

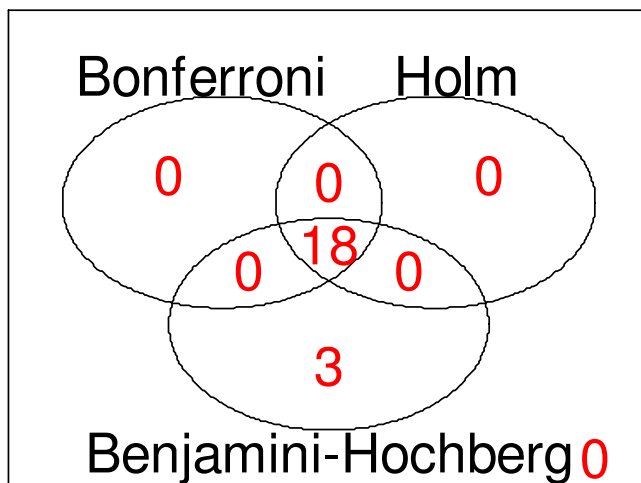
Voorbeeld tentamen Statistiek 3 BIN (2013)

Opgave 1. Kiezen van een statistische toets voor DEG's

- a. [10 pt] Juiste analyses: **1-sample t-toets** ($\mu = 0$) EN/OF **Wilcoxon's signed-rank toets** ($\mu = 0$) want: van onafhankelijke 6 replica's checken of afwijkt van $M = 0$ (geen up/down regulatie).
- b. [10 pt] Juiste analyses: **1-way ANOVA** EN/OF **Kruskal-Wallis toets** want: 3 groepen met verschillende aantallen replica's, dus M berekening kan niet...
- c. [10 pt] Juiste analyses: **gepaarde t-toets** EN/OF **Wilcoxon's signed rank toets** want: 3 gepaarde R/G metingen (per array 1).
- d. [10 pt] p-waarden: 0.0002816505 0.0004379076, dus beiden DEG ($p < 0.05$)!

Opgave 2. Bloomococcus en antibiotica

- a. [10 pt] Juiste parametrische analyse: **1-way ANOVA** want: 3 onafhankelijke sets met waarden (allen van verschillende arrays), gegeven dat varianties gelijk zijn.
- b. [10 pt] Verwacht aantal FP = $\alpha * N = 0.05 * 2000 = 100$ genen
- c. [10 pt] $N = 113$ DEG's
- d. [20 pt] **Bonferroni**: $N = 18$ DEG's,
Genen: "Gene.1", "Gene.18", "Gene.35", "Gene.52", "Gene.53", "Gene.54", "Gene.131",
"Gene.245", "Gene.255", "Gene.389", "Gene.644", "Gene.707", "Gene.777", "Gene.987",
"Gene.1007", "Gene.1009", "Gene.1515", "Gene.1616".
Holm: $N = 18$ DEG's,
Genen: "Gene.1", "Gene.18", "Gene.35", "Gene.52", "Gene.53", "Gene.54", "Gene.131",
"Gene.245", "Gene.255", "Gene.389", "Gene.644", "Gene.707", "Gene.777", "Gene.987",
"Gene.1007", "Gene.1009", "Gene.1515", "Gene.1616".
Benjamini-Hochberg: $N = 21$ DEG's,
Genen: "Gene.1", "Gene.18", "Gene.35", "Gene.52", "Gene.53", "Gene.54", "Gene.131",
"Gene.245", "Gene.255", "Gene.389", "Gene.401", "Gene.644", "Gene.707", "Gene.777",
"Gene.801", "Gene.987", "Gene.1001", "Gene.1007", "Gene.1009", "Gene.1515".
- e. [10 pt] Venn diagram met library "limma":



- f. [10 pt] Kruistabel voor Gene Enrichment Analysis:

	FUNCTION	
Holm	yes	no
yes	1	17
no	150	1832

Fisher test:

Fisher's Exact Test for Count Data

data: XTab

p-value = 1

alternative hypothesis: true odds ratio is not equal to 1

95 percent confidence interval:

0.01707826 4.64824134

sample estimates:

odds ratio

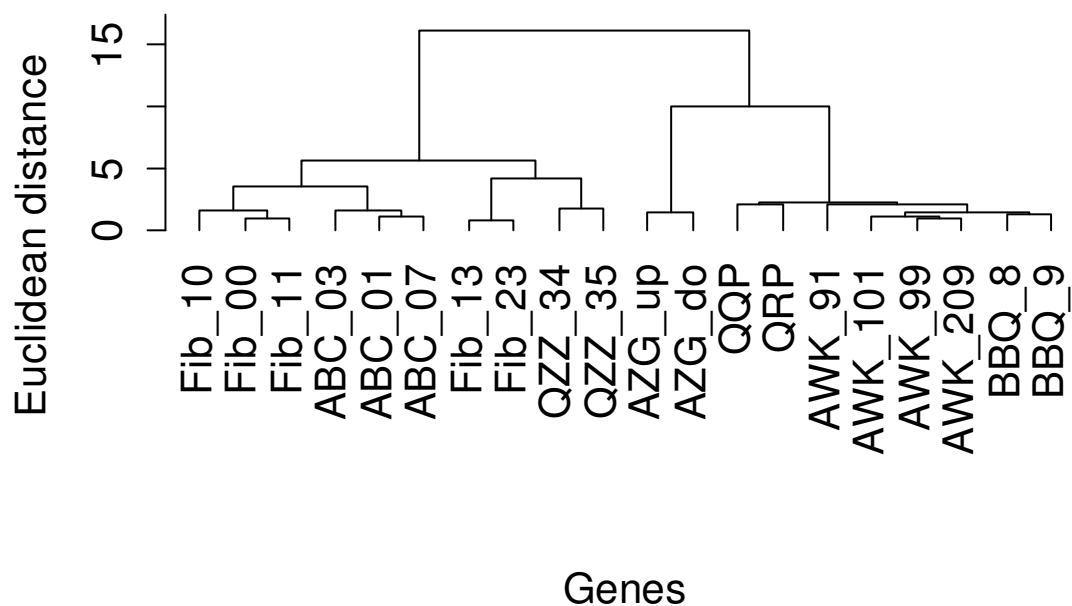
0.7185117

p-waarde > 0.05, dus geen significante enrichment/depletion!

Opgave 3. Ziekte van Bloom

- a. [10 pt]

Hierarchical clustering (Euclidean/Average linkage)



- b. [5 pt]

Cluster 1: "ABC_01" "ABC_03" "ABC_07" "Fib_00" "Fib_10" "Fib_11"

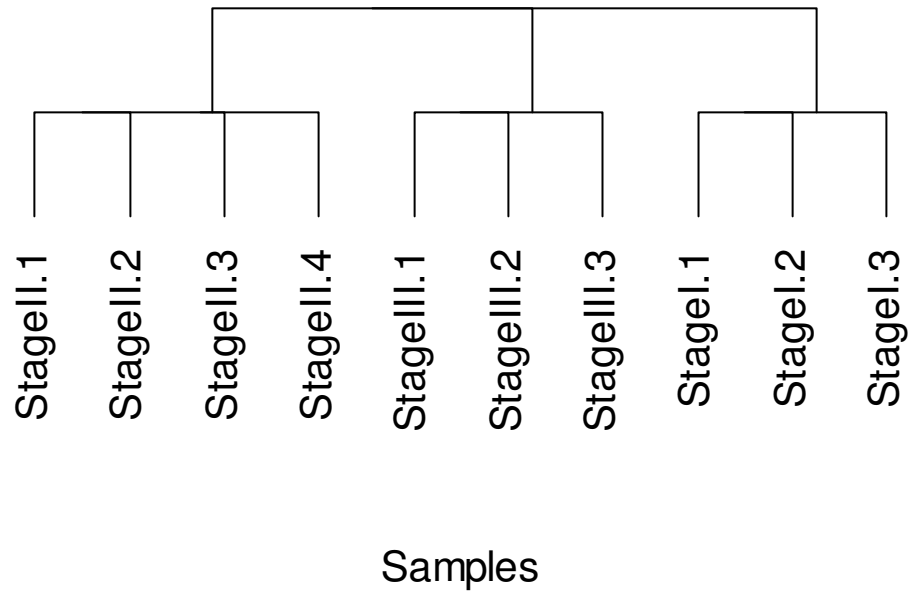
Cluster 2: "QZZ_34" "QZZ_35" "Fib_13" "Fib_23"

Cluster 3: "AZG_up" "AZG_do"

Cluster 4: "QQP" "QRP" "BBQ_8" "BBQ_9" "AWK_91" "AWK_99" "AWK_101" "AWK_209"

C.

K-means clustering (k = 3)



d.

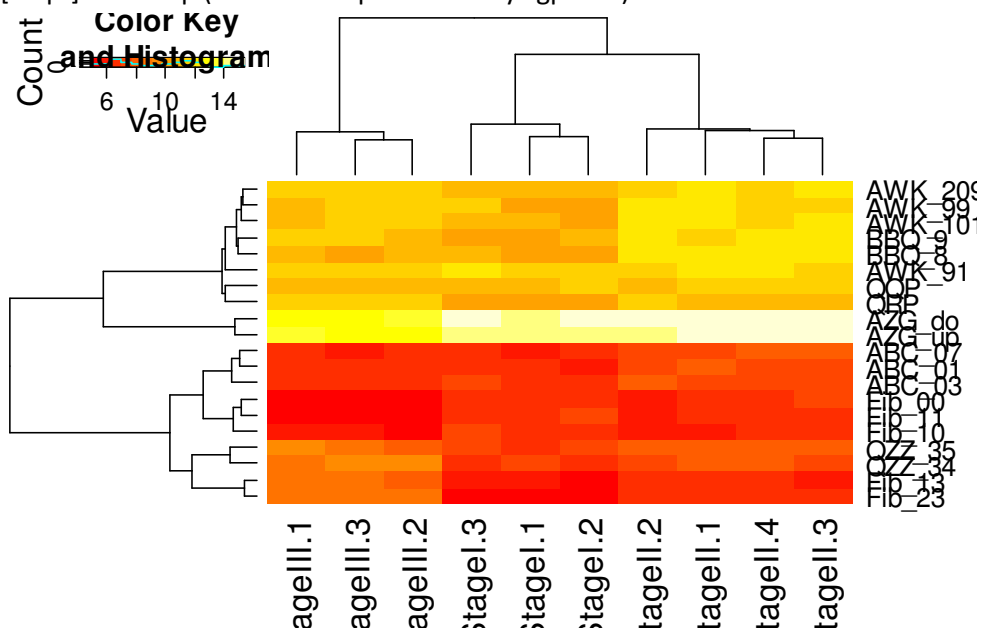
Cluster 1: "Stagell.1" "Stagell.2" "Stagell.3" "Stagell.4"

Cluster 2: "StageIII.1" "StageIII.2" "StageIII.3"

Cluster 3: "Stagel.1" "Stagel.2" "Stagel.3"

Dit klopt met de biologische indeling!

e.



Opgave 4. Gene Enrichment Analysis

a. [10 pt] Fisher test:

```
Fisher's Exact Test for Count Data
```

```
data: X.a
```

```
p-value = 0.03184
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
1.040642 3.524649
```

```
sample estimates:
```

```
odds ratio
```

```
1.915255
```

p-waarde < 0.05, en OR > 1, dus een significante enrichment!

b. [10 pt] Fisher test:

```
Fisher's Exact Test for Count Data
```

```
data: X.b
```

```
p-value = 0.001232
```

```
alternative hypothesis: true odds ratio is not equal to 1
```

```
95 percent confidence interval:
```

```
0.1930969 0.6741107
```

```
sample estimates:
```

```
odds ratio
```

```
0.3567828
```

p-waarde < 0.05, en OR < 1, dus een significante depletion!