An Exploration of the Use of Trastuzumab and Bevacizumab to treat additional types of Cancers

A Project completed under the Figure One Lab Internship Emulator

Key Scientific Question

Using scRNA-seq data from available cancer cell lines, how will you explore the use of the FDA-approved antibody therapies Trastuzumab and Bevacizumab in the treatment of additional cancers?
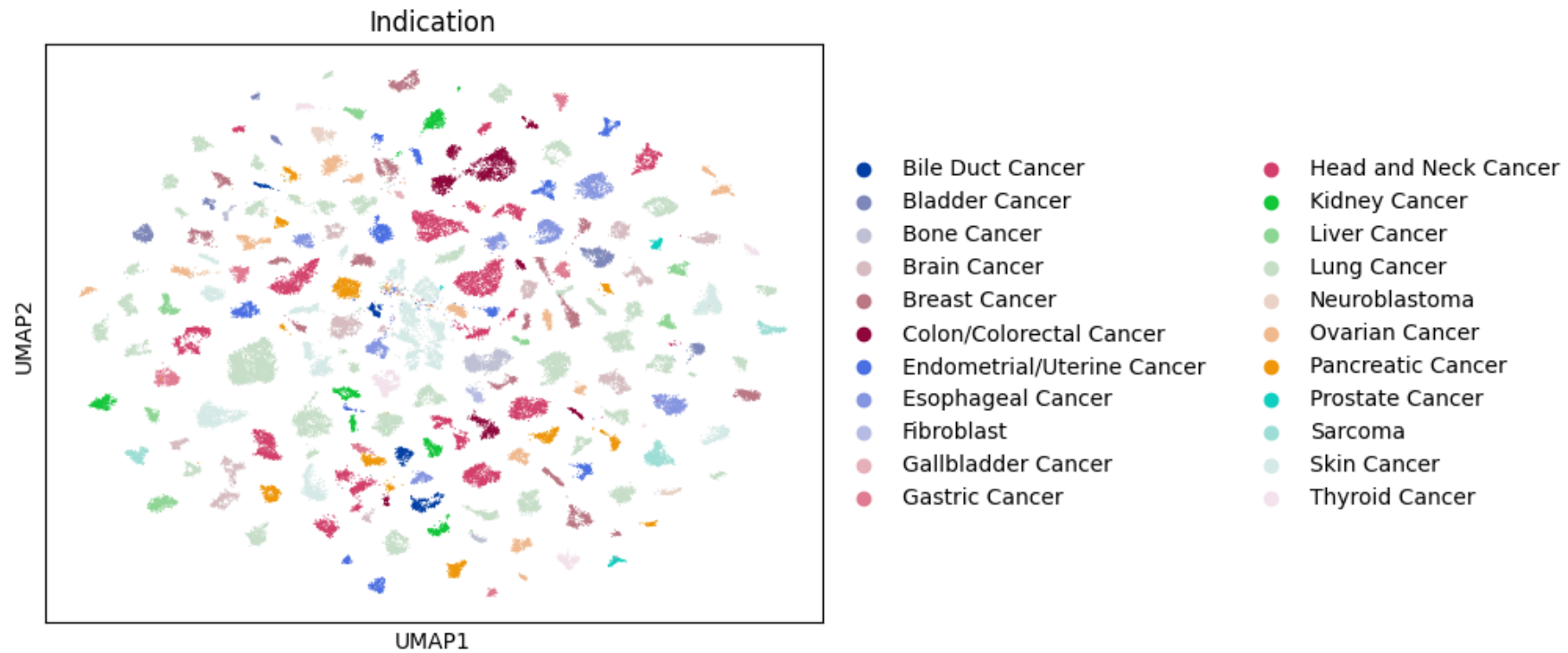
Overview

- Trastuzumab and Bevacizumab are drugs consisting of antibodies produced in a laboratory that can help the body's immune system to fight cancer.

- Trastuzumab targets HER2 (human epidermal growth factor receptor 2), encoded by the ERBB2 gene.

- Bevacizumab targets VEGFA (vascular endothelial growth factor A), encoded by the VEGFA gene.
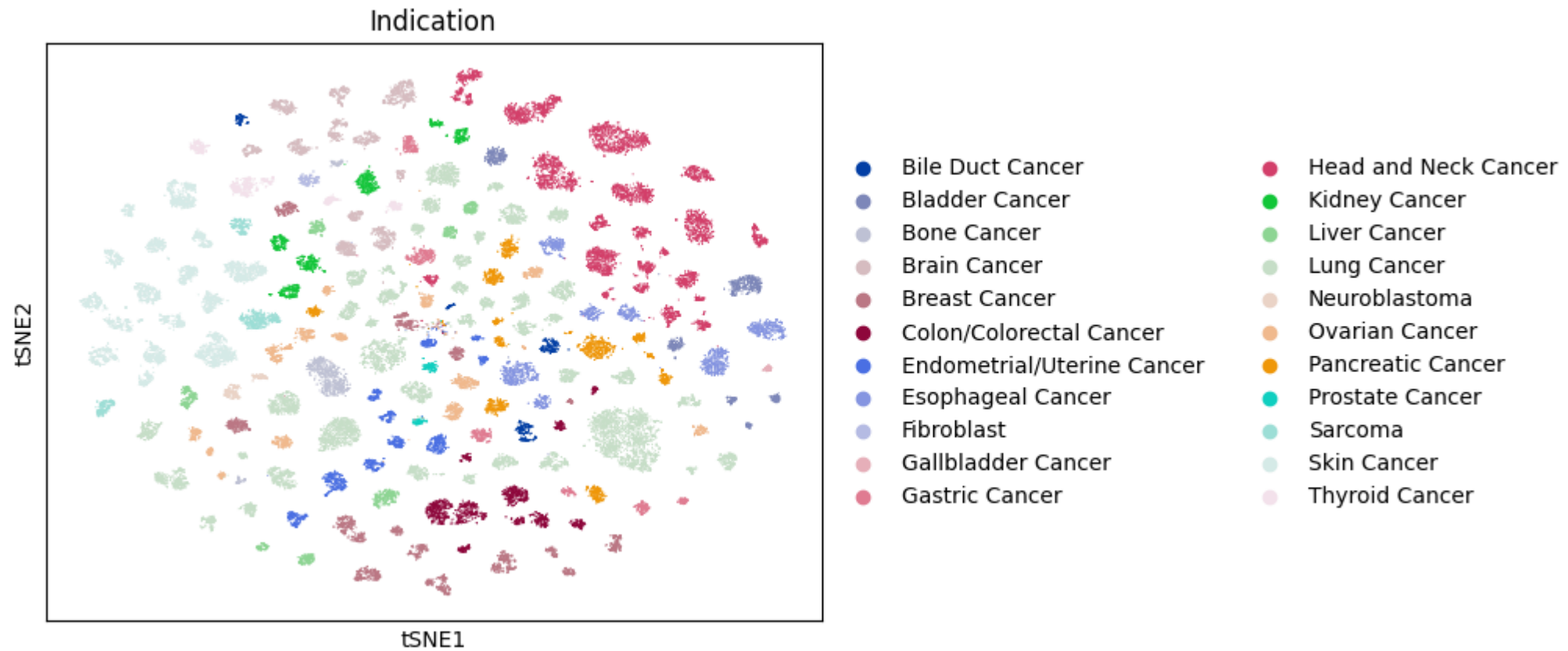
Methods

- Preprocessing, annotation and initial exploration of the data was done using code provided in *240701_kinker_anndata.ipynb* and *240702_kinker_scanpy.ipynb,* two Jupyter notebooks created by Mr. Dean Lee.

- The data was further analysed using differential gene expression analysis, t-tests, one-way ANOVA tests and Tukey's HSD tests.

- The analysis was conducted in Python, using the *scanpy* and *anndata* packages.
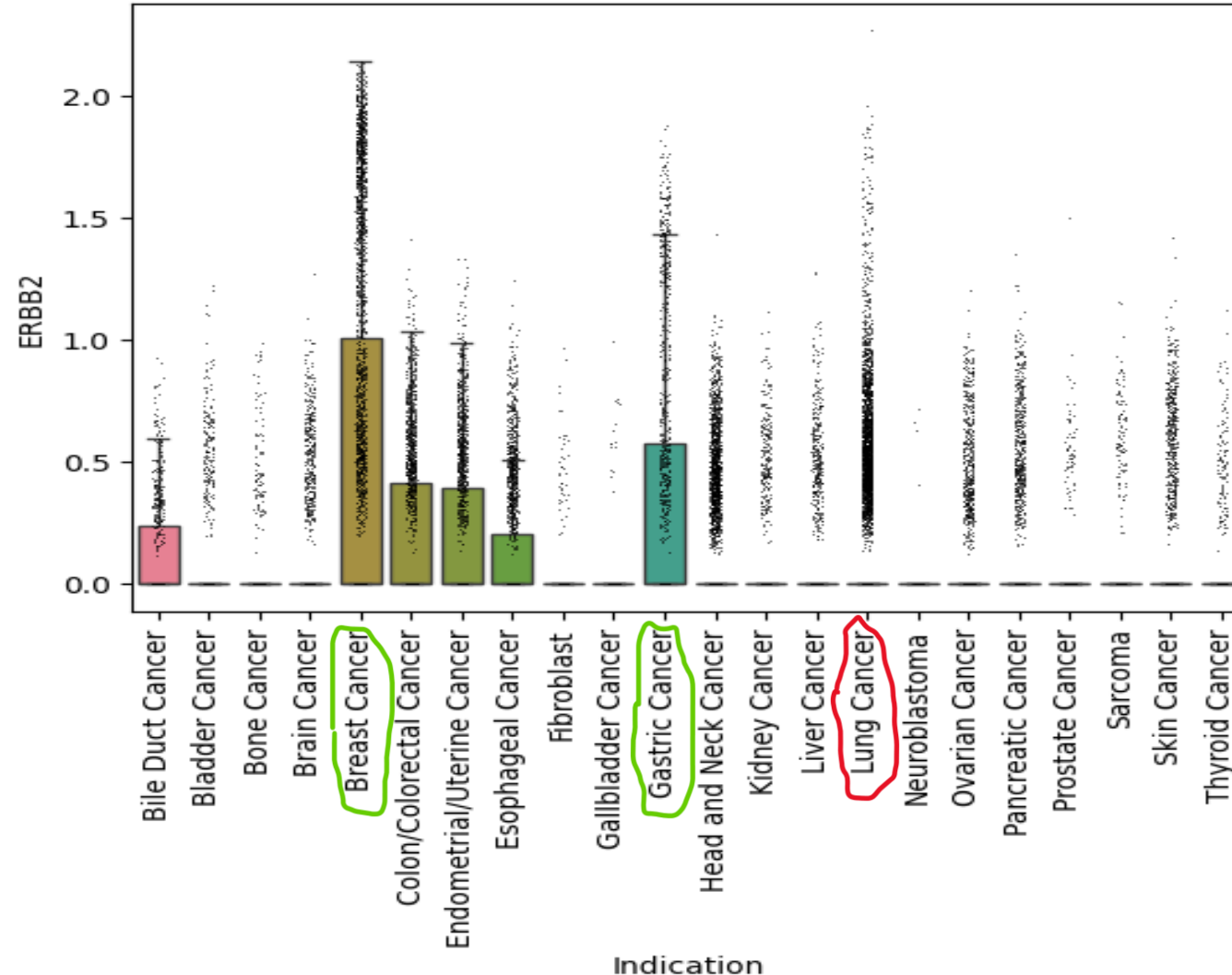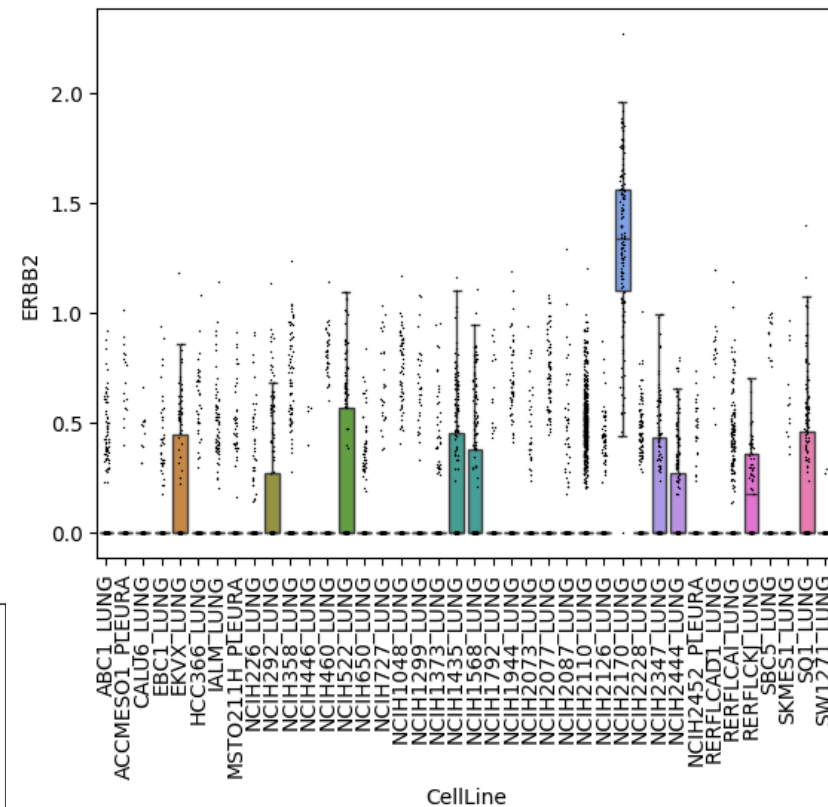
Plots

UMAP clustering

tSNE plot

Indication

Legend:
- Bile Duct Cancer
- Bladder Cancer
- Bone Cancer
- Brain Cancer
- Breast Cancer
- Colon/Colorectal Cancer
- Endometrial/Uterine Cancer
- Esophageal Cancer
- Fibroblast
- Gallbladder Cancer
- Gastric Cancer
- Head and Neck Cancer
- Kidney Cancer
- Liver Cancer
- Lung Cancer
- Neuroblastoma
- Ovarian Cancer
- Pancreatic Cancer
- Prostate Cancer
- Sarcoma
- Skin Cancer
- Thyroid Cancer

Trastuzumab Target Exploration

Expression of ERBB2 gene by Indication

Cancers for which the drug is currently approved have been outlined in green. Additional cancers – lung cancer, in this case – in which the ERBB2 gene appears to be strongly expressed have been outlined in red. This may indicate a potential application of the drug in the treatment of this type of cancer.
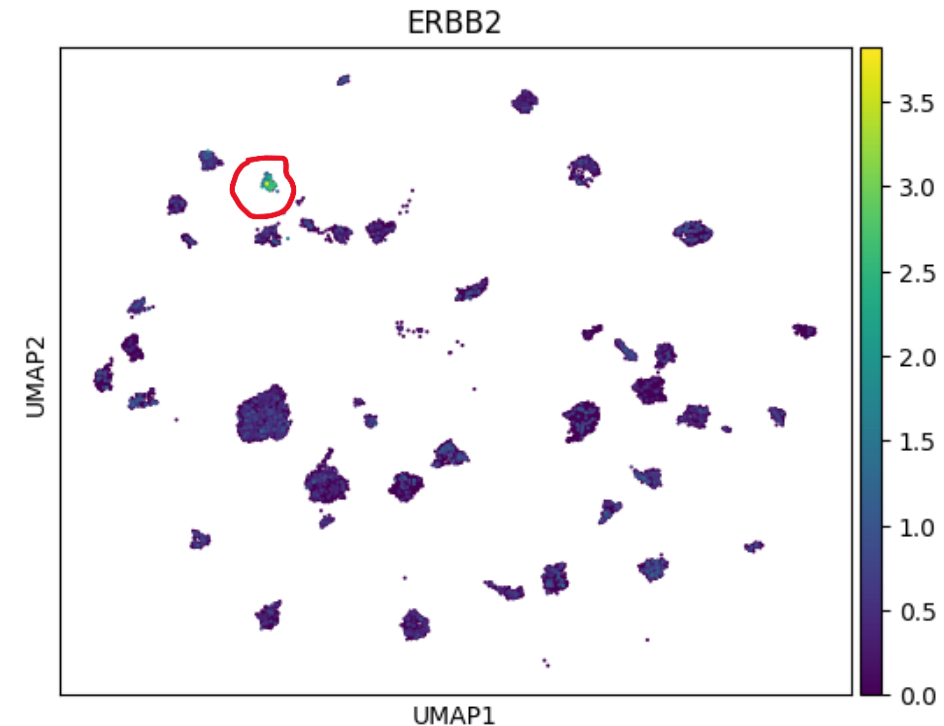
Expression of ERBB2 by
Cell Line for Breast, Gastric
and Lung Cancers
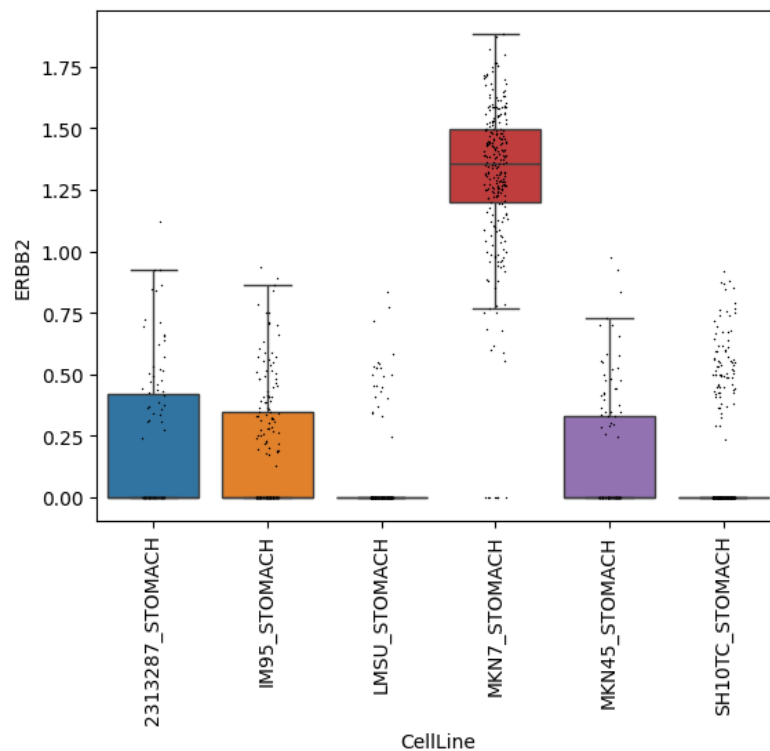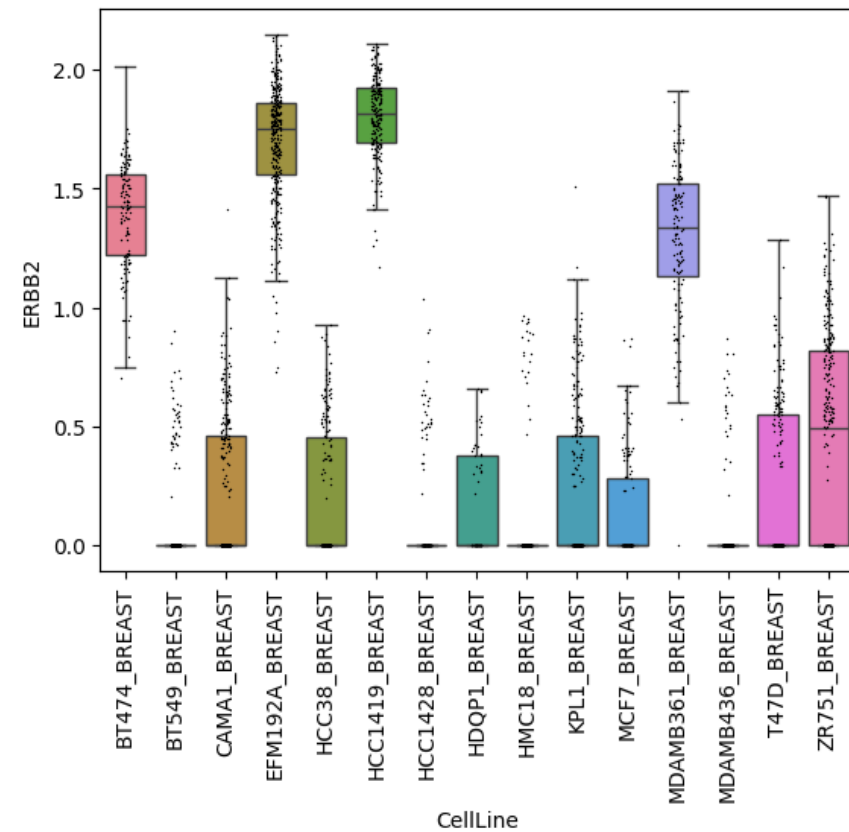
Mapping of Lung Cancer Cell Lines to Varying Levels of ERBB2 Expression
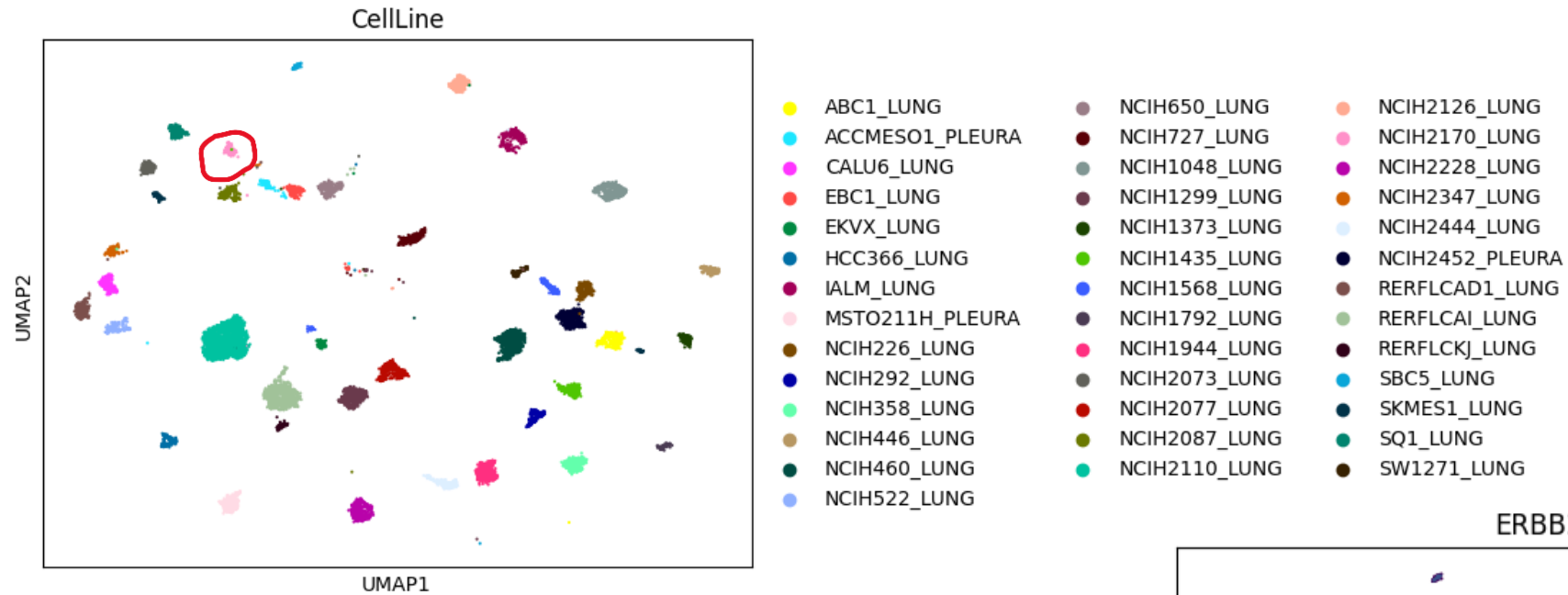
Clusters of cells that express a significantly higher level of the ERBB2 gene compared to other clusters have been encircled in red in both plots. A comparison of the two UMAPs will indicate that the NCIH2170_LUNG cell line appears to have a significantly higher expression level for this gene.

Differential Gene Expression Analysis

Differential gene expression analysis revealed that the NCIH2170_LUNG cell line had a significantly higher level of ERBB2 expression.

p-value = 2.532832e-59

adjusted p-value = 9.444313e-58

Statistical Tests

Three statistical tests – the t-test, the one-way ANOVA and the pairwise Tukey's HSD test – were conducted in order to determine if the expression of ERBB2 differed among the lung cancer cell lines that were analysed. The statistic and p-values of the first two tests are provided below:

- t-test

T statistic: 38.66725312879398
P value: 2.551458288533391e-77

- ANOVA
F statistic: 4005.561594678734
P value: 0.0

Conclusions

From multiple lines of evidence, it may be inferred that the ERBB2 gene, which is targeted by the drug Trastuzumab, is differentially expressed in certain cell lines of lung cancer.

These results appear to indicate that the potential use and regulatory approval of this drug in the treatment of lung cancer should be explored through further analysis.

Bevacizumab Target Exploration

Expression of VEGFA gene by Indication

Cancers for which the drug is currently approved have been outlined in green. Additional cancers – head and neck cancer, in this case – in which the VEGFA gene appears to be strongly expressed have been outlined in red. This may indicate a potential application of the drug in the treatment of this type of cancer.
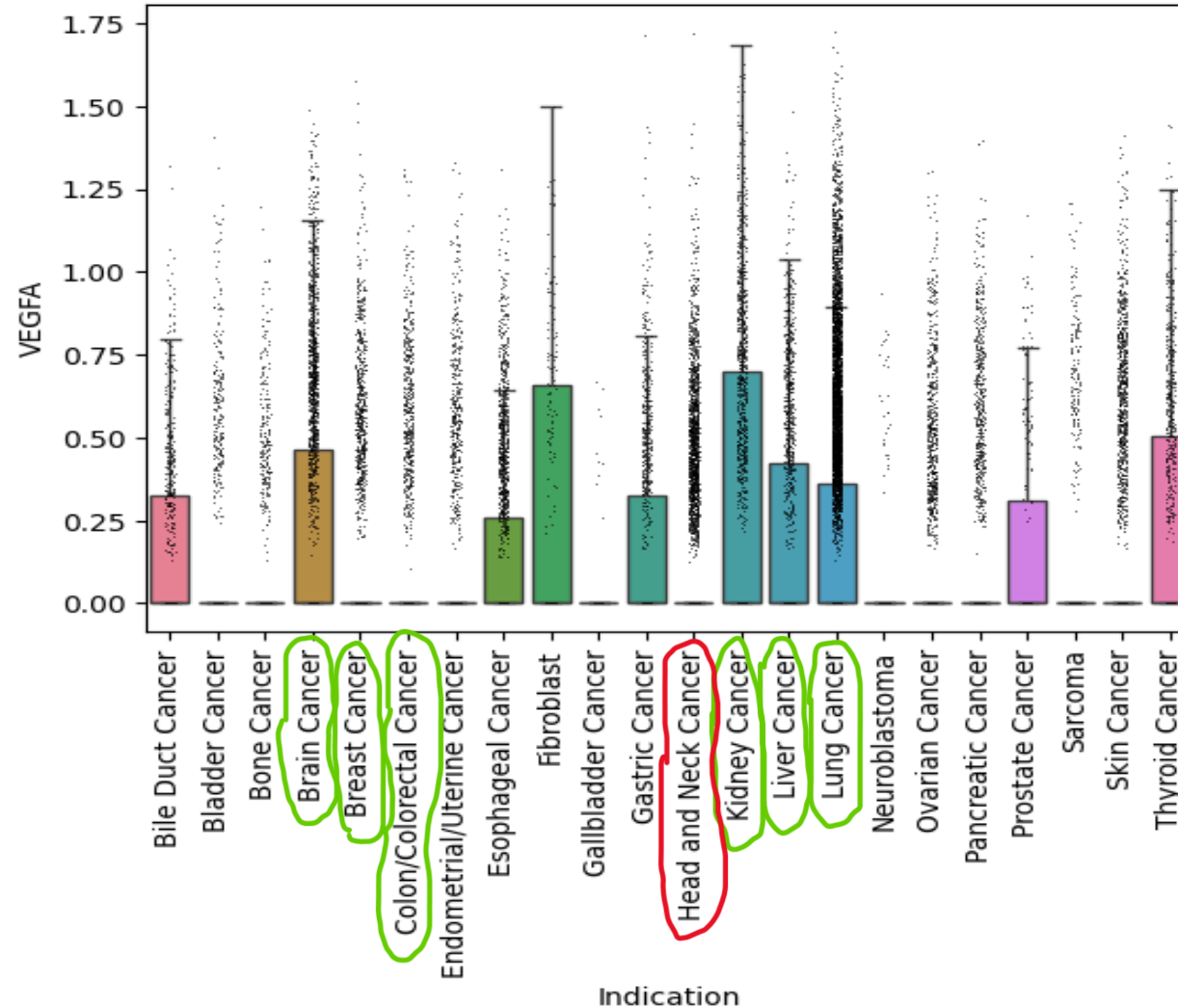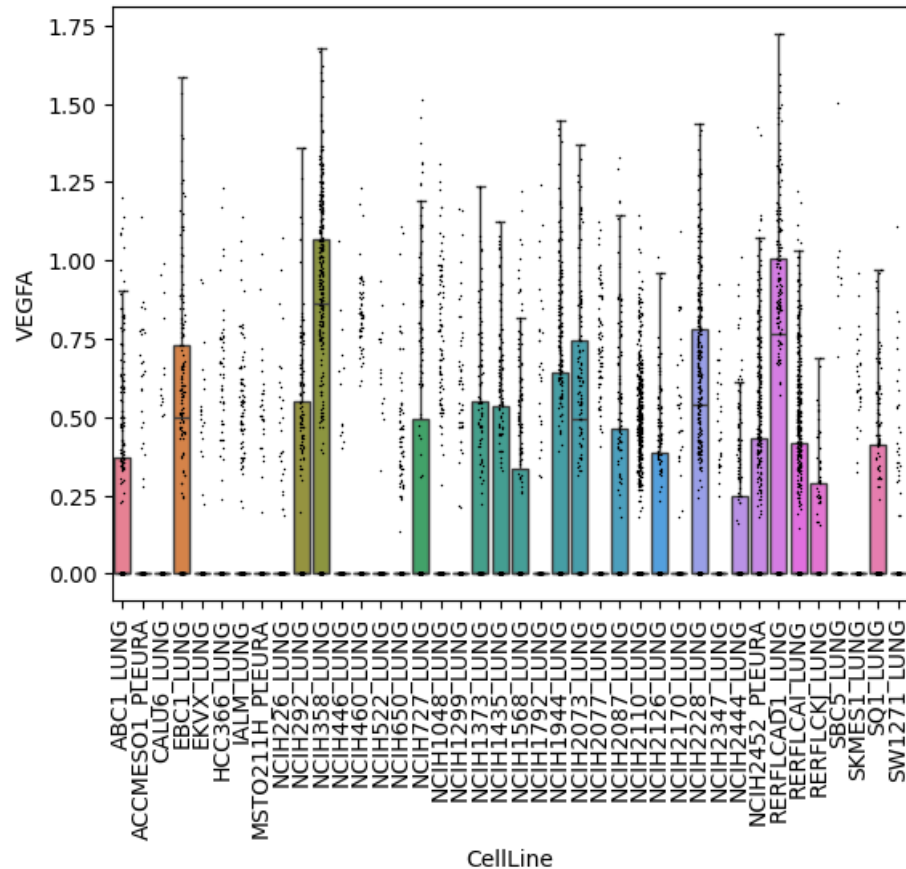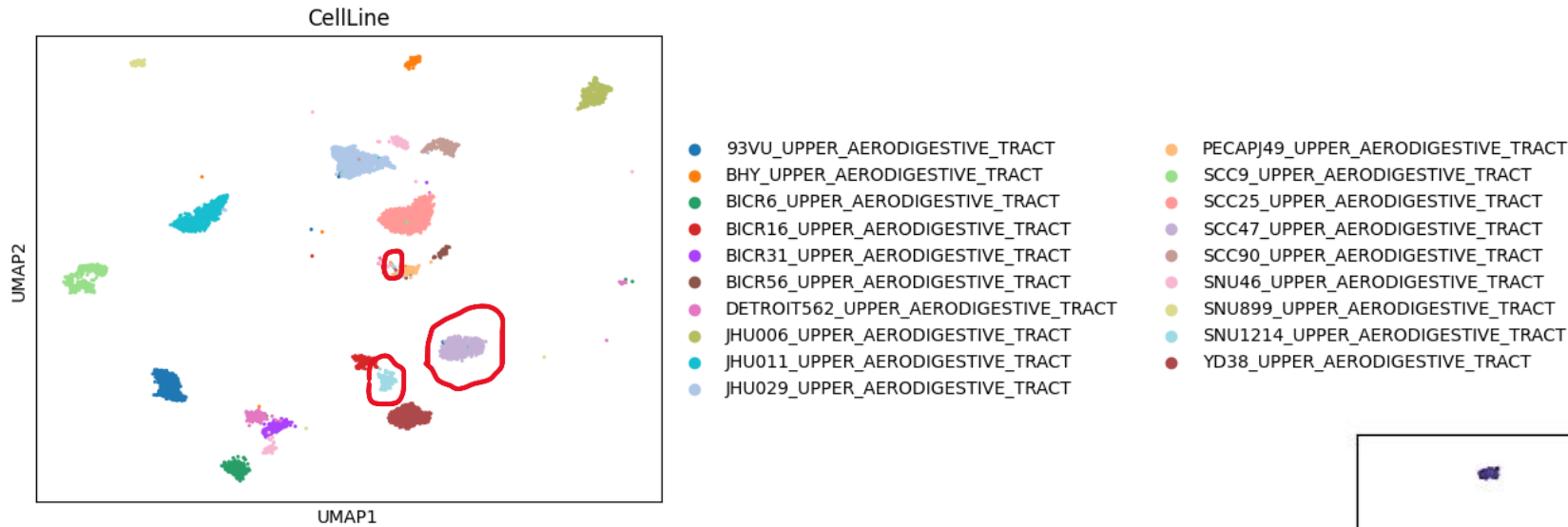
Expression of VEGFA by Cell Line for Lung, Kidney and Head and Neck Cancers
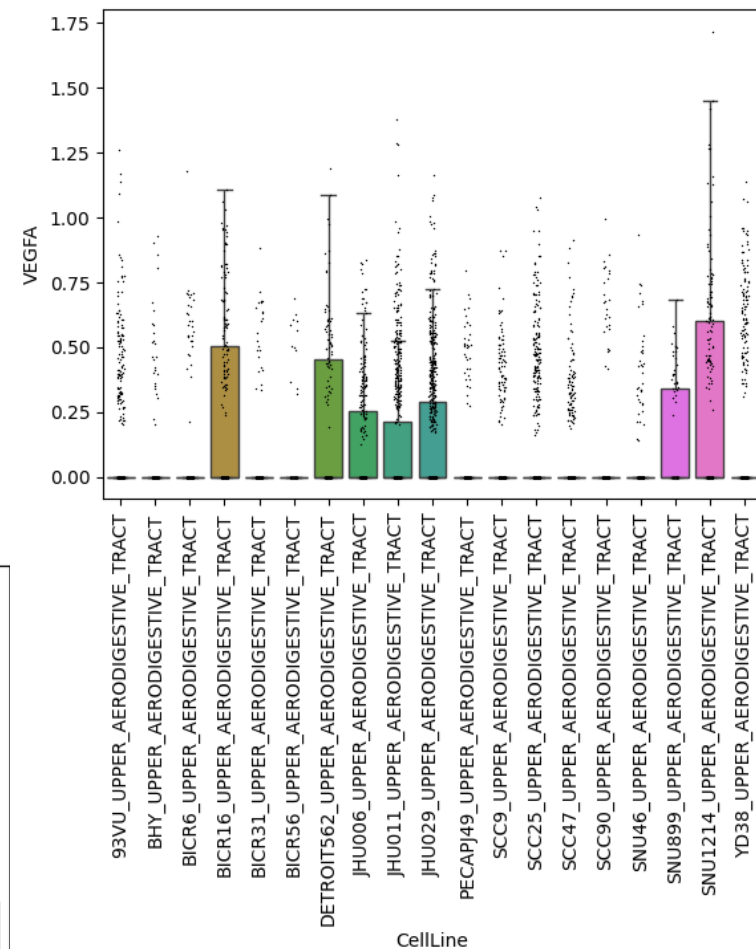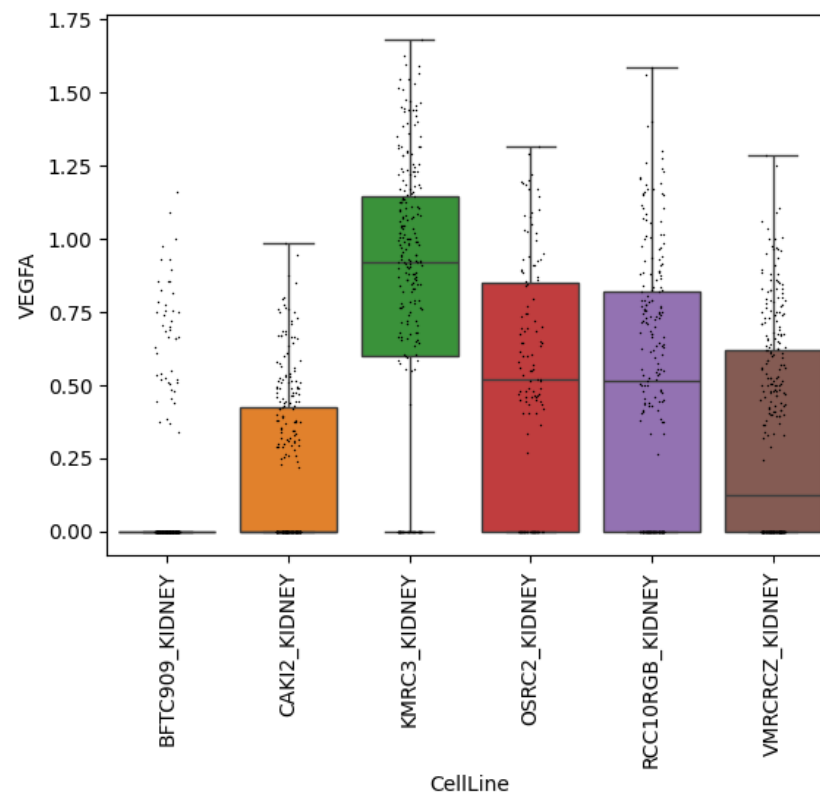
Mapping of Head and Neck Cancer Cell Lines to Varying Levels of VEGFA Expression

CellLine

- 93VU_UPPER_AERODIGESTIVE_TRACT
- BHY_UPPER_AERODIGESTIVE_TRACT
- BICR6_UPPER_AERODIGESTIVE_TRACT
- BICR16_UPPER_AERODIGESTIVE_TRACT
- BICR31_UPPER_AERODIGESTIVE_TRACT
- BICR56_UPPER_AERODIGESTIVE_TRACT
- DETROIT562_UPPER_AERODIGESTIVE_TRACT
- JHU006_UPPER_AERODIGESTIVE_TRACT
- JHU011_UPPER_AERODIGESTIVE_TRACT
- JHU029_UPPER_AERODIGESTIVE_TRACT
- PECAPJ49_UPPER_AERODIGESTIVE_TRACT
- SCC9_UPPER_AERODIGESTIVE_TRACT
- SCC25_UPPER_AERODIGESTIVE_TRACT
- SCC47_UPPER_AERODIGESTIVE_TRACT
- SCC90_UPPER_AERODIGESTIVE_TRACT
- SNU46_UPPER_AERODIGESTIVE_TRACT
- SNU899_UPPER_AERODIGESTIVE_TRACT
- SNU1214_UPPER_AERODIGESTIVE_TRACT
- YD38_UPPER_AERODIGESTIVE_TRACT

Clusters of cells that express a significantly higher level of the VEGFA gene compared to other clusters have been encircled in red in both plots. It should be noted that, in this case, there appear to be several clusters with reasonably high expression levels for this gene, and that, unlike in the previous case of ERBB2, there does not appear to be a single cluster with a significantly high expression level than the other clusters.. A comparison of the two UMAPs will indicate that the SCC47_UPPER_AERODIGESTIVE_TRACT and SNU1214_UPPER_AERODIGESTIVE_TRACT cell lines appear to have a significantly higher expression level for this gene.
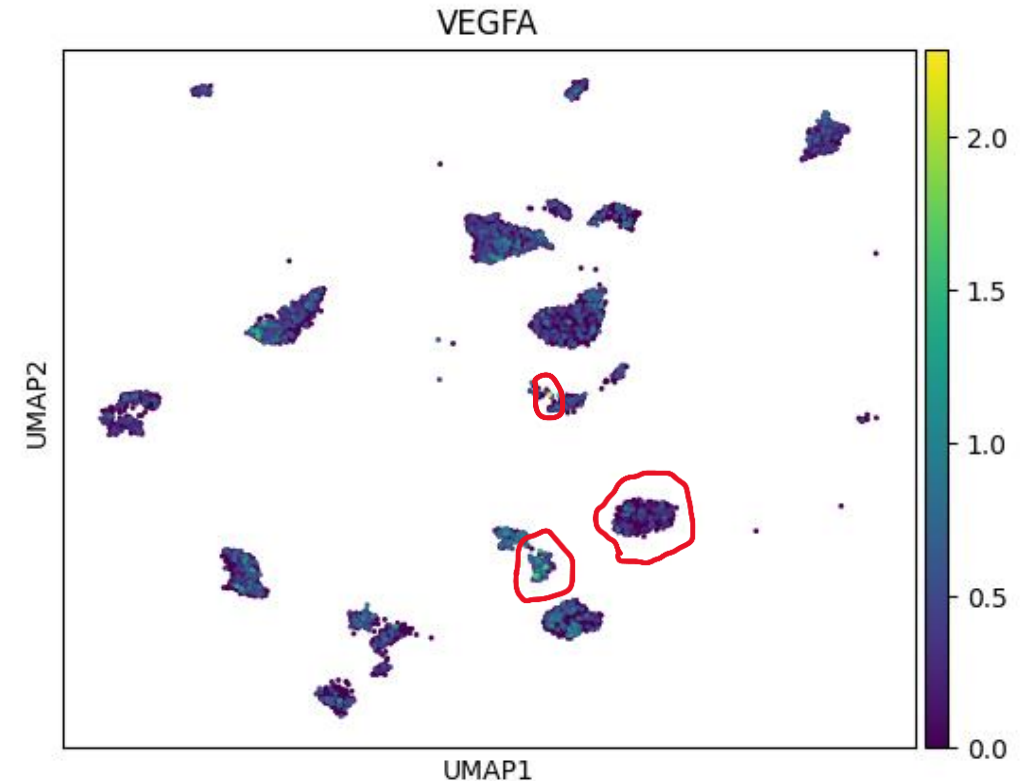
VEGFA

Differential Gene Expression Analysis

Differential gene expression analysis revealed that the SCC47_UPPER_AERODIGESTIVE_TRACT and SNU1214_UPPER_AERODIGESTIVE_TRACT cell lines had significantly higher levels of VEGFA expression.

- SCC47_UPPER_AERODIGESTIVE_TRACT

p-value = 3.476825e-17

adjusted p-value = 2.478338e-16

- SNU1214_UPPER_AERODIGESTIVE_TRACT

p-value = 1.172390e-10

adjusted p-value = 8.290419e-10

Statistical Tests

Three statistical tests – the t-test, the one-way ANOVA and the pairwise Tukey's HSD test – were conducted in order to determine if the expression of VEGFA differed among the two cell lines under consideration. The statistic and p-values of the first two tests are provided below:

- t-test

SCC47_UPPER_AERODIGESTIVE_TRACT

T statistic: -8.12978787965765
P value: 1.475578100728705e-15

SNU1214_UPPER_AERODIGESTIVE_TRACT

T statistic: 7.257565460715087
P value: 9.732026231693352e-12

- ANOVA
SCC47_UPPER_AERODIGESTIVE_TRACT
F statistic: 32.69481930676174
P value: 1.1220137923608294e-08

SNU1214_UPPER_AERODIGESTIVE_TRACT
F statistic: 148.4034830848608
P value: 8.405728969806294e-34

Conclusions

From multiple lines of evidence, it may be inferred that the VEGFA gene, which is targeted by the drug Bevacizumab, is differentially expressed in certain cell lines of head and neck cancer.

These results appear to indicate that the potential use and regulatory approval of this drug in the treatment of head and neck cancer should be explored through further analysis.

Caveats

- There are certain caveats for the use of cell culture lines that should be considered in the interpretation of the results of this analysis.

- Cells in culture lines tend to experience genetic drift over a period of time. They may not always accurately reflect the gene expression patterns of cancer cells *in vivo*.

Acknowledgements

This project was completed under the Figure One Lab Internship Emulator, organised by Mr Dean Lee on LinkedIn. The initial stages of the analysis were built off of code written by Mr Lee.

I express my gratitude to Mr Lee for providing a freely accessible platform for aspiring bioinformaticians to hone their skills.