

```

[1] "# of discriminatory genes = 75"

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

[1] "fold change for nondiscriminatory genes: 0.1"

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 $\lambda_2 = (0.1, 0.2, \dots, 2)$ and $\tau = (0.1, 0.2, \dots, 2)$:

```

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

[1]      1.00      0.20      0.30 19289.13

```

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

```

[1] "Mean pi: 0.610869565217391" "Mean pi: 0.389130434782609"

[1] "First 3 genes:"

      [,1]      [,2]
[1,] 3.371353 4.352187
[2,] 4.597608 5.593849
[3,] 4.107702 5.101011

[1] "Last 3 genes:"

      [,1]      [,2]
[98,] 6.317181 6.485090
[99,] 5.870248 6.036294
[100,] 6.295512 6.466286

[1] "Mean % of nondiscriminatory genes: 0.2418"

[1] "Mean ARI: 1"

[1] "Mean sensitivity: 0.999733333333333"

[1] "Mean false positive rate: 0.0336"

```