Composing HTML documents with hwriter

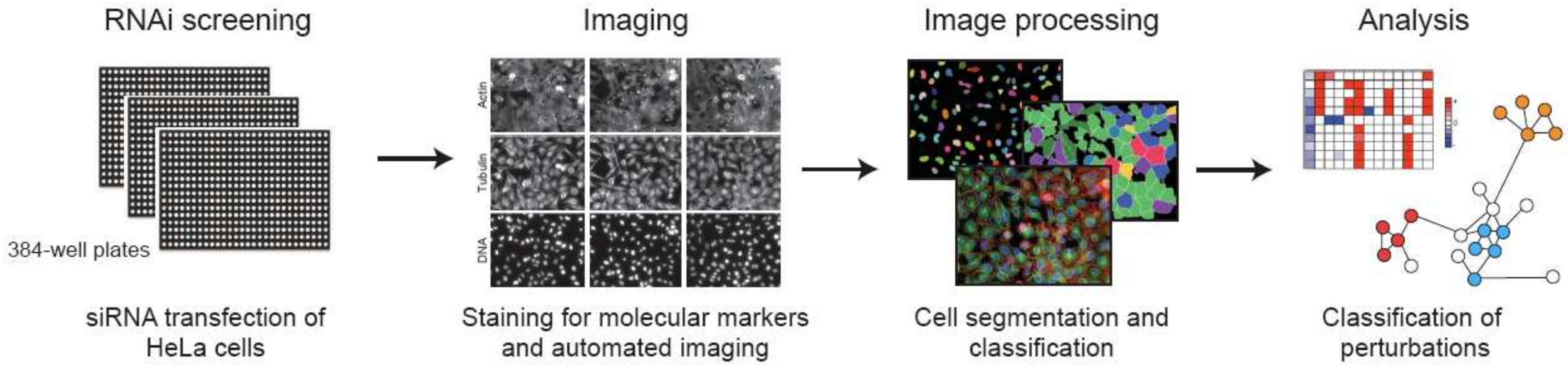
(with an application to cellular imaging)

Gregoire Pau, Wolfgang Huber
European Molecular Biology Laboratory
Heidelberg, Germany
gregoire.pau@embl.de

EMBL-EBI

Motivation

- Reporting analysis of complex data sets
  - Large and/or heterogeneous data
  - Example: analysis of genome-wide imaging screens
  - 22184 genes, 91384 images, 6×10^6 cells, 181 features per cell



- HTML reports
  - Universal format, distributable, easy to browse
  - Supports tables, images and hyperlinks
  - Multiple pages, interactive

The package hwriter

- No HTML support in base R
- Development of the package hwriter [R journal, April 2009]
  - Available on CRAN
  - Simple syntax
  - Powerful (document composition, CSS styling, Javascript)
  - Fast

- Core function
  - hwrite(x, page=NULL, …)
  - Writes the object x in page, using the formatting arguments '…'
  - x is an R object (vector, matrix, character string…)
  - page can be a filename, a file connection or omitted

Formatting arguments

- Simple example

```
> hwrite('Hello world !', 'doc.html')
```

Hello world !

- Using hyperlink with link

```
> hwrite('Hello world !', 'doc.html',
  link='http://cran.r-project.org/')
```

Hello world !

- Using CSS style with style

```
> hwrite('Hello world !', 'doc.html',
  style='color: red; font-size: 20pt;
  font-family: Gill Sans Ultra Bold')
```

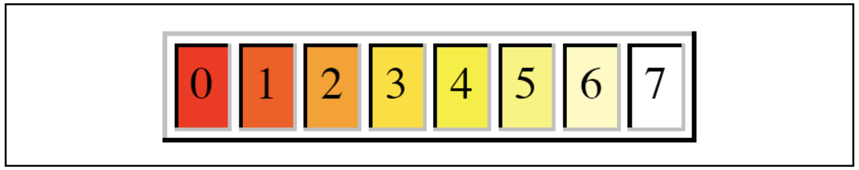
Hello world !

- Vector arguments are recycled using R recycling rules
- Support many other formatting arguments (CSS class, margins, padding, Javascript, alignment…)

Vectors and matrices

- Vector

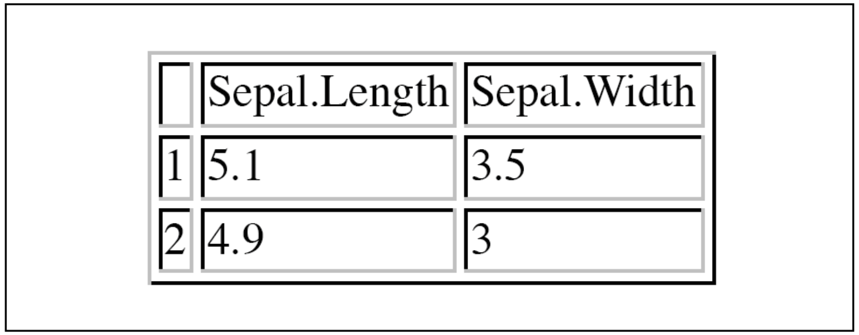
```
> colors = rgb(colorRamp(c('red', 'yellow',
  'white'))((0:7)/7), max=255)
> hwrite(0:7, 'doc.html', bgcolor=colors)
```



0	1	2	3	4	5	6	7

- Matrix

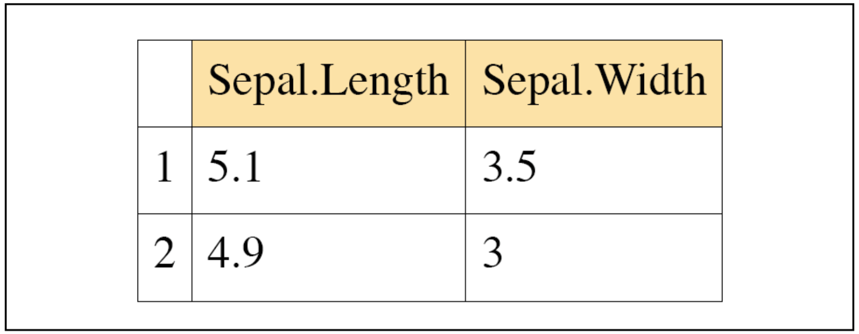
```
> hwrite(iris[1:2, 1:2], 'doc.html')
```



	Sepal.Length	Sepal.Width
1	5.1	3.5
2	4.9	3

- Matrix with colors and CSS style

```
> hwrite(iris[1:2, 1:2], 'doc.html',
  row.bgcolor='#ffdc98', table.style=
  'border-collapse:collapse',
  table.cellpadding='5px')
```



	Sepal.Length	Sepal.Width
1	5.1	3.5
2	4.9	3

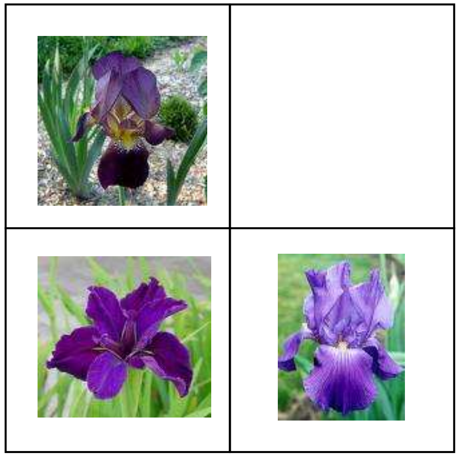
Images

- Using hwriteImage

```
> hwriteImage('iris1.jpg', 'doc.html')
```

- Matrix of images

```
> m = matrix(c('iris1.jpg', 'iris2.jpg', NA,
  'iris3.jpg'), nr=2)
> hwriteImage(m, 'doc.html', width='50%')
```

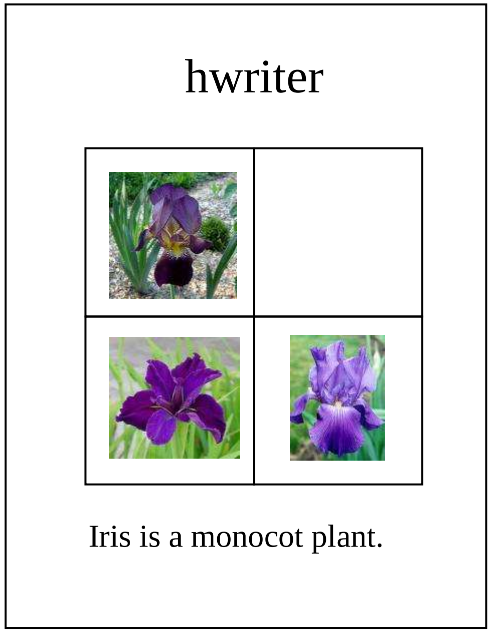


- Supports recycling, hyperlinking, CSS style, borders, padding, margins, alignment, background color…

Composing HTML pages

- Open a new HTML page with openPage
- Append elements into the page
- Close page with closePage

```
> p = openPage('doc.html')
> hwrite('hwriter', p, center=TRUE, heading=1)
> hwriteImage('iris1.jpg', p)
> hwrite('Iris is a monocot plant.', p)
> closePage(p)
```



hwriter

Iris is a monocot plant.

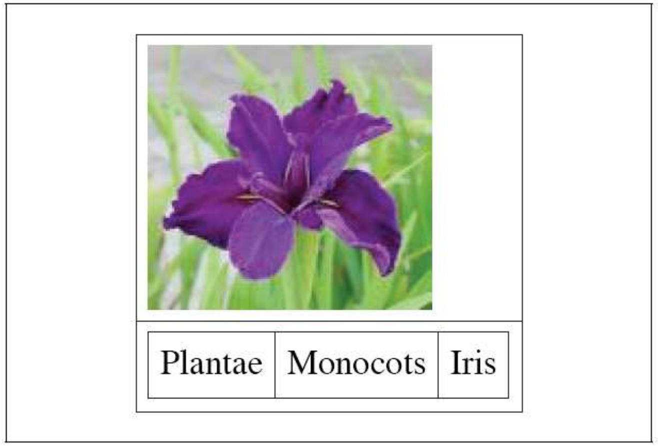
Nesting HTML elements

- Every hwrite call returns an HTML code part
- Code parts can be nested into tables to compose the layout

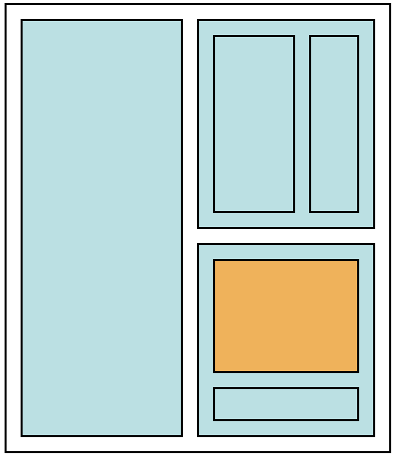
```
> cap = hwrite(c('Plantae','Monocots','Iris'),
  table.style='border-collapse:collapse',
  table.cellpadding='5px')
> print(cap)
```

```
<table style="border-collapse:collapse"
cellpadding="5px" border="1"><tr><td>
Plantae</td><td>Monocots</td><td>Iris
</td></tr></table>
```

```
> m = matrix(c(hwriteImage('iris1.jpg'), cap))
> hwrite(m, 'doc.html', table.cellpadding=
 '5px', table.style='border-collapse:collapse')
```

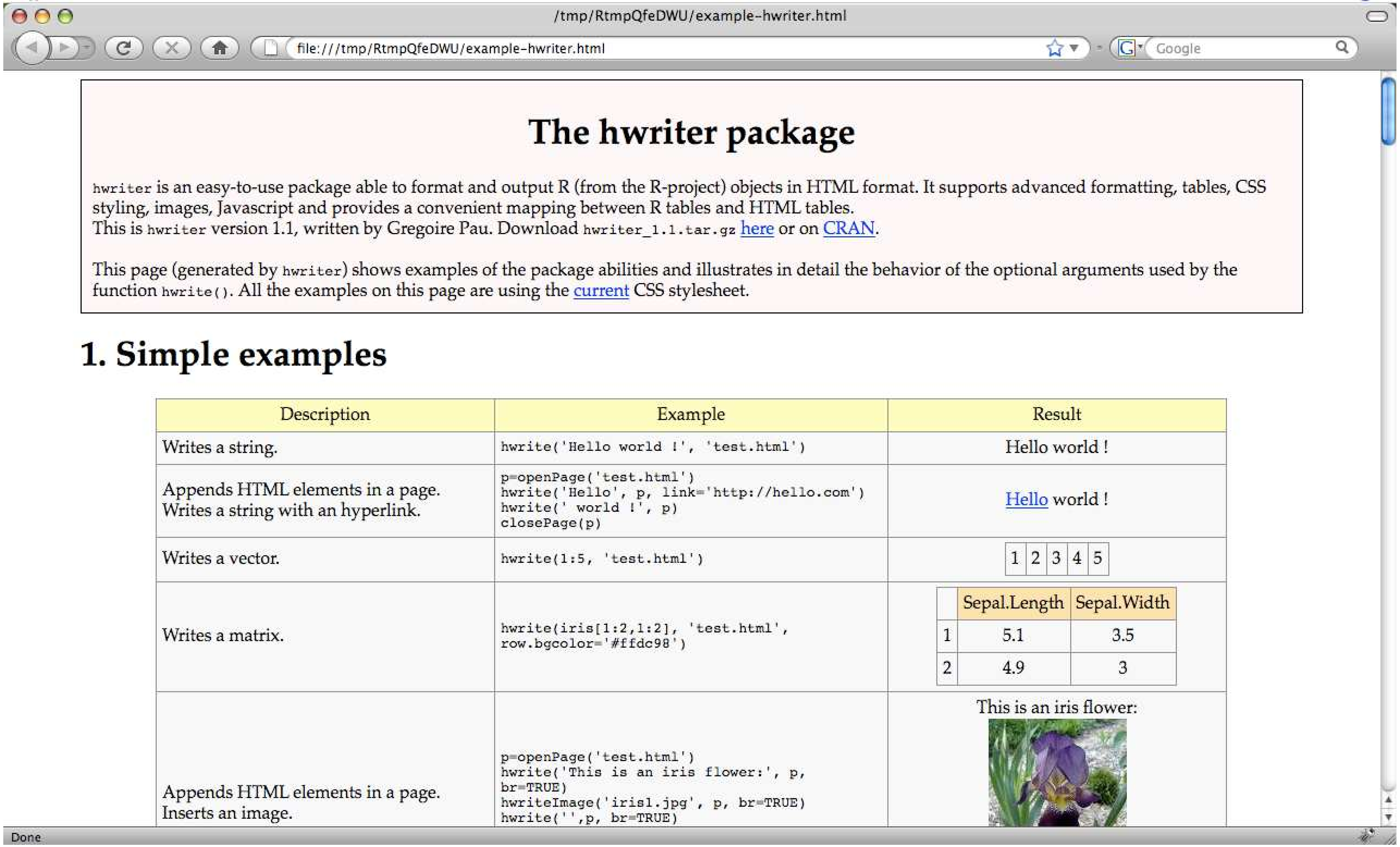


Plantae	Monocots	Iris



Documentation

```
> example(hwriter)
```



The hwriter package

`hwriter` is an easy-to-use package able to format and output R (from the R-project) objects in HTML format. It supports advanced formatting, tables, CSS styling, images, Javascript and provides a convenient mapping between R tables and HTML tables.
This is `hwriter` version 1.1, written by Gregoire Pau. Download `hwriter_1.1.tar.gz` here or on CRAN.

This page (generated by `hwriter`) shows examples of the package abilities and illustrates in detail the behavior of the optional arguments used by the function `hwrite()`. All the examples on this page are using the current CSS stylesheet.

1. Simple examples

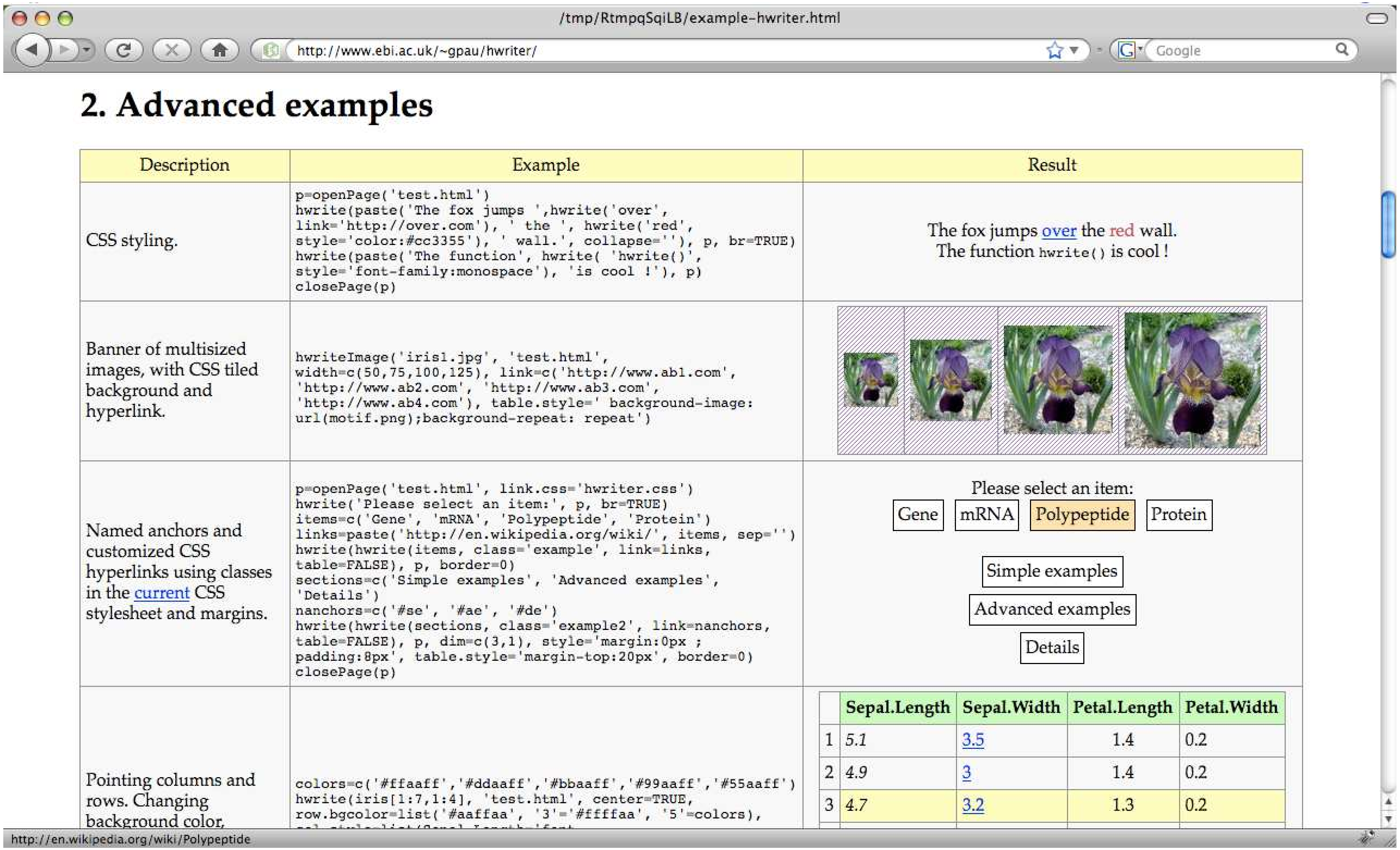
Description	Example	Result
Writes a string.	`hwrite('Hello world !', 'test.html')`	Hello world !

Appends HTML elements in a page.
Writes a string with an hyperlink.
	`p=openPage('test.html')`
	`hwrite('Hello', p, link='http://hello.com')`
	`hwrite(' world !', p)`
	`closePage(p)`
		Hello world !

Writes a vector.	`hwrite(1:5, 'test.html')`	1 2 3 4 5
Writes a matrix.	`hwrite(iris[1:2,1:2], 'test.html',` `row.bgcolor='#ffdc98')`	Sepal.Length Sepal.Width 1 5.1 3.5 2 4.9 3

Appends HTML elements in a page.
Inserts an image.
	`p=openPage('test.html')`
	`hwrite('This is an iris flower:', p,`
	`br=TRUE)`
	`hwriteImage('iris1.jpg', p, br=TRUE)`
	`hwrite('',p, br=TRUE)`
		This is an iris flower:

Documentation



2. Advanced examples

Description	Example	Result

CSS styling.

```
p=openPage('test.html')
hwrite(paste('The fox jumps ',hwrite('over',
link='http://over.com'), ' the ', hwrite('red',
style='color:#cc3355'), ' wall.', collapse=''), p, br=TRUE)
hwrite(paste('The function', hwrite( 'hwrite()',
style='font-family:monospace'), 'is cool !'), p)
closePage(p)
```

The fox jumps over the red wall.
The function `hwrite()` is cool !

Banner of multisized images, with CSS tiled background and hyperlink.

```
hwriteImage('iris1.jpg', 'test.html',
width=c(50,75,100,125), link=c('http://www.ab1.com',
'http://www.ab2.com', 'http://www.ab3.com',
'http://www.ab4.com'), table.style=' background-image:
url(motif.png);background-repeat: repeat')
```

Named anchors and customized CSS hyperlinks using classes in the current CSS stylesheet and margins.

```
p=openPage('test.html', link.css='hwriter.css')
hwrite('Please select an item:', p, br=TRUE)
items=c('Gene', 'mRNA', 'Polypeptide', 'Protein')
links=paste('http://en.wikipedia.org/wiki/', items, sep='')
hwrite(hwrite(items, class='example', link=links,
table=FALSE), p, border=0)
sections=c('Simple examples', 'Advanced examples',
'Details')
nanchors=c('#se', '#ae', '#de')
hwrite(hwrite(sections, class='example2', link=nanchors,
table=FALSE), p, dim=c(3,1), style='margin:0px ;
padding:8px', table.style='margin-top:20px', border=0)
closePage(p)
```

Please select an item:

Gene mRNA Polypeptide Protein

Simple examples

Advanced examples

Details

Pointing columns and rows. Changing background color,

```
colors=c('#ffaaff','#ddaaff','#bbaaff','#99aaff','#55aaff')
hwrite(iris[1:7,1:4], 'test.html', center=TRUE,
row.bgcolor=list('#aaffaa', '3'='#ffffaa', '5'=colors),
```

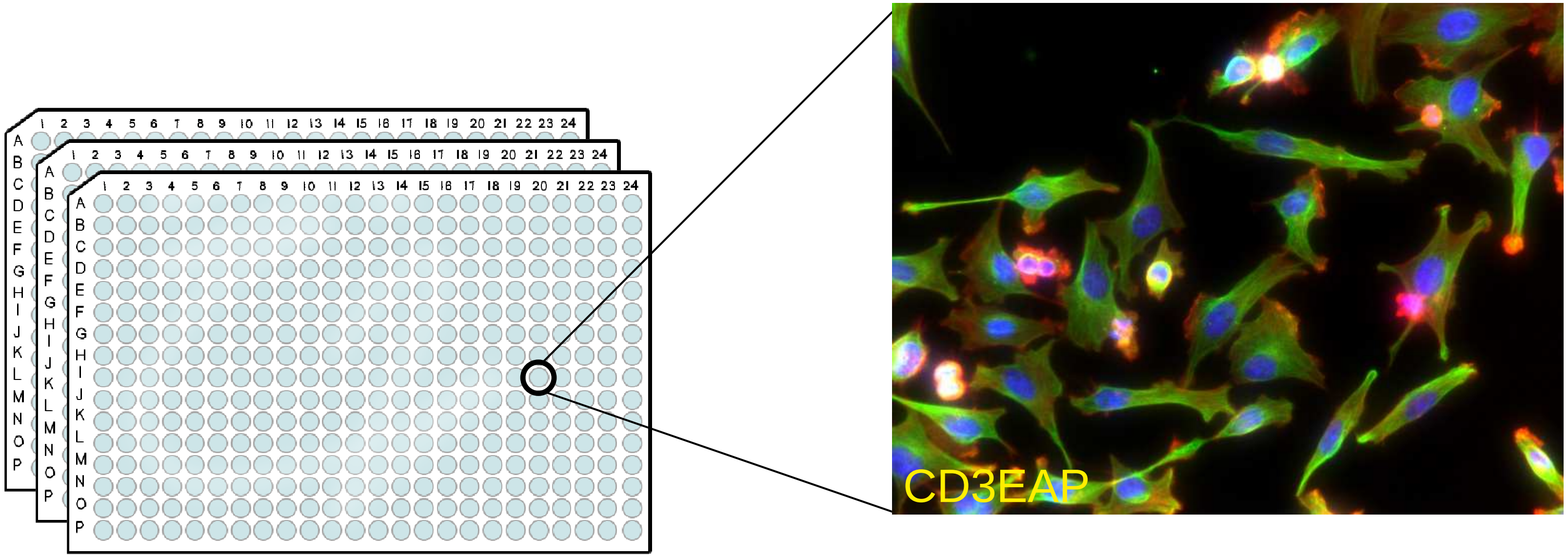
	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width
1	5.1	3.5	1.4	0.2
2	4.9	3	1.4	0.2
3	4.7	3.2	1.3	0.2

http://en.wikipedia.org/wiki/Polypeptide

Reporting analysis of high-throughput cellular assays

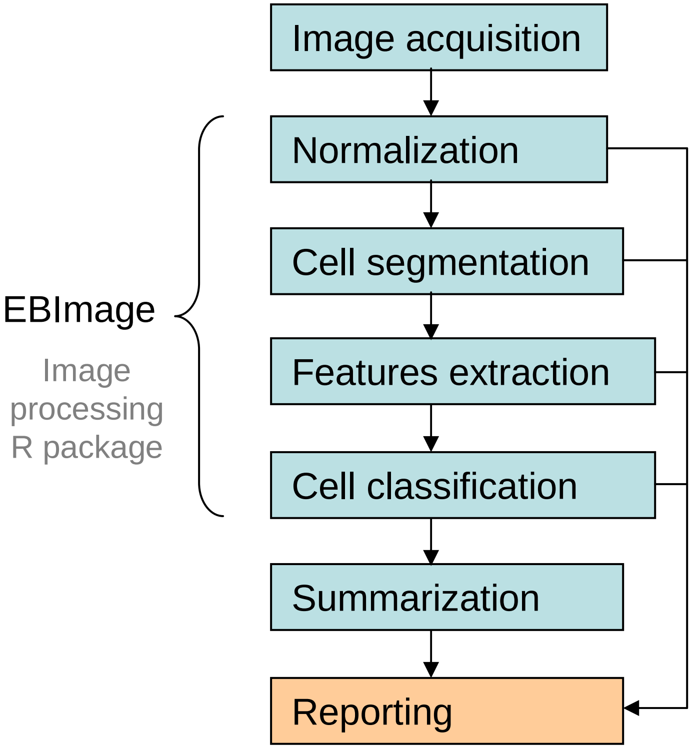
Experimental setup

- Genome-wide RNAi knockdowns (22839 genes)
- Human HeLa cells incubated for 48 h and fixed
- Readout
  - imageHTS: fluorescence microscopy images
  - CellHTS: scalar values (ATP levels from luciferase assay)
  - Mitocheck: Time-lapse video sequences



imageHTS

imageHTS workflow



EBImage

Image processing R package

Image acquisition

Normalization

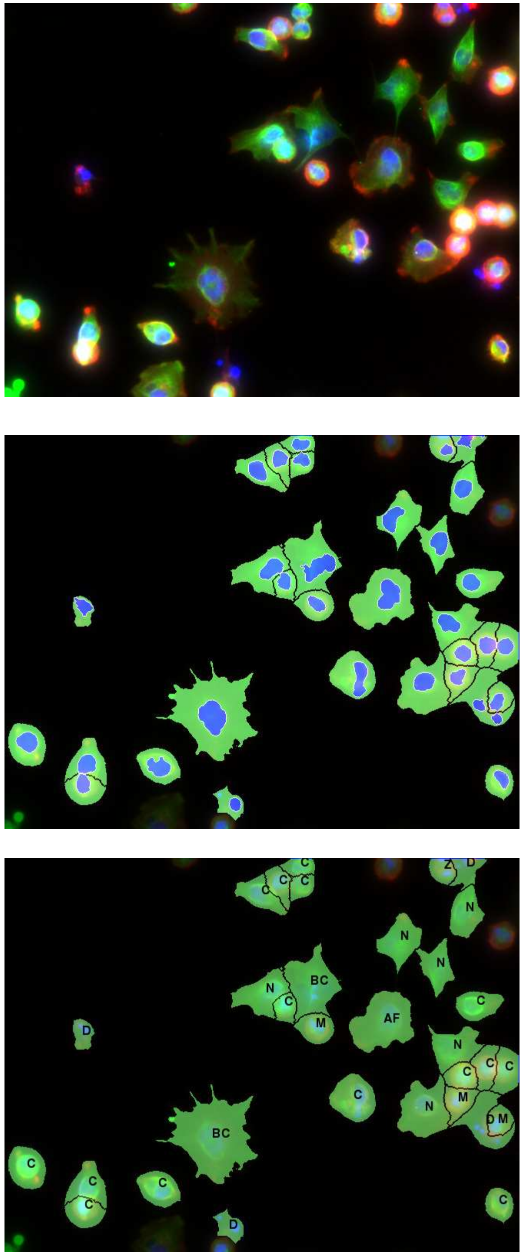
Cell segmentation

Features extraction

Cell classification

Summarization

Reporting



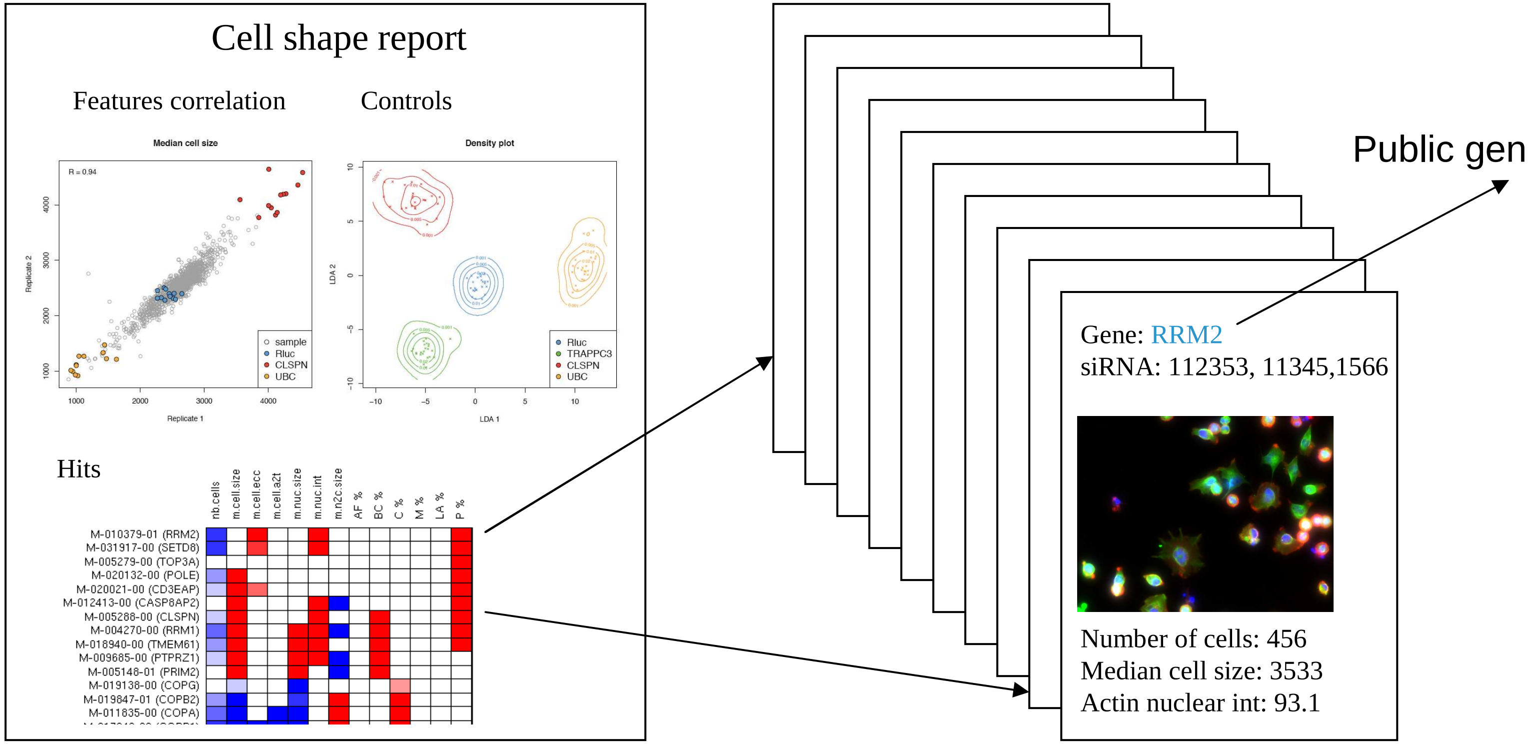
181 features

Cells

```
   g.x        g.y      g.s g.p    g.pdm    g.pdsd    g.effr
123.1391   3.288660  194  67  9.241719  4.165079  7.85825
206.7460   9.442248  961 153 20.513190  7.755419 17.48987
502.9589   7.616438  219  60  8.286918  1.954156  8.34924
 20.1919  22.358418 1568 157 22.219461  3.139197 22.34076
344.7959  45.501992 2259 233 35.158966 15.285795 26.81533
188.2611  50.451863 2711 249 28.732680  6.560911 29.37580
269.7996  46.404036 2131 180 26.419631  5.529232 26.04454
106.6127  58.364243 1348 143 21.662879  6.555683 20.71428
218.5582  77.299007 1913 215 25.724580  6.706719 24.67644
 19.1766  81.840147 1908 209 26.303760  7.864686 24.64417
  6.3558  62.017647  340  68 10.314127  2.397136 10.40314
 58.9873  86.034128 2139 214 27.463158  6.525559 26.09338
245.1087  94.387405 1048 123 18.280901  2.894758 18.26441
411.2741 109.198678 2572 225 28.660816  7.914664 28.61281
167.8151 107.966014 1942 160 24.671533  2.534342 24.86277
281.7084 121.609892 2871 209 31.577270  6.470767 30.23024
479.2334 143.098241 1649 183 23.913630  6.116630 22.91054
186.5930 146.693122 2079 199 27.280908  6.757808 25.72481
356.7303 148.253418 3145 285 34.746206 11.297632 31.63992
449.2436 147.798319  119  37  5.873578  1.563250  6.15458
```

Reporting

- Generation of a hierarchical website
- Several levels
  - General (screen quality, number of genes tested…) [1 page]
  - Gene (image, hyperlink to public databases…) [22839 pages]
  - Cell types (distribution of cellular features…) [12 pages]



Cell shape report

Features correlation

Controls

Hits

	nb.cells	m.cell.size	m.cell.ecc	m.cell.a2t	m.nuc.size	m.nuc.int	m.n2c.size	AF %	BC %	C %	M %	LA %	P %
M-010379-01 (RRM2)	🟦		🟥			🟥							🟥
M-031917-00 (SETD8)	🟦		🟥			🟥							🟥
M-005279-00 (TOP3A)													🟥
M-020132-00 (POLE)	🟦	🟥											🟥
M-020021-00 (CD3EAP)	🟦	🟥	🟥										🟥
M-012413-00 (CASP8AP2)		🟥				🟥	🟦						🟥
M-005288-00 (CLSPN)	🟦	🟥				🟥			🟥				🟥
M-004270-00 (RRM1)	🟦	🟥			🟥	🟥	🟦		🟥				🟥
M-018940-00 (TMEM61)	🟦	🟥			🟥	🟥			🟥				🟥
M-009685-00 (PTPRZ1)	🟦	🟥			🟥	🟥	🟦		🟥				
M-005148-01 (PRIM2)		🟥			🟥		🟦		🟥				
M-019138-00 (COPG)		🟦			🟦					🟥			
M-019847-01 (COPB2)	🟦	🟦			🟦		🟥			🟥			
M-011835-00 (COPA)	🟦	🟦		🟦	🟦		🟥			🟥			

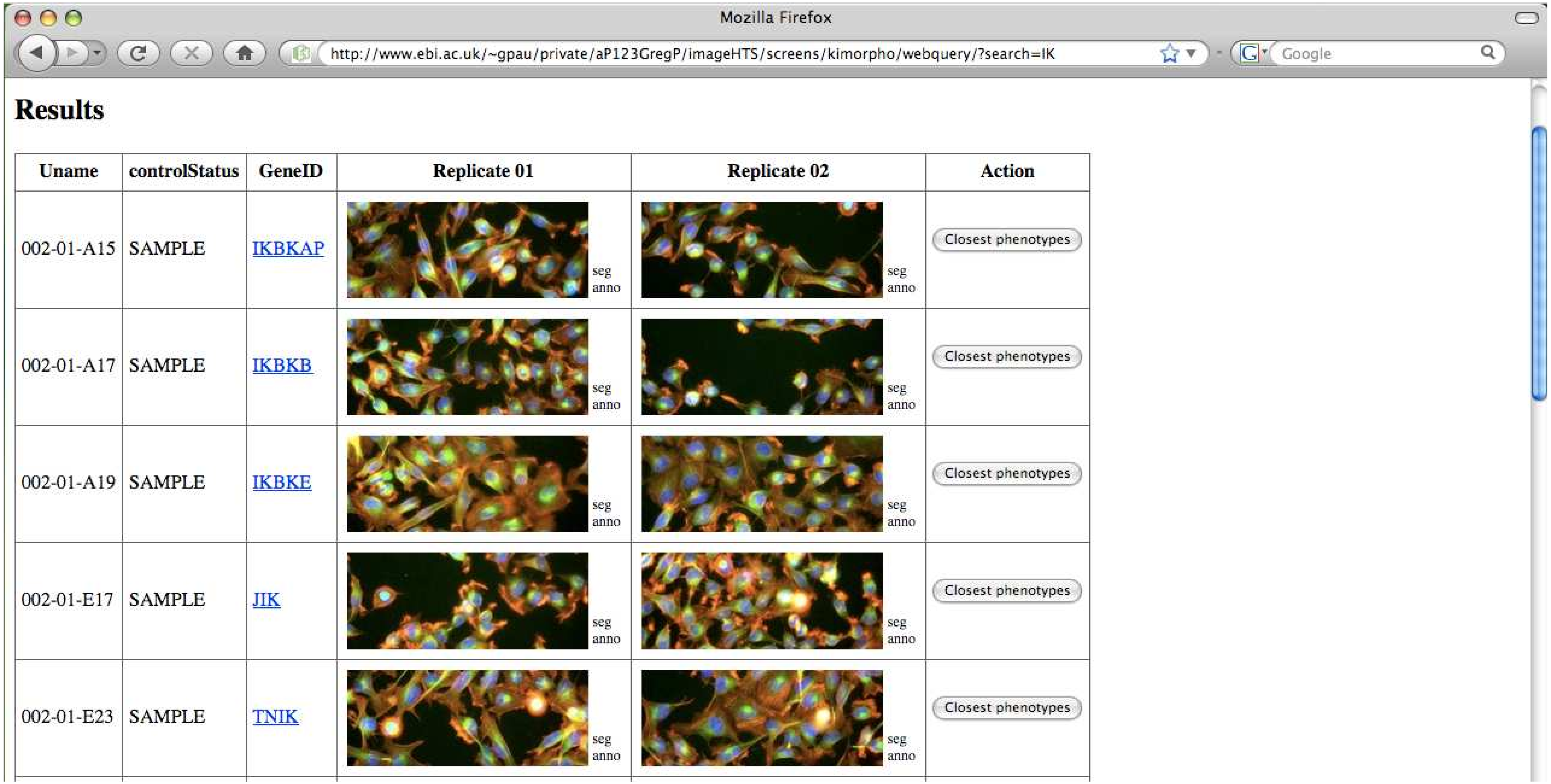
Public gene database

Gene: RRM2
siRNA: 112353, 11345,1566

Number of cells: 456
Median cell size: 3533
Actin nuclear int: 93.1

Phenotypic distance

- $d^2(x_1, x_2) = (x_1 - x_2)^t M (x_1 - x_2)$, where M is positive definite
- Fitting M according to $L = ||Y - d(X_1, X_2)||_2 + \gamma ||M||_1$
- Generation of a page showing phenotypes close to IKBKAP

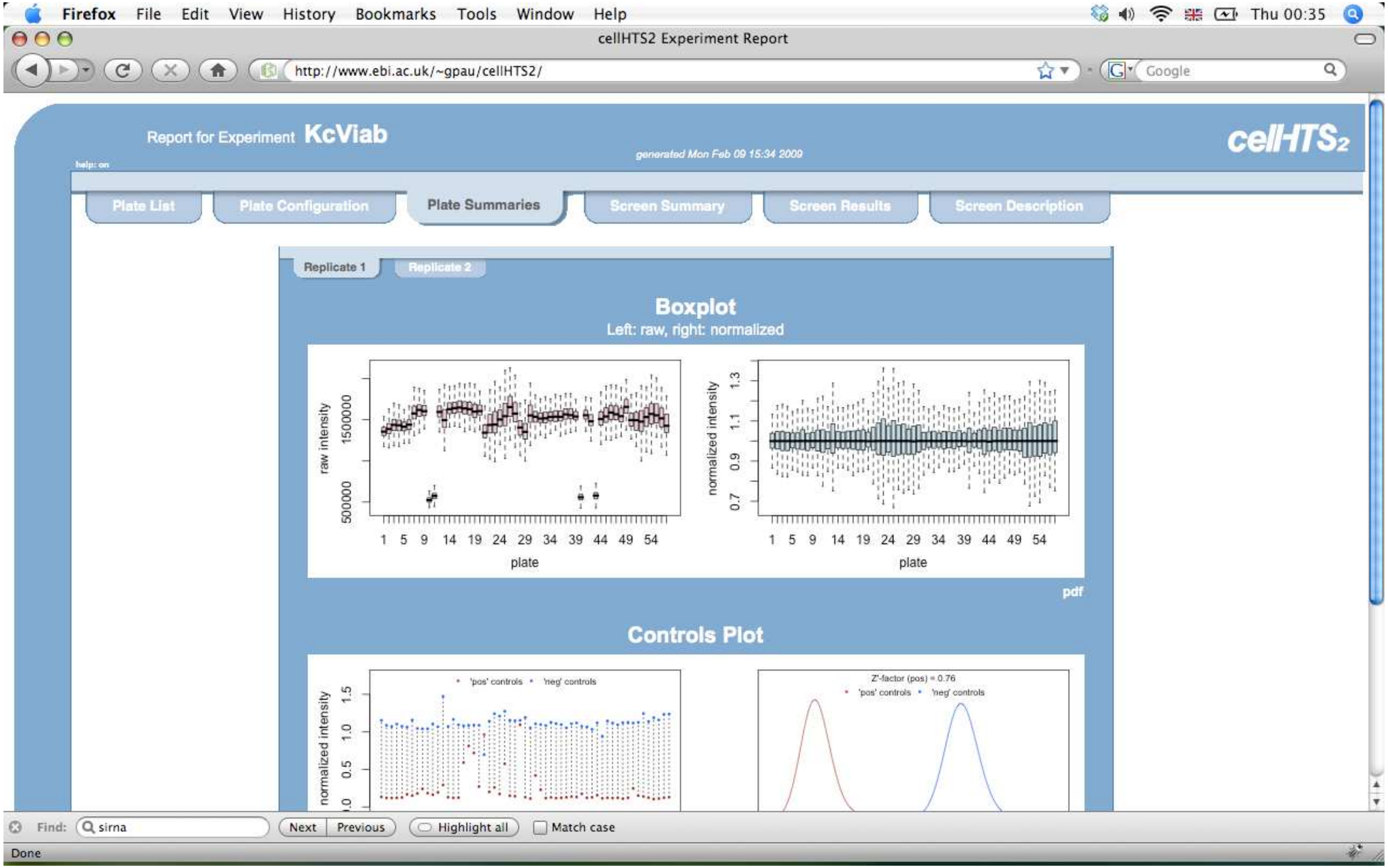


Results

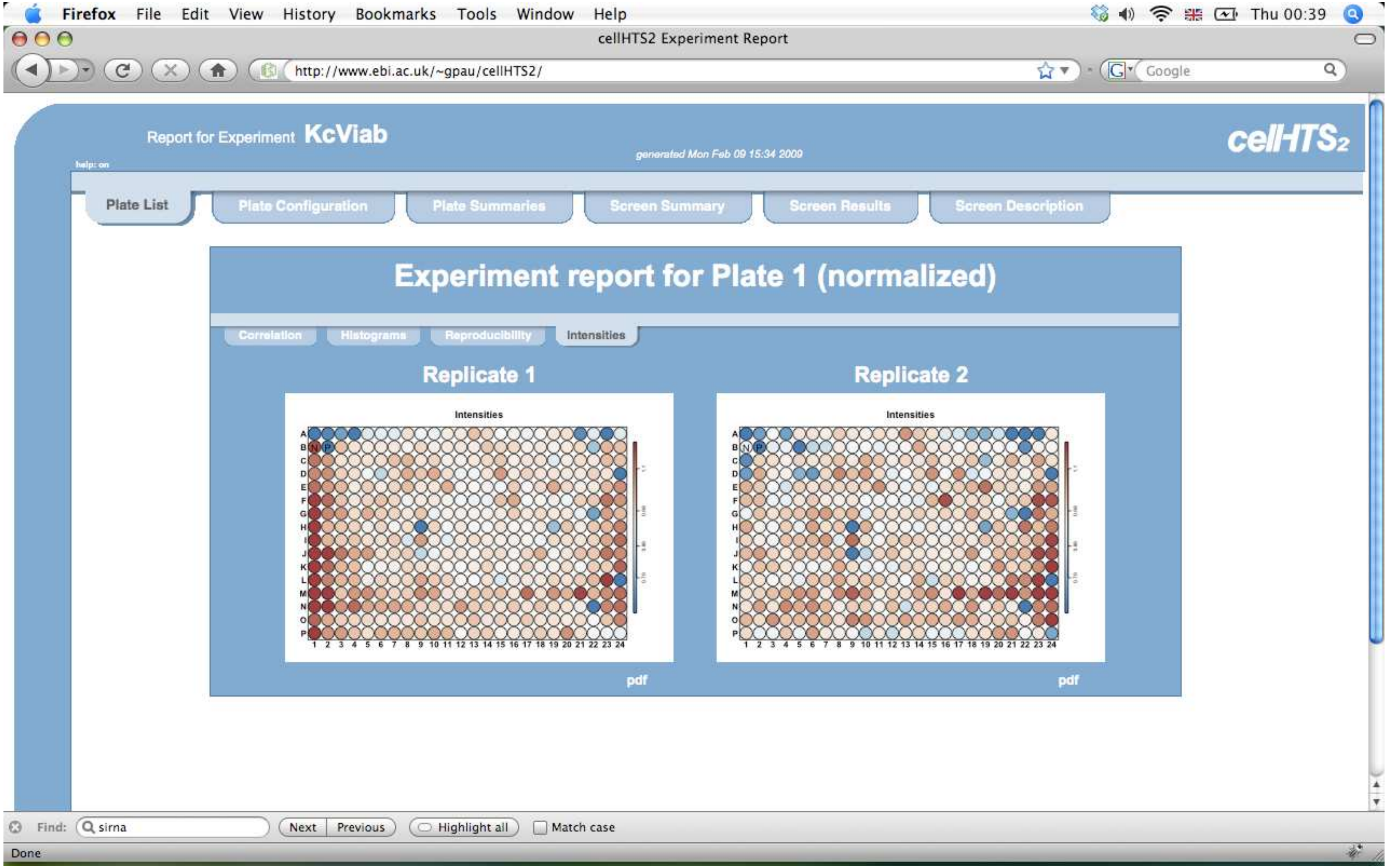
Uname	controlStatus	GeneID	Replicate 01	Replicate 02	Action
002-01-A15	SAMPLE	IKBKAP	seg anno	seg anno	Closest phenotypes
002-01-A17	SAMPLE	IKBKB	seg anno	seg anno	Closest phenotypes
002-01-A19	SAMPLE	IKBKE	seg anno	seg anno	Closest phenotypes
002-01-E17	SAMPLE	JIK	seg anno	seg anno	Closest phenotypes
002-01-E23	SAMPLE	TNIK	seg anno	seg anno	Closest phenotypes

CellHTS2

- Bioconductor package [Florian Hahne et al.]

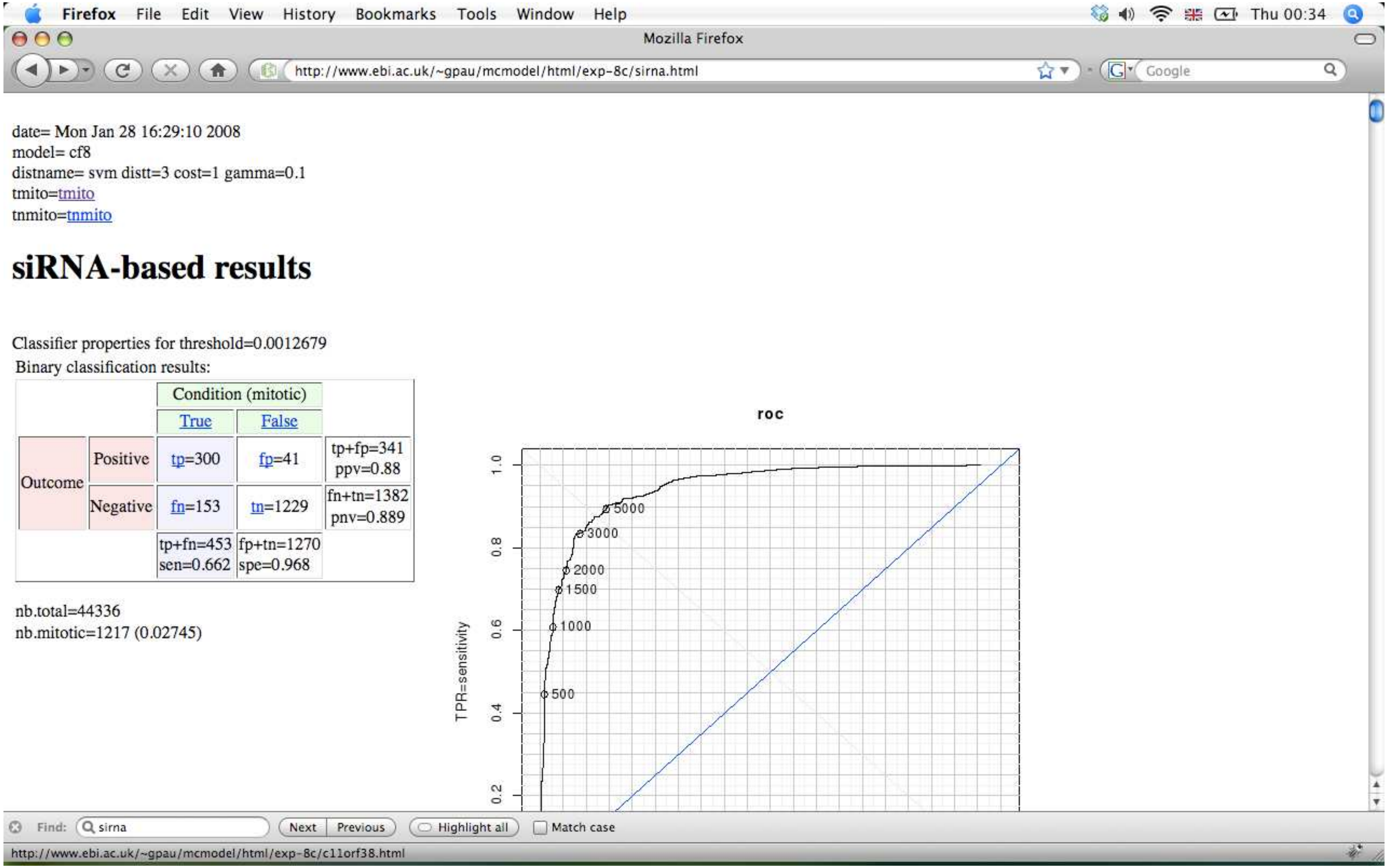


CellHTS2


Firefox File Edit View History Bookmarks Tools Window Help
Thu 00:39
cellHTS2 Experiment Report
http://www.ebi.ac.uk/~gpau/cellHTS2/
Google
Report for Experiment KcViab
help: on
generated Mon Feb 09 15:34 2009
cellHTS2
Plate List
Plate Configuration
Plate Summaries
Screen Summary
Screen Results
Screen Description
Experiment report for Plate 1 (normalized)
Correlation
Histograms
Reproducibility
Intensities
Replicate 1
Replicate 2
Intensities
A B C D E F G H I J K L M N O P
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
N
P
pdf
Find: sirna
Next
Previous
Highlight all
☐ Match case
Done

Mitocheck

- Exhaustive identification of genes involved in cell division



date= Mon Jan 28 16:29:10 2008
model= cf8
distname= svm distt=3 cost=1 gamma=0.1
tmito=tmito
tnmito=tnmito

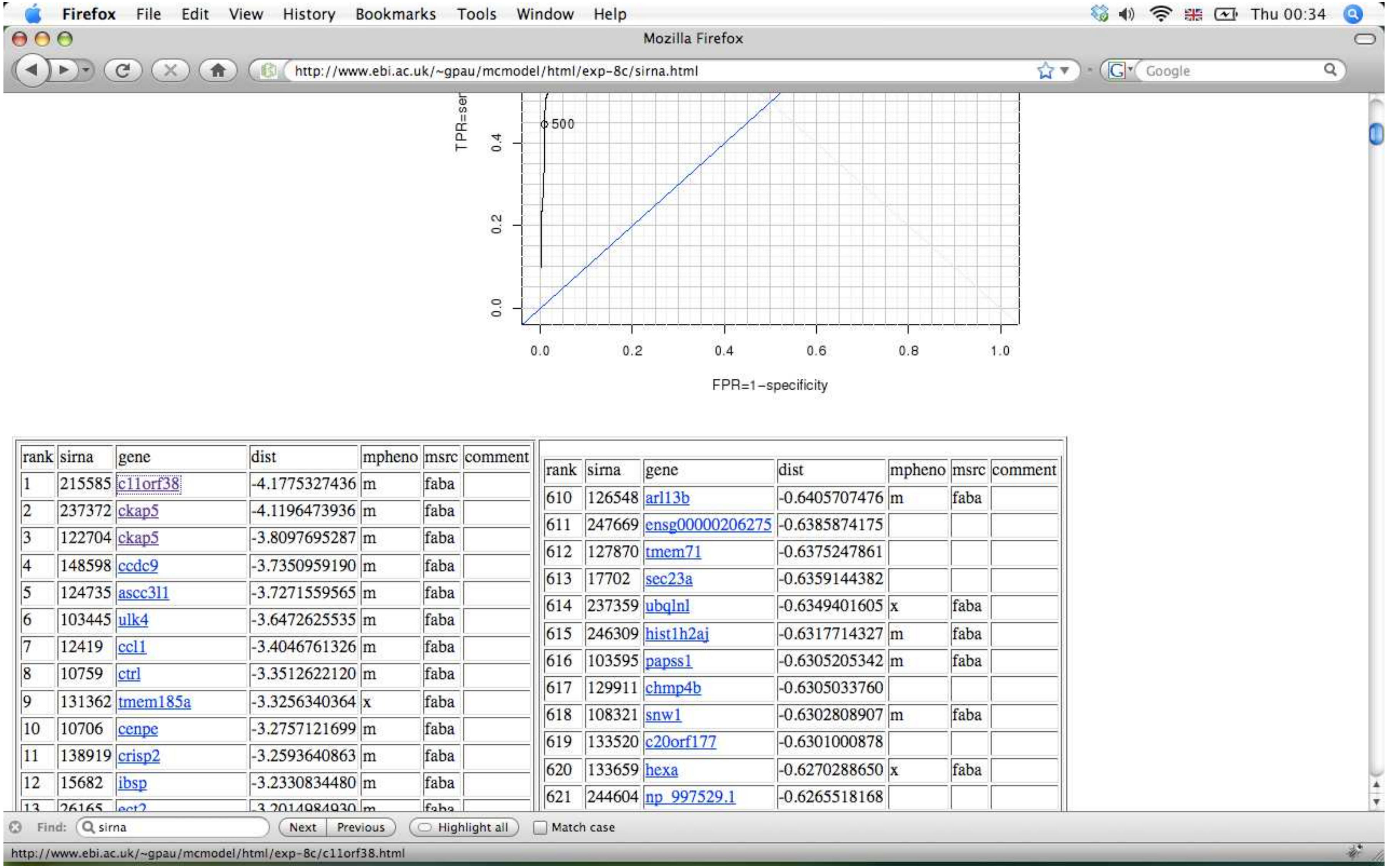
siRNA-based results

Classifier properties for threshold=0.0012679
Binary classification results:

		Condition (mitotic)		
		True	False	
Outcome	Positive	tp=300	fp=41	tp+fp=341 ppv=0.88
	Negative	fn=153	tn=1229	fn+tn=1382 pnv=0.889
		tp+fn=453 sen=0.662	fp+tn=1270 spe=0.968	

nb.total=44336
nb.mitotic=1217 (0.02745)

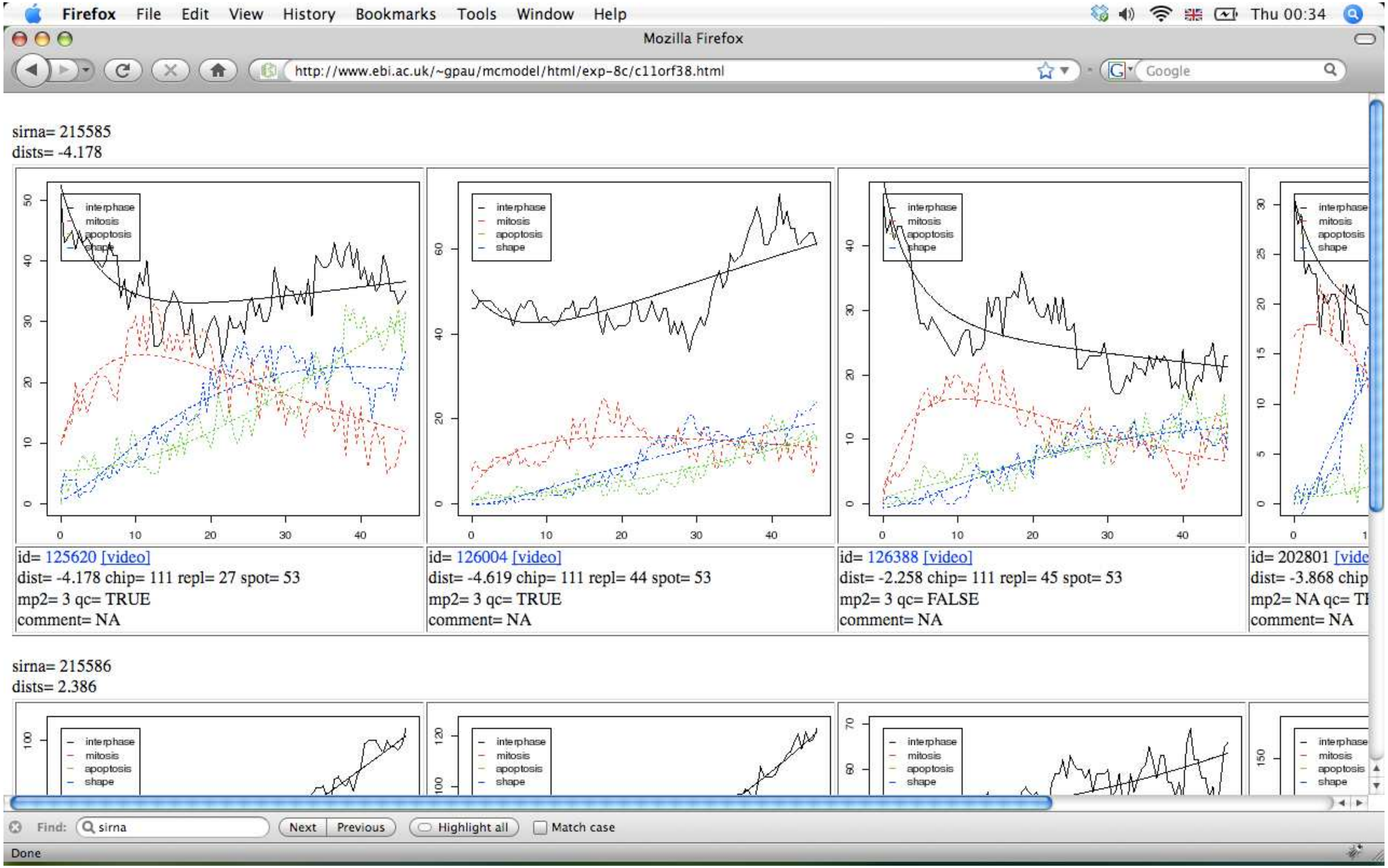
Mitocheck



rank	sirna	gene	dist	mpheno	msrc	comment
1	215585	c11orf38	-4.1775327436	m	faba	
2	237372	ckap5	-4.1196473936	m	faba	
3	122704	ckap5	-3.8097695287	m	faba	
4	148598	ccdc9	-3.7350959190	m	faba	
5	124735	ascc3l1	-3.7271559565	m	faba	
6	103445	ulk4	-3.6472625535	m	faba	
7	12419	ccl1	-3.4046761326	m	faba	
8	10759	ctrl	-3.3512622120	m	faba	
9	131362	tmem185a	-3.3256340364	x	faba	
10	10706	cenpe	-3.2757121699	m	faba	
11	138919	crisp2	-3.2593640863	m	faba	
12	15682	ibsp	-3.2330834480	m	faba	
13	26165	ect2	-3.2014984930	m	faba	

rank	sirna	gene	dist	mpheno	msrc	comment
610	126548	arl13b	-0.6405707476	m	faba	
611	247669	ensg00000206275	-0.6385874175			
612	127870	tmem71	-0.6375247861			
613	17702	sec23a	-0.6359144382			
614	237359	ubqlnl	-0.6349401605	x	faba	
615	246309	hist1h2aj	-0.6317714327	m	faba	
616	103595	papss1	-0.6305205342	m	faba	
617	129911	chmp4b	-0.6305033760			
618	108321	snw1	-0.6302808907	m	faba	
619	133520	c20orf177	-0.6301000878			
620	133659	hexa	-0.6270288650	x	faba	
621	244604	np_997529.1	-0.6265518168			

Mitocheck



sirna= 215585
dists= -4.178

id= 125620 [video] dist= -4.178 chip= 111 repl= 27 spot= 53 mp2= 3 qc= TRUE comment= NA	id= 126004 [video] dist= -4.619 chip= 111 repl= 44 spot= 53 mp2= 3 qc= TRUE comment= NA	id= 126388 [video] dist= -2.258 chip= 111 repl= 45 spot= 53 mp2= 3 qc= FALSE comment= NA	id= 202801 [vide dist= -3.868 chip mp2= NA qc= TI comment= NA

sirna= 215586
dists= 2.386

Conclusion

- Simple, powerful and fast generation of web pages
- Useful to report complex, large and hierarchical information
- Freely available on CRAN
- Used by several packages
  - cellHTS2
  - arrayQualityExpress
  - ShortRead
  - imageHTS (soon available)

- Thanks for your attention !