```
[1] "# of discriminatory genes = 99"

[1] "fold change for discriminatory genes: 1"

[1] "fold change for nondiscriminatory genes: 0.5"

choose_k
  2
100
```

The optimal number of clusters K is set as the most frequently found K from the 100 simulations tabulated above. It is found by using the BIC criterion after running the unpenalized EM algorithm on K spanning from 2 to 8.

Using the last set of simulated counts, I ran a grid search across varying tuning parameters. As done in Pan et al, I fixed $\lambda_1 = 1$, and searched over $\lambda_2 = (0.1, 0.2, ..., .2)$ and $\tau = (0.1, 0.2, ..., 2)$:

```
[1] "lambda1, lambda2, tau, BIC:"

        [,1] [,2] [,3]      [,4]
 [1,]    1  0.1  0.1 18579.56
 [2,]    1  0.1  0.2 18579.56
 [3,]    1  0.1  0.3 18579.56
 [4,]    1  0.1  0.4 18579.56
 [5,]    1  0.1  0.5 18579.56
 [6,]    1  0.1  0.6 18579.56
 [7,]    1  0.1  0.7 18579.56
 [8,]    1  0.2  0.1 18579.56
 [9,]    1  0.2  0.2 18579.56
[10,]    1  0.2  0.3 18579.56
[11,]    1  0.2  0.4 18579.56
[12,]    1  0.2  0.5 18579.56
[13,]    1  0.2  0.6 18579.56
[14,]    1  0.2  0.7 18579.56
[15,]    1  0.3  0.1 18579.56
[16,]    1  0.3  0.2 18579.56
[17,]    1  0.3  0.3 18579.56
[18,]    1  0.4  0.1 18579.56
[19,]    1  0.4  0.2 18579.56
[20,]    1  0.4  0.5 18579.56
[21,]    1  0.5  0.1 18579.56
[22,]    1  0.6  0.2 18579.56
```

The results of the final run based on optimal tuning parameters are below:
Below are the summary of results:

```
[1] "Mean pi: 0.520869565217391" "Mean pi: 0.479130434782609"
```

```
[1] "First 3 genes:"

         [,1]     [,2]
[1,] 3.365798 4.385404
[2,] 4.583266 5.613659
[3,] 4.099911 5.121616

[1] "Last 3 genes:"

           [,1]     [,2]
 [98,] 5.879908 6.920768
 [99,] 5.437310 6.470980
[100,] 6.119160 6.639156

[1] "Mean % of nondiscriminatory genes: 0"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 1"

[1] "Mean false positive rate: 1"
```