT bioinformatics

Hierarchical clustering

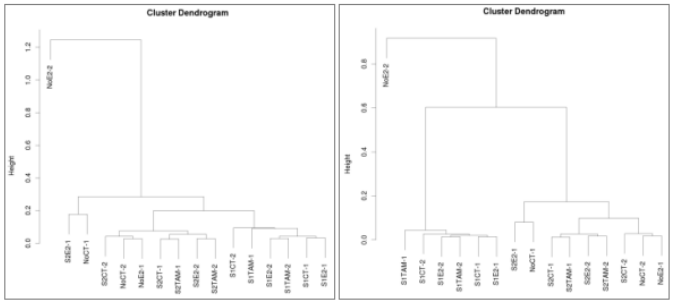
The Hierarchical Clustering plot is computed in two steps: first it computes an expression measure distance between all pairs of arrays and then it creates the tree from these distances. The distance absolute values are of interest as well as the groups of arrays that emerged from this analysis.

Left : Raw data // Right : Normalized data.

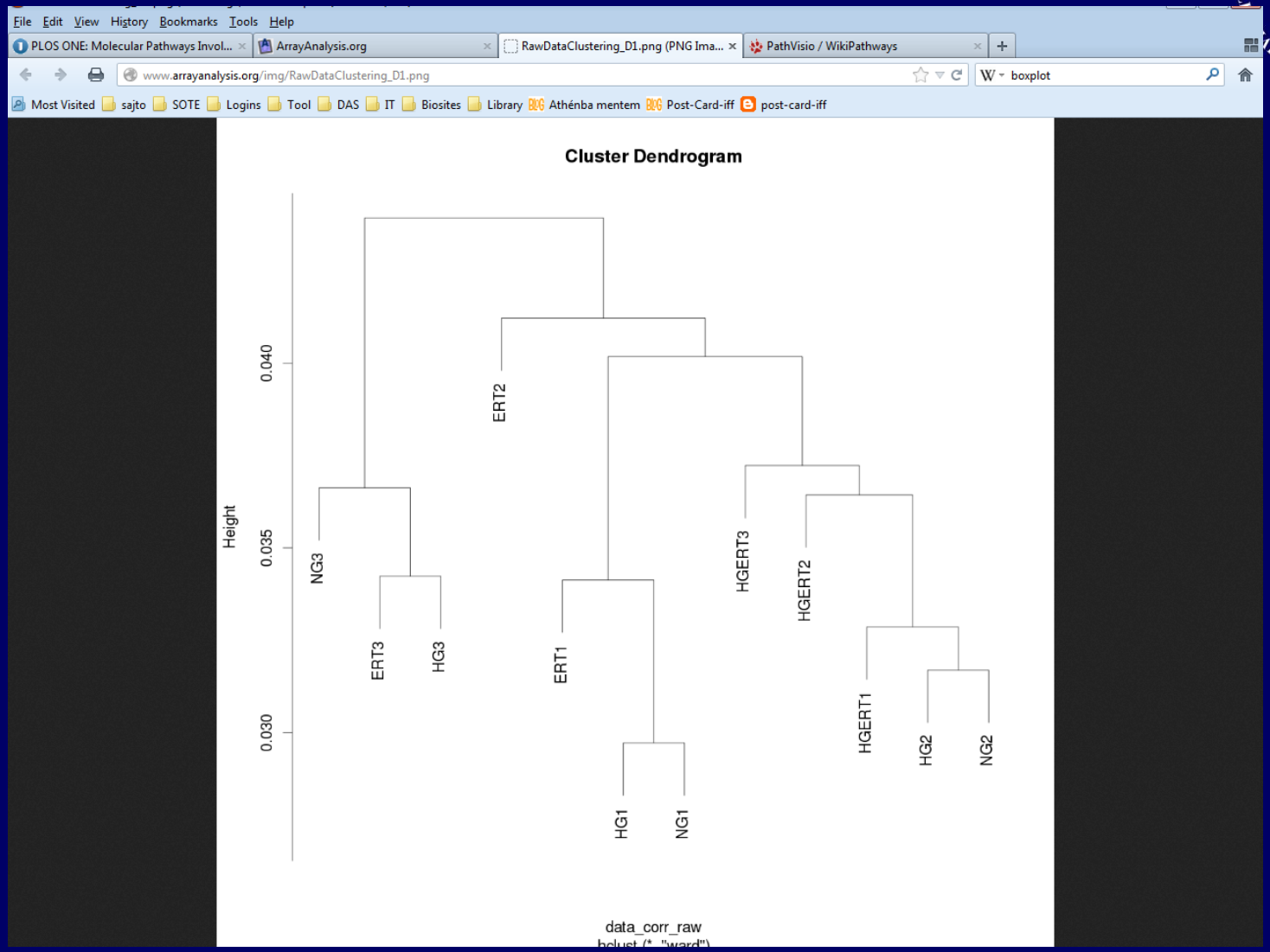
Dataset #1: The distance value ranges are very low (0.02 for raw data and less than 0.1 after normalization). The expression values are very homogeneous between arrays.

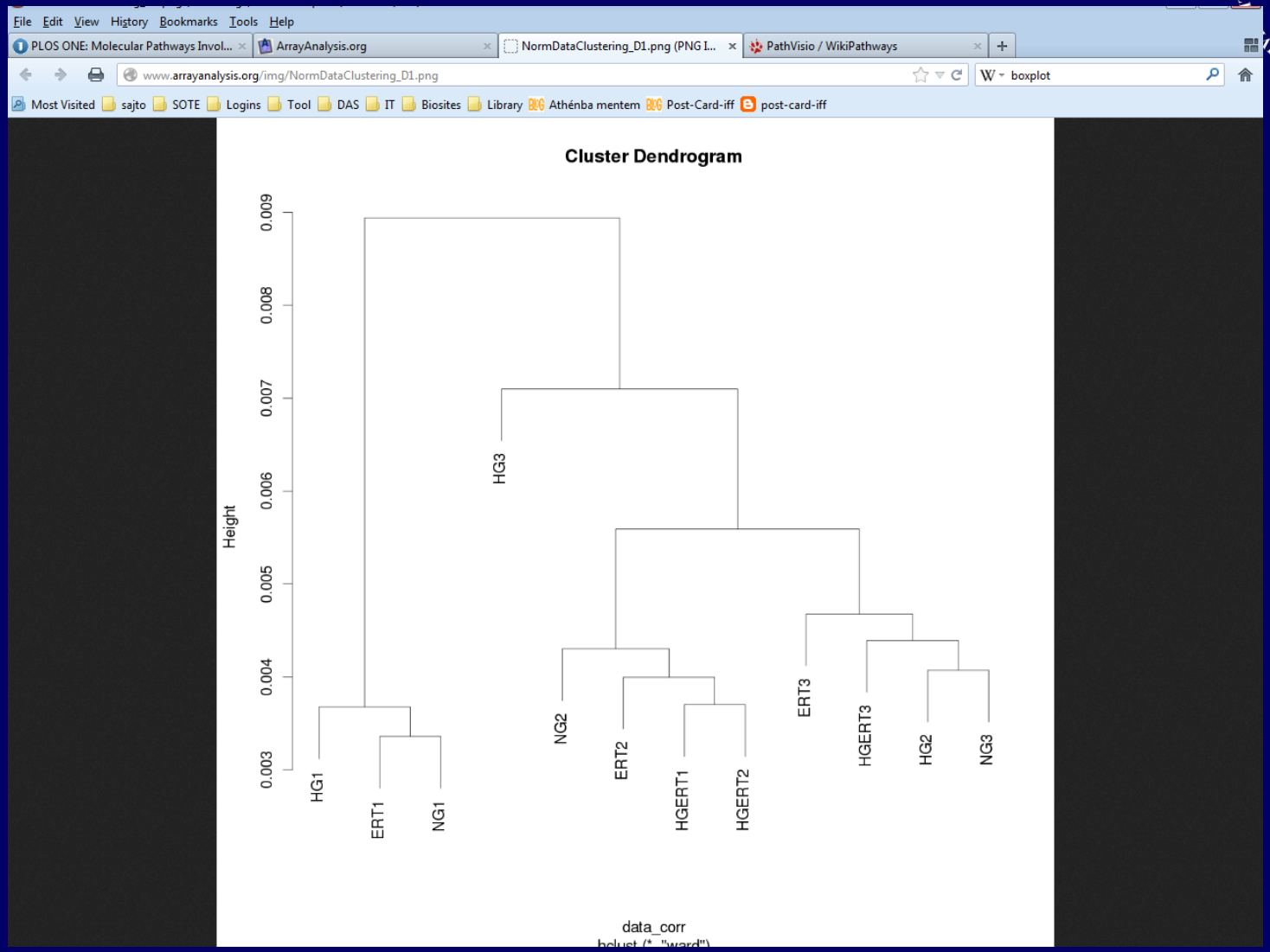


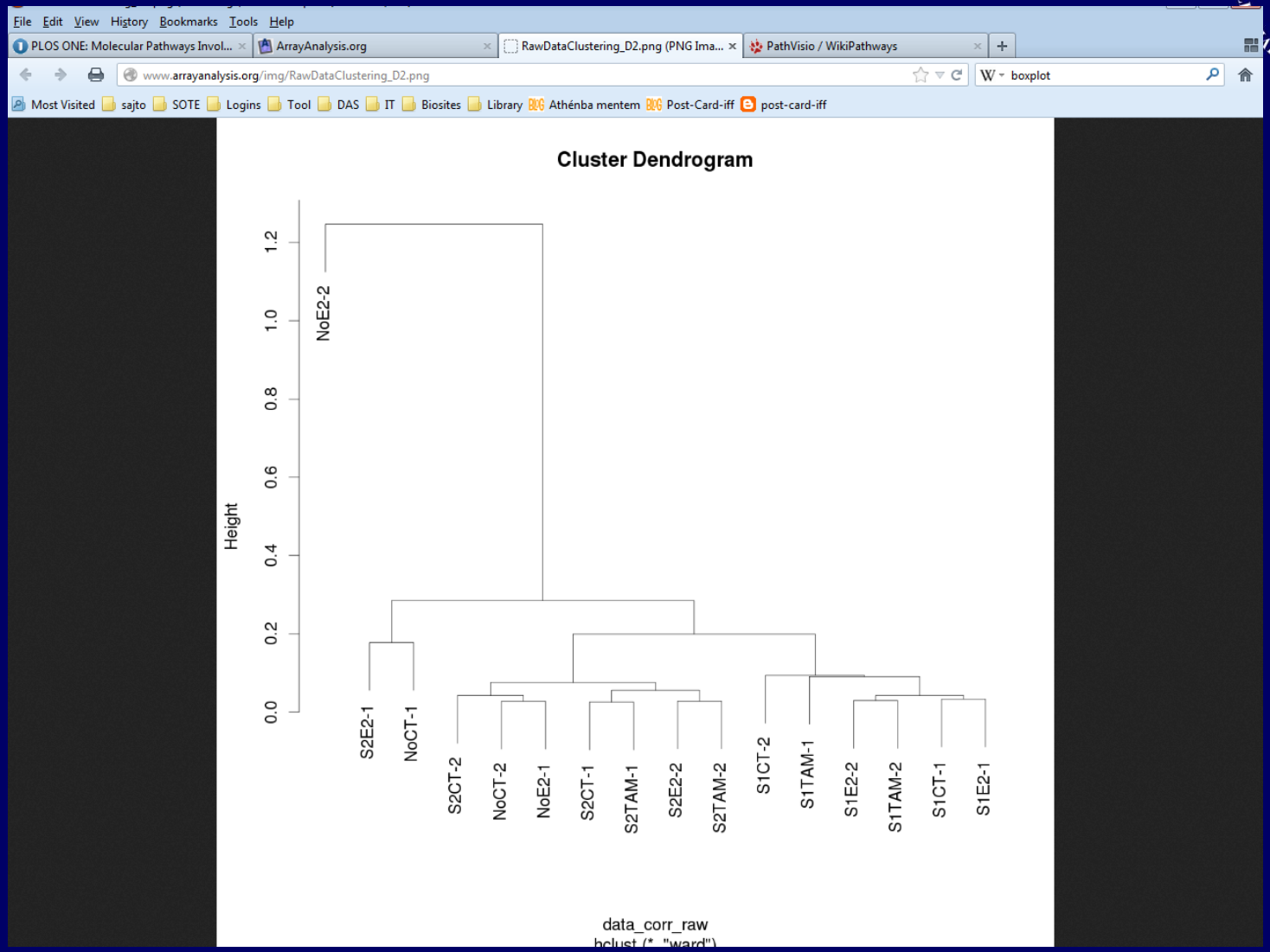
Dataset #2: The distance ranges are higher for this dataset (1.2 for raw data and around 0.8 after normalization): we clearly see here that array NoE2-2 is very different from all other arrays even after normalization which confirms what most of the quality control graphs conluded.

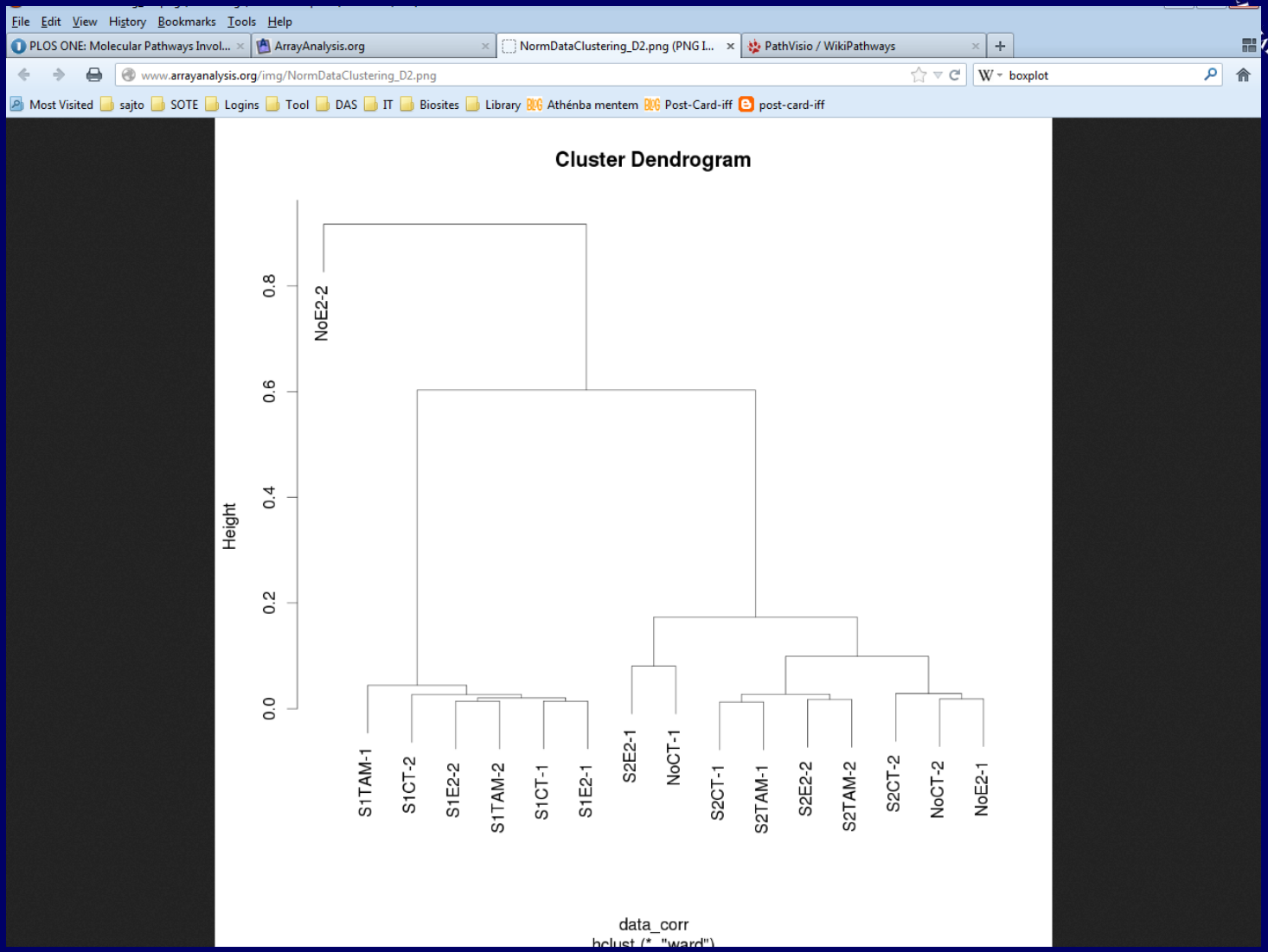


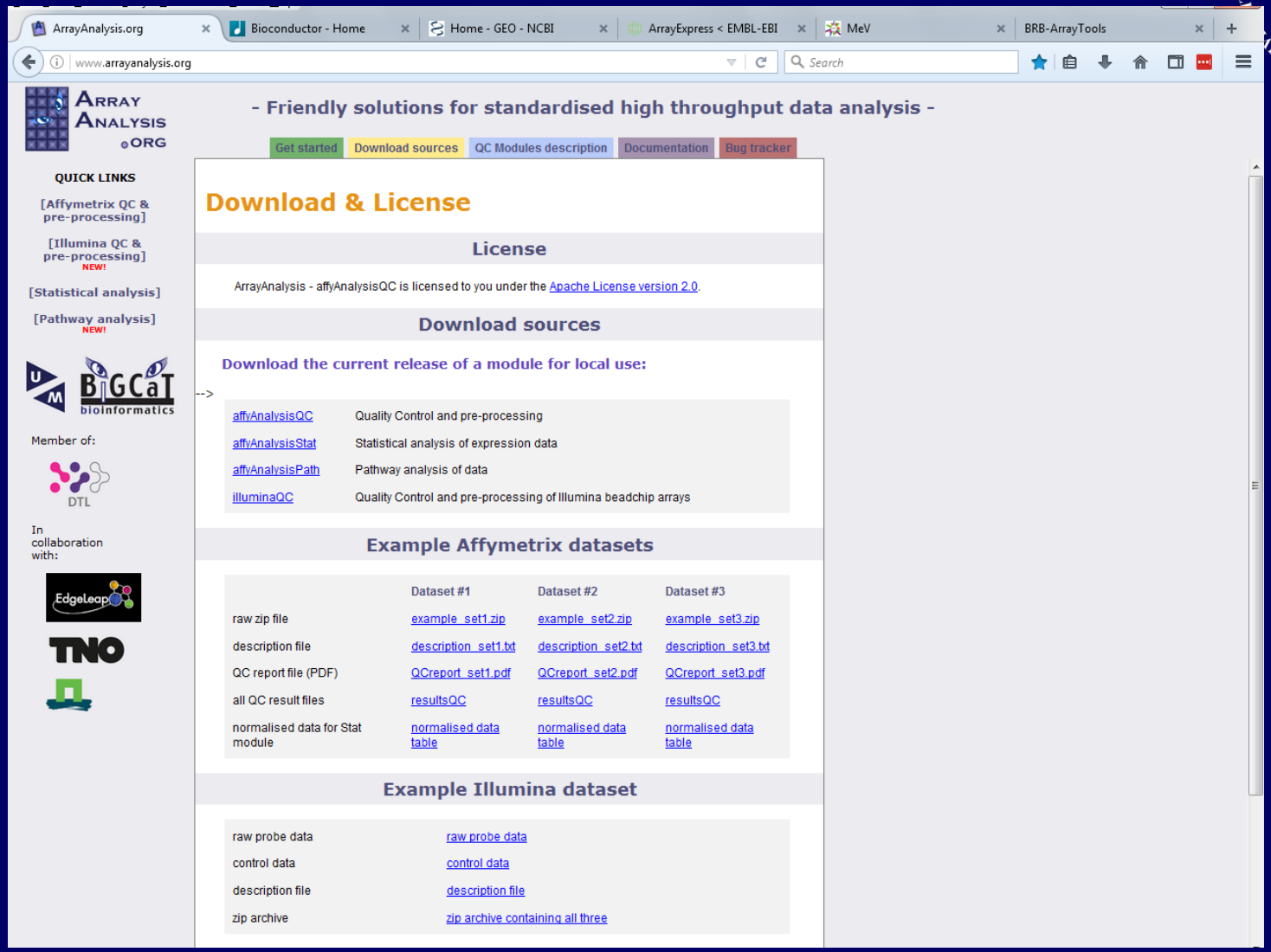
[Technical documentation of the function]

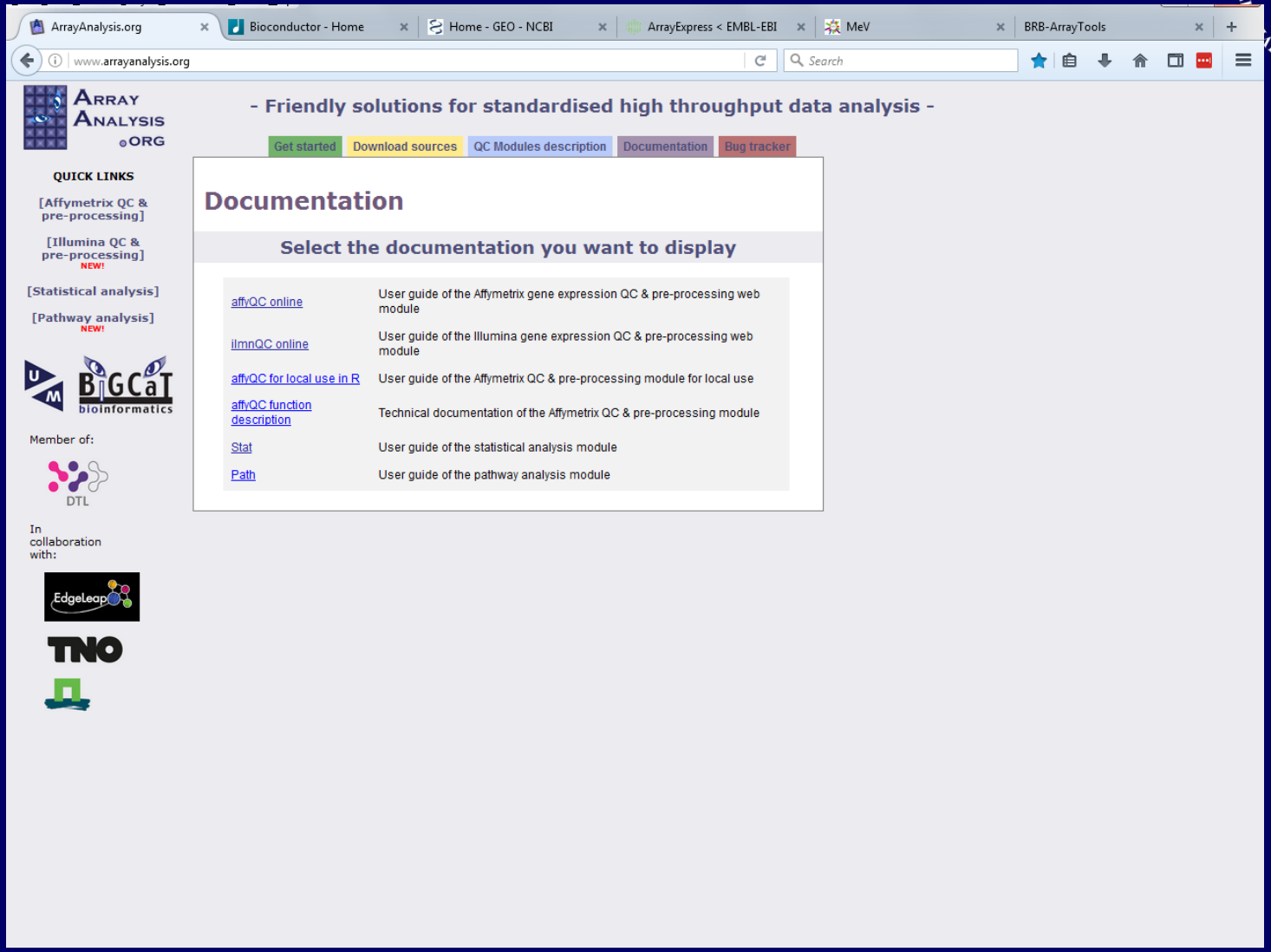

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataClustering_D1.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataClustering_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Cluster Dendrogram
Height
0.040
0.035
0.030
NG3
ERT3
HG3
ERT2
ERT1
HG1
NG1
HGERT3
HGERT2
HGERT1
HG2
NG2
data_corr_raw

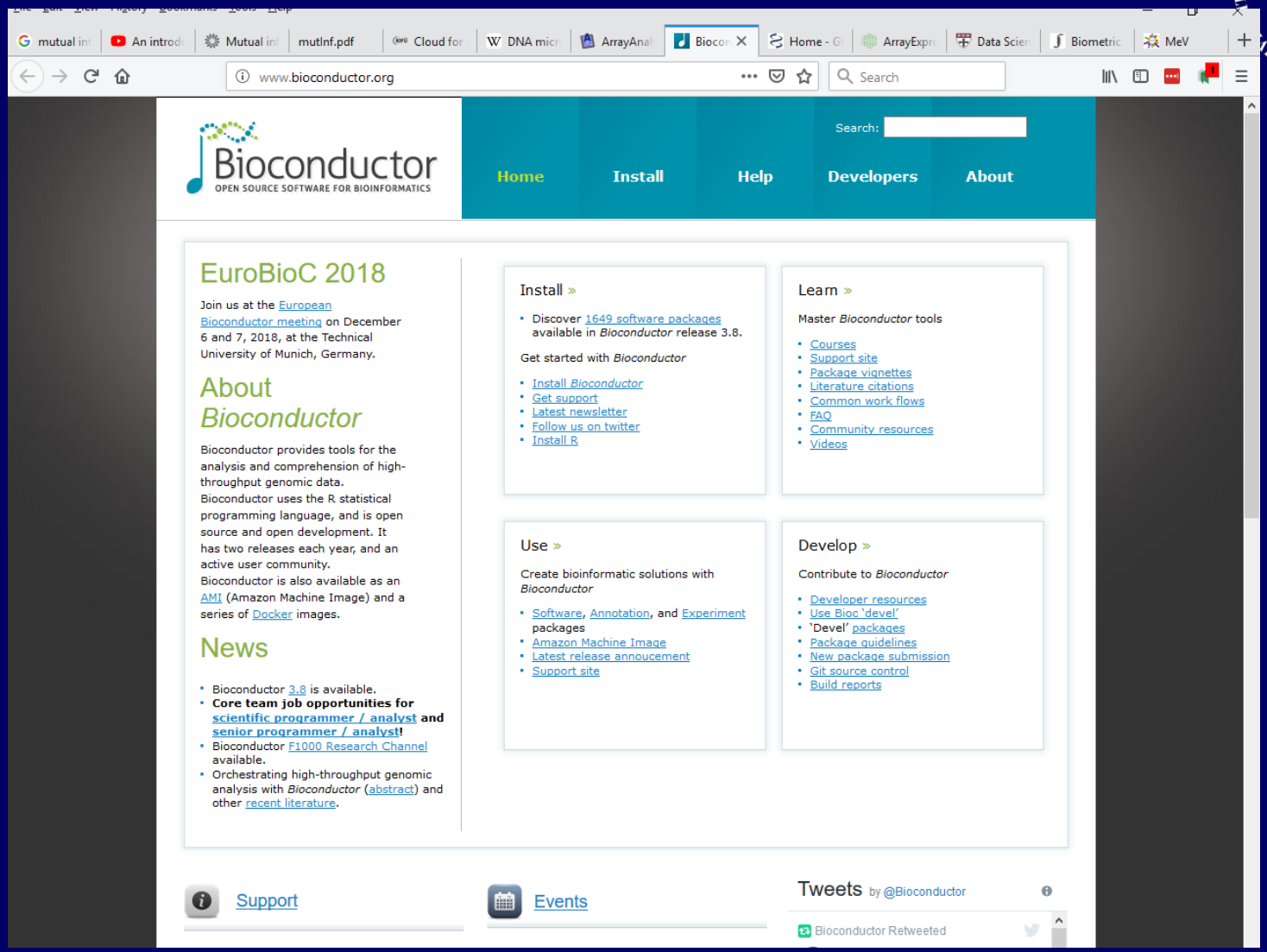

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataClustering_D1.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataClustering_D1.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Cluster Dendrogram
Height
0.009
0.008
0.007
0.006
0.005
0.004
0.003
HG3
HG1
ERT1
NG1
NG2
ERT2
HGERT1
HGERT2
ERT3
HGERT3
HG2
NG3
data_corr
hclust (*, "ward")

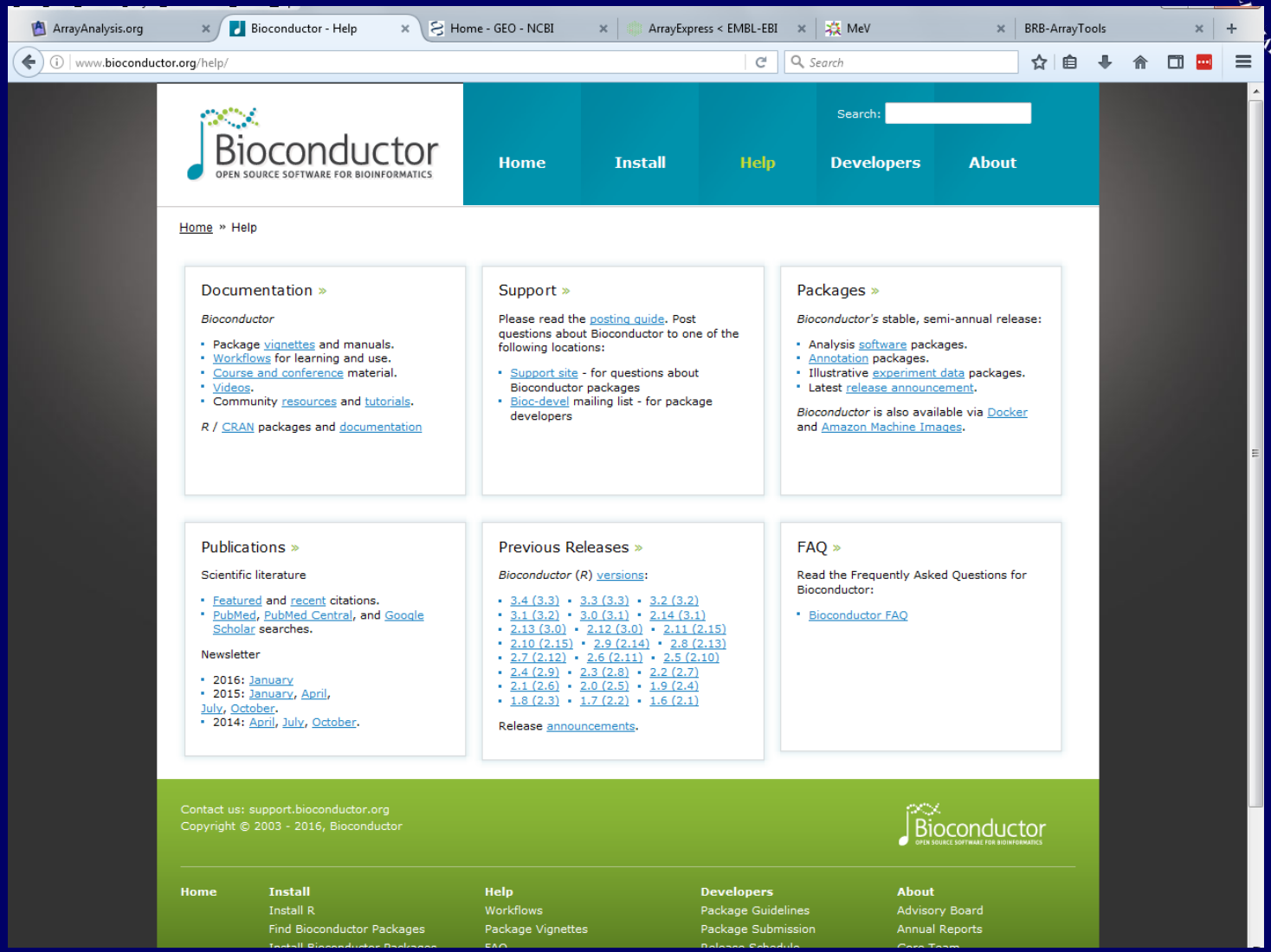

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
RawDataClustering_D2.png (PNG Ima...
PathVisio / WikiPathways
www.arrayanalysis.org/img/RawDataClustering_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Cluster Dendrogram
Height
0.0
0.2
0.4
0.6
0.8
1.0
1.2
NoE2-2
S2E2-1
NoCT-1
S2CT-2
NoCT-2
NoE2-1
S2CT-1
S2TAM-1
S2E2-2
S2TAM-2
S1CT-2
S1TAM-1
S1E2-2
S1TAM-2
S1CT-1
S1E2-1
data_corr_raw
hclust (*, "ward")

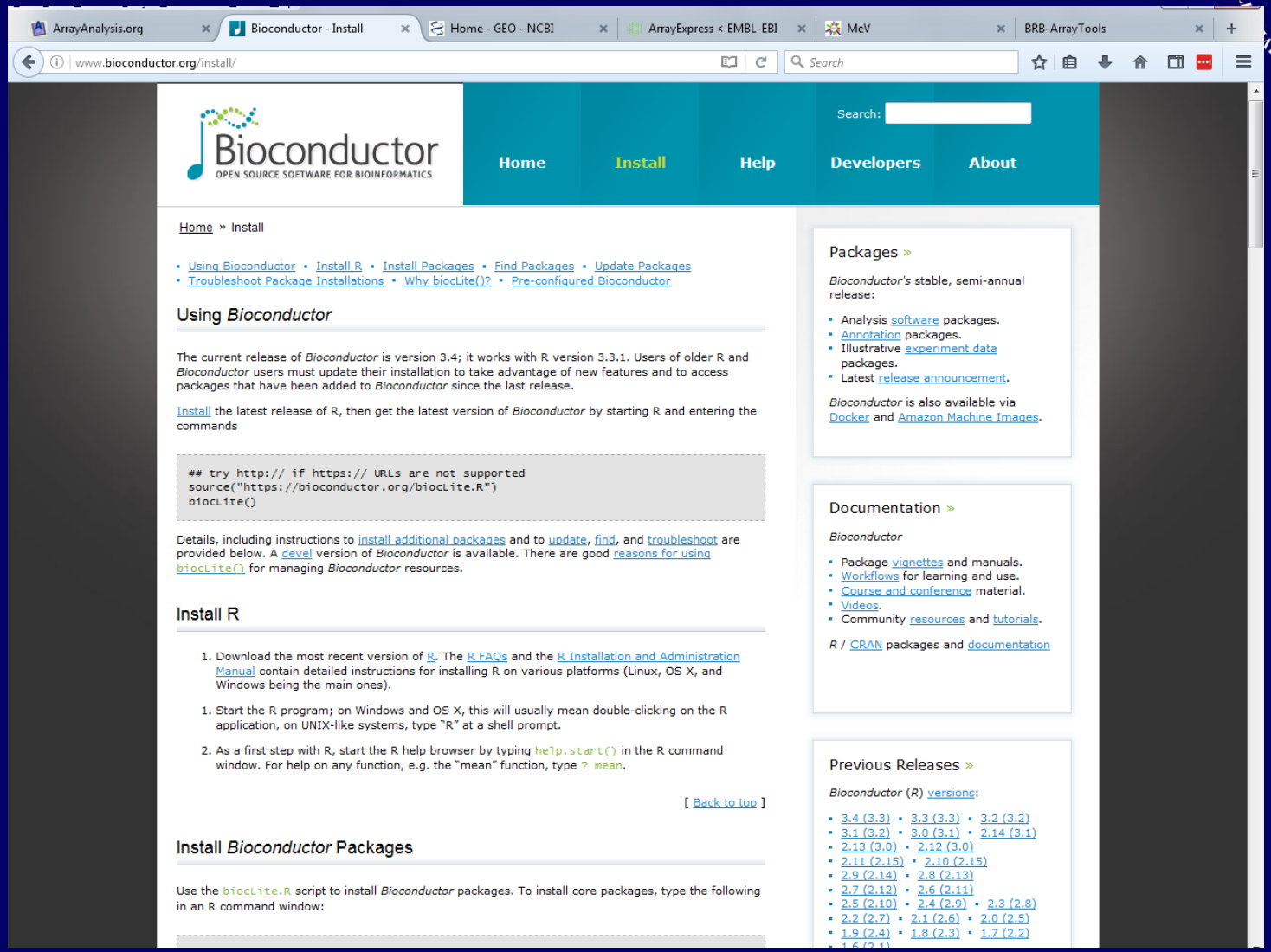

File Edit View History Bookmarks Tools Help
PLOS ONE: Molecular Pathways Invol...
ArrayAnalysis.org
NormDataClustering_D2.png (PNG I...
PathVisio / WikiPathways
www.arrayanalysis.org/img/NormDataClustering_D2.png
boxplot
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
Cluster Dendrogram
Height
0.0
0.2
0.4
0.6
0.8
NoE2-2
S1TAM-1
S1CT-2
S1E2-2
S1TAM-2
S1CT-1
S1E2-1
S2E2-1
NoCT-1
S2CT-1
S2TAM-1
S2E2-2
S2TAM-2
S2CT-2
NoCT-2
NoE2-1
data_corr
hclust (*, "ward")

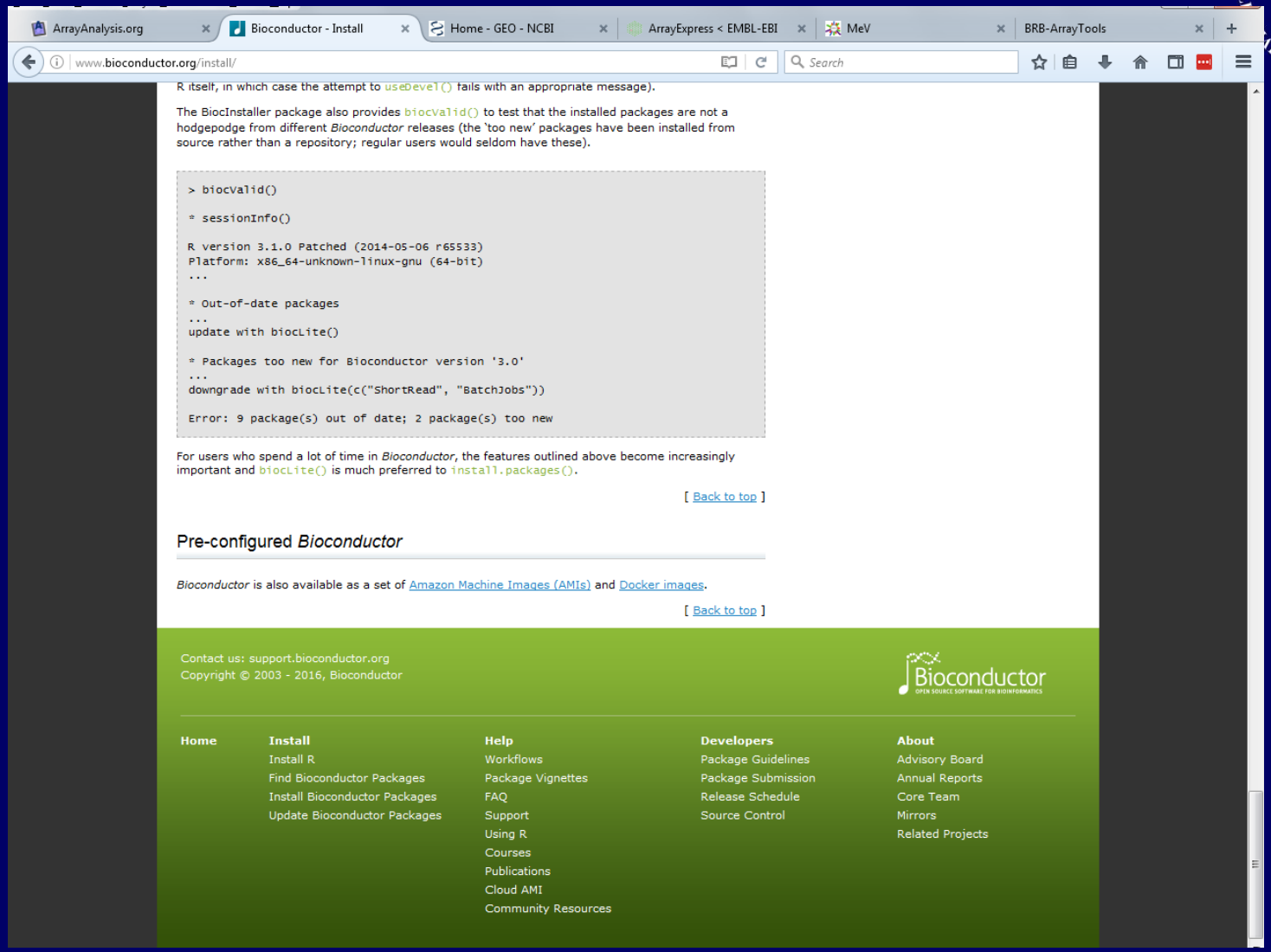

ArrayAnalysis.org
Bioconductor - Home
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.arrayanalysis.org
Search
ARRAY ANALYSIS .ORG
- Friendly solutions for standardised high throughput data analysis -
Get started
Download sources
QC Modules description
Documentation
Bug tracker
QUICK LINKS
[Affymetrix QC & pre-processing]
[Illumina QC & pre-processing]
NEW!
[Statistical analysis]
[Pathway analysis]
NEW!
BiGCaT bioinformatics
Member of:
DTL
In collaboration with:
EdgeLeap
TNO
Download & License
License
ArrayAnalysis - affyAnalysisQC is licensed to you under the Apache License version 2.0.
Download sources
Download the current release of a module for local use:
affyAnalysisQC Quality Control and pre-processing
affyAnalysisStat Statistical analysis of expression data
affyAnalysisPath Pathway analysis of data
illuminaQC Quality Control and pre-processing of Illumina beadchip arrays
Example Affymetrix datasets
Dataset #1 Dataset #2 Dataset #3
raw zip file example_set1.zip example_set2.zip example_set3.zip
description file description_set1.txt description_set2.txt description_set3.txt
QC report file (PDF) QCreport_set1.pdf QCreport_set2.pdf QCreport_set3.pdf
all QC result files resultsQC resultsQC resultsQC
normalised data for Stat module normalised data table normalised data table normalised data table
Example Illumina dataset
raw probe data raw probe data
control data control data
description file description file
zip archive zip archive containing all three

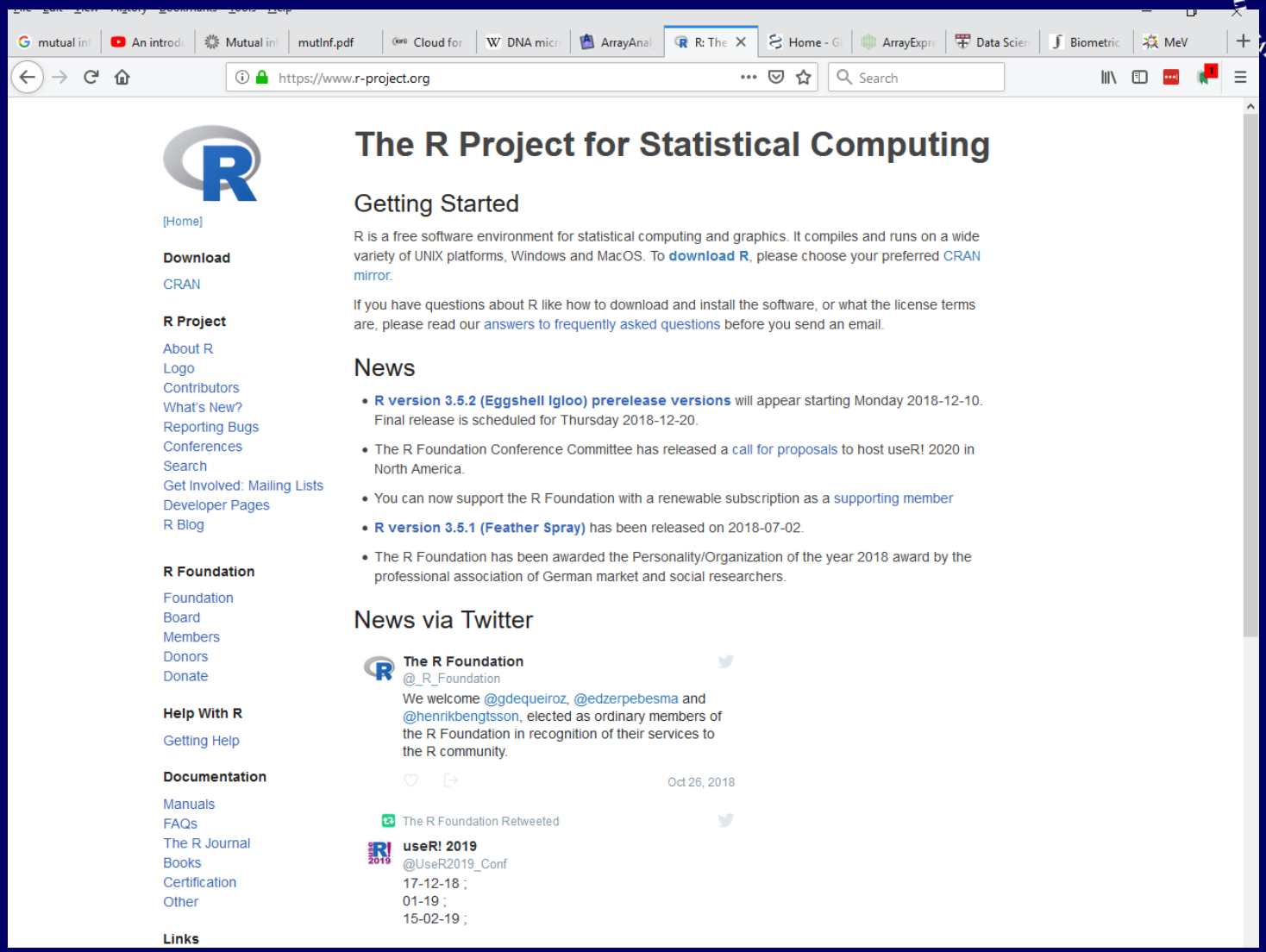

ArrayAnalysis.org
Bioconductor - Home
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.arrayanalysis.org
Search
ARRAY ANALYSIS .ORG
- Friendly solutions for standardised high throughput data analysis -
Get started
Download sources
QC Modules description
Documentation
Bug tracker
QUICK LINKS
[Affymetrix QC & pre-processing]
[Illumina QC & pre-processing]
NEW!
[Statistical analysis]
[Pathway analysis]
NEW!
BiGCaT bioinformatics
Member of:
DTL
In collaboration with:
EdgeLeap
TNO
Documentation
Select the documentation you want to display
affyQC online
User guide of the Affymetrix gene expression QC & pre-processing web module
illmnQC online
User guide of the Illumina gene expression QC & pre-processing web module
affyQC for local use in R
User guide of the Affymetrix QC & pre-processing module for local use
affyQC function description
Technical documentation of the Affymetrix QC & pre-processing module
Stat
User guide of the statistical analysis module
Path
User guide of the pathway analysis module

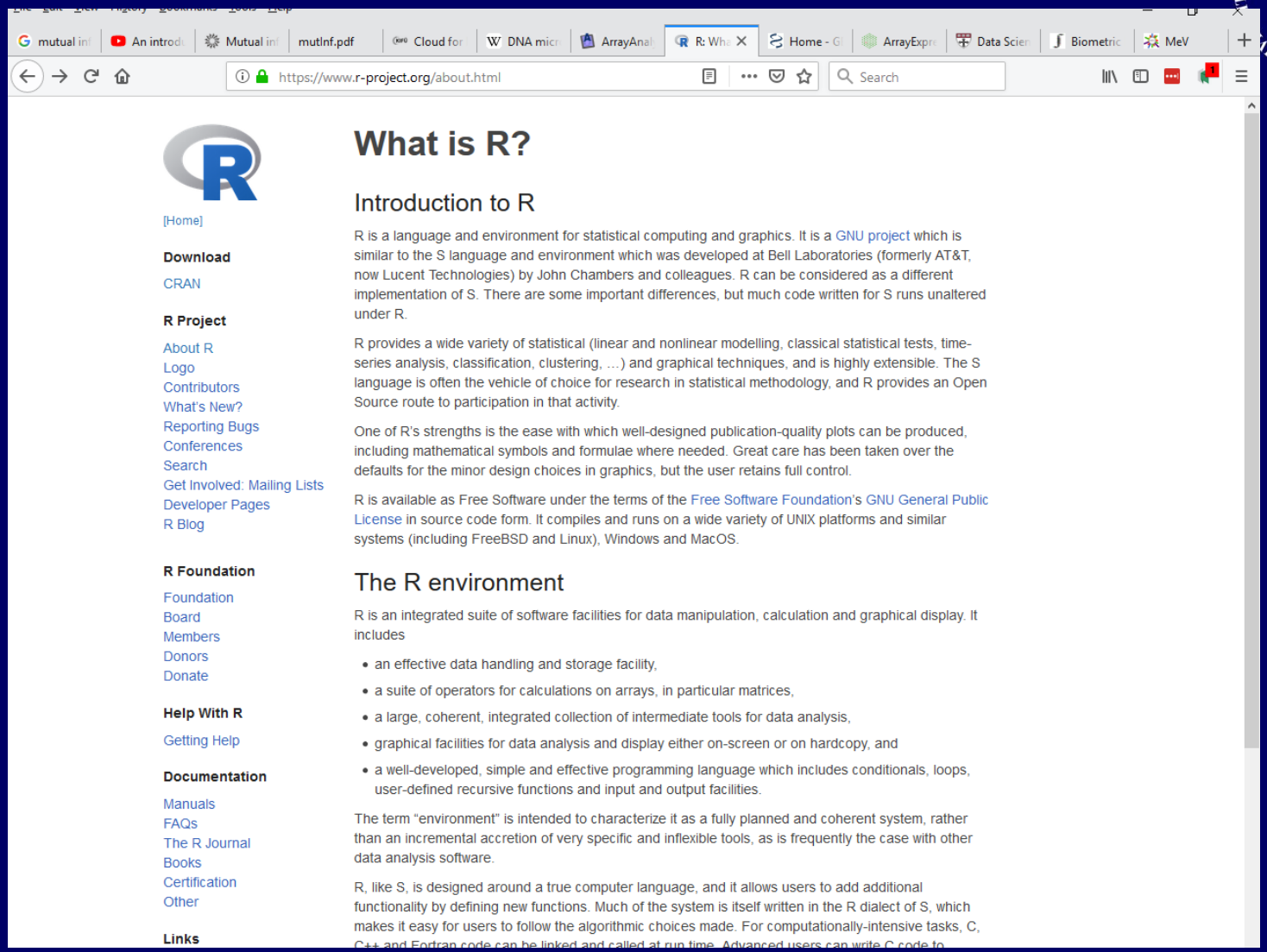

www.bioconductor.org
Search
Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS
Search:
Home
Install
Help
Developers
About
EuroBioC 2018
Join us at the European Bioconductor meeting on December 6 and 7, 2018, at the Technical University of Munich, Germany.
About Bioconductor
Bioconductor provides tools for the analysis and comprehension of high-throughput genomic data.
Bioconductor uses the R statistical programming language, and is open source and open development. It has two releases each year, and an active user community.
Bioconductor is also available as an AMI (Amazon Machine Image) and a series of Docker images.
News
Bioconductor 3.8 is available.
Core team job opportunities for scientific programmer / analyst and senior programmer / analyst!
Bioconductor F1000 Research Channel available.
Orchestrating high-throughput genomic analysis with Bioconductor (abstract) and other recent literature.
Install »
Discover 1649 software packages available in Bioconductor release 3.8.
Get started with Bioconductor
Install Bioconductor
Get support
Latest newsletter
Follow us on twitter
Install R
Learn »
Master Bioconductor tools
Courses
Support site
Package vignettes
Literature citations
Common work flows
FAQ
Community resources
Videos
Use »
Create bioinformatic solutions with Bioconductor
Software, Annotation, and Experiment packages
Amazon Machine Image
Latest release annoucement
Support site
Develop »
Contribute to Bioconductor
Developer resources
Use Bioc 'devel'
'Devel' packages
Package guidelines
New package submission
Git source control
Build reports
Support
Events
Tweets by @Bioconductor
Bioconductor Retweeted

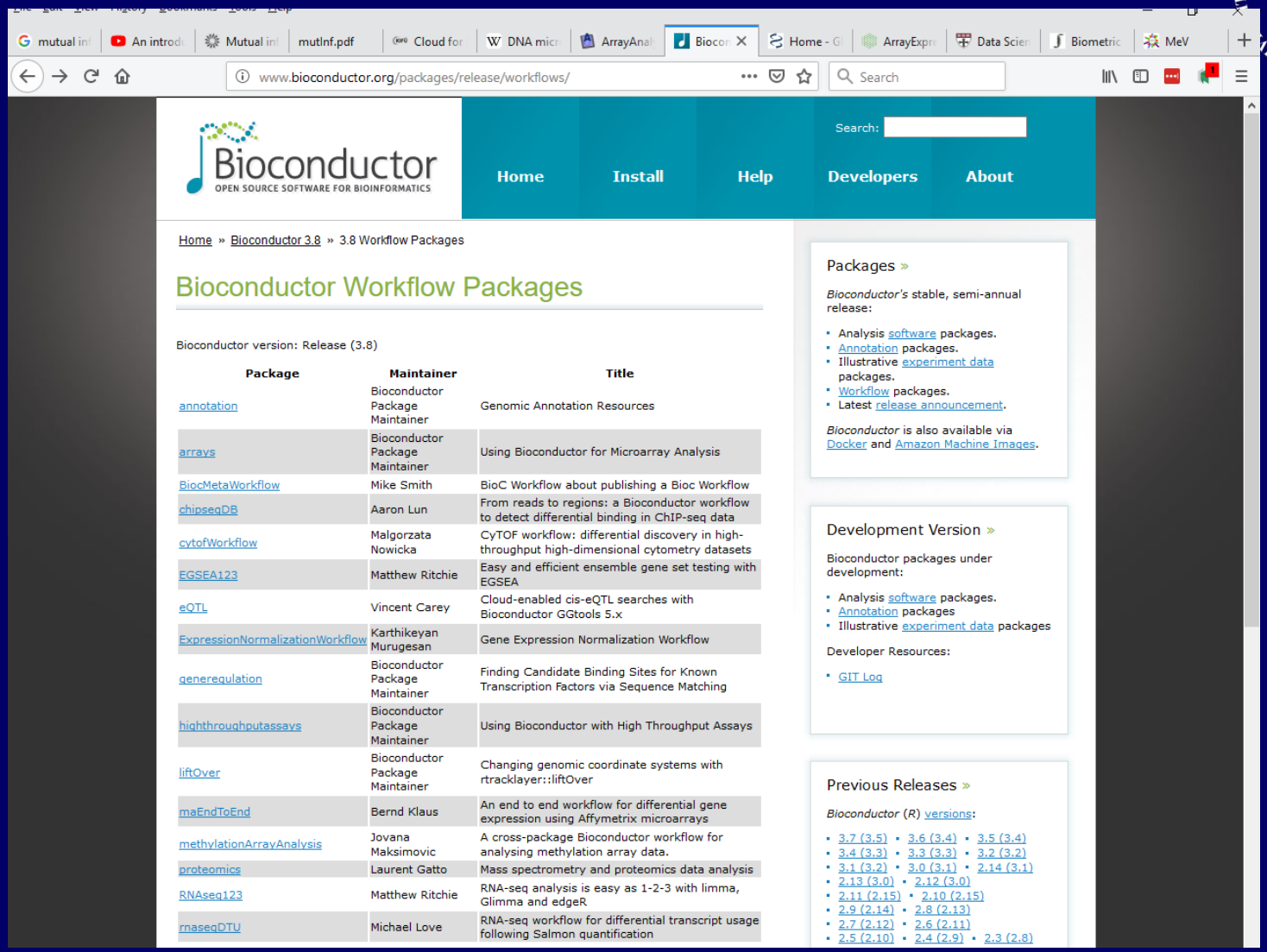

ArrayAnalysis.org
Bioconductor - Help
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.bioconductor.org/help/
Search
Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS
Search:
Home
Install
Help
Developers
About
Home » Help
Documentation »
Bioconductor
Package vignettes and manuals.
Workflows for learning and use.
Course and conference material.
Videos.
Community resources and tutorials.
R / CRAN packages and documentation
Support »
Please read the posting guide. Post questions about Bioconductor to one of the following locations:
Support site - for questions about Bioconductor packages
Bioc-devel mailing list - for package developers
Packages »
Bioconductor's stable, semi-annual release:
Analysis software packages.
Annotation packages.
Illustrative experiment data packages.
Latest release announcement.
Bioconductor is also available via Docker and Amazon Machine Images.
Publications »
Scientific literature
Featured and recent citations.
PubMed, PubMed Central, and Google Scholar searches.
Newsletter
2016: January
2015: January, April, July, October.
2014: April, July, October.
Previous Releases »
Bioconductor (R) versions:
3.4 (3.3) • 3.3 (3.3) • 3.2 (3.2)
3.1 (3.2) • 3.0 (3.1) • 2.14 (3.1)
2.13 (3.0) • 2.12 (3.0) • 2.11 (2.15)
2.10 (2.15) • 2.9 (2.14) • 2.8 (2.13)
2.7 (2.12) • 2.6 (2.11) • 2.5 (2.10)
2.4 (2.9) • 2.3 (2.8) • 2.2 (2.7)
2.1 (2.6) • 2.0 (2.5) • 1.9 (2.4)
1.8 (2.3) • 1.7 (2.2) • 1.6 (2.1)
Release announcements.
FAQ »
Read the Frequently Asked Questions for Bioconductor:
Bioconductor FAQ
Contact us: support.bioconductor.org
Copyright © 2003 - 2016, Bioconductor
Bioconductor
Home
Install
Install R
Find Bioconductor Packages
Help
Workflows
Package Vignettes
Developers
Package Guidelines
Package Submission
About
Advisory Board
Annual Reports


ArrayAnalysis.org
Bioconductor - Install
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.bioconductor.org/install/
Search
Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS
Home
Install
Help
Developers
About
Search:
Home » Install
Using Bioconductor
Install R
Install Packages
Find Packages
Update Packages
Troubleshoot Package Installations
Why biocLite()?
Pre-configured Bioconductor
Using Bioconductor
The current release of Bioconductor is version 3.4; it works with R version 3.3.1. Users of older R and Bioconductor users must update their installation to take advantage of new features and to access packages that have been added to Bioconductor since the last release.
Install the latest release of R, then get the latest version of Bioconductor by starting R and entering the commands
try http:// if https:// URLs are not supported
source("https://bioconductor.org/biocLite.R")
biocLite()
Details, including instructions to install additional packages and to update, find, and troubleshoot are provided below. A devel version of Bioconductor is available. There are good reasons for using biocLite() for managing Bioconductor resources.
Install R
1. Download the most recent version of R. The R FAQs and the R Installation and Administration Manual contain detailed instructions for installing R on various platforms (Linux, OS X, and Windows being the main ones).
1. Start the R program; on Windows and OS X, this will usually mean double-clicking on the R application, on UNIX-like systems, type "R" at a shell prompt.
2. As a first step with R, start the R help browser by typing help.start() in the R command window. For help on any function, e.g. the "mean" function, type ? mean.
[Back to top]
Install Bioconductor Packages
Use the biocLite.R script to install Bioconductor packages. To install core packages, type the following in an R command window:
Packages »
Bioconductor's stable, semi-annual release:
Analysis software packages.
Annotation packages.
Illustrative experiment data packages.
Latest release announcement.
Bioconductor is also available via Docker and Amazon Machine Images.
Documentation »
Bioconductor
Package vignettes and manuals.
Workflows for learning and use.
Course and conference material.
Videos.
Community resources and tutorials.
R / CRAN packages and documentation
Previous Releases »
Bioconductor (R) versions:
3.4 (3.3) • 3.3 (3.3) • 3.2 (3.2)
3.1 (3.2) • 3.0 (3.1) • 2.14 (3.1)
2.13 (3.0) • 2.12 (3.0)
2.11 (2.15) • 2.10 (2.15)
2.9 (2.14) • 2.8 (2.13)
2.7 (2.12) • 2.6 (2.11)
2.5 (2.10) • 2.4 (2.9) • 2.3 (2.8)
2.2 (2.7) • 2.1 (2.6) • 2.0 (2.5)
1.9 (2.4) • 1.8 (2.3) • 1.7 (2.2)


ArrayAnalysis.org
Bioconductor - Install
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
www.bioconductor.org/install/
Search
R itself, in which case the attempt to useDevel() fails with an appropriate message).
The BiocInstaller package also provides biocValid() to test that the installed packages are not a hodgepodge from different Bioconductor releases (the 'too new' packages have been installed from source rather than a repository; regular users would seldom have these).
> biocValid()
* sessionInfo()
R version 3.1.0 Patched (2014-05-06 r65533)
Platform: x86_64-unknown-linux-gnu (64-bit)
...
* Out-of-date packages
...
update with biocLite()
* Packages too new for Bioconductor version '3.0'
...
downgrade with biocLite(c("ShortRead", "BatchJobs"))
Error: 9 package(s) out of date; 2 package(s) too new
For users who spend a lot of time in Bioconductor, the features outlined above become increasingly important and biocLite() is much preferred to install.packages().
[Back to top]
Pre-configured Bioconductor
Bioconductor is also available as a set of Amazon Machine Images (AMIs) and Docker images.
[Back to top]
Contact us: support.bioconductor.org
Copyright © 2003 - 2016, Bioconductor
Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS
Home
Install
Install R
Find Bioconductor Packages
Install Bioconductor Packages
Update Bioconductor Packages
Help
Workflows
Package Vignettes
FAQ
Support
Using R
Courses
Publications
Cloud AMI
Community Resources
Developers
Package Guidelines
Package Submission
Release Schedule
Source Control
About
Advisory Board
Annual Reports
Core Team
Mirrors
Related Projects


https://www.r-project.org
Search
The R Project for Statistical Computing
Getting Started
R is a free software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. To download R, please choose your preferred CRAN mirror.
If you have questions about R like how to download and install the software, or what the license terms are, please read our answers to frequently asked questions before you send an email.
News
R version 3.5.2 (Eggshell Igloo) prerelease versions will appear starting Monday 2018-12-10. Final release is scheduled for Thursday 2018-12-20.
The R Foundation Conference Committee has released a call for proposals to host useR! 2020 in North America.
You can now support the R Foundation with a renewable subscription as a supporting member
R version 3.5.1 (Feather Spray) has been released on 2018-07-02.
The R Foundation has been awarded the Personality/Organization of the year 2018 award by the professional association of German market and social researchers.
News via Twitter
The R Foundation
@_R_Foundation
We welcome @gdequeiroz, @edzerpebesma and @henrikbengtsson, elected as ordinary members of the R Foundation in recognition of their services to the R community.
Oct 26, 2018
The R Foundation Retweeted
useR! 2019
@UseR2019_Conf
17-12-18 ;
01-19 ;
15-02-19 ;
[Home]
Download
CRAN
R Project
About R
Logo
Contributors
What's New?
Reporting Bugs
Conferences
Search
Get Involved: Mailing Lists
Developer Pages
R Blog
R Foundation
Foundation
Board
Members
Donors
Donate
Help With R
Getting Help
Documentation
Manuals
FAQs
The R Journal
Books
Certification
Other
Links


https://www.r-project.org/about.html
Search
[Home]
Download
CRAN
R Project
About R
Logo
Contributors
What's New?
Reporting Bugs
Conferences
Search
Get Involved: Mailing Lists
Developer Pages
R Blog
R Foundation
Foundation
Board
Members
Donors
Donate
Help With R
Getting Help
Documentation
Manuals
FAQs
The R Journal
Books
Certification
Other
Links
What is R?
Introduction to R
R is a language and environment for statistical computing and graphics. It is a GNU project which is similar to the S language and environment which was developed at Bell Laboratories (formerly AT&T, now Lucent Technologies) by John Chambers and colleagues. R can be considered as a different implementation of S. There are some important differences, but much code written for S runs unaltered under R.
R provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering, ...) and graphical techniques, and is highly extensible. The S language is often the vehicle of choice for research in statistical methodology, and R provides an Open Source route to participation in that activity.
One of R's strengths is the ease with which well-designed publication-quality plots can be produced, including mathematical symbols and formulae where needed. Great care has been taken over the defaults for the minor design choices in graphics, but the user retains full control.
R is available as Free Software under the terms of the Free Software Foundation's GNU General Public License in source code form. It compiles and runs on a wide variety of UNIX platforms and similar systems (including FreeBSD and Linux), Windows and MacOS.
The R environment
R is an integrated suite of software facilities for data manipulation, calculation and graphical display. It includes
an effective data handling and storage facility,
a suite of operators for calculations on arrays, in particular matrices,
a large, coherent, integrated collection of intermediate tools for data analysis,
graphical facilities for data analysis and display either on-screen or on hardcopy, and
a well-developed, simple and effective programming language which includes conditionals, loops, user-defined recursive functions and input and output facilities.
The term "environment" is intended to characterize it as a fully planned and coherent system, rather than an incremental accretion of very specific and inflexible tools, as is frequently the case with other data analysis software.
R, like S, is designed around a true computer language, and it allows users to add additional functionality by defining new functions. Much of the system is itself written in the R dialect of S, which makes it easy for users to follow the algorithmic choices made. For computationally-intensive tasks, C,



www.bioconductor.org/packages/release/workflows/

Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS

Home | Install | Help | Developers | About

Search:

Home » Bioconductor 3.8 » 3.8 Workflow Packages

Bioconductor Workflow Packages

Bioconductor version: Release (3.8)

Package	Maintainer	Title
annotation	Bioconductor Package Maintainer	Genomic Annotation Resources
arrays	Bioconductor Package Maintainer	Using Bioconductor for Microarray Analysis
BiocMetaWorkflow	Mike Smith	BioC Workflow about publishing a Bioc Workflow
chipseqDB	Aaron Lun	From reads to regions: a Bioconductor workflow to detect differential binding in ChIP-seq data
cytofWorkflow	Malgorzata Nowicka	CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets
EGSEA123	Matthew Ritchie	Easy and efficient ensemble gene set testing with EGSEA
eQTL	Vincent Carey	Cloud-enabled cis-eQTL searches with Bioconductor GGtools 5.x
ExpressionNormalizationWorkflow	Karthikeyan Murugesan	Gene Expression Normalization Workflow
generegulation	Bioconductor Package Maintainer	Finding Candidate Binding Sites for Known Transcription Factors via Sequence Matching
highthroughputassays	Bioconductor Package Maintainer	Using Bioconductor with High Throughput Assays
liftOver	Bioconductor Package Maintainer	Changing genomic coordinate systems with rtracklayer::liftOver
maEndToEnd	Bernd Klaus	An end to end workflow for differential gene expression using Affymetrix microarrays
methylationArrayAnalysis	Jovana Maksimovic	A cross-package Bioconductor workflow for analysing methylation array data.
proteomics	Laurent Gatto	Mass spectrometry and proteomics data analysis
RNAseq123	Matthew Ritchie	RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR
rnaseqDTU	Michael Love	RNA-seq workflow for differential transcript usage following Salmon quantification

Packages »

Bioconductor's stable, semi-annual release:

- Analysis software packages.
- Annotation packages.
- Illustrative experiment data packages.
- Workflow packages.
- Latest release announcement.

Bioconductor is also available via Docker and Amazon Machine Images.

Development Version »

Bioconductor packages under development:

- Analysis software packages.
- Annotation packages
- Illustrative experiment data packages

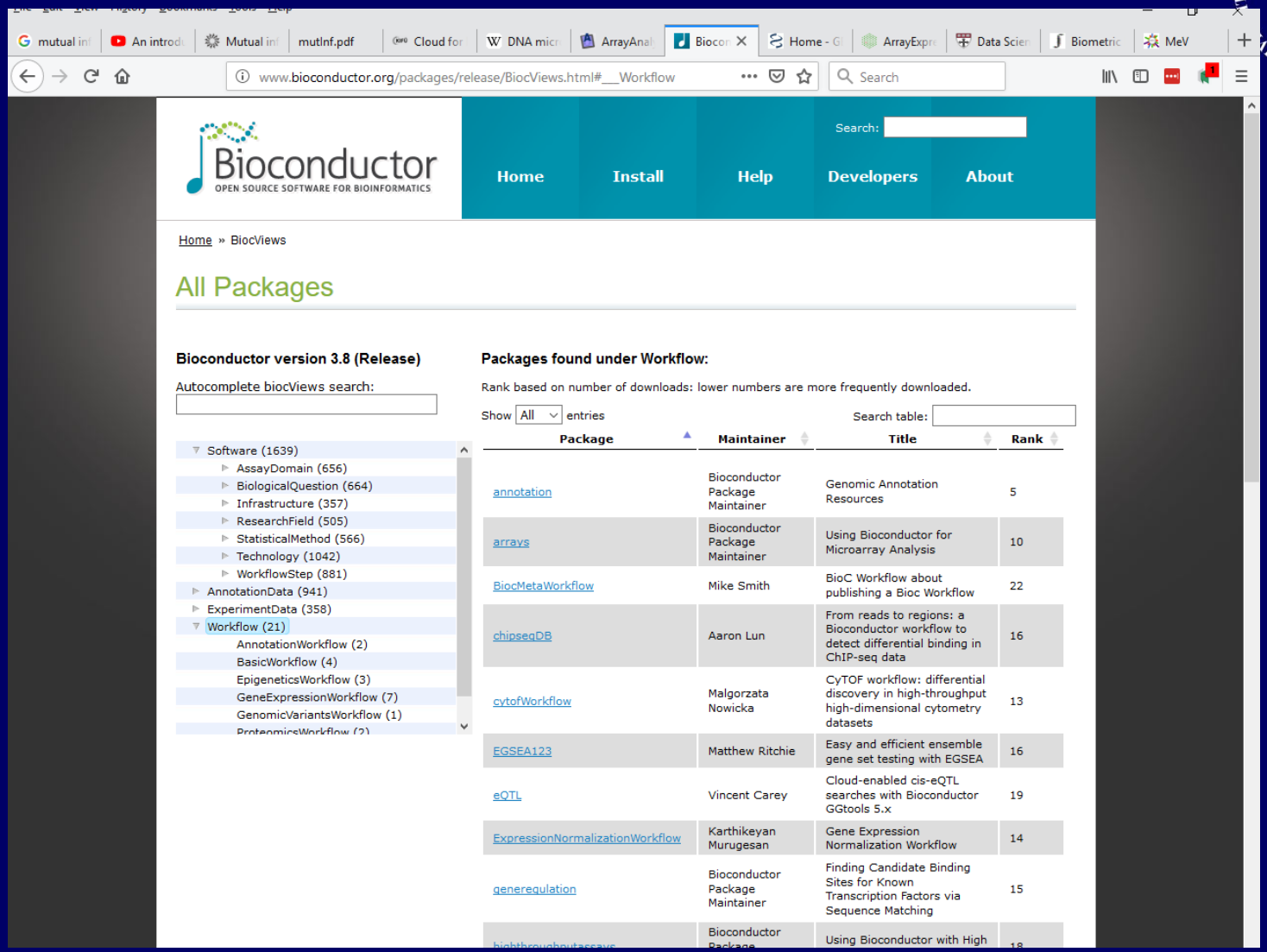
Developer Resources:

- GIT Log

Previous Releases »

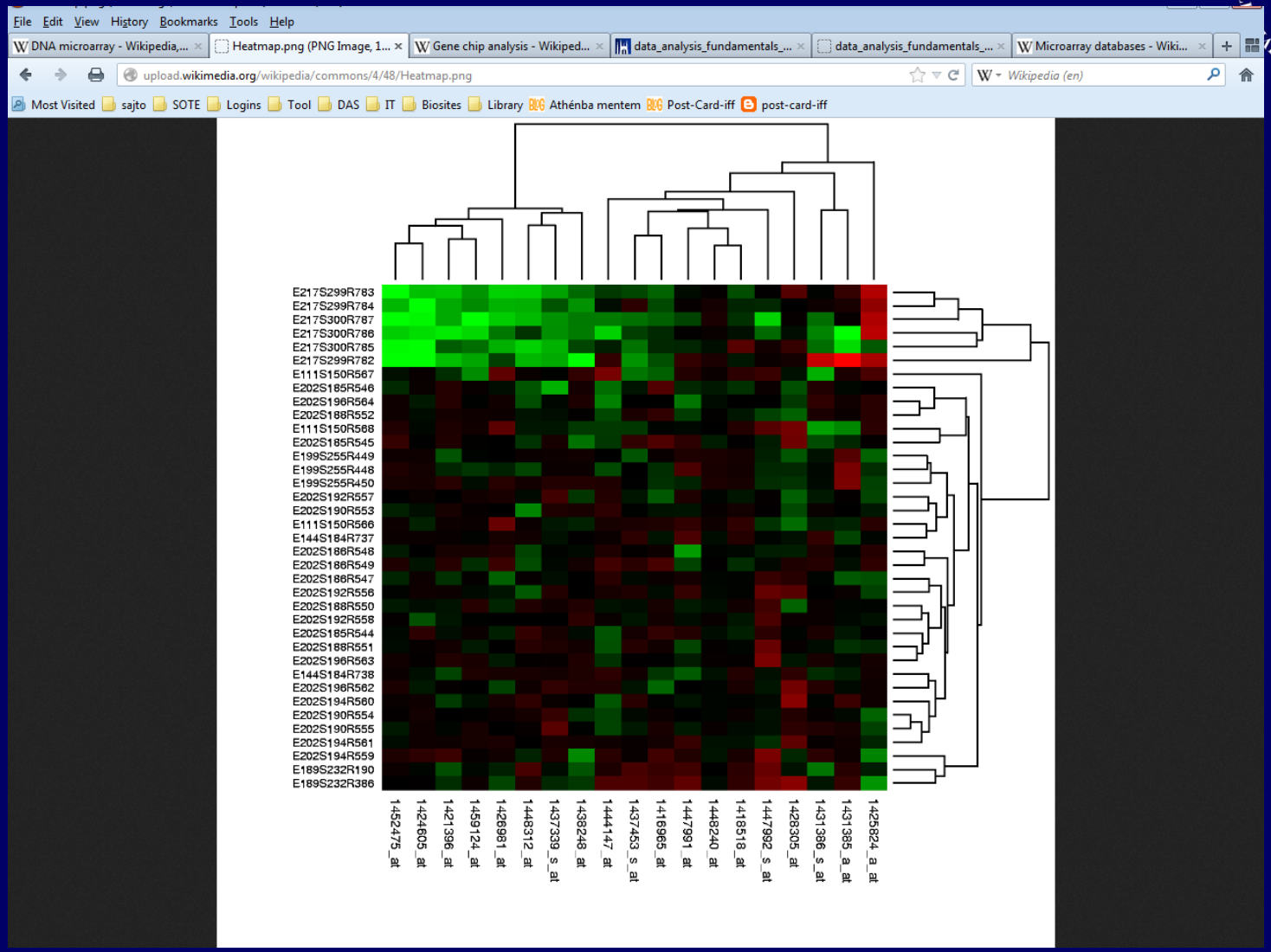
Bioconductor (*R*) versions:

- 3.7 (3.5) • 3.6 (3.4) • 3.5 (3.4)
- 3.4 (3.3) • 3.3 (3.3) • 3.2 (3.2)
- 3.1 (3.2) • 3.0 (3.1) • 2.14 (3.1)
- 2.13 (3.0) • 2.12 (3.0)
- 2.11 (2.15) • 2.10 (2.15)
- 2.9 (2.14) • 2.8 (2.13)
- 2.7 (2.12) • 2.6 (2.11)
- 2.5 (2.10) • 2.4 (2.9) • 2.3 (2.8)


www.bioconductor.org/packages/release/BiocViews.html#___Workflow
Search
Bioconductor
OPEN SOURCE SOFTWARE FOR BIOINFORMATICS
Home
Install
Help
Developers
About
Search:
Home » BiocViews
All Packages
Bioconductor version 3.8 (Release)
Autocomplete biocViews search:
Software (1639)
AssayDomain (656)
BiologicalQuestion (664)
Infrastructure (357)
ResearchField (505)
StatisticalMethod (566)
Technology (1042)
WorkflowStep (881)
AnnotationData (941)
ExperimentData (358)
Workflow (21)
AnnotationWorkflow (2)
BasicWorkflow (4)
EpigeneticsWorkflow (3)
GeneExpressionWorkflow (7)
GenomicVariantsWorkflow (1)
Packages found under Workflow:
Rank based on number of downloads: lower numbers are more frequently downloaded.
Show All entries
Search table:
Package
Maintainer
Title
Rank
annotation
Bioconductor Package Maintainer
Genomic Annotation Resources
5
arrays
Bioconductor Package Maintainer
Using Bioconductor for Microarray Analysis
10
BiocMetaWorkflow
Mike Smith
BioC Workflow about publishing a Bioc Workflow
22
chipseqDB
Aaron Lun
From reads to regions: a Bioconductor workflow to detect differential binding in ChIP-seq data
16
cytofWorkflow
Malgorzata Nowicka
CyTOF workflow: differential discovery in high-throughput high-dimensional cytometry datasets
13
EGSEA123
Matthew Ritchie
Easy and efficient ensemble gene set testing with EGSEA
16
eQTL
Vincent Carey
Cloud-enabled cis-eQTL searches with Bioconductor GGtools 5.x
19
ExpressionNormalizationWorkflow
Karthikeyan Murugesan
Gene Expression Normalization Workflow
14
generegulation
Bioconductor Package Maintainer
Finding Candidate Binding Sites for Known Transcription Factors via Sequence Matching
15

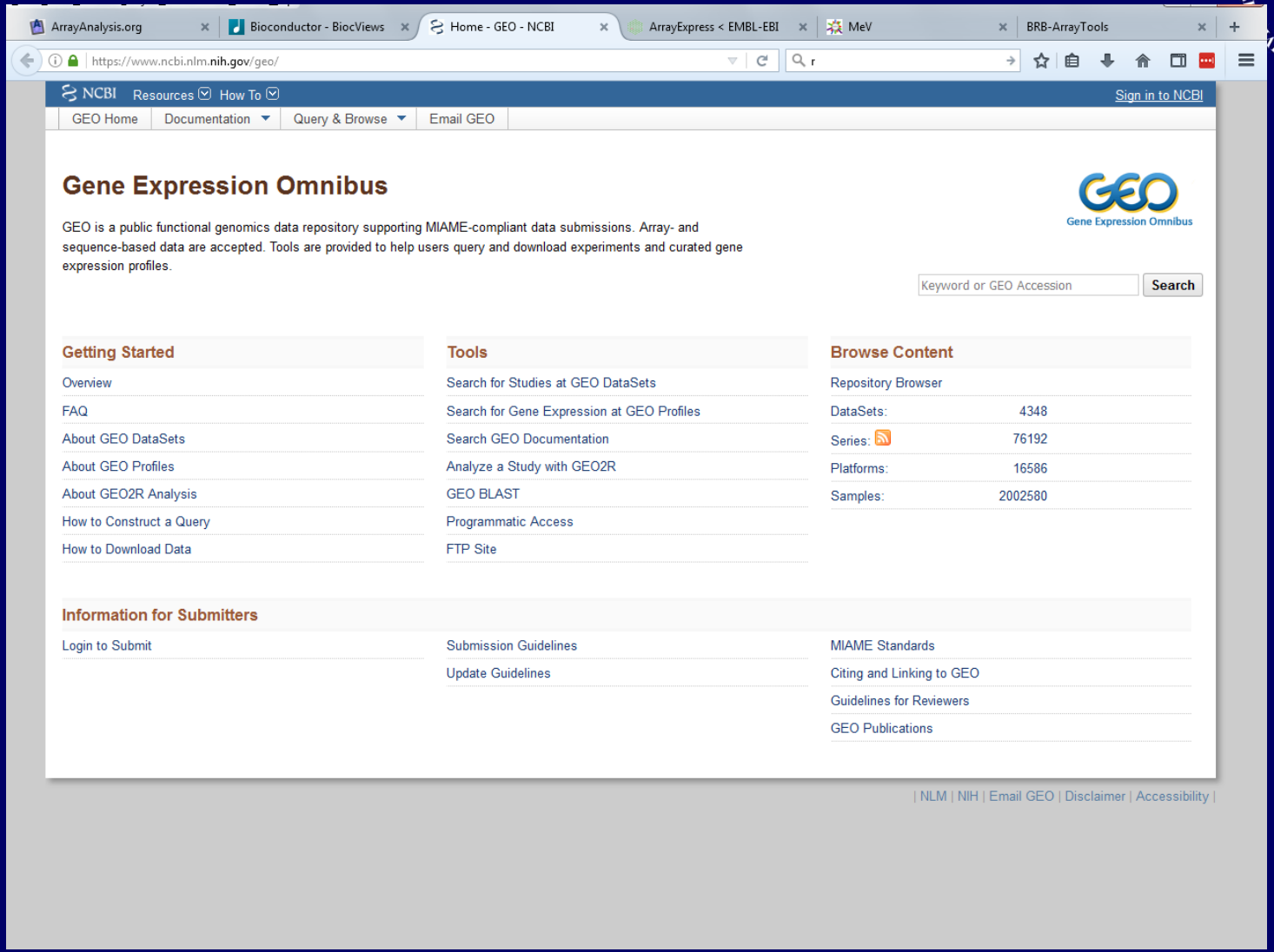
További lépések

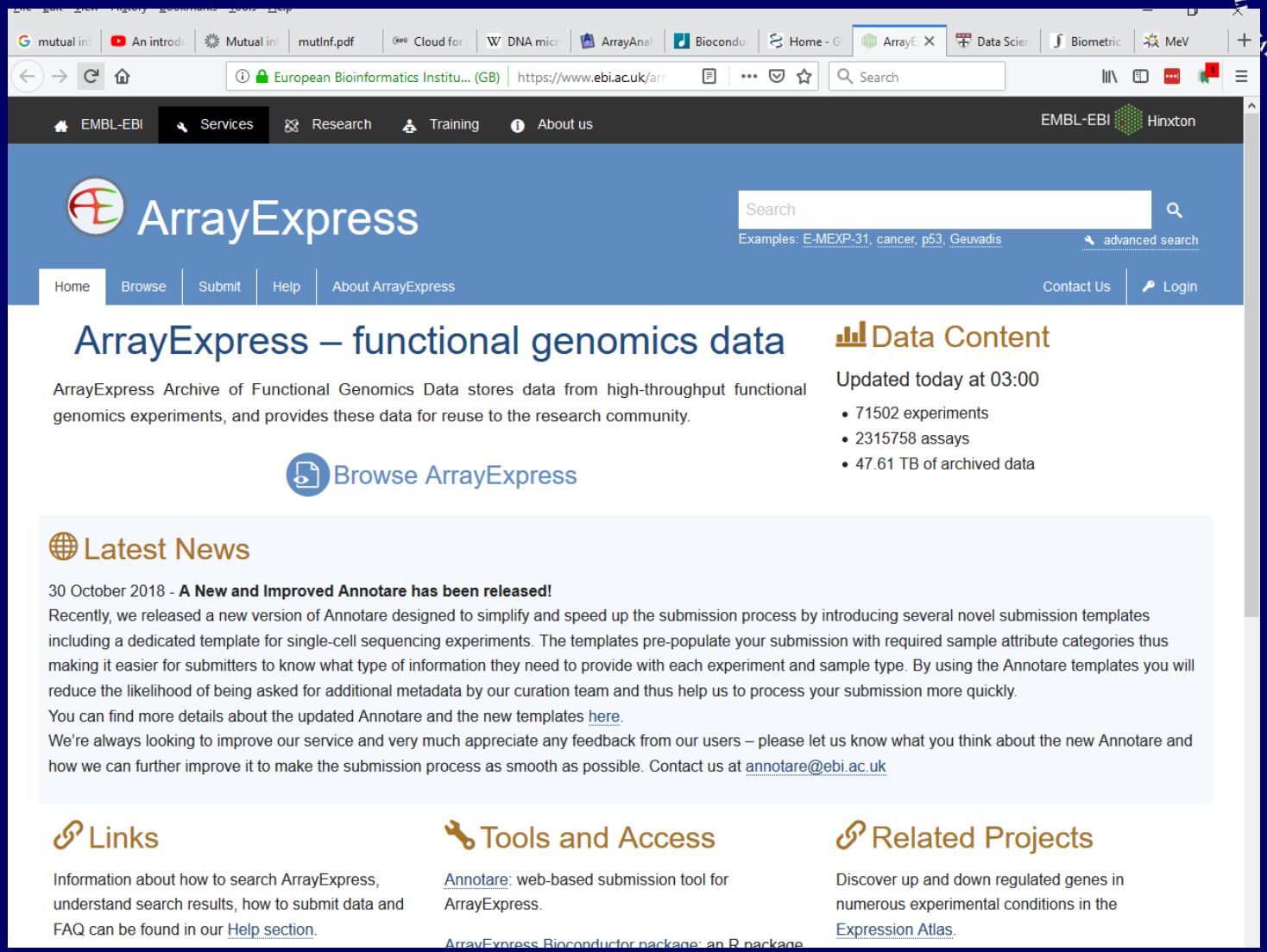
- Az adatok elemzése
  - Az adatok az expressziós szint változásokat adják meg
  - Ezek klaszterekbe és gén-listákba rendezhetők
  - Ezek értelmezése rendszerbiológiai feladat
- Szabványosítás
  - A gén-chip-ek adatai nem hasonlíthatók össze közvetlenül – annotálás (GO)
  - A kísérletek szintén nem összevethetők – MIAME

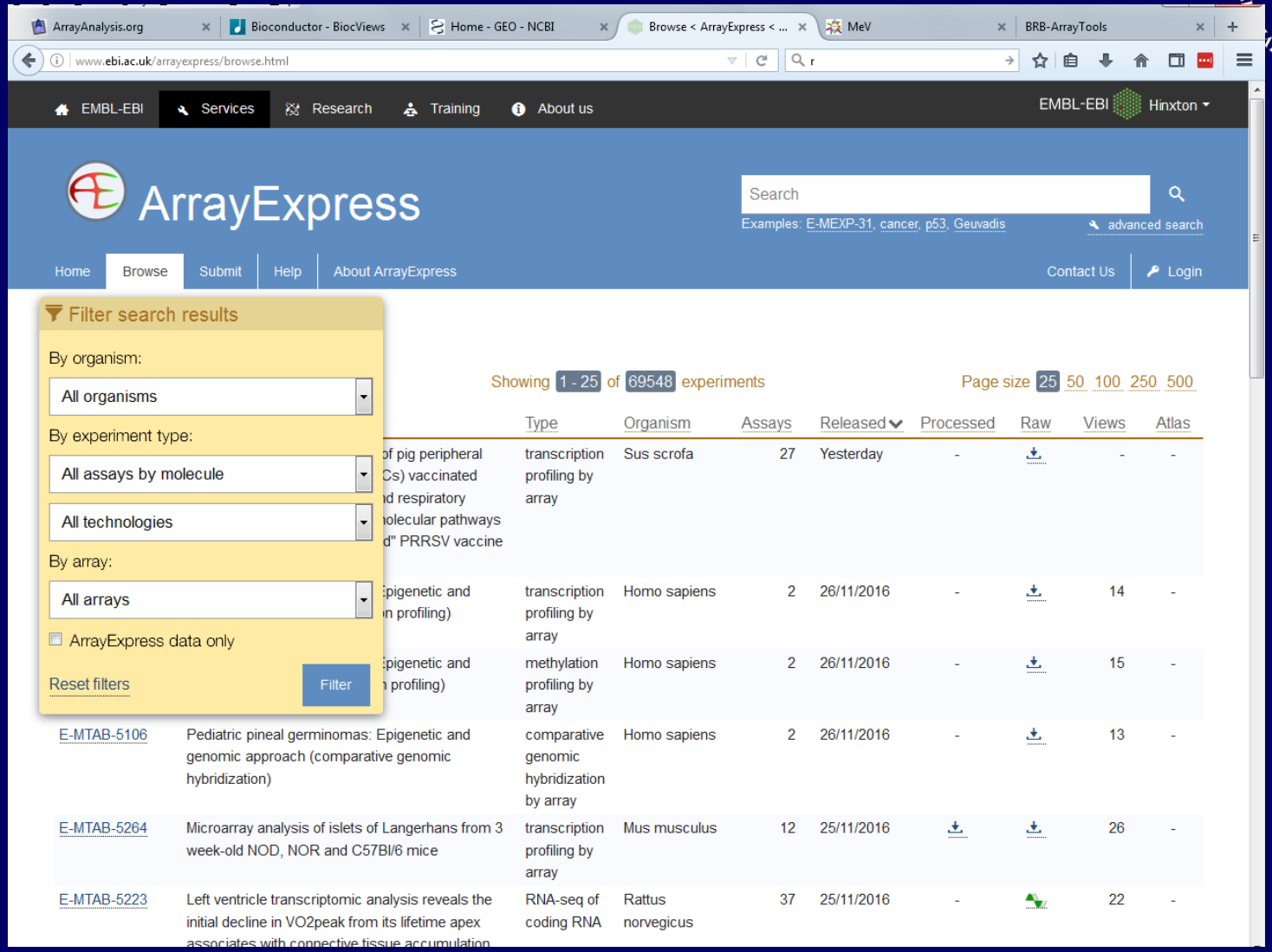

File Edit View History Bookmarks Tools Help
DNA microarray - Wikipedia,...
Heatmap.png (PNG Image, 1...
Gene chip analysis - Wikiped...
data_analysis_fundamentals_...
data_analysis_fundamentals_...
Microarray databases - Wiki...
upload.wikimedia.org/wikipedia/commons/4/48/Heatmap.png
Wikipedia (en)
Most Visited sajto SOTE Logins Tool DAS IT Biosites Library Athénba mentem Post-Card-iff post-card-iff
E217S299R783
E217S299R784
E217S300R787
E217S300R786
E217S300R785
E217S299R782
E111S150R567
E202S185R546
E202S196R564
E202S188R552
E111S150R568
E202S185R545
E199S255R449
E199S255R448
E199S255R450
E202S192R557
E202S190R553
E111S150R566
E144S184R737
E202S188R548
E202S188R549
E202S188R547
E202S192R556
E202S188R550
E202S192R558
E202S185R544
E202S188R551
E202S196R563
E144S184R738
E202S196R562
E202S194R560
E202S190R554
E202S190R555
E202S194R561
E202S194R559
E189S232R190
E189S232R386
1-452475_at
1-424605_at
1-421396_at
1-459124_at
1-426981_at
1-448312_at
1-437339_s_at
1-438248_at
1-444147_at
1-437453_s_at
1-416965_at
1-447991_at
1-448240_at
1-418518_at
1-447992_s_at
1-428305_at
1-431386_s_at
1-431385_a_at
1-425824_a_at

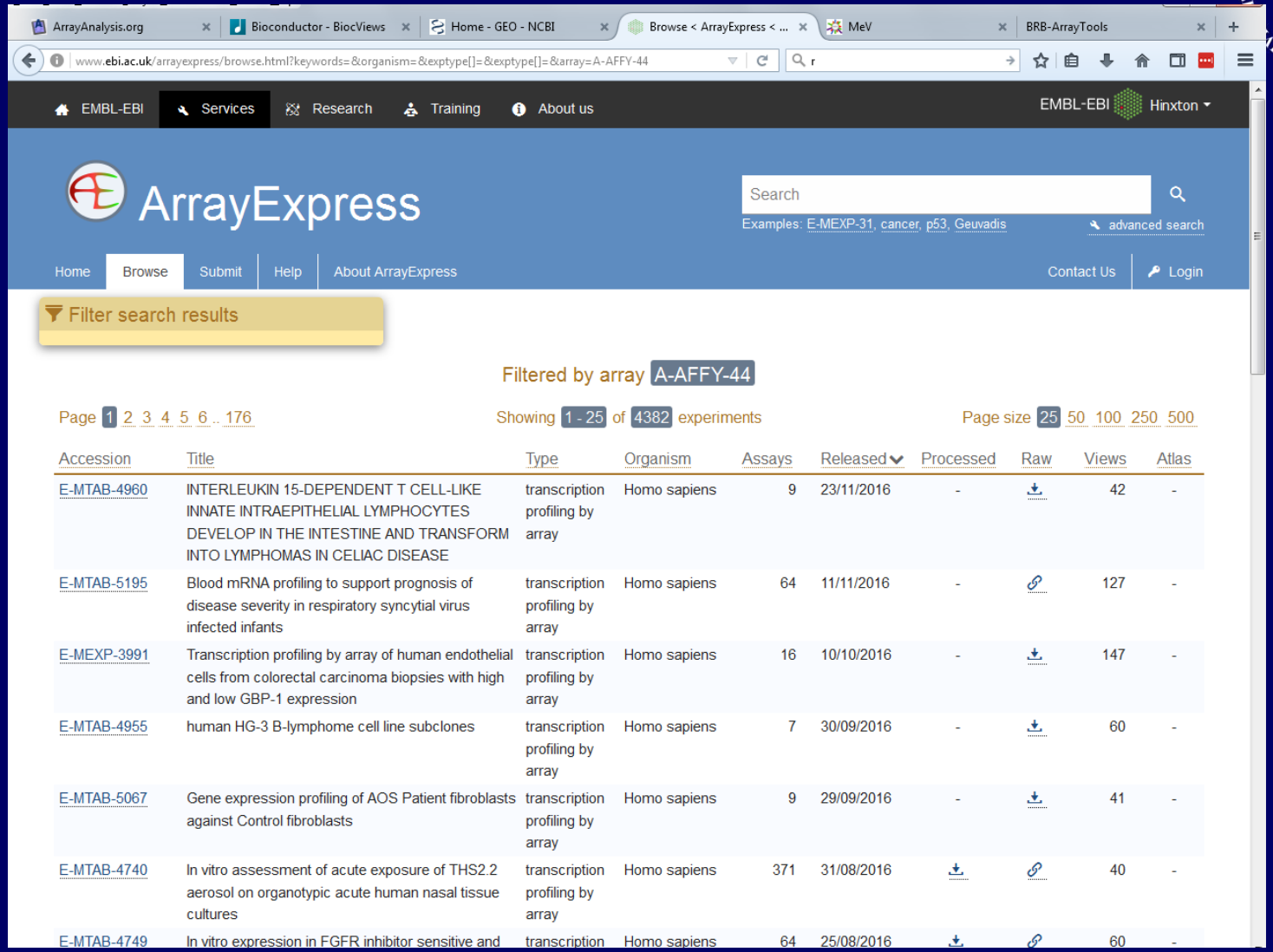
Kapcsolódó adatbázisok

- A gén-chip adatok nyers adatait adatbázisba gyüjtik
- ArrayExpress, GEO: nagyrészt átfedő tartalom, de két független kezdeményezés
- Web:
  - http://www.ncbi.nlm.nih.gov/geo/
  - http://www.ebi.ac.uk/arrayexpress/


ArrayAnalysis.org
Bioconductor - BiocViews
Home - GEO - NCBI
ArrayExpress < EMBL-EBI
MeV
BRB-ArrayTools
https://www.ncbi.nlm.nih.gov/geo/
NCBI Resources How To
Sign in to NCBI
GEO Home
Documentation
Query & Browse
Email GEO
Gene Expression Omnibus
GEO is a public functional genomics data repository supporting MIAME-compliant data submissions. Array- and sequence-based data are accepted. Tools are provided to help users query and download experiments and curated gene expression profiles.
Keyword or GEO Accession
Search
Getting Started
Overview
FAQ
About GEO DataSets
About GEO Profiles
About GEO2R Analysis
How to Construct a Query
How to Download Data
Tools
Search for Studies at GEO DataSets
Search for Gene Expression at GEO Profiles
Search GEO Documentation
Analyze a Study with GEO2R
GEO BLAST
Programmatic Access
FTP Site
Browse Content
Repository Browser
DataSets: 4348
Series: 76192
Platforms: 16586
Samples: 2002580
Information for Submitters
Login to Submit
Submission Guidelines
Update Guidelines
MIAME Standards
Citing and Linking to GEO
Guidelines for Reviewers
GEO Publications
| NLM | NIH | Email GEO | Disclaimer | Accessibility |


EMBL-EBI
Services
Research
Training
About us
EMBL-EBI Hinxton
ArrayExpress
Search
Examples: E-MEXP-31, cancer, p53, Geuvadis
advanced search
Home
Browse
Submit
Help
About ArrayExpress
Contact Us
Login
ArrayExpress – functional genomics data
ArrayExpress Archive of Functional Genomics Data stores data from high-throughput functional genomics experiments, and provides these data for reuse to the research community.
Browse ArrayExpress
Data Content
Updated today at 03:00
71502 experiments
2315758 assays
47.61 TB of archived data
Latest News
30 October 2018 - A New and Improved Annotare has been released!
Recently, we released a new version of Annotare designed to simplify and speed up the submission process by introducing several novel submission templates including a dedicated template for single-cell sequencing experiments. The templates pre-populate your submission with required sample attribute categories thus making it easier for submitters to know what type of information they need to provide with each experiment and sample type. By using the Annotare templates you will reduce the likelihood of being asked for additional metadata by our curation team and thus help us to process your submission more quickly.
You can find more details about the updated Annotare and the new templates here.
We're always looking to improve our service and very much appreciate any feedback from our users – please let us know what you think about the new Annotare and how we can further improve it to make the submission process as smooth as possible. Contact us at annotare@ebi.ac.uk
Links
Information about how to search ArrayExpress, understand search results, how to submit data and FAQ can be found in our Help section.
Tools and Access
Annotare: web-based submission tool for ArrayExpress.
Related Projects
Discover up and down regulated genes in numerous experimental conditions in the Expression Atlas.


ArrayAnalysis.org
Bioconductor - BiocViews
Home - GEO - NCBI
Browse < ArrayExpress < ...
MeV
BRB-ArrayTools
www.ebi.ac.uk/arrayexpress/browse.html
EMBL-EBI
Services
Research
Training
About us
EMBL-EBI
Hinxton
ArrayExpress
Search
Examples: E-MEXP-31, cancer, p53, Geuvadis
advanced search
Home
Browse
Submit
Help
About ArrayExpress
Contact Us
Login
Filter search results
By organism:
All organisms
By experiment type:
All assays by molecule
All technologies
By array:
All arrays
ArrayExpress data only
Reset filters
Filter
Showing 1 - 25 of 69548 experiments
Page size 25 50 100 250 500
Type
Organism
Assays
Released
Processed
Raw
Views
Atlas
transcription profiling by array
Sus scrofa
27
Yesterday
transcription profiling by array
Homo sapiens
2
26/11/2016
14
methylation profiling by array
Homo sapiens
2
26/11/2016
15
E-MTAB-5106
Pediatric pineal germinomas: Epigenetic and genomic approach (comparative genomic hybridization)
comparative genomic hybridization by array
Homo sapiens
2
26/11/2016
13
E-MTAB-5264
Microarray analysis of islets of Langerhans from 3 week-old NOD, NOR and C57Bl/6 mice
transcription profiling by array
Mus musculus
12
25/11/2016
26
E-MTAB-5223
Left ventricle transcriptomic analysis reveals the initial decline in VO2peak from its lifetime apex associates with connective tissue accumulation
RNA-seq of coding RNA
Rattus norvegicus
37
25/11/2016
22



ArrayAnalysis.org | Bioconductor - BiocViews | Home - GEO - NCBI | Browse < ArrayExpress < ... | MeV | BRB-ArrayTools

www.ebi.ac.uk/arrayexpress/browse.html?keywords=&organism=&exptype[]=&exptype[]=&array=A-AFFY-44

EMBL-EBI | Services | Research | Training | About us | EMBL-EBI Hinxton

ArrayExpress

Search

Examples: E-MEXP-31, cancer, p53, Geuvadis | advanced search

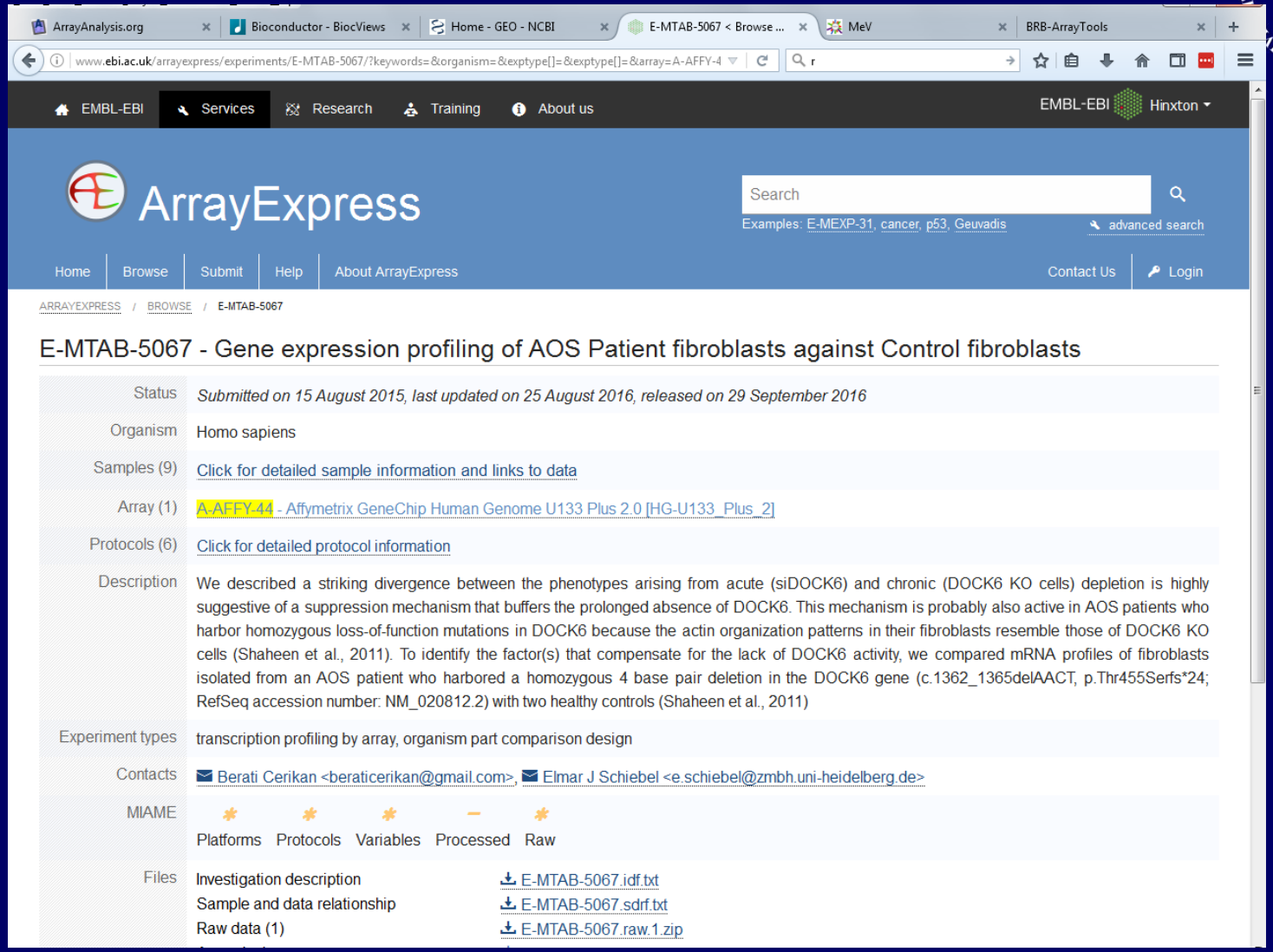
Home | Browse | Submit | Help | About ArrayExpress | Contact Us | Login

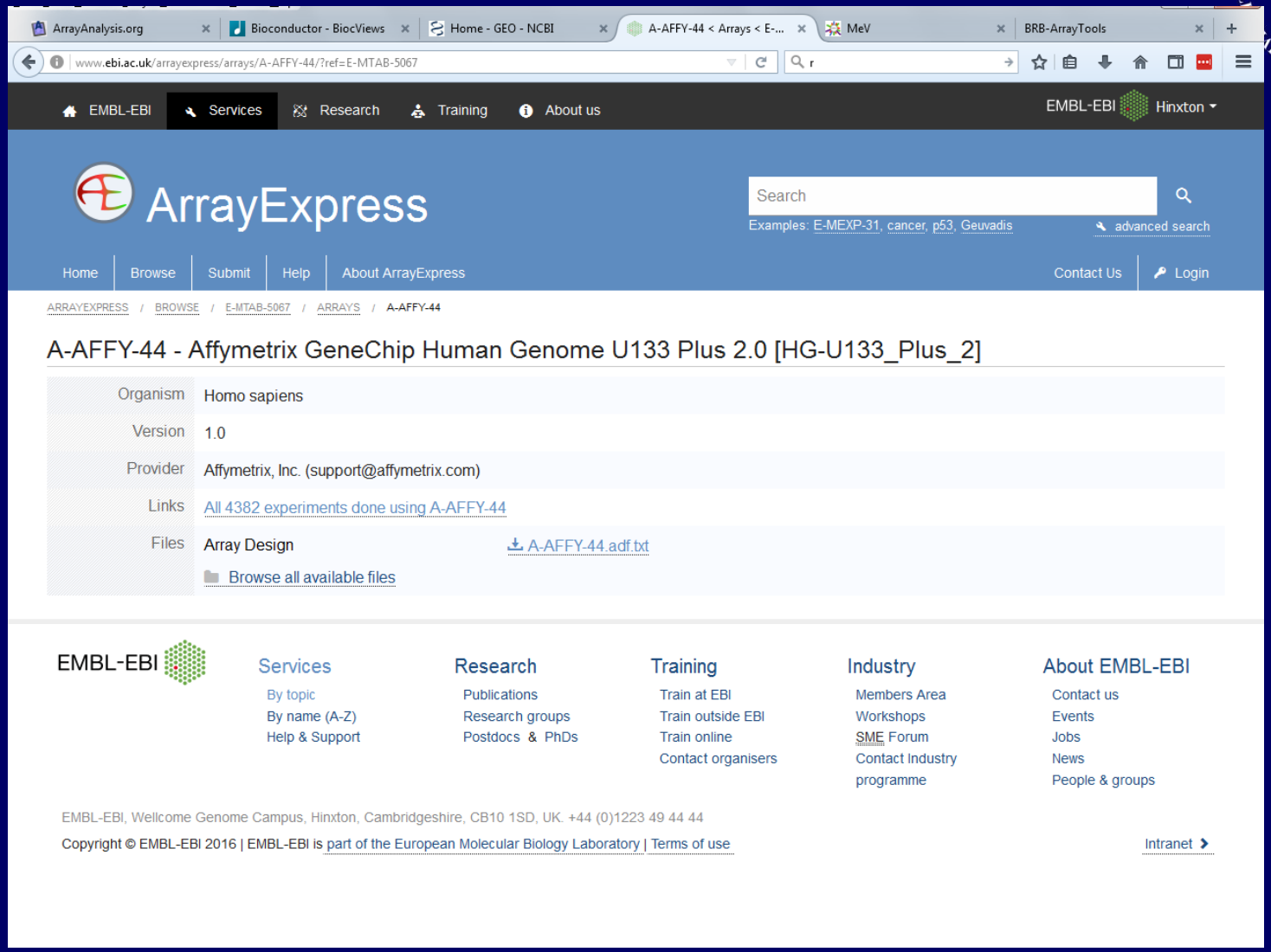
Filter search results

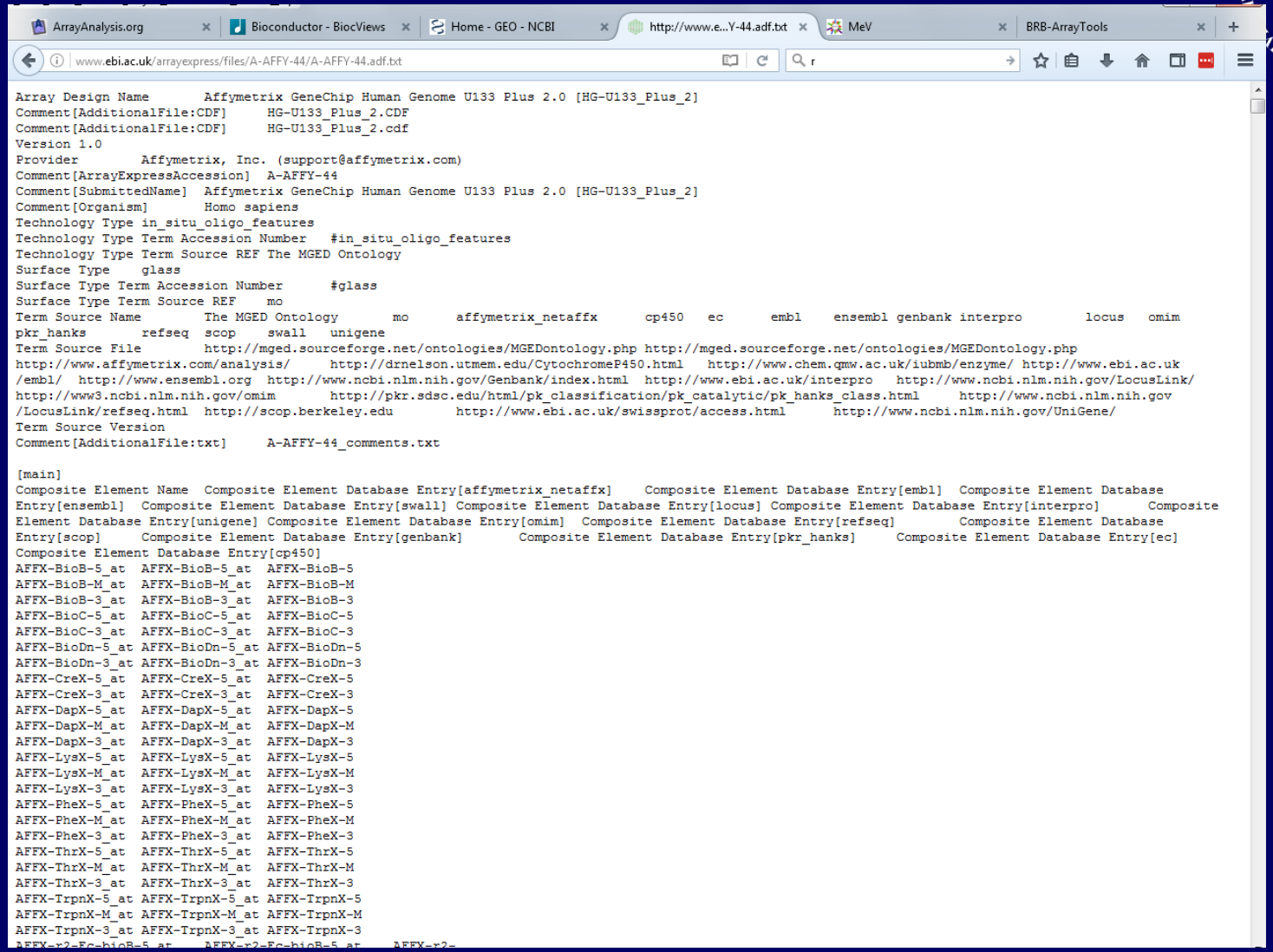
Filtered by array A-AFFY-44

Page 1 2 3 4 5 6 .. 176 | Showing 1 - 25 of 4382 experiments | Page size 25 50 100 250 500

Accession	Title	Type	Organism	Assays	Released	Processed	Raw	Views	Atlas
E-MTAB-4960	INTERLEUKIN 15-DEPENDENT T CELL-LIKE INNATE INTRAEPITHELIAL LYMPHOCYTES DEVELOP IN THE INTESTINE AND TRANSFORM INTO LYMPHOMAS IN CELIAC DISEASE	transcription profiling by array	Homo sapiens	9	23/11/2016	-		42	-
E-MTAB-5195	Blood mRNA profiling to support prognosis of disease severity in respiratory syncytial virus infected infants	transcription profiling by array	Homo sapiens	64	11/11/2016	-		127	-
E-MEXP-3991	Transcription profiling by array of human endothelial cells from colorectal carcinoma biopsies with high and low GBP-1 expression	transcription profiling by array	Homo sapiens	16	10/10/2016	-		147	-
E-MTAB-4955	human HG-3 B-lymphome cell line subclones	transcription profiling by array	Homo sapiens	7	30/09/2016	-		60	-
E-MTAB-5067	Gene expression profiling of AOS Patient fibroblasts against Control fibroblasts	transcription profiling by array	Homo sapiens	9	29/09/2016	-		41	-
E-MTAB-4740	In vitro assessment of acute exposure of THS2.2 aerosol on organotypic acute human nasal tissue cultures	transcription profiling by array	Homo sapiens	371	31/08/2016			40	-
E-MTAB-4749	In vitro expression in FGFR inhibitor sensitive and	transcription	Homo sapiens	64	25/08/2016			60	-

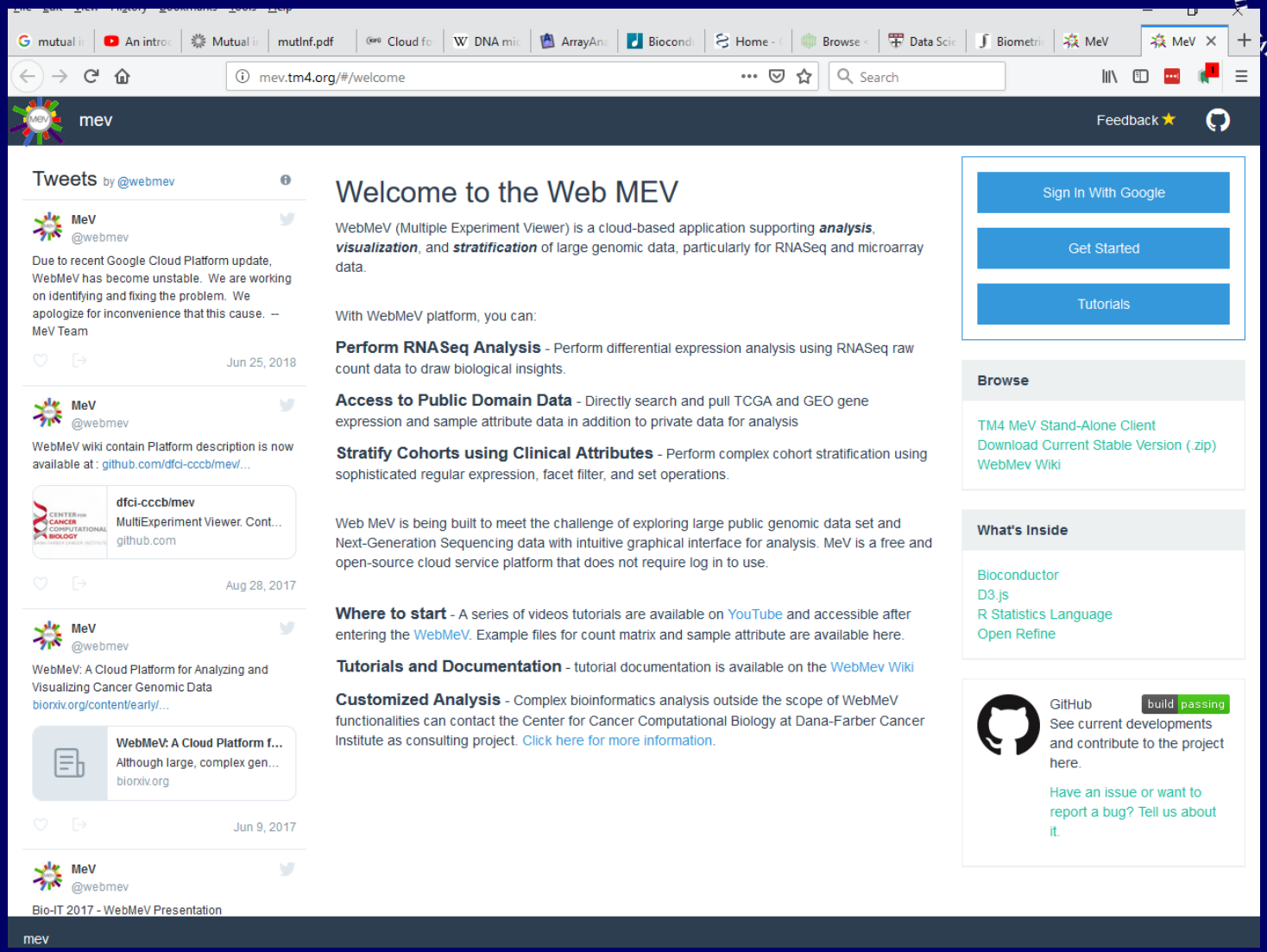

ArrayAnalysis.org
Bioconductor - BiocViews
Home - GEO - NCBI
E-MTAB-5067 < Browse ...
MeV
BRB-ArrayTools
www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-5067/?keywords=&organism=&exptype[]=&exptype[]=&array=A-AFFY-4
EMBL-EBI
Services
Research
Training
About us
EMBL-EBI
Hinxton
ArrayExpress
Search
Examples: E-MEXP-31, cancer, p53, Geuvadis
advanced search
Home
Browse
Submit
Help
About ArrayExpress
Contact Us
Login
ARRAYEXPRESS / BROWSE / E-MTAB-5067
E-MTAB-5067 - Gene expression profiling of AOS Patient fibroblasts against Control fibroblasts
Status
Submitted on 15 August 2015, last updated on 25 August 2016, released on 29 September 2016
Organism
Homo sapiens
Samples (9)
Click for detailed sample information and links to data
Array (1)
A-AFFY-44 - Affymetrix GeneChip Human Genome U133 Plus 2.0 [HG-U133_Plus_2]
Protocols (6)
Click for detailed protocol information
Description
We described a striking divergence between the phenotypes arising from acute (siDOCK6) and chronic (DOCK6 KO cells) depletion is highly suggestive of a suppression mechanism that buffers the prolonged absence of DOCK6. This mechanism is probably also active in AOS patients who harbor homozygous loss-of-function mutations in DOCK6 because the actin organization patterns in their fibroblasts resemble those of DOCK6 KO cells (Shaheen et al., 2011). To identify the factor(s) that compensate for the lack of DOCK6 activity, we compared mRNA profiles of fibroblasts isolated from an AOS patient who harbored a homozygous 4 base pair deletion in the DOCK6 gene (c.1362_1365delAACT, p.Thr455Serfs*24; RefSeq accession number: NM_020812.2) with two healthy controls (Shaheen et al., 2011)
Experiment types
transcription profiling by array, organism part comparison design
Contacts
Berati Cerikan <beraticerikan@gmail.com>, Elmar J Schiebel <e.schiebel@zmbh.uni-heidelberg.de>
MIAME
Platforms
Protocols
Variables
Processed
Raw
Files
Investigation description
E-MTAB-5067.idf.txt
Sample and data relationship
E-MTAB-5067.sdrf.txt
Raw data (1)
E-MTAB-5067.raw.1.zip

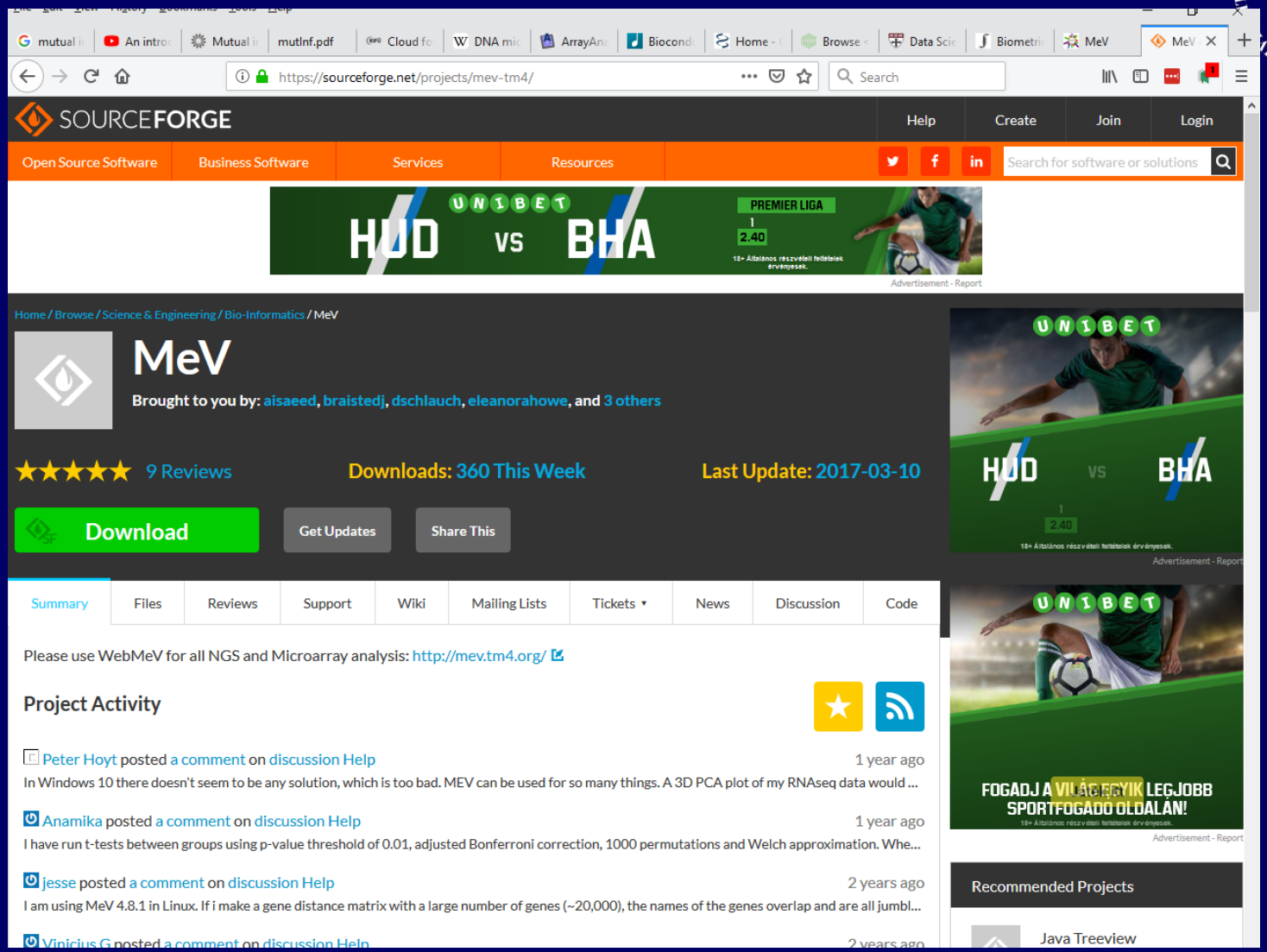

ArrayAnalysis.org
Bioconductor - BiocViews
Home - GEO - NCBI
A-AFFY-44 < Arrays < E-...
MeV
BRB-ArrayTools
www.ebi.ac.uk/arrayexpress/arrays/A-AFFY-44/?ref=E-MTAB-5067
EMBL-EBI
Services
Research
Training
About us
EMBL-EBI Hinxton
ArrayExpress
Search
Examples: E-MEXP-31, cancer, p53, Geuvadis
advanced search
Home
Browse
Submit
Help
About ArrayExpress
Contact Us
Login
ARRAYEXPRESS / BROWSE / E-MTAB-5067 / ARRAYS / A-AFFY-44
A-AFFY-44 - Affymetrix GeneChip Human Genome U133 Plus 2.0 [HG-U133_Plus_2]
Organism
Homo sapiens
Version
1.0
Provider
Affymetrix, Inc. (support@affymetrix.com)
Links
All 4382 experiments done using A-AFFY-44
Files
Array Design
A-AFFY-44.adf.txt
Browse all available files
EMBL-EBI
Services
By topic
By name (A-Z)
Help & Support
Research
Publications
Research groups
Postdocs & PhDs
Training
Train at EBI
Train outside EBI
Train online
Contact organisers
Industry
Members Area
Workshops
SME Forum
Contact Industry programme
About EMBL-EBI
Contact us
Events
Jobs
News
People & groups
EMBL-EBI, Wellcome Genome Campus, Hinxton, Cambridgeshire, CB10 1SD, UK. +44 (0)1223 49 44 44
Copyright © EMBL-EBI 2016 | EMBL-EBI is part of the European Molecular Biology Laboratory | Terms of use
Intranet

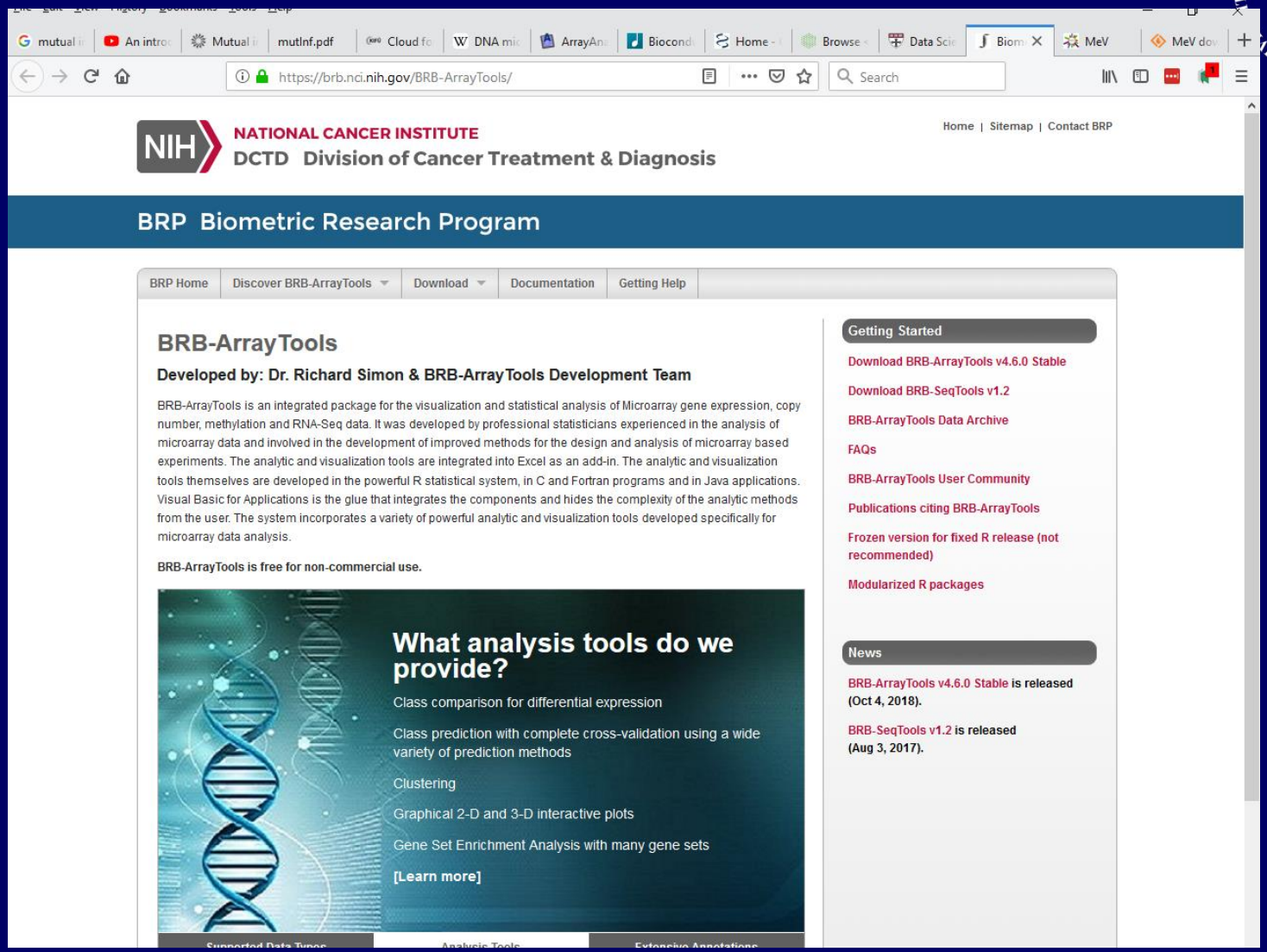

ArrayAnalysis.org
Bioconductor - BiocViews
Home - GEO - NCBI
http://www.e...Y-44.adf.txt
MeV
BRB-ArrayTools
www.ebi.ac.uk/arrayexpress/files/A-AFFY-44/A-AFFY-44.adf.txt
Array Design Name Affymetrix GeneChip Human Genome U133 Plus 2.0 [HG-U133_Plus_2]
Comment[AdditionalFile:CDF] HG-U133_Plus_2.CDF
Comment[AdditionalFile:CDF] HG-U133_Plus_2.cdf
Version 1.0
Provider Affymetrix, Inc. (support@affymetrix.com)
Comment[ArrayExpressAccession] A-AFFY-44
Comment[SubmittedName] Affymetrix GeneChip Human Genome U133 Plus 2.0 [HG-U133_Plus_2]
Comment[Organism] Homo sapiens
Technology Type in_situ_oligo_features
Technology Type Term Accession Number #in_situ_oligo_features
Technology Type Term Source REF The MGED Ontology
Surface Type glass
Surface Type Term Accession Number #glass
Surface Type Term Source REF mo
Term Source Name The MGED Ontology mo affymetrix_netaffx cp450 ec embl ensembl genbank interpro locus omim pkr_hanks refseq scop swall unigene
Term Source File http://mged.sourceforge.net/ontologies/MGEDontology.php http://mged.sourceforge.net/ontologies/MGEDontology.php http://www.affymetrix.com/analysis/ http://drnelson.utmem.edu/CytochromeP450.html http://www.chem.qmw.ac.uk/iubmb/enzyme/ http://www.ebi.ac.uk/embl/ http://www.ensembl.org http://www.ncbi.nlm.nih.gov/Genbank/index.html http://www.ebi.ac.uk/interpro http://www.ncbi.nlm.nih.gov/LocusLink/ http://www3.ncbi.nlm.nih.gov/omim http://pkr.sdsc.edu/html/pk_classification/pk_catalytic/pk_hanks_class.html http://www.ncbi.nlm.nih.gov/LocusLink/refseq.html http://scop.berkeley.edu http://www.ebi.ac.uk/swissprot/access.html http://www.ncbi.nlm.nih.gov/UniGene/
Term Source Version
Comment[AdditionalFile:txt] A-AFFY-44_comments.txt
[main]
Composite Element Name Composite Element Database Entry[affymetrix_netaffx] Composite Element Database Entry[embl] Composite Element Database Entry[ensembl] Composite Element Database Entry[swall] Composite Element Database Entry[locus] Composite Element Database Entry[interpro] Composite Element Database Entry[unigene] Composite Element Database Entry[omim] Composite Element Database Entry[refseq] Composite Element Database Entry[scop] Composite Element Database Entry[genbank] Composite Element Database Entry[pkr_hanks] Composite Element Database Entry[ec] Composite Element Database Entry[cp450]
AFFX-BioB-5_at AFFX-BioB-5_at AFFX-BioB-5
AFFX-BioB-M_at AFFX-BioB-M_at AFFX-BioB-M
AFFX-BioB-3_at AFFX-BioB-3_at AFFX-BioB-3
AFFX-BioC-5_at AFFX-BioC-5_at AFFX-BioC-5
AFFX-BioC-3_at AFFX-BioC-3_at AFFX-BioC-3
AFFX-BioDn-5_at AFFX-BioDn-5_at AFFX-BioDn-5
AFFX-BioDn-3_at AFFX-BioDn-3_at AFFX-BioDn-3
AFFX-CreX-5_at AFFX-CreX-5_at AFFX-CreX-5
AFFX-CreX-3_at AFFX-CreX-3_at AFFX-CreX-3
AFFX-DapX-5_at AFFX-DapX-5_at AFFX-DapX-5
AFFX-DapX-M_at AFFX-DapX-M_at AFFX-DapX-M
AFFX-DapX-3_at AFFX-DapX-3_at AFFX-DapX-3
AFFX-LysX-5_at AFFX-LysX-5_at AFFX-LysX-5
AFFX-LysX-M_at AFFX-LysX-M_at AFFX-LysX-M
AFFX-LysX-3_at AFFX-LysX-3_at AFFX-LysX-3
AFFX-PheX-5_at AFFX-PheX-5_at AFFX-PheX-5
AFFX-PheX-M_at AFFX-PheX-M_at AFFX-PheX-M
AFFX-PheX-3_at AFFX-PheX-3_at AFFX-PheX-3
AFFX-ThrX-5_at AFFX-ThrX-5_at AFFX-ThrX-5
AFFX-ThrX-M_at AFFX-ThrX-M_at AFFX-ThrX-M
AFFX-ThrX-3_at AFFX-ThrX-3_at AFFX-ThrX-3
AFFX-TrpnX-5_at AFFX-TrpnX-5_at AFFX-TrpnX-5
AFFX-TrpnX-M_at AFFX-TrpnX-M_at AFFX-TrpnX-M
AFFX-TrpnX-3_at AFFX-TrpnX-3_at AFFX-TrpnX-3

Gén-chip elemző alkalmazások

- Ingyenes és felhasználóbarát
- Mev:
  - http://mev.tm4.org/#/welcome
- BRB-tools:
  - https://brb.nci.nih.gov/BRB-ArrayTools/


mev.tm4.org/#/welcome
mev
Feedback
Tweets by @webmev
MeV
@webmev
Due to recent Google Cloud Platform update, WebMeV has become unstable. We are working on identifying and fixing the problem. We apologize for inconvenience that this cause. -- MeV Team
Jun 25, 2018
MeV
@webmev
WebMeV wiki contain Platform description is now available at : github.com/dfci-cccb/mev/...
dfci-cccb/mev
MultiExperiment Viewer. Cont...
github.com
Aug 28, 2017
MeV
@webmev
WebMeV: A Cloud Platform for Analyzing and Visualizing Cancer Genomic Data
biorxiv.org/content/early/...
WebMeV: A Cloud Platform f...
Although large, complex gen...
biorxiv.org
Jun 9, 2017
MeV
@webmev
Bio-IT 2017 - WebMeV Presentation
Welcome to the Web MEV
WebMeV (Multiple Experiment Viewer) is a cloud-based application supporting analysis, visualization, and stratification of large genomic data, particularly for RNASeq and microarray data.
With WebMeV platform, you can:
Perform RNASeq Analysis - Perform differential expression analysis using RNASeq raw count data to draw biological insights.
Access to Public Domain Data - Directly search and pull TCGA and GEO gene expression and sample attribute data in addition to private data for analysis
Stratify Cohorts using Clinical Attributes - Perform complex cohort stratification using sophisticated regular expression, facet filter, and set operations.
Web MeV is being built to meet the challenge of exploring large public genomic data set and Next-Generation Sequencing data with intuitive graphical interface for analysis. MeV is a free and open-source cloud service platform that does not require log in to use.
Where to start - A series of videos tutorials are available on YouTube and accessible after entering the WebMeV. Example files for count matrix and sample attribute are available here.
Tutorials and Documentation - tutorial documentation is available on the WebMev Wiki
Customized Analysis - Complex bioinformatics analysis outside the scope of WebMeV functionalities can contact the Center for Cancer Computational Biology at Dana-Farber Cancer Institute as consulting project. Click here for more information.
Sign In With Google
Get Started
Tutorials
Browse
TM4 MeV Stand-Alone Client
Download Current Stable Version (.zip)
WebMev Wiki
What's Inside
Bioconductor
D3.js
R Statistics Language
Open Refine
GitHub
build passing
See current developments and contribute to the project here.
Have an issue or want to report a bug? Tell us about it.
mev


SOURCEFORGE
Help
Create
Join
Login
Open Source Software
Business Software
Services
Resources
Search for software or solutions
UNIBET
HUD vs BHA
PREMIER LIGA
2.40
Advertisement - Report
Home / Browse / Science & Engineering / Bio-Informatics / MeV
MeV
Brought to you by: aisaeed, braistedj, dschlauch, eleanorahowe, and 3 others
9 Reviews
Downloads: 360 This Week
Last Update: 2017-03-10
Download
Get Updates
Share This
Summary
Files
Reviews
Support
Wiki
Mailing Lists
Tickets
News
Discussion
Code
Please use WebMeV for all NGS and Microarray analysis: http://mev.tm4.org/
Project Activity
Peter Hoyt posted a comment on discussion Help
1 year ago
In Windows 10 there doesn't seem to be any solution, which is too bad. MEV can be used for so many things. A 3D PCA plot of my RNAseq data would ...
Anamika posted a comment on discussion Help
1 year ago
I have run t-tests between groups using p-value threshold of 0.01, adjusted Bonferroni correction, 1000 permutations and Welch approximation. Whe...
jesse posted a comment on discussion Help
2 years ago
I am using MeV 4.8.1 in Linux. If i make a gene distance matrix with a large number of genes (~20,000), the names of the genes overlap and are all jumbl...
Vinicius G posted a comment on discussion Help
2 years ago
FOGADJ A VILÁG EGYIK LEGJOBB SPORTFOGADÓ OLDALÁN!
Recommended Projects
Java Treeview


https://brb.nci.nih.gov/BRB-ArrayTools/
NATIONAL CANCER INSTITUTE
DCTD Division of Cancer Treatment & Diagnosis
Home | Sitemap | Contact BRP
BRP Biometric Research Program
BRP Home
Discover BRB-ArrayTools
Download
Documentation
Getting Help
BRB-ArrayTools
Developed by: Dr. Richard Simon & BRB-ArrayTools Development Team
BRB-ArrayTools is an integrated package for the visualization and statistical analysis of Microarray gene expression, copy number, methylation and RNA-Seq data. It was developed by professional statisticians experienced in the analysis of microarray data and involved in the development of improved methods for the design and analysis of microarray based experiments. The analytic and visualization tools are integrated into Excel as an add-in. The analytic and visualization tools themselves are developed in the powerful R statistical system, in C and Fortran programs and in Java applications. Visual Basic for Applications is the glue that integrates the components and hides the complexity of the analytic methods from the user. The system incorporates a variety of powerful analytic and visualization tools developed specifically for microarray data analysis.
BRB-ArrayTools is free for non-commercial use.
What analysis tools do we provide?
Class comparison for differential expression
Class prediction with complete cross-validation using a wide variety of prediction methods
Clustering
Graphical 2-D and 3-D interactive plots
Gene Set Enrichment Analysis with many gene sets
[Learn more]
Getting Started
Download BRB-ArrayTools v4.6.0 Stable
Download BRB-SeqTools v1.2
BRB-ArrayTools Data Archive
FAQs
BRB-ArrayTools User Community
Publications citing BRB-ArrayTools
Frozen version for fixed R release (not recommended)
Modularized R packages
News
BRB-ArrayTools v4.6.0 Stable is released (Oct 4, 2018).
BRB-SeqTools v1.2 is released (Aug 3, 2017).

ArrayAnalysis.org | Bioconductor - BiocViews | Home - GEO - NCBI | http://www.e...Y-44.adf.txt | GitHub - dfci-cccb/mev: ... | BRB-ArrayTools

https://brb.nci.nih.gov/BRB-ArrayTools/features.html

NIH NATIONAL CANCER INSTITUTE

DCTD: BIOMETRIC RESEARCH PROGRAM

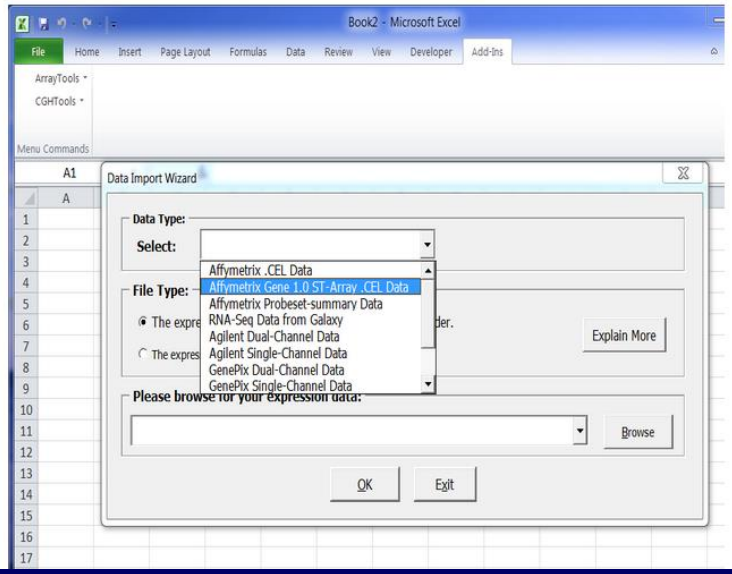
BRB-ArrayTools

Developed by: Richard Simon & BRB-ArrayTools Development Team

Home | Discover BRB-ArrayTools | Download | Documentation | Getting Help

Supported Data Types

- Expression data
  - Affymetrix 3'-IVT arrays
  - Affymetrix genechip ST 1.0 arrays
  - Agilent single and dual channel arrays
  - Illumina expression arrays
  - GenePix data
  - mAdb Archive data
  - GDS datasets from



Mit tanultunk ma?

- A gén-chip (micro-array) technológia nagyon elterjedt és hatékony
- A kiválasztott rendszer globális állapotát jellemzi
- Bioinformatika nélkül elképzelhetetlen az adatok kiértékelése
- Nyilvános adatbázisokon keresztül rengeteg eredmény érhető el

Feladat

A GEO vagy az ArrayExpress adatbazisban keress számodra érdekes kisérletet.