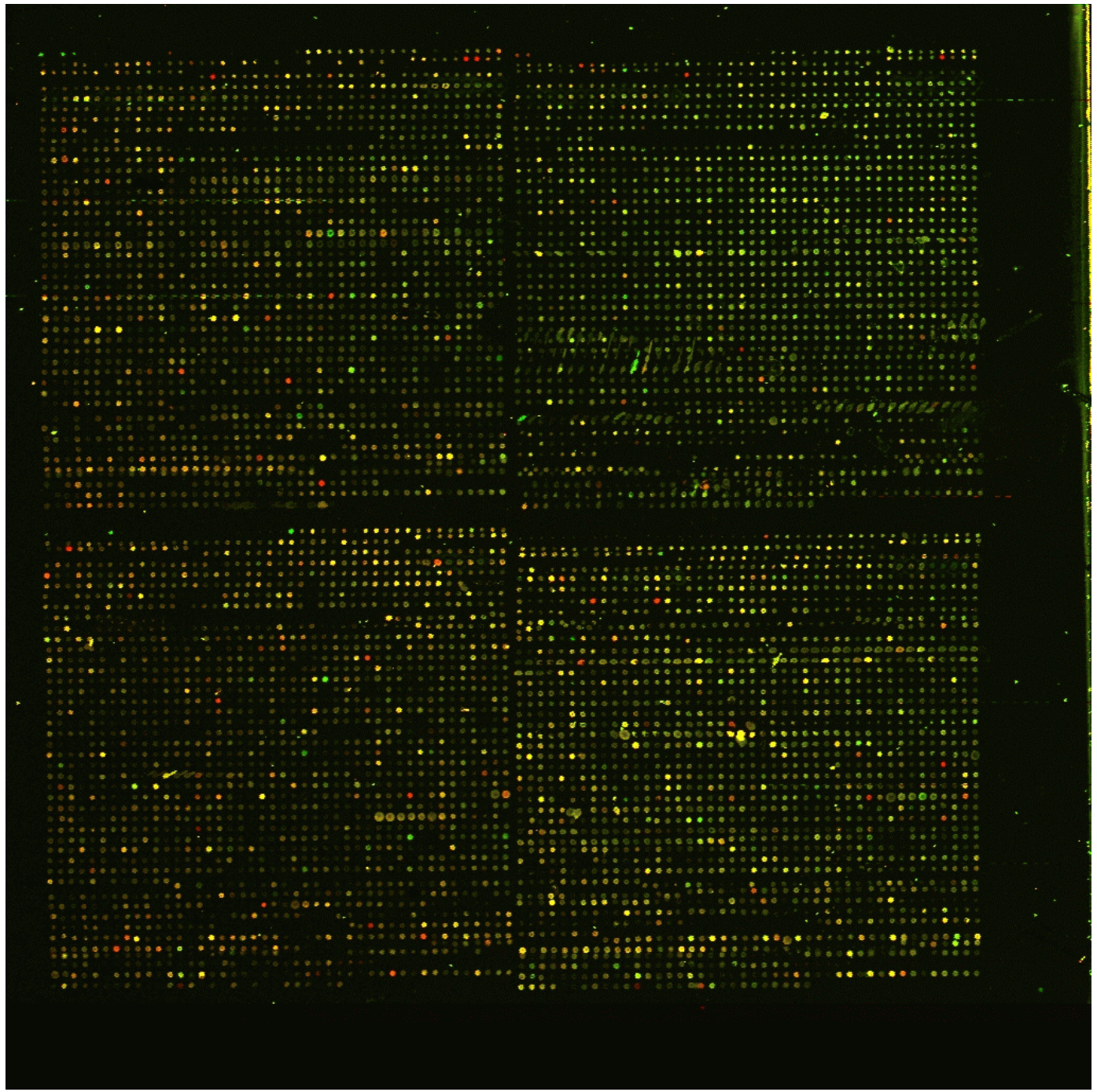


COMP364: Gene Expression Microarray Data Visualization

Jérôme Waldispühl, McGill University



Yeast Cell Cycle Data

Download the gene expression file at:

<http://genome-www.stanford.edu/cellcycle/data/rawdata/combined.txt>

It is a tab-delimited data-file.

Each column is an experiment and reports the expression level of the gene at different time points.

The first row stores the gene names.

Exercise 1: Use a csv reader to read and store the file content.

Exercise 2: Print the file content as a heat-map.

CSV reader

```
import csv
```

```
fh = open(<filename>,'r')
```

```
reader = csv.reader(fh,delimiter='\t')
```

“reader” is an iterable that allows you to read your file line
by line. For instance you can count the number of line:

```
rowcmpt=0
```

```
for line in reader:
```

```
    rowcmpt +=1
```

CSV reader

To read one single line, you can use `.next()`.
The following read the first line (the header) and move the
iterable to the next line.

```
fh = open(<filename>,'r')  
reader = csv.reader(fh,delimiter='\t')  
expListIDs = reader.next()
```

```
# print its content
```

```
print expListIDs
```

Question: How many experiments does this file contain?

Exercises

Exercise 1: Use a csv reader to read and store the file content.

Exercise 2: Print the file content as a heat-map.

Tips:

1. Read the file and compute the number of row & column first.
2. Create a matrix (use numpy!) with the appropriate size.
3. Re-open the file and fill the matrix.