8/3/09

Using cellHTS2 for cell based assays

Florian Hahne
Fred Hutchinson Cancer Research Center

FRED HUTCHINSON CANCER RESEARCH CENTER

A LIFE OF SCIENCE

cellHTS2: purpose

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- analyze high throughput cell-based assays with low complexity readouts
  e.g. plate reader assays, luminescence assays…
- from raw data to annotated hit list
- data preprocessing, normalization
- data quality assessment
- replicate scoring, annotation
- analysis audit trail

cellHTS2: assay types

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- single color assays
- multi color assays
- activator or inhibitor type assays
- dual way assays
- independent of the instrument, as long as the output is ASCII text
- not coupled to a particular screening library, or to particular organisms

cellHTS2: limitations

Assumes that every well contains the same type of reagents:

the same controls (can be circumvented but becomes very awkward…)

the same library

cellHTS2: publication

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

File Edit View History Bookmarks Tools Help

http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=

Google

Analysis of cell-based RNAi s... | Analysis of cell-based R...

PubMed Central Search Journal List

BioMed Central | Genome Biology | this article | search | submit a manuscript | register

Journal List > Genome Biol > v.7(7); 2006

Abstract
Full Text
Supplementary Material
PDF (655K)
Contents
Archive

Related material:
PubMed related arts
GO

PubMed articles by:
Boutros, M.
Brás, L.
Huber, W.

Top

Genome Biol. 2006; 7(7): R66. PMCID: PMC1779553
Published online 2006 July 25. doi: 10.1186/gb-2006-7-7-r66.

Analysis of cell-based RNAi screens

Reviewed by Michael Boutros,[*1] Lígia P Brás,[2,3] and Wolfgang Huber[*2]

[1]Signaling and Functional Genomics, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany

[2]EMBL - European Bioinformatics Institute, Cambridge CB10 1SD, UK

[3]Centre for Chemical and Biological Engineering, IST, Technical University of Lisbon, Av. Rovisco Pais, P-1049-001 Lisbon, Portugal

[*]Corresponding author.

Michael Boutros: m.boutros@dkfz.de; Lígia P Brás: ligia@ebi.ac.uk; Wolfgang Huber: huber@ebi.ac.uk

Received March 27, 2006; Revised June 7, 2006; Accepted July 25, 2006.

This article has been cited by other articles in PMC.

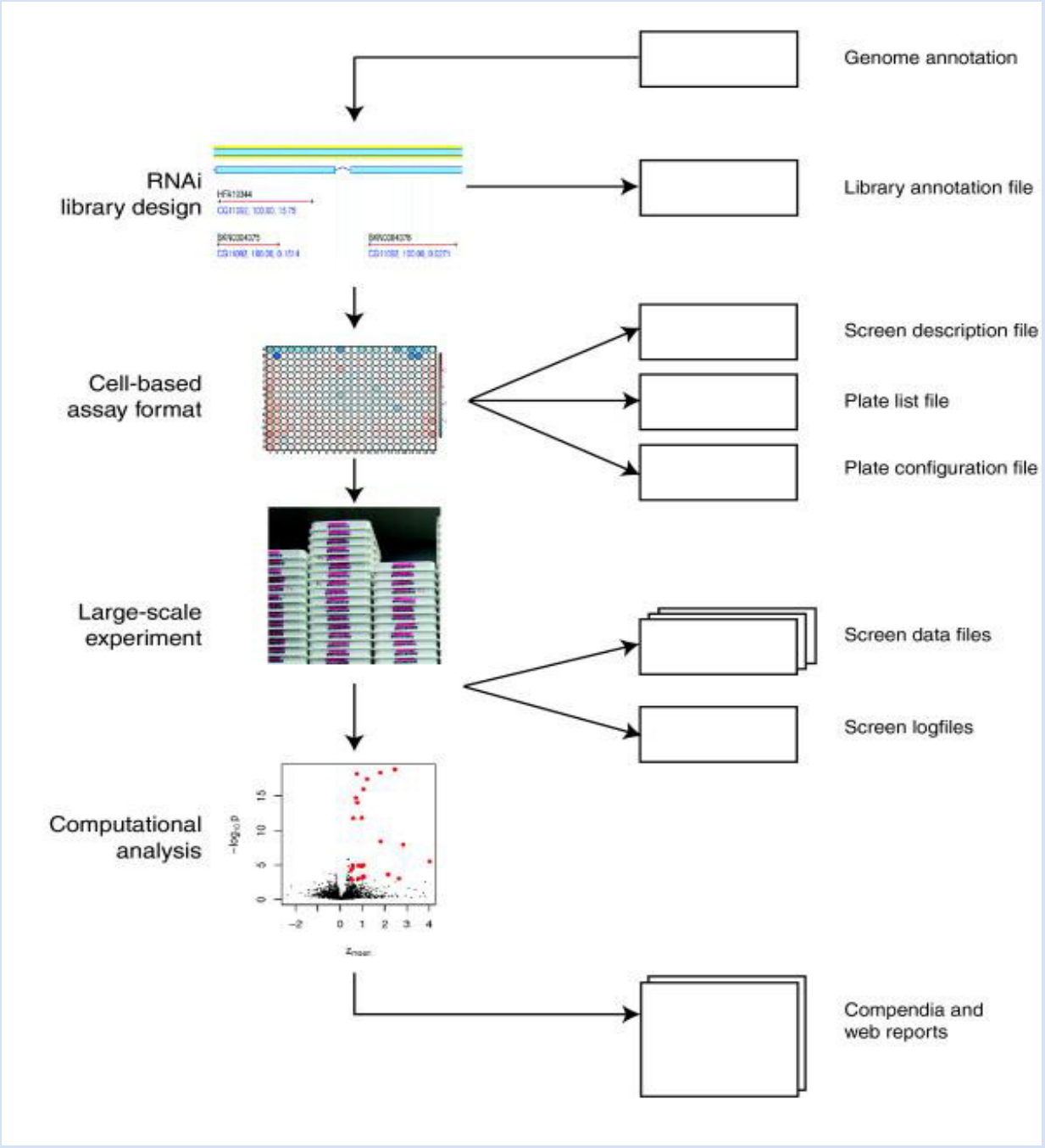
Abstract

RNA interference (RNAi) screening is a powerful technology for functional

http://www.pubmedcentral.nih.gov/articlerender.fcgi?tool=pubmed&pubmedid=16869968#supplementary-material-sec

cellHTS2: structure

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE



Data structure: screening data

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

Numeric values x_{ijk}

- i = wells (e.g. 20,000)
- j = different reporters (e.g. 2)
- k = different assays (e.g. 5)

Metadata about wells

- p_i = plate in which is well i
- r_i = row (within plate) of well i
- c_i = column (within plate) of well I
- siRNA sequence, target gene, ...

Metadata about reporters

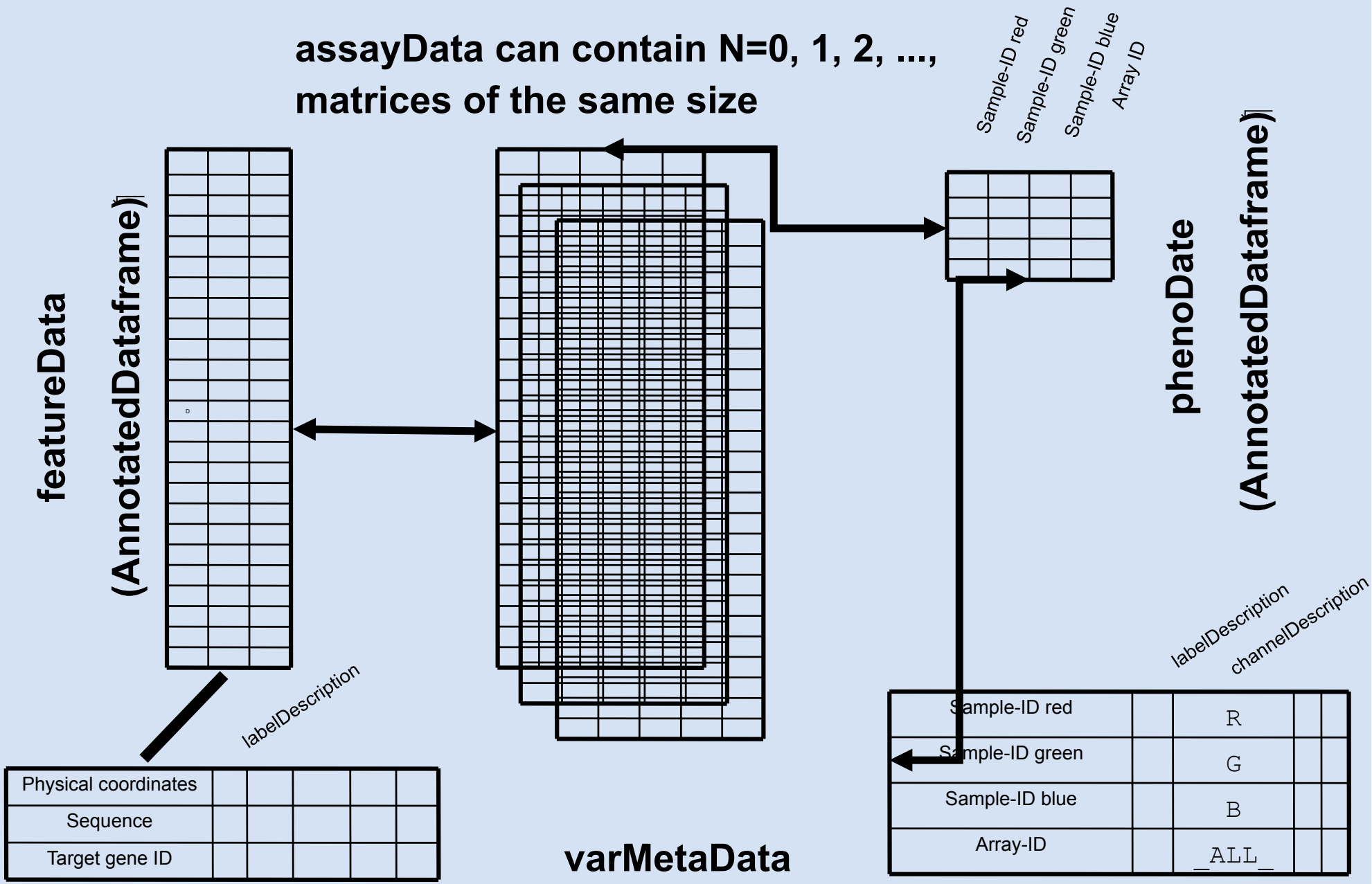
- Fluc, Rluc, ...

Metadata about assays k

- replicate number
- different variants of the assay (e.g. Wnt1_LRP6_Frz8, β-Catenin, LRP6ΔE1-4, Wnt1_R-Spondin3, Wnt3a)
- date it was done

Data structure: NChannelSet

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE



Data import: function *readPlateList()*

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- reads data from most established instruments, but the can handle arbitrary data formats through user-defined import functions

- automatically detects plate layout

- raw data files need to be specified using a plate list file

    columns Filename, Plate, Replicate, (Batch), …

- can read collated data from a single file

Data analysis: four stages

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

1) Configuration
   controls, screen design, flagging

2) Annotation
   features, e.g. target genes

3) Normalization
   between and within plates, data transformation

4) Replicate scoring and summarization
   standardization and reduction to a single value per feature (e.g. z-scores)

human readable audit trail in form of an interactive HTML document can be produced at each of these steps

Configuration: function *configure()*

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- screen level: screen description file (MIAME)
- plate level: plate configuration file

  regular expressions as annotation rules

Plate	Well	Content
*	*	sample
*	A0[1-2]	other
*	B01	neg
*	B02	pos

  visualization of plate layouts
  configurationAsScreenPlot()

- flagging: screen log file

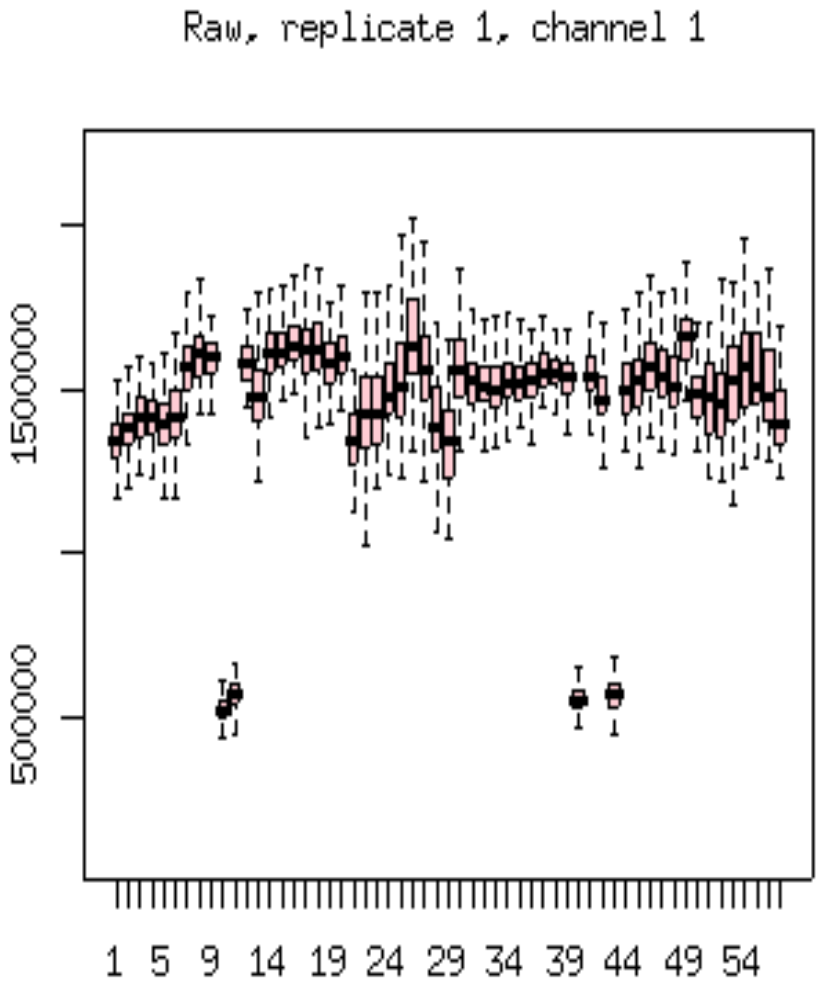
Annotation: function *annotate()*

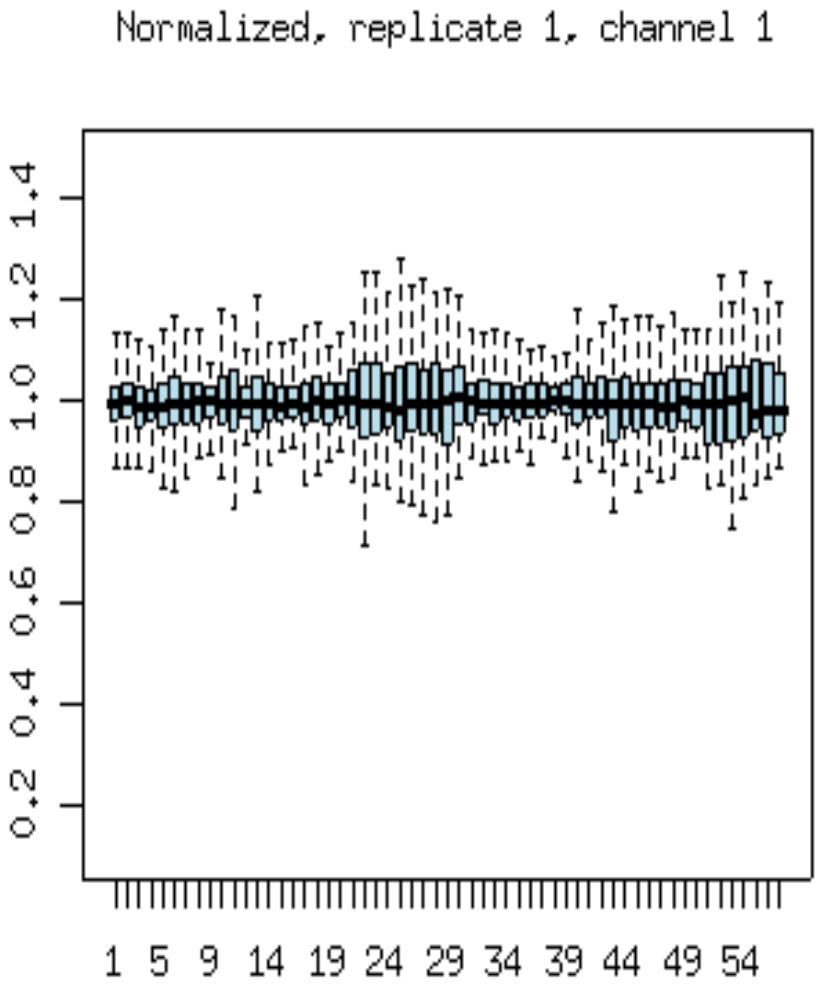
FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- description of features in annotation file
- mandatory columns: Plate, Well, GeneID
- optional column GeneSymbol: Human-readable name

Normalization: between plates

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE





Normalization: *estimating parameters*

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

From which data points:

- **Based on the intensities of the controls**

  if they work uniformly well across all plates

- **Based on the intensities of the samples**

  invoke assumptions such as "most genes have no effect", or "same distribution of effect sizes"

Which estimator:

mean *vs* median *vs* shorth

standard deviation *vs* MAD vs IQR

No universally optimal answer, it depends on the data.
In the best case, it doesn't matter.

Normalization: between plates

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- by estimator of location for each plate:
  mean, median, shorth, percent control
- data transformation (optional):
  log or linear scale
- Variance adjustment (optional):
  by plate, by batch, by experiment
- function *normalizePlates()*

Normalization: *spatial within plates*

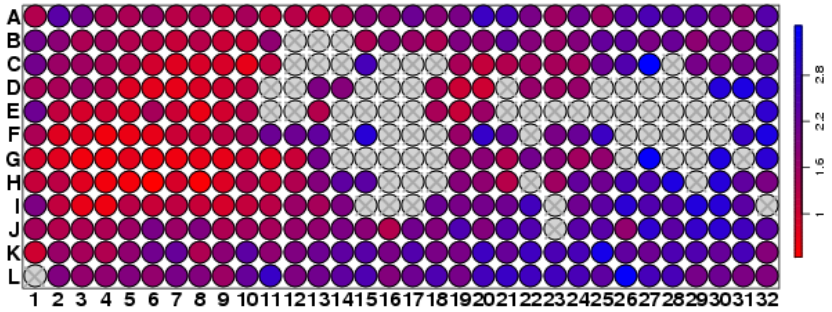
FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

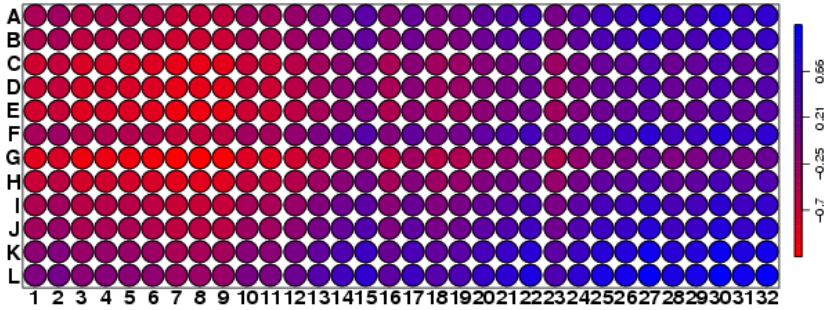
_B_-score:

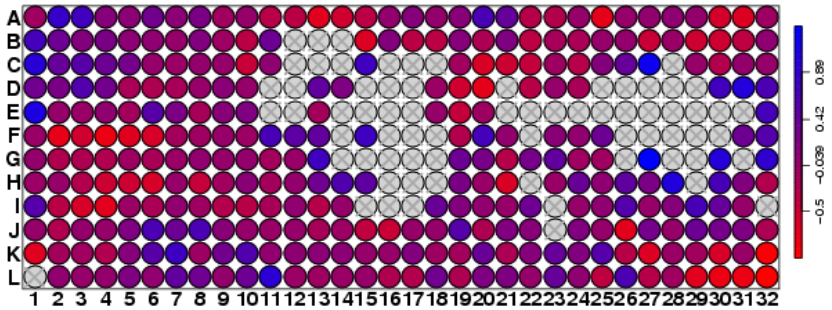
two-way medianpolish

r^{th} row
c^{th} column
i^{th} plate

Malo et al., Nat. Biotech. 2006







Scoring and summarization:

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

- *scoreReplicates()*
  z-score, normalized percent inhibition
- *summarizeReplicates()*
  e.g. mean, max, min
- *summarizeChannels()*

HTML report: function *writeReport()*

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

report

Acknowledgements

FRED HUTCHINSON CANCER RESEARCH CENTER
A LIFE OF SCIENCE

Ligia Bras

Wolfgang Huber

Michael Boutros

Thomas Horn, Tina Büchling, Dorothee Nickles, Dierk Ingelfinger

Elin Axelsson

Gregoire Pau

Martin Morgan