**Association Between mRNA Expression and Copy-Number Alterations for FAT1 and EGFR among Three Cancer Types**

By

Jianwei Chen

Date: 12/29

**Statistical Methods**

Pearson and Spearman correlation coefficients were calculated for the association between copy-number alterations and mRNA expression and tested with Wald’s test for each of the two genes (FAT1 and EGFR) among Head and Neck Squamous Cell Carcinoma (HNSCC), Lung Squamous Cell Carcinoma (Lung SCC), and Cervical Squamous Cell Carcinoma (Cervical SCC), respectively. Copy-number was treated as numeric when calculated for Pearson and Spearman correlation coefficients.

General Linear Model (GLM) was performed to compare the means of mRNA and log2(mRNA) expression between all the copy-number alteration categories which are treated as categorical. Box plots for the distribution of mRNA expression and log2(mRNA) for each level of copy-number alteration categories were presented. F-test was used for the overall comparison. The significance level was set at 0.05. SAS (V9.4) software was used for data management and analysis.

**Results**

1. **HNSCC**

1.1 FAT1 of HNSCC

Table 1.1.1 Pearson correlation

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| FAT1 | 0.28023 | <0.0001 |

Table 1.1.2 Spearman correlation

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| FAT1 | 0.25439 | <0.0001 |

Table 1.1.3 Frequency of FAT1 copy number

| FAT1CNA | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| --- | --- | --- | --- | --- |
| -2 | 34 | 6.85 | 34 | 6.85 |
| -1 | 171 | 34.48 | 205 | 41.33 |
| 0 | 252 | 50.81 | 457 | 92.14 |
| 1 | 36 | 7.26 | 493 | 99.40 |
| 2 | 3 | 0.60 | 496 | 100.00 |

Figure 1.1.1 Boxplot of mRNA in each of level of copy number for FAT1 gene

Figure 1.1.2 Boxplot of log2(mRNA) in each of level of copy number for FAT1 gene 

1.2 EGFR of HNSCC

Table 1.2.1 Pearson correlation coefficient of EGFR

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| EGFR | 0.43751 | 0.0001 |

Table 1.2.2 Spearman correlation coefficient of EGFR

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| EGFR | 0.41173 | <0.0001 |

Table 1.2.3 Frequency of EGFR copy number

| EGFRCNA | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| --- | --- | --- | --- | --- |
| -1 | 20 | 4.03 | 20 | 4.03 |
| 0 | 266 | 53.63 | 286 | 57.66 |
| 1 | 157 | 31.65 | 443 | 89.31 |
| 2 | 53 | 10.69 | 496 | 100.00 |

Figure 1.2.1 Boxplot of mRNA in each of level of copy number for EGFR gene



Figure 1.2.2 Boxplot of log2(mRNA) in each of level of copy number for EGFR gene 

1. **Lung SCC**

2.1 FAT1 of Lung SCC

Table 2.1.1 Pearson correlation coefficient among Lung SCC

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| FAT1 | 0.24886 | <0.0001 |

Table 2.1.2 Spearman correlation coefficient

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| FAT1 | 0.25767 | <0.0001 |

Table 2.1.3 Frequency of FAT1 copy number among Lung SCC

| FAT1CNA | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| --- | --- | --- | --- | --- |
| -2 | 27 | 5.76 | 27 | 5.76 |
| -1 | 285 | 60.77 | 312 | 66.52 |
| 0 | 130 | 27.72 | 442 | 94.24 |
| 1 | 24 | 5.12 | 466 | 99.36 |
| 2 | 3 | 0.64 | 469 | 100.00 |

Figure 2.1.1 Boxplot of mRNA in each of level of copy number for FAT1 among Lung SCC



Figure 2.1.2 Boxplot of log2(mRNA) in each of level of copy number for FAT1 among Lung SCC



2.2 EGFR of Lung SCC

Table 2.2.1 Pearson correlation of EGFR of Lung SCC

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| EGFR | 0.39638 | <0.0001 |

Table 2.2.2 Spearman correlation of EGFR of Lung SCC

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| EGFR | 0.41173 | <0.0001 |

Table 2.2.3 Frequency of EGFR copy number of Lung SCC

| EGFRCNA | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| --- | --- | --- | --- | --- |
| -2 | 2 | 0.43 | 2 | 0.43 |
| -1 | 49 | 10.45 | 51 | 10.87 |
| 0 | 176 | 37.53 | 227 | 48.40 |
| 1 | 213 | 45.42 | 440 | 93.82 |
| 2 | 29 | 6.18 | 469 | 100.00 |

Figure 2.2.1 Boxplot of mRNA in each of level of copy number for EGFR among Lung SCC



Figure 2.2.2 Boxplot of log2(mRNA) in each of level of copy number for EGFR among Lung SCC



1. **Cervical SCC**

3.1 FAT1 of Cervical SCC

Table 3.1.1 Pearson correlation coefficient

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| FAT1 | 0.28609 | <0.0001 |

Table 3.1.2 Spearman correlation coefficient

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| FAT1 | 0.21310 | 0.0004 |

Table 3.1.3 Frequency of FAT1 copy number

| FAT1CNA | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| --- | --- | --- | --- | --- |
| -2 | 9 | 3.24 | 9 | 3.24 |
| -1 | 99 | 35.61 | 108 | 38.85 |
| 0 | 148 | 53.24 | 256 | 92.09 |
| 1 | 22 | 7.91 | 278 | 100.00 |

Figure 3.1.1 Boxplot between FAT1 copy number and mRNA



Figure 3.2.2 Boxplot between FAT1 copy number and log2(mRNA)



3.2 EGFR of Cervical SCC

Table 3.2.1 Pearson correlation of EGFR

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| EGFR | 0.39813 | <0.0001 |

Table 3.2.2 Spearman correlation of EGFR

|  |  |  |
| --- | --- | --- |
| Gene | Sample estimate | P-value |
| EGFR | 0.27976 | <0.0001 |

Table 3.2.3 Frequency of EGFR copy number

| EGFRCNA | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| --- | --- | --- | --- | --- |
| -1 | 21 | 7.55 | 21 | 7.55 |
| 0 | 207 | 74.46 | 228 | 82.01 |
| 1 | 44 | 15.83 | 272 | 97.84 |
| 2 | 6 | 2.16 | 278 | 100.00 |

Figure 3.2.1 Boxplot between EGFR copy number and mRNA



Figure 3.2.2 Boxplot between EGFR copy number and log2(mRNA)



**Conclusions**

From the results, we could see that mRNA and copy-number alterations were positively and significantly correlated for FAT1 and EGFR among each of the 3 cancer types. GLM test results demonstrated that the means of mRNA expression were significantly different across the different levels of copy number among each of the three cancer types.