

An `exprSet` for the Iyer genomic time series database

VJ Carey, stvjc@channing.harvard.edu

October 16, 2006

1 Overview

Iyer, Eisen et al (Science 1999, v283 83-87) report a cDNA-chip based experiment to illustrate the transcriptional response of fibroblasts to serum. The original data are archived in full at genome-www.stanford.edu/serum/data.html. This package provides access to a subset of the data leading to Figure 2 of their paper. It would be worthwhile to provide high-level objects representing the entire dataset, and this will be taken up in the future.

2 The `Iyer517` `exprSet`

To get access to the data, install the `Iyer517` package and then attach it:

```
> library(Iyer517)
```

A summary of the key dataset is:

```
> show(Iyer517)
```

```
ExpressionSet (storageMode: lockedEnvironment)
assayData: 517 features, 19 samples
  element names: exprs
phenoData
  sampleNames: EOHR, E15MIN, ..., EUNSYNC (19 total)
  varLabels and varMetadata:
    time.hrs: time, NA=Unsync
    cycloheximide: cycloheximide present
featureData
  featureNames: W95909, AA045003, ..., W88650 (517 total)
  varLabels and varMetadata: none
experimentData: use 'experimentData(object)'
pubMedIds: 9872747
Annotation [1] ""
```

The first few expression records are:

```
> exprs(Iyer517)[1:4, 1:6]
```

	EOHR	E15MIN	E30MIN	E1HR	E2HR	E4HR
W95909	1	0.72	0.10	0.57	1.08	0.66
AA045003	1	1.58	1.05	1.15	1.22	0.54
AA044605	1	1.10	0.97	1.00	0.90	0.67
W88572	1	0.97	1.00	0.85	0.84	0.72

Note that columns 1 to 13 correspond to sampling times in the absence of cycloheximide (an inhibitor of protein synthesis) and columns 14 to 19 correspond to sampling times in the presence of cycloheximide. The tags UNSYN and UNSYNC are sampling from cells in exponential replication.

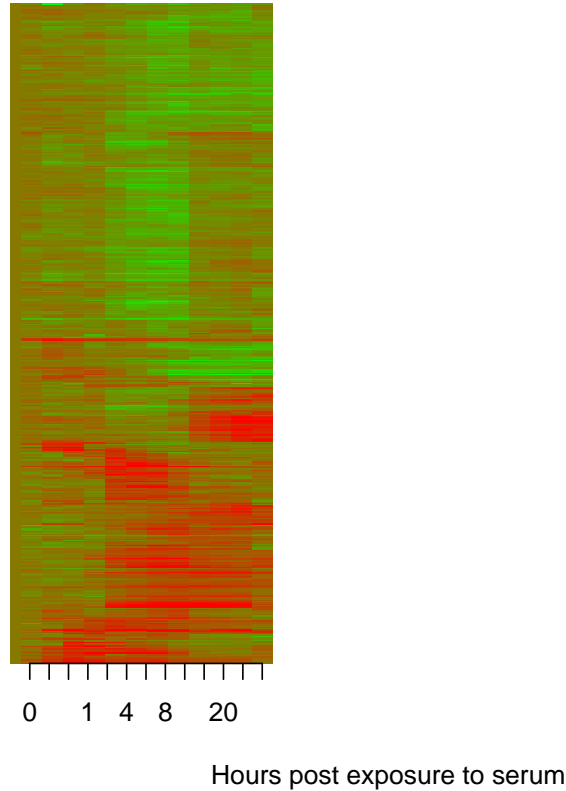
3 Replication of some findings

3.1 Figure 2

To reproduce Figure 2 we need a color scheme and some transformations. The following seems to do a reasonable job:

```
> chg <- seq(0.1, 8, 0.01)
> mycol <- rgb(chg/8, 1 - chg/8, 0)
> CEX <- exprs(Iyer517)
> CEX[CEX > 8] <- 8

> image(t(log10(CEX[517:1, 1:13])), col = mycol, xlim = c(0, 3),
+       axes = FALSE, xlab = "Hours post exposure to serum")
> axis(1, at = (1:13)/13, lab = c("0", ".25", ".5", "1", "2", "4",
+   "6", "8", "12", "16", "20", "24", "u"), cex = 0.3)
```



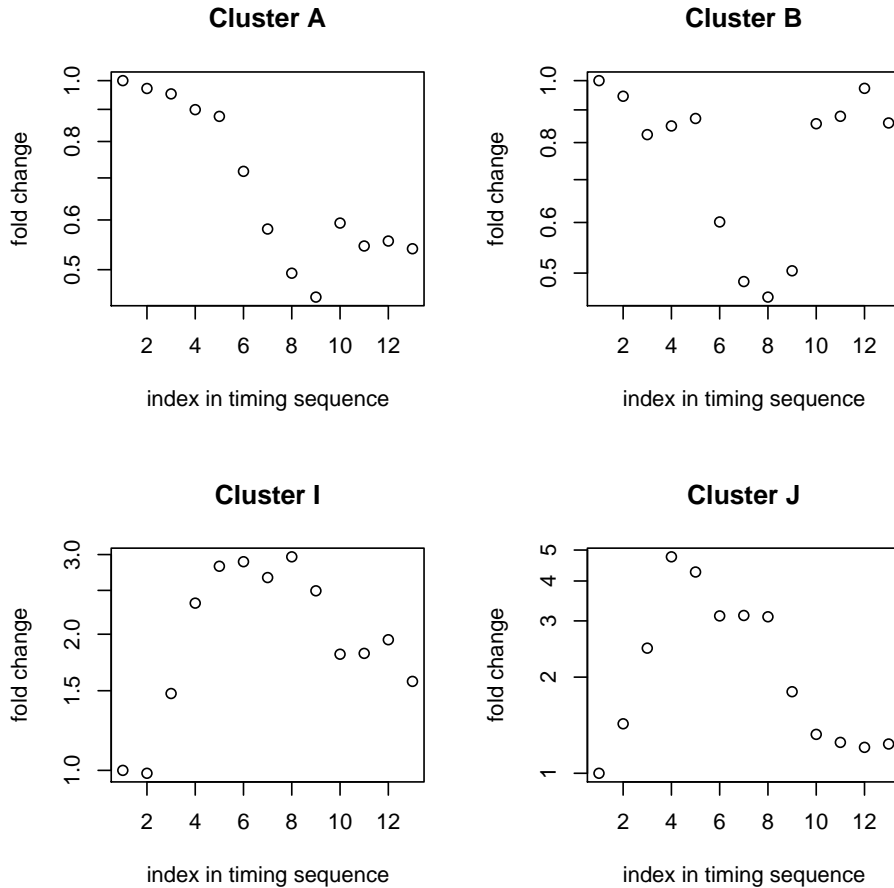
However, the time 0 column of Figure 2 in the paper shows some variability. This is hard to square with the caption indicating that the data depicted are ratios relative to time 0.

3.2 The mean within-cluster trajectories

To orient it seems we need clusters contiguous to the boundaries of the image matrix, because there are gaps of unspecified length between many of the clusters.

```
> par(mfrow = c(2, 2))
> plot(apply((CEX[1:100, 1:13]), 2, mean), main = "Cluster A",
+       log = "y", ylab = "fold change", xlab = "index in timing sequence")
> plot(apply((CEX[101:242, 1:13]), 2, mean), main = "Cluster B",
+       log = "y", ylab = "fold change", xlab = "index in timing sequence")
> plot(apply((CEX[483:499, 1:13]), 2, mean), main = "Cluster I",
+       log = "y", ylab = "fold change", xlab = "index in timing sequence")
```

```
> plot(apply((CEX[500:517, 1:13]), 2, mean), main = "Cluster J",
+       log = "y", ylab = "fold change", xlab = "index in timing sequence")
```



The trajectories are very similar to those reported in the paper.

4 Extended annotation

An effort has been made to incorporate GO tags into this data resource.

```
> data(IyerAnnotated)
> print(IyerAnnotated[1:5, ])
```

Iclust	GB	seqno	locusid	G01	G02	G03	G04	G05
1	N	W95909	1 80298	<NA>	<NA>	<NA>	<NA>	<NA>
2	A	AA045003	2 6414	GO:0008430	<NA>	<NA>	<NA>	<NA>
3	A	AA044605	3 5300	GO:0003755	GO:0005515	<NA>	<NA>	<NA>
4	A	W88572	4 2037	GO:0005200	<NA>	<NA>	<NA>	<NA>
5	A	AA029909	5 51747	<NA>	<NA>	<NA>	<NA>	<NA>

At the time of construction, at most 5 GO tags had been associated with any probes in the dataset, and a large number of probes lacked both Locus Link and GO tags.