The effect of gene CD99 between male and female patients

Introduction:

The condition of colorectal cancer occurs when cells in the colon or rectum grow out of control. Colorectal cancer was relatively uncommon before 1950, but it has since become a common malignancy in Western countries, accounting for around 10% of cancer-related death (Kuipers et al., 2015). According to the Mayo Clinic data analysis, there are over 200,000 cases in the United States per year. Under this circumstance, it is important to discover the original cause of colorectal cancer and to find the correlated treatment. There are many reasons to cause colorectal cancer, including external and internal. For external reasons, population aging accompanied by a high prevalence of bad food habits, smoking, low physical activity, and obesity is one of the causes (Kuipers et al., 2015). However, internal reasons seem to be more complex and hard to deal with, and those are mostly due to some genes mutation. CD99 (MIC2; single-chain type-1 glycoprotein) is a 32-kDa transmembrane protein that is extensively O-glycosylated and found on leukocytes and active endothelium (Huijbers et al., 2019). With more evidence appearing, CD99 shows a high-level expression in many cancer, including colorectal cancer (Kuipers et al., 2015).

Thus, in this experiment, TCGA (The Cancer Genome Atlas —- a cancer genomics program that examines over 20,000 primary cancer and matched normal samples spanning 33 cancer types) is used as the database and the patient's samples are separated into two categories: male and female. The CD99 gene is analyzed under the R environment. According to one of the articles, it concludes that men have higher levels of CD99, and data reveal that men are more likely than women to get brain cancer (Copper et al., 2018). Therefore, this conclusion is utilized

as this experiment hypothesis and the question is posted that what's the difference in CD99 gene expression between males and females.

Method:

TCGA is used as an experiment database, and the R is used to analyze data. Firstly, the CD99 gene counts are examined between males and females. In R, we accessed the clinic data and sum_exp through the package of TCGAbiolinks and SummarizedExperiment. Then we find the Ensembl gene ID of gene CD99 and use it to get the gene counts in each patient. We draw the boxplot with the gender on the x-axis and the counts of gene CD99 on the y-axis. After that, we use the library of survival and survminer, and the clinic data to create the Kaplan-Meier survival plot to examine the survival time between males and females. In addition, we draw the survival group between high-expression CD99 gene patients and low-expression CD99 gene patients. Then, we connect these graphs together to create the survival plots of different levels of expression of male patients and female patients. Finally, we try to find the mutation rate of CD99 between male groups and female groups. The clinic data and maf_object which is created by using maftools are conducted to draw the lollipopPlot of the CD99 gene compared with males and females.

Results:

For the first experiment, we found that the gene counts in male groups are slighter greater than those in females (Figure 1). When conducting the mutation analysis, we found that the CD99 has a small mutation in patients' samples and the mutation only occurs in males (Figure 2).

For the survival analysis, men's overall survival probability is higher than that of women's (Figure 3). The survival probability between CD99 high-expression patients and CD99 low-expression patients shows a difference. It is obvious to see that CD99 high-expression patients have a higher survival probability over years than low-expression patients (Figure 4). When dividing patients into males and females, we found that there is no significant difference between CD99 various levels of expression in male groups (Figure 5). Whereas, female patients who have high CD99 gene expression show a higher survival probability than low CD99 gene expression in female patients (Figure 6).
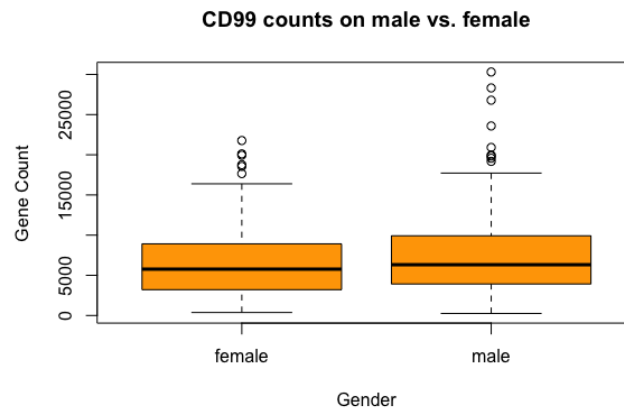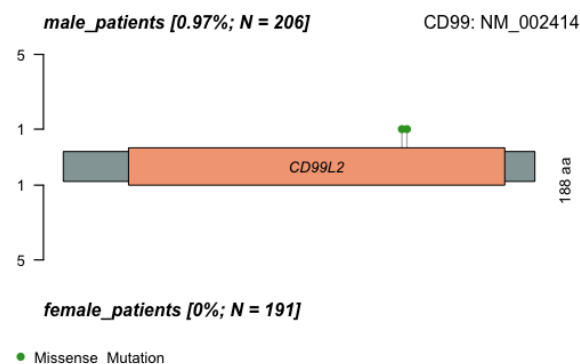

Figure 1: CD99 counts between females and males

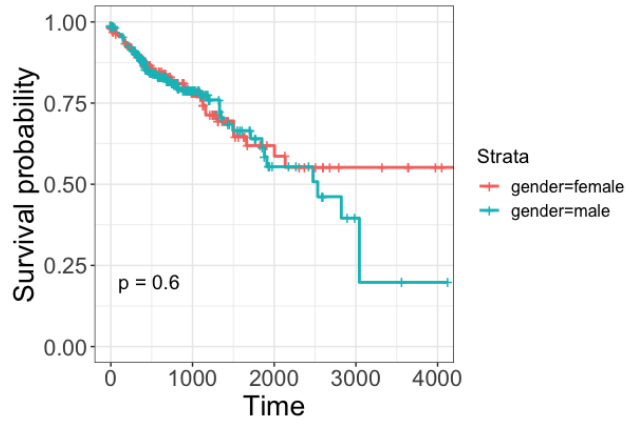**Figure 2**: Mutation Distribution in males and females

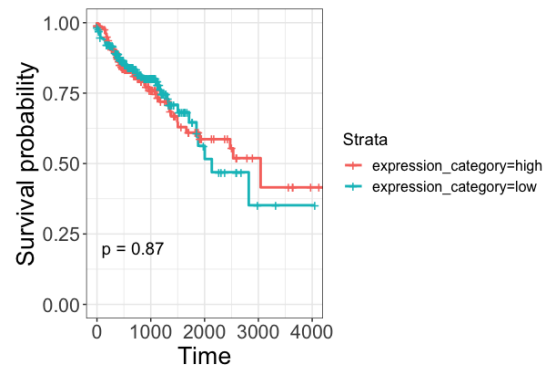**Figure 3**: Survival probability between males and females

Figure 4: Survival probability between CD99 high expression patients
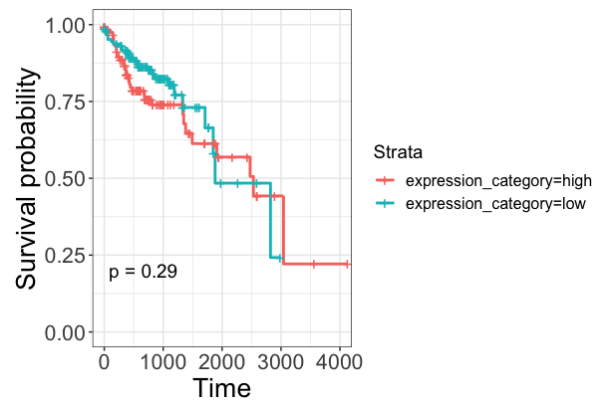
and CD99 low expression patients

Figure 5: Survival probability between CD99 high expression male patients

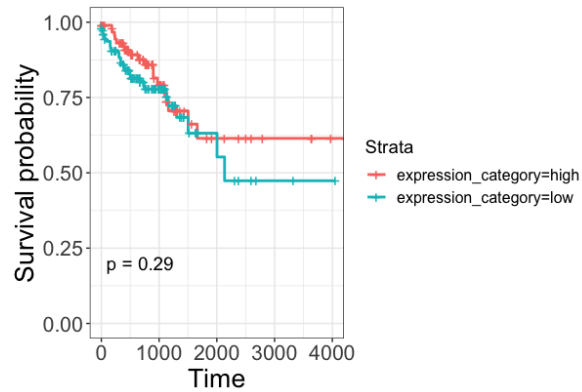and CD99 low expression male patients

Figure 6: Survival probability between CD99 high expression female patients

and CD99 low expression female patients

Discussion:

For the first gene counts experiment, our conclusion is proved by Taylor Cooper's project that states CD99 is upregulated in men (Copper et al., 2018). In addition, due to our sample size is not big enough and samples' uncertain characteristics, it is not credible to conclude that men's overall survival probability is higher than that of women's, leading to the following figures all having some deviations. However, we could still see some patterns that high expression of CD99 may serve as an element in cancer treatment and it is already demonstrated in Elisabeth's article which concludes that Targeting Tumor Vascular CD99 inhibits tumor growth (Huijbers et al., 2019). CD99 has few mutations in sample size and the only mutation is missense mutation which only occurs in male patients. This conclusion is proved in Taylor Cooper's project that illustrates

that brain cancer is more common in men than in women due to the CD99 gene mutation (Copper et al., 2018).

However, the experiments still have some limitations that need further research. For example, After getting the conclusion of CD99 mutation only occurs in male groups, it is better to use this conclusion in other databases to verify whether truly females do not have CD99 mutation or not. Also, we need to further study the CD99 mutation related to cancer. The article, only shows that the CD99 mutation may cause brain cancer. What about other cancer types, such as colorectal cancer? These questions are valuable and need to be solved in the future.

Citation

1. Kuipers, E. J., Grady, W. M., Lieberman, D., Seufferlein, T., Sung, J. J., Boelens, P. G., van de Velde, C. J. H., & Watanabe, T. (2015, November 5). *Colorectal cancer*. Nature reviews. Disease primers. Retrieved March 3, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4874655/

2. Huijbers, E. J. M., van der Werf, I. M., Faber, L. D., Sialino, L. D., van der Laan, P., Holland, H. A., Cimpean, A. M., Thijssen, V. L. J. L., van Beijnum, J. R., & Griffioen, A. W. (2019, April 2). *Targeting tumor vascular CD99 inhibits tumor growth*. Frontiers in immunology. Retrieved March 3, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6455290/

3. *The cancer genome atlas program*. National Cancer Institute. (n.d.). Retrieved March 3, 2022, from https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga

4. *Differential expression of CD99 in males and females*. Biology Computes. (2018, November 6). Retrieved March 3, 2022, from https://gtbinf.wordpress.com/2018/11/04/differential-expression-of-cd99-in-males-and-females/

5. Huijbers, E. J. M., van der Werf, I. M., Faber, L. D., Sialino, L. D., van der Laan, P., Holland, H. A., Cimpean, A. M., Thijssen, V. L. J. L., van Beijnum, J. R., & Griffioen, A. W. (2019, April 2). *Targeting tumor vascular CD99 inhibits tumor growth*. Frontiers in immunology. Retrieved March 4, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6455290/