egfr_luise_TCGA_validation

Mumtahena Rahman

December 22, 2014

Now checking the PCA plot for batch adjustment

Now running ASSIGN with 500 genes

Now let's correlate with drug response

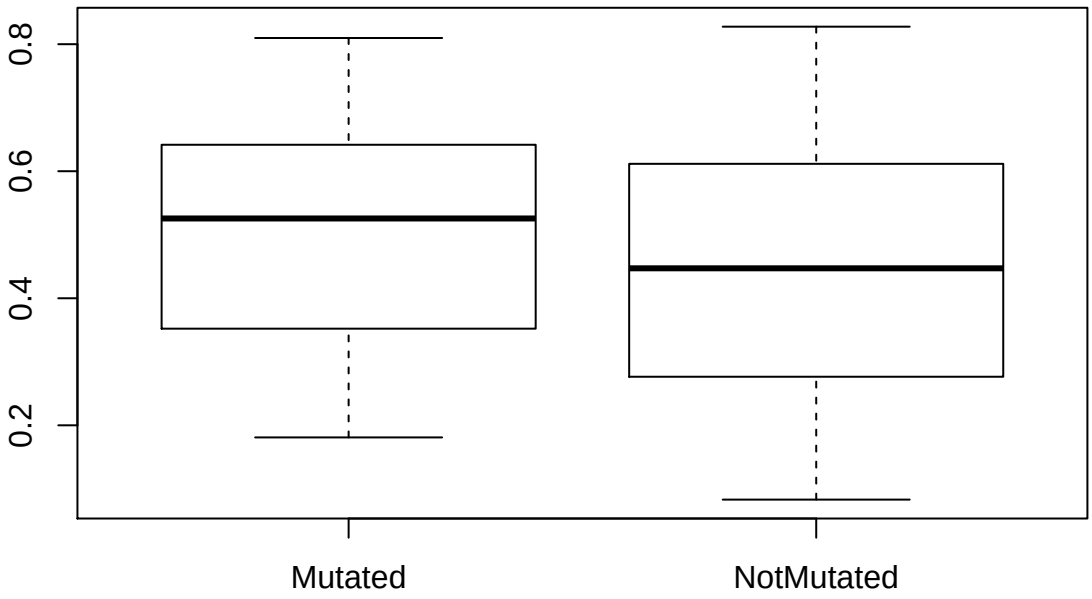
```
## [1] 170  10
```

Staring to validate in TCGA GBM cancer type. Not much luck with the mutation only data. However, mutation and amplification data together was robust enough for EGFR signature validation. I think EGFR signature consisting of 15 genes are better predictor both in ICBP and TCGA GBM datasets.

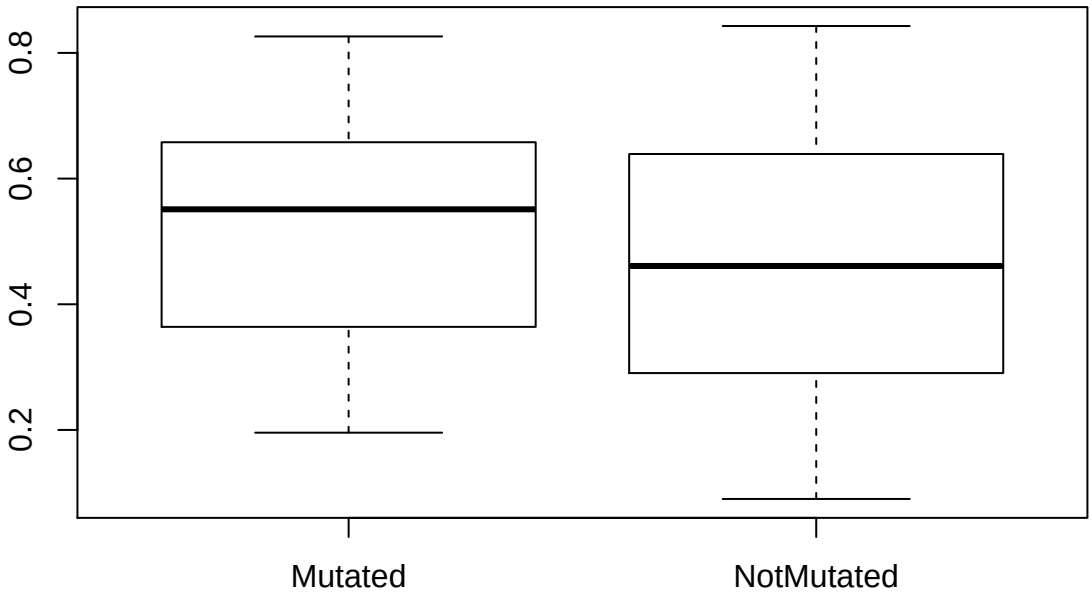
```
## [1] 3
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 0.7581, df = 136.6, p-value = 0.4497
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03736468  0.08382662
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.4844259                0.4611949
##
## [1] 4
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 0.6075, df = 110.267, p-value = 0.5448
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03513116  0.06619223
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.4522139                0.4366834
##
## [1] 5
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 1.4646, df = 136.131, p-value = 0.1453
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.0156940  0.1053148
```

```
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.4862884                0.4414780
##
## [1] 6
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 0.6665, df = 138.605, p-value = 0.5062
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03813491  0.07691406
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.4582057                0.4388162
##
## [1] 7
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 1.8119, df = 114.088, p-value = 0.07264
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.005360662  0.120219591
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.5047095                0.4472800
##
## [1] 8
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 0.8865, df = 135.855, p-value = 0.3769
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03374054  0.08857380
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.5078052                0.4803886
##
## [1] 9
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 0.6115, df = 111.623, p-value = 0.5421
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03537406  0.06695723
## sample estimates:
##    mean in group Mutated mean in group NotMutated
```

```
##               0.5061157               0.4903241
##
## [1] 10
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 1.557, df = 137.879, p-value = 0.1218
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.01295791  0.10896277
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##               0.5227727               0.4747703
##
## [1] 11
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 0.8233, df = 139.346, p-value = 0.4117
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.03404285  0.08262502
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##               0.5078503               0.4835592
##
## [1] 12
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, i] by V2
## t = 1.8615, df = 114.237, p-value = 0.06525
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.003819262  0.122839495
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##               0.5245672               0.4650571
```




```
##
##  Welch Two Sample t-test
##
## data:  combined_egfr_status[, 7] by V2
## t = 1.8119, df = 114.088, p-value = 0.07264
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.005360662  0.120219591
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.5047095                0.4472800
```




```
##
##  Welch Two Sample t-test
##
```

```
## data:  combined_egfr_status[, 12] by V2
## t = 1.8615, df = 114.237, p-value = 0.06525
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.003819262  0.122839495
## sample estimates:
##    mean in group Mutated mean in group NotMutated
##                0.5245672                0.4650571

## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap10/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -2.2462, df = 146.448, p-value = 0.02619
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.131940107 -0.008433454
## sample estimates:
## mean in group 0 mean in group 1
##       0.4329265       0.5031133
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap100/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -1.9582, df = 151.605, p-value = 0.05204
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.0952743892  0.0004246026
## sample estimates:
## mean in group 0 mean in group 1
##       0.4250030       0.4724279
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap15/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -3.2945, df = 155.006, p-value = 0.001222
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.16237912 -0.04064512
## sample estimates:
## mean in group 0 mean in group 1
##       0.4062605       0.5077726
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap20/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -1.8959, df = 150.249, p-value = 0.0599
```

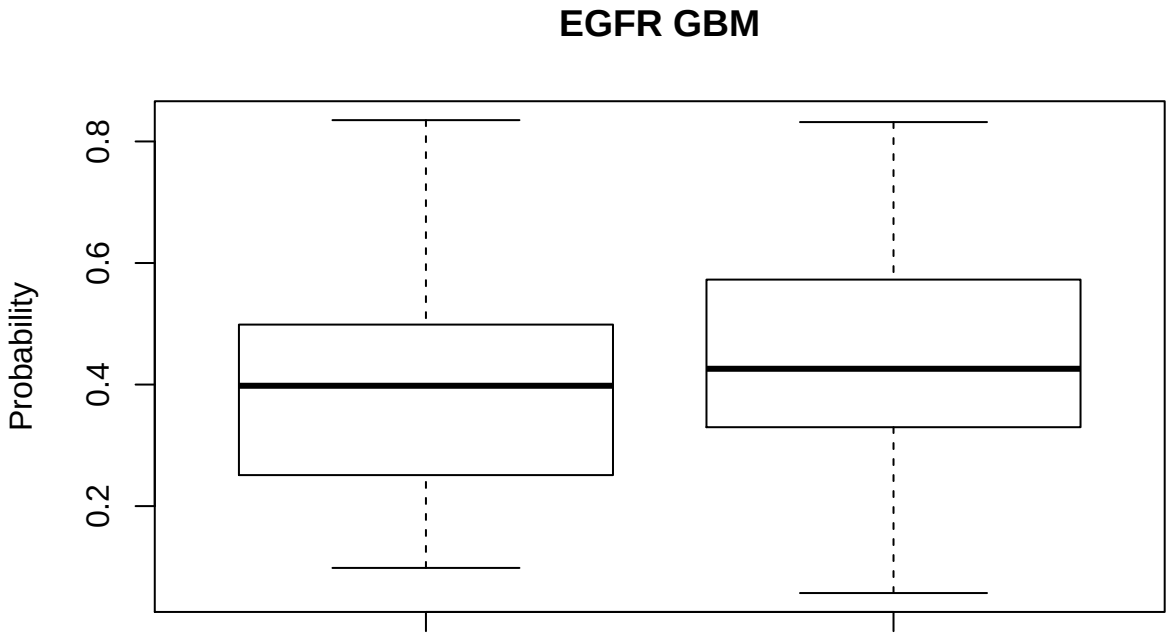
```
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.116097207  0.002399667
## sample estimates:
## mean in group 0 mean in group 1
##       0.4173597       0.4742085
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap5/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -3.8072, df = 153.775, p-value = 0.0002026
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.17200348 -0.05448241
## sample estimates:
## mean in group 0 mean in group 1
##       0.4046440       0.5178869
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap10/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -2.4075, df = 145.78, p-value = 0.01731
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.13768790 -0.01353991
## sample estimates:
## mean in group 0 mean in group 1
##       0.4504716       0.5260855
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap100/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -1.9684, df = 150.36, p-value = 0.05086
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.0969898712  0.0001842127
## sample estimates:
## mean in group 0 mean in group 1
##       0.4782827       0.5266856
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap15/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -3.3769, df = 153.503, p-value = 0.0009288
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
```

```
##  -0.16692986 -0.04370608
## sample estimates:
## mean in group 0 mean in group 1
##       0.4388074       0.5441254
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap20/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -1.9909, df = 147.645, p-value = 0.04833
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.1210593120 -0.0004509956
## sample estimates:
## mean in group 0 mean in group 1
##       0.4618397       0.5225948
##
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap5/pathway_activity_testset.csv"
##
##  Welch Two Sample t-test
##
## data:  GBMmergeMutation_AmpsALL[, i] by GBMmergeMutation_AmpsALL[, 12]
## t = -3.8632, df = 153.05, p-value = 0.0001649
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.17543422 -0.05671539
## sample estimates:
## mean in group 0 mean in group 1
##       0.4218723       0.5379471
```

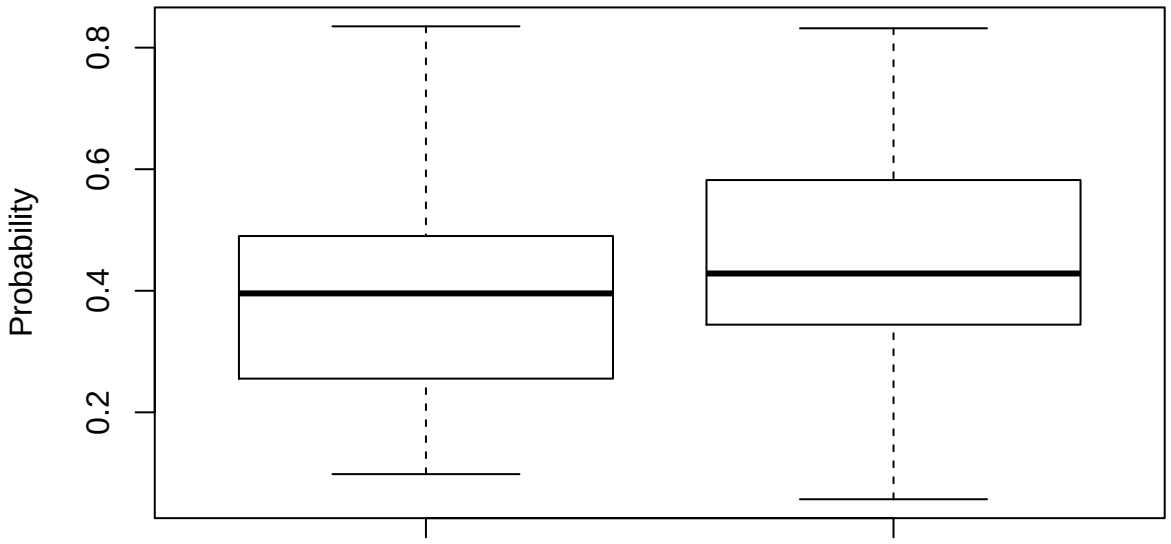
Now finally let's validate EGFR signature with proteomic data. 15 gene signature is still the best...

```
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap10/pathway_activity_testset.csv"
## [1] 0.03477265
## [1] 0.05688628
## [1] 0.0481631
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap100/pathway_activity_testset.csv"
## [1] 0.07063226
## [1] 0.08231324
## [1] 0.06253627
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap15/pathway_activity_testset.csv"
## [1] 0.1621338
## [1] 0.2207077
## [1] 0.2066836
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap20/pathway_activity_testset.csv"
## [1] -0.003783769
## [1] 0.02583441
## [1] 0.02636722
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/adap5/pathway_activity_testset.csv"
## [1] 0.1104467
## [1] 0.1505573
## [1] 0.1428622
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap10/pathway_activity_testset.csv"
```

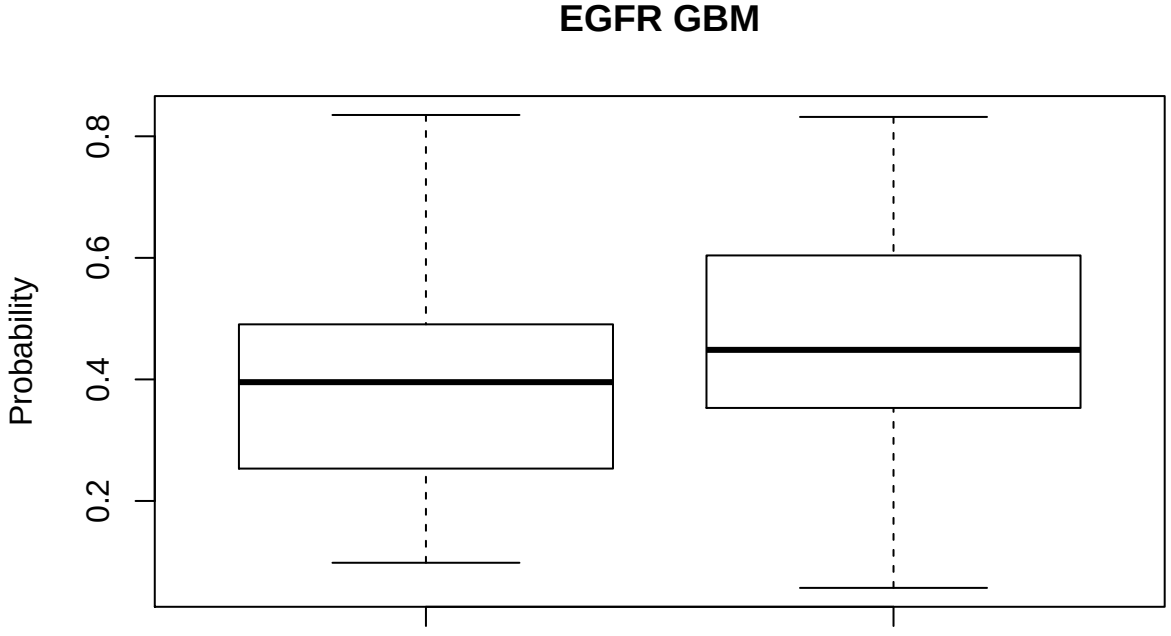
```
## [1] 0.03868406
## [1] 0.05990665
## [1] 0.05062194
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap100/pathway_activity_testset.csv"
## [1] 0.07225069
## [1] 0.0844144
## [1] 0.06456557
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap15/pathway_activity_testset.csv"
## [1] 0.1681949
## [1] 0.2266978
## [1] 0.2152353
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap20/pathway_activity_testset.csv"
## [1] -0.002257334
## [1] 0.02534639
## [1] 0.02775566
## [1] "/Users/mumtahenarahman/Desktop/tmp/egfr/tcga/nonadap5/pathway_activity_testset.csv"
## [1] 0.1143849
## [1] 0.1569323
## [1] 0.1486903
```




```
##
##  Welch Two Sample t-test
##
## data:  mergeProteinAll[mergeProteinAll$EGFR < 0, 4] and mergeProteinAll[mergeProteinAll$EGFR > 0, 4]
## t = -1.0807, df = 59.226, p-value = 0.2842
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.14484737  0.04325188
## sample estimates:
## mean of x mean of y
## 0.4072314 0.4580292
```




```
##
##  Welch Two Sample t-test
##
## data:  mergeProteinAll[mergeProteinAll$EGFR_pY1068 < 0, 4] and mergeProteinAll[mergeProteinAll$EGFR_
## t = -1.3482, df = 67.744, p-value = 0.1821
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.15278745  0.02958027
## sample estimates:
## mean of x mean of y
## 0.4034948 0.4650984
```




```
##
```

```
##  Welch Two Sample t-test
##
## data:  mergeProteinAll[mergeProteinAll$EGFR_pY1173 < 0, 4] and mergeProteinAll[mergeProteinAll$EGFR_
## t = -1.7565, df = 69.996, p-value = 0.08337
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.16912643  0.01072759
## sample estimates:
## mean of x mean of y
## 0.3972638 0.4764632
```

This analysis was run on Tue Dec 23 00:58:27 2014