Hello everyone. Today I am going to be talking about analyzing sex-biased gene expression in autoimmune diseases.

For those of you who don't know me, I'm Vidya Vedula, and I'm a junior. I'm excited to tell you a little bit of what I did this past semester.

So first, I'm going to talk about the background of my project.

My project is about autoimmune diseases, so I think it's appropriate to know a few things about them. What are they?

Autoimmune diseases are when your immune system has decided that your healthy cells are foreign and starts to attack them.

An exact cause of autoimmune diseases hasn't been pinpointed yet. However, scientists have suspected many causes such as: sex, twice as many women suffer from autoimmune dusease compared to men, and this has; race, some autoimmune diseases are common in some races. For example, Systemic Lupus Erythematosus is much more common in African Americans and Hispanics compared to Caucasians. Genetics also play a big role. Some autoimmune diseases run in the family, and while it might not be the same disease passed down, the susceptibility is passed down. The "western diet", high in cholesterol, fat, and sugar has also been suspected. These foods can cause inflammation, which can set off an immune system response, but this hasn't been proven yet. . Another suspect is the "hygiene hypothesis". Due to the technology of vaccines, many children today aren't as exposed to germs as before. Due to this, the immune system might have an overreaction to the most harmless substances.

What are some examples of autoimmune diseases? Some examples are multiple sclerosis, which is when your immune system eats away at the protective layer of your neurons; rheumatoid arthritis, which is when your immune system attacks joint linings, causing painful swelling; systemic lupus erythematosus or SLE, which is when your immune system attacks its own tissues such as the skin, brain, and blood vessels, causing painful inflammation; and type 1 diabetes, which is when the pancreas produces an inadequate amount of insulin. As you can see, most of these diseases are very known to us. That leads us to wonder, how common are autoimmune diseases?

In the US, 23.5 million people have autoimmune diseases. An additional 8 million people have autoantibodies, antibodies which are made against the body which make them more susceptible to autoimmune diseases. Also, autoimmune diseases are one of the top 10 leading causes of death in girls and women.

There is also a huge sex disparty in autoimmune diseases. In all of the people who suffer from autoimmune diseases, women make up 75% while men only make up 25%.
This leads us to the question, what are some genes that contribute to the disparity in autoimmune diseases? A lot of previous research was done before about the higher prevalence of autoimmune diseases in women. One proposition to explain this are hormonal factors. Research has shown that sex hormones should bind to estrogen receptors expressed by the immune cells in order to

balance immune function. However, only a few cells express estrogen receptors. Another piece of evidence is that T cells increased cytokine production, which are molecules important in cell signaling, when exposed to estrogen. Estrogen can also negatively affect the apoptosis of B cells, the antibody making cell, causing clones of autoreactive cells. As we all know, the hormonal levels in a female fluctuate constantly, so the excess of hormones can cause immune system responses. Prolactin is another hormone that has been suspected. One study was done on the effects of excess and lack of prolactin in female mice. The study found that mice with a prolactin-inhibitor had longer longevities and had more antibodies that detect SLE. On the other hand, mice with glands that produce more prolactin had accelerated mortality and proteins in their urine, which is a key symptom of SLE.

Next I'm going to be talking about the methods and materials of my project.

To compute statistically, I used R and RStudio. This has helped me to process and analyze large amounts of data.

The dataset I used is a database called DICE which contains information about the donor's sex, race, ethnicity, and the count of different immune cells per 1 million transcripts. Some donors' count was done again after a few days. This data was collected by using RNA-Seq.
Here is an image of the dataset for your clarification. As you can see, there are different columns containing information about each donor. The right-most column contains the information for classical monocytes, but there are many more cell types and only one was shown here for your clarification.

RNA-Seq is a sequencing technique to quantify RNA in a sample. This technology was declared the Breakthrough of the Year by Science in 2018 when it was used to map the genome of zebrafish and african clawed frog.

First, mRNA is converted into small fragments of cDNA. Sequencing adapters are added to the ends of the cDNA fragments, and a short sequence is obtained from each cDNA fragment. They are then aligned with the reference genome or transcriptome and classified as either exonic reads, junction reads, or poly(A) end reads. These are then used to generate a base-resolution expression profile.

We are now halfway through the presentation. Now, I have my dataset, so let's talk about the code I used. The first step I did was to obtain the essential information.

For this, I selected just the columns needed for my analysis to keep it simple, which includes the Donor ID, sex, and all of the cell types. For this, I used the command shown on the right, where .data represents the dataset.

The next step is to filter the sexes.

For this, I divided the dataset into male and female. For this, I used the command "filter".
The next step was to find the average of the cell types.

For this, I calculated the mean of each cell type for each divided dataset. For this, I used the command “colMeans”.

The next step was to find the difference between the sexes.

Instinctively, for this step, I subtracted the male average from the female average for each cell type and took the absolute value of that difference.

Finally, the last thing to do is to arrange the differences.

After using the “arrange” command, this is the table with the differences between the sexes arranged in ascending order.

Now, we’ll be talking about the results.

This graph shows the average count in females across the cell types. As you can see, classical monocytes have the highest count, and Naive TREG cells have the fewest count.

This graph shows the average count in males across the cell types. As with the graph in females, classical monocytes have the highest count, and Naive TREG cells have the fewest count.
This graph shows the difference in average female count and average male count. As you can see, NK Cells and Naive CD4+ T cells have the largest difference.

Now, I am moving to the discussion and analysis.

As I said before, NK Cells and Naive CD4+ T cells have the highest difference in counts between the sexes, so these are the cells I will be focusing on those. First, I’ll be talking about NK cells.

NK cells are a type of lymphocyte that defends the body from infections. One of their defining functions are attacking aging, infected, and cancerous cells. This picture here shows how a NK cell attacks a cancerous cell. After the activation of the receptor, cytotoxic molecules are released into the intracellular space that was formed because the NKcell and cancer cell got joined as shown above. The cytotoxic molecules contain proteins that end up killing the cancer cells. There are two different NK cell types. One is CD56 NK Cells which express killer immunoglobulin-like receptors or KIR, which serve as the key regulators of the cell function.

The second type is CD56bright Nk cells, which don’t express KIR.

What is the relation between NK cells and autoimmune diseases? Human NK cells express 3 natural cytotoxicity receptors, receptors that receive cytotoxic elements. These 3 NCRs are NKp46, NKp44, and NKp30, and all three of these are genes that have been researched and associated with autoimmune diseases. NK cells can also promote inflammation by giving out cytotoxic effects against other types of cells with NCRs, which respond by either killing the cell or releasing cytokines.

In terms of the sex disparity we see, women have a lower number of NK cells, which has been associated with autoimmune diseases.

Now we move on to the other type of cells, Naive CD4+ T cells. These cells are differentiated in the thymus, meaning that they were confirmed to be CD4+ cells. They underwent positive and negative processes of central selection, meaning they eliminated any developing cells that are reactive to self.

These types of cells do not have research that suggests there is as strong correlation with autoimmune diseases as NK cells. Anyways, there is still research that has been done. Naive CD4+ T cells have been found to be functionally impaired in people with rheumatoid arthritis. They have also been observed to be in a higher count in women.

As I said, autoimmune diseases are one of the top 10 leading causes of death among women. By determining which genes play the largest role in this disparity, the scientific community can come up with ideas to diminish it. For example, if it turns out that this disparity is due to the overexpression of a gene, we can design an inhibitor. Solving this problem will benefit everyone. This semester, I have identified some different cell types that could be associated in the sex disparity of autoimmune diseases.

For my future plans, the next thing to do is to narrow down on specific genes. After this, I would like to explore the correlation with race. It is known that some races have higher autoimmune rates than other races. I think it would be interesting to explore that more. Another correlation I would like to explore is if the counts changed after a few days. As you remember, some people had their blood drawn on two days, and seeing whether the count changed would be interesting.

I would like to thank Mr. Mohlhenrich for his support and encouragement for this project.

These are my references.

Thank you, and does anyone have any questions?