

# Lab

You've learned about the SEA-AD database, the basics of transcriptomic (gene expression) data, and how to navigate the Cell×Gene explorer. Now, it's time to apply that knowledge to investigate molecular changes linked to Alzheimer's disease.

Our research question is: **What gene expression changes are associated with Alzheimer's disease in the microglia of the middle temporal gyrus?**

That's a pretty focused question—so where did it come from? Here's the background:

- Microglia have been strongly implicated in the progression of Alzheimer's disease (Keren-Shaul et al. 2017)
- The middle temporal gyrus (MTG) is a brain region key to memory processing and among the first areas to show AD-related dysfunction (Papeo, Agostini, and Lingnau 2019; Chen et al. 2022).
- The SEA-AD team's own analysis (Gabbito et al. 2024) found that microglia exhibit some of the most significant gene expression changes associated with AD. This gives us a chance to explore one of their major findings firsthand.
- By focusing only on microglia, we'll load a smaller subset of the SEA-AD database – important because the site can be unstable or slow when handling the full dataset.

## Make Predictions

Before we begin, make some predictions about what we might find.

**Q1. The human genome contains roughly 20,000 genes. How many of these do you think might be differentially regulated – that is, showing meaningful changes in expression – in cells affected by Alzheimer's disease?**

1. One or two genes
2. Tens to hundreds genes
3. Thousands of genes
4. Nearly all the genes

**Q2: Think about the complexity of Alzheimer's disease and the number of cellular pathways it might impact. Would you expect just a handful of genes to be involved, or a much broader shift in gene expression?**

**Q3. When considering the impact of Alzheimer's disease on gene expression, how would you describe the changes you expect to see?**

1. Qualitative differences – clear, distinct changes like the “on/off” pattern we saw with XIST expression.
2. More subtle changes – shifts in the level of expression rather than outright presence or absence, meaning the differences are a matter of degree rather than kind.

**Q4: Reflect on whether disease-related gene expression changes are likely to be dramatic and categorical, or more nuanced and gradual.**

### Start a Lab Notebook

During this lab we're going to ask you to explore a complex dataset, and as you explore we'll ask you to take screen shots and make figures.

To keep a record of your work, start a single document where you can make figures (a Google Slide, a PowerPoint). When prompted to make figures, add each figure into this document (1 figure per page). Be sure to take notes or add captions to your figures so you know which figure is which. You can then submit a single file with all your work at the end of the lab.

### Getting started

We're going to load the SEA-AD data again, but this time we'll focus exclusively on the **microglia** from the **Middle Temporal Gyrus (MTG)**.

Instead of selecting the top item on the list, scroll down near the bottom and find “**Microglia-PVM-MTG: Seattle Alzheimer's Disease Atlas (SEA-AD)**”. Click **Explore** on that option.

<https://cellxgene.cziscience.com/collections/1ca90a2d-2943-483d-b678-b809bf464c30>

Make sure you select the **microglia from the MTG** – not the last dataset on the list, which is for the dorsolateral prefrontal cortex (DLPFC).



Figure 1

You should now be in the basic SEA-AD explorer, but zoomed in to show only the microglia from the Middle Temporal Gyrus (MTG). Your UMAP will look much simpler, displaying just the microglia cluster from the larger SEA-AD dataset. Your screen will probably look something like this:

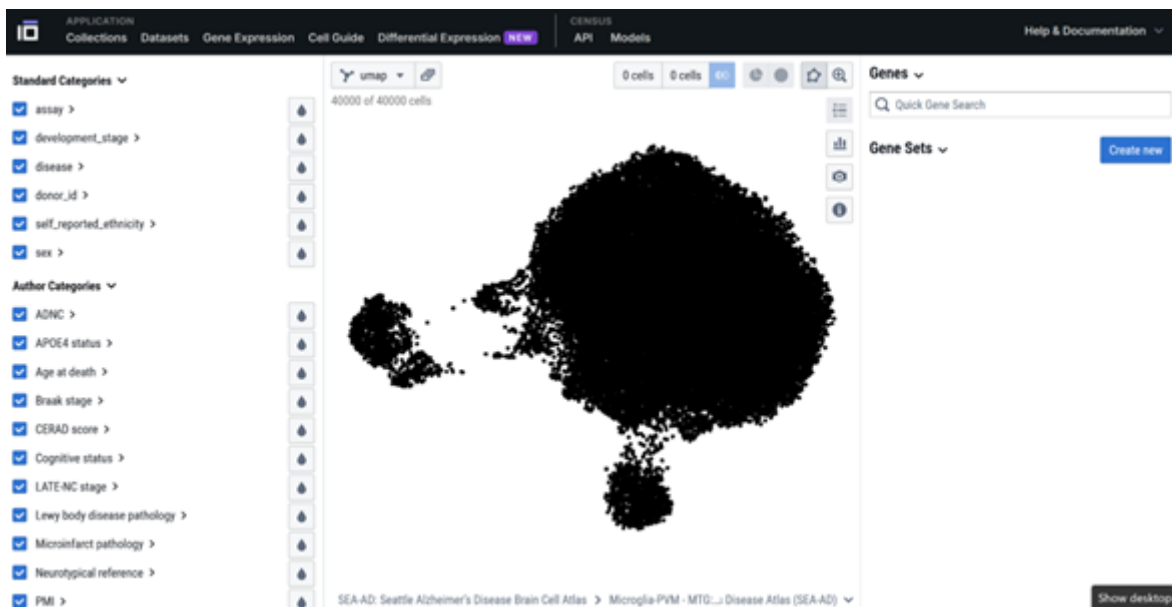


Figure 2

**Q5: How many individual cells are in this dataset?**

1. 0 cells
2. 40,000 cells
3. Impossible to tell

**Q6: Think back to the pre-lab. How many individual donors are in this study?**

1. Just 1, all of these cells are from one human
2. About 80 donors; each had many cells analyzed
3. About 40,000 - each donor had just 1 cell analyzed

### Explore 1 Gene: APOE

Let's start by looking for a difference in gene expression in a gene already known to be involved in Alzheimer's Disease: APOE.

1. Enter APOE in the Quick Gene Search Textbox

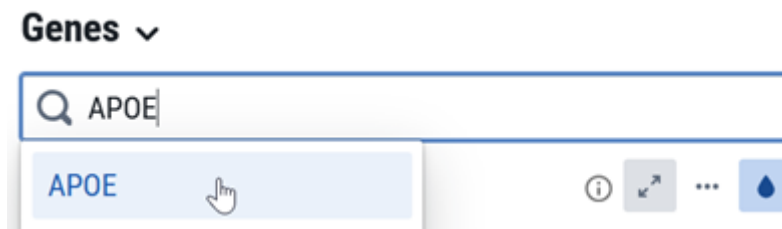


Figure 3

2. Use the water-drop tool to color code the cells by **APOE** expression. Then, expand the **APOE** gene entry to view its overall expression histogram. Take a moment to interpret what this visualization tells you about **APOE** expression in microglia. What patterns or levels of expression do you notice?

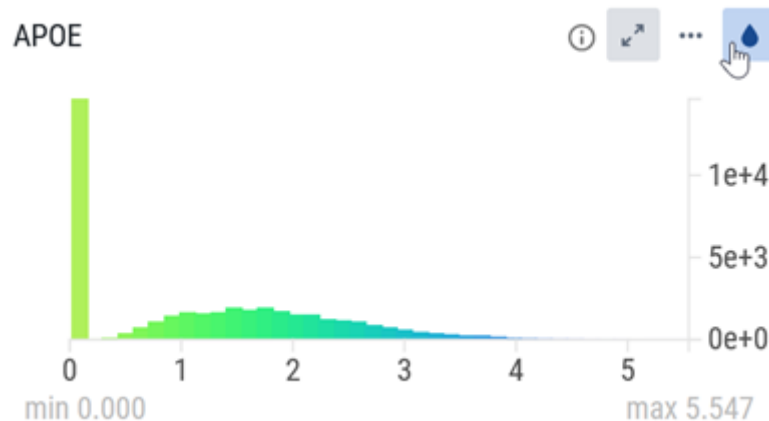


Figure 4

3. Next, investigate whether **APOE** expression differs in microglia from Alzheimer's donors compared to healthy donors. Expand the **disease** category to reveal histograms for each group. Do the histograms look noticeably different?

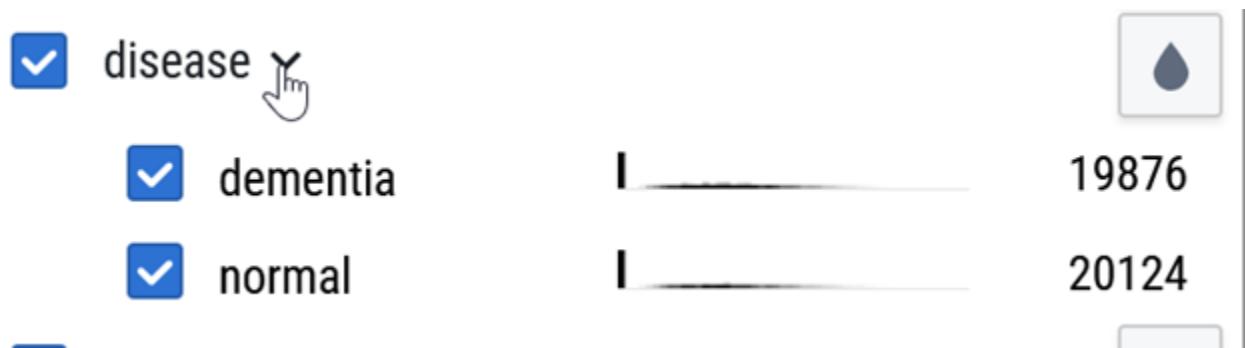


Figure 5

4. Use the check and uncheck buttons to toggle **APOE** expression visualization for each group individually. Each time you switch groups, be sure to move your mouse pointer back to the UMAP to avoid unintended highlighting of other cells.

**Q7:** Based on what you see, do the expression patterns appear different between the two groups? Do these histograms seem very different to you?

**Q8:** Create a figure comparing the two UMAPs of normal versus AD donors. The figure should look something like what you see below. This will be Figure 1

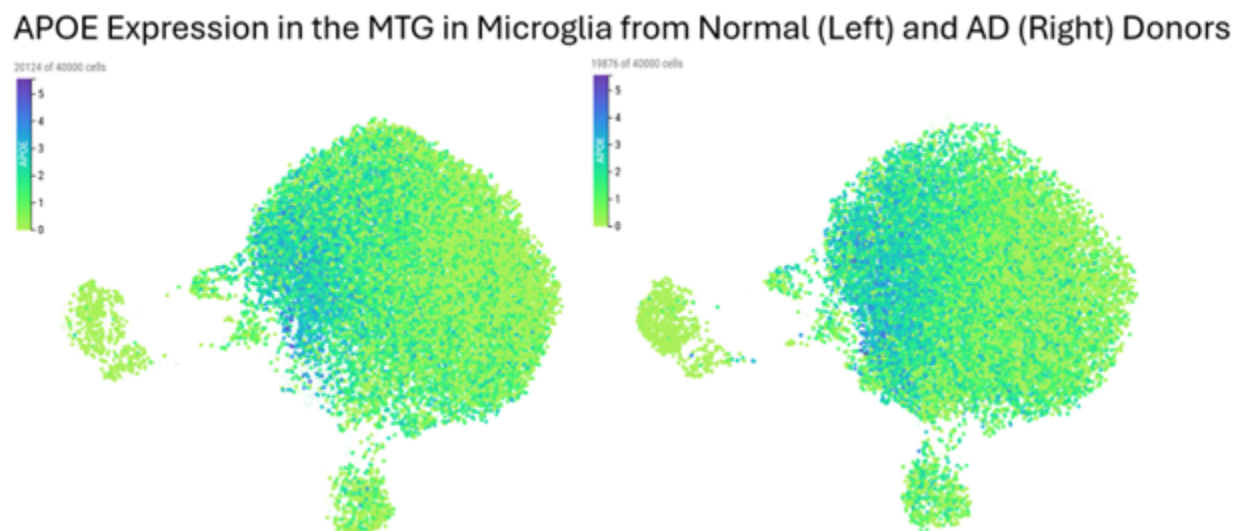


Figure 6

**Q9:** Do the APOE gene expression UMAPs look very different between microglia from Alzheimer's donors and healthy donors?

1. Yeah, microglia from AD donors have way more APOE expression.
2. Yeah, microglia from AD donors have way less APOE expression.
3. Uhhhh... I really don't see that much of a difference!

If you said you don't see a huge difference, you're right! While it would be great to download the raw data and run formal statistical tests, this difference doesn't pass the "ocular test" – meaning it's not an obvious change you can easily spot just by looking.

### Test *\*all\** the genes

We could keep exploring genes one by one to find those differentially expressed in microglia from AD donors, but with over 20,000 genes in the human genome, that would take ages! There has to be a better way.

Let's leverage statistics and computing power to compare gene expression across every gene – one at a time – and identify those with the most statistically reliable differences.

To do this, we first need to define two groups in the Cell×Gene explorer: (1) microglia from healthy donors and (2) microglia from donors with Alzheimer's disease. Once these groups are set, the explorer will handle the heavy computational work of comparing all genes between them.

Let's get started!

1. Define the set of microglia from normal donors.

In the disease roll down, uncheck the dementia box and check the **normal** box so that only microglia from the normal donors are selected. This should provide 20,124 cells.



Figure 7

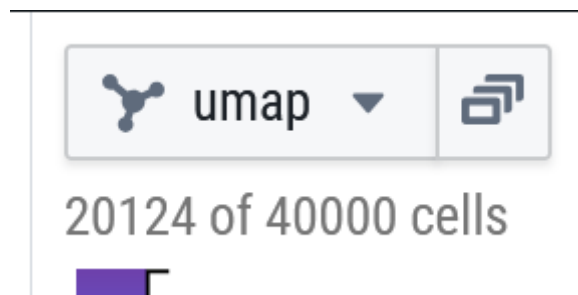


Figure 8

Now, click the Population 1 button (the left “0 cells” button) – this tells the Cell×Gene explorer that this set of cells is the first set that you want to use to compare gene expression. You'll see that it tells you that you've selected 20,124 cells for this group.

2. Define the set of microglia from the AD donors

In the disease roll down, you'll now check the **dementia** box (that's the donors with Alzheimer's) and uncheck the normal box. You'll now have 19,876 microglia selected.

Now click the Population 2 button (be careful not to re-click the Population 1 button). This should lock in the 19,876 microglia from the Alzheimer's donors for comparison, and it will make the comparison button turn blue. You can now ask the Cell×Gene explorer to compare gene expression across these two sets of microglia.

3. Conduct the differential gene expression analysis:

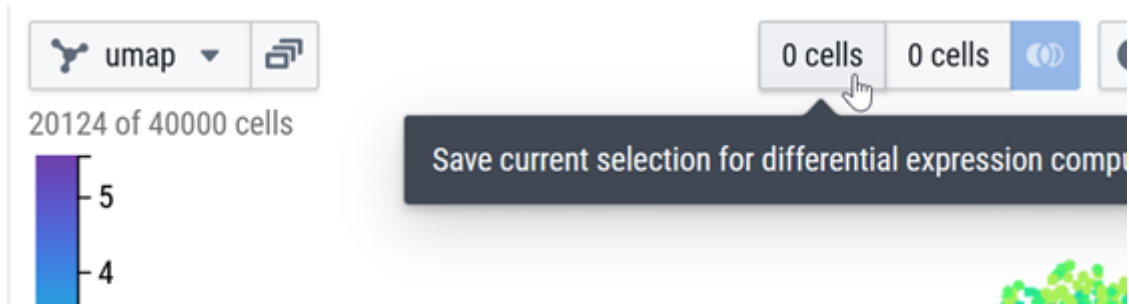


Figure 9

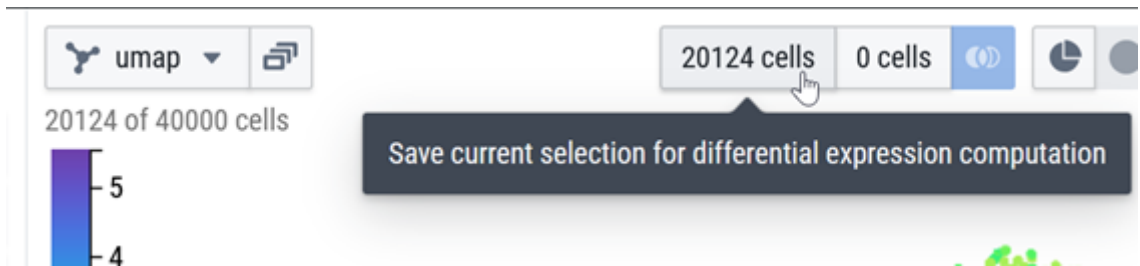


Figure 10

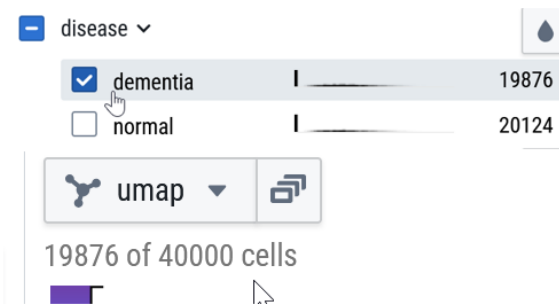


Figure 11

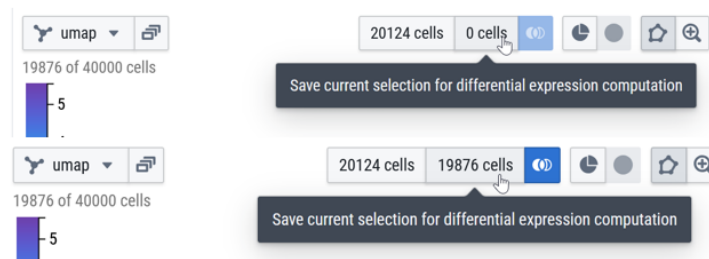


Figure 12

Click the blue **Gene Expression Comparison** button. The Cell×Gene explorer will process the data, which might take a moment.

Be aware – sometimes the tool crashes here, so you may need to restart if that happens.

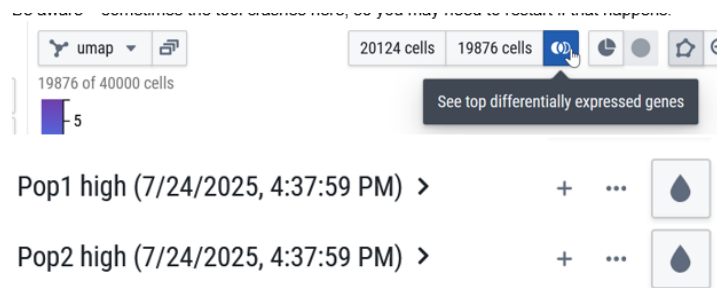


Figure 13

If all goes well, you'll see two gene lists appear on the right side:

**Pop 1 high:** Since Population 1 is the cells from healthy donors, this list shows genes with higher expression in normal microglia compared to Alzheimer's microglia.

**Pop 2 high:** This list contains genes with higher expression in microglia from Alzheimer's disease donors.

4. Rename our output to make it easier to understand

For the **Pop 1 high** list, click the three-dot menu (...) and select “**Edit gene set name and description.**” In the popup, rename it to “**Lower in Alzheimer's Microglia**”

Remember, this list contains genes that are higher in microglia from healthy donors – but it's often easier to think of them as being **lower in Alzheimer's microglia.**

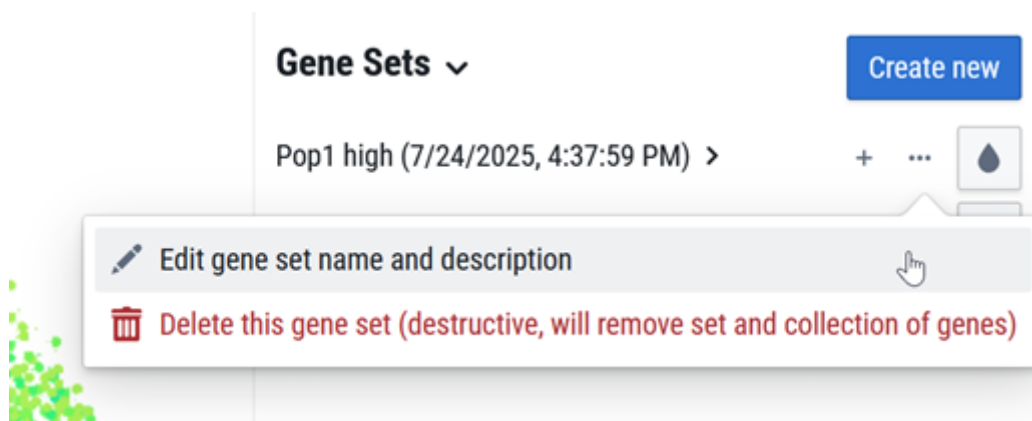


Figure 14

Now, rename the **Pop 2 high** list. Click the three-dot menu (...) for that list, select “**Edit gene set name and description,**” and in the popup, rename it to “**Higher in Alzheimer's Microglia**”

You should now have something like this on your screen:

**What is happening when you push that button? Let's talk about Stats!**

The Cell×Gene tool uses Welch's t-test to identify differentially expressed genes between groups of cells. A Welch's t-test is a statistical test used to compare the means of two independent groups, particularly when the variances of the groups are unequal. It's an alternative to the student's t-test (which you may have heard

Edit gene set name and description

Rename Lower in Alzheimer's Microglia

Lower in Alzheimer's Microglia

Edit description

Cancel

Edit gene set name and description

Figure 15

Gene Sets

Create new

Lower in Alzheimer's Microglia
+ ...

Pop2 high (8/27/2025, 11:14:18 AM)
+ ...

Edit gene set name and description

Delete this gene set (destructive, will remove set and collection of genes)

Figure 16

Edit gene set name and description

Rename Pop2 high (8/27/2025, 11:14:18 AM)

Higher in Alzheimer's Microglia

Edit description

Cancel

Edit gene set name and description

Figure 17





Figure 18

of), which assumes equal variances. Welch’s t-test provides a more reliable comparison when the assumption is violated.

We call this extensive data-profiling a **screen** – it means we can rapidly check for lots of possibilities to “screen” for interesting genes. This is a great way to discover new and unexpected things. It’s also an “unbiased” technique because we test *every* gene, not just the ones we already think might be important. The drawback, though, is that when we screen, we radically increase our risk of over-fitting – finding patterns that might have been true in this sample but which might not really generalize to all Alzheimer’s patients. In the full SEA-AD project, the authors took steps to avoid this issue (like checking other data sets and with other methods)

Thus, while a differential gene analysis is a great tool, it’s important to know that this method is useful for preliminary investigations and identifying potential markers. It’s recommended to follow up with more robust methods for formal analysis and most especially to test for the same patterns in **independent samples** to see if they hold up. While the Welch’s t-test is relatively fast and efficient, it may not be as accurate as other, more computationally intensive statistical tests when dealing with single-cell RNA sequencing data (e.g. [Ritchie et al., 2015](#)).

### Explore the Most Up-Regulated Gene, MT-ND3

Let’s start with the genes up-regulated in microglia from Alzheimer’s donors:

- Expand the “Higher in Alzheimer’s Microglia” list



Figure 19

- The first gene listed is MT-ND3. Use the eye-drop tool to color-code the UMAP for MT-ND3 expression. Also, expand this gene to see its overall expression histogram.
- Explore the expression profile for MT-ND3. Look at its overall expression histogram. Think about it: What is this histogram telling you? Do all cells express this gene?
- Compare expression between microglia from normal and AD donors:
- Roll down the **disease** filter. Compare the by-group histograms. Think about it: Do they seem that different?

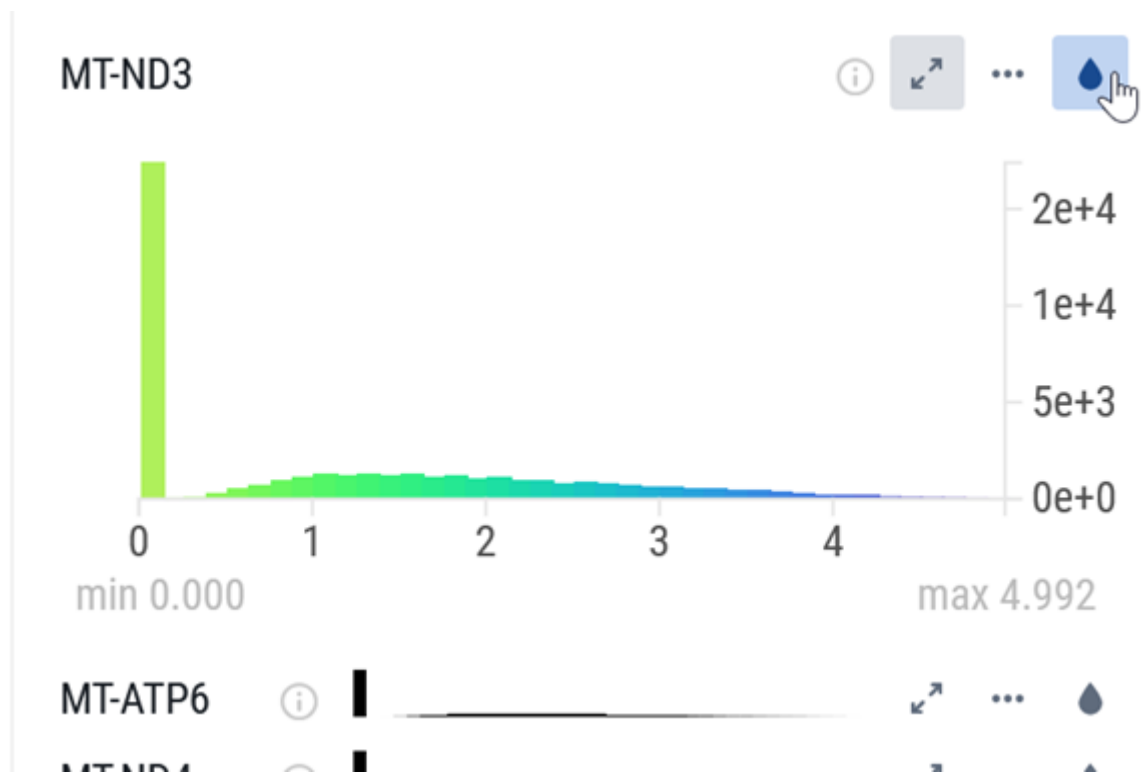


Figure 20

<input checked="" type="checkbox"/>	dementia	19876
<input type="checkbox"/>	normal	20124

Figure 21

- Isolate down to AD (dementia) or normal microglia, screen shot each, and line them up to compare.

**Q.10** Be sure to label your images so you know which image comes from which group as you will be turning this in. This will be part of Figure 2. Your figure should end up looking something like this:

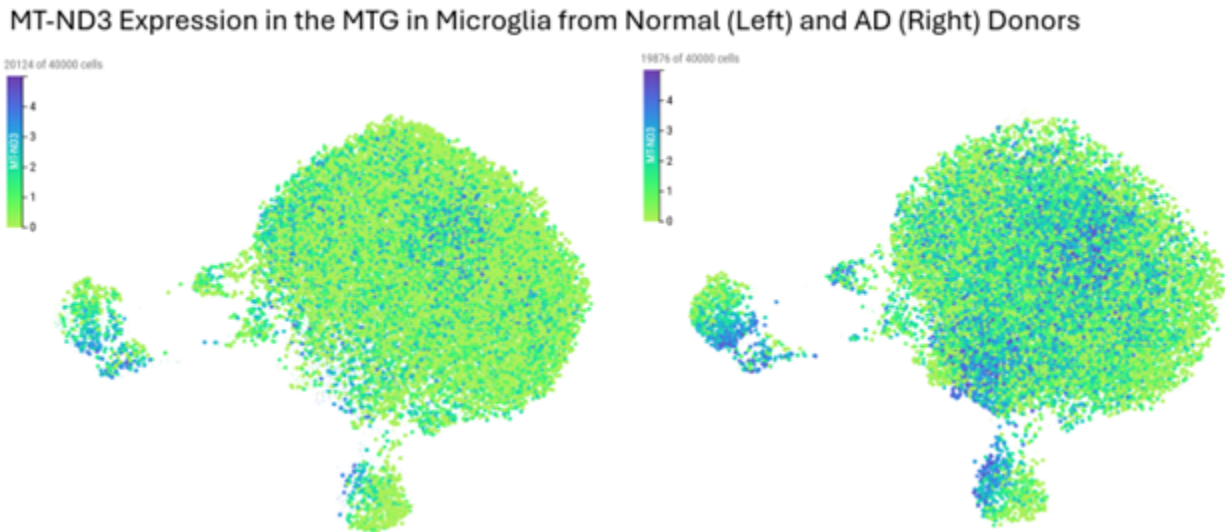


Figure 22

Think about it!

- Can you see a clear difference? Is it dramatic/qualitative or more a matter of degree/quantitative? Are all AD microglia “abnormal” (different in expression compared to control) or only some?
- What does this gene do? Use the (i) information button to find out more about it. Note: information is not always available for all genes. If no info is listed, look it up here: <https://www.ncbi.nlm.nih.gov/gene/> or here <https://www.genecards.org/>
- What, exactly, do mitochondria do?

**Q11:** Write, in your own words, about the difference in expression you are seeing for MT-ND3 based on your created Figure 1. Would you call this a dramatic change in expression or a subtle change in expression? Does it seem like most microglia have higher expression in AD or just some?

### Explore More of the Up-Regulated Genes

Keep exploring the up-regulated list. Go on to the 2nd and 3rd most up-regulated (MT-ATP6) and create another figure comparing gene expression differences between normal and AD.

**Q12:** Be sure to label your images so you know which image comes from which group as you will be turning this in. This will be the 2nd part of Figure 2 comparing the top two up-regulated genes (MT-ND3 and MT-ATP6)

Now, look at the remaining genes on the list. Think about it: Do you notice some commonality in the functions of these genes? What happens to the degree of difference as you work through the list? Do the differences get more noticeable or less? Why? How far down the list before you don’t really feel confident about a difference?

**Q13:** Based on your inspection of the list, write a short overall summary in your own words. Try to express yourself clearly and succinctly (briefly) and without jargon. Your

paragraph should help another reader understand:

- What is the association between AD and gene expression in MTG microglia?
- Are there widespread differences or are they more subtle?
- Are there consistencies in the types of genes related to AD?

**Q14:** So we will have it for later, write down the top 10 up-regulated genes. This would work best in a spreadsheet (Excel or Google Sheets). Be sure to label what your list is.

Note: The gene names are sometimes cut off in the Cell×Gene explorer, but if you hover over the name you'll get the whole thing:



Figure 23

### Explore the Most Down-Regulated Genes

Now explore the down-regulated genes:

- Browse through the list
- Make specific gene expression comparison figures (well labelled) for the **top two down-regulated genes**
- Work down the list to where you no longer feel very confident about the differences, though you don't have to make figures past the first two genes

**Q15:** Be sure to label your images so you know which image comes from which group as you will be turning this in. This will be the Figure 3 comparing the top two down-regulated genes

**Q16:** Write up a short paragraph summarizing your exploration of the down-regulated genes

**Q17:** So we'll have it for later, also write down the names of the top 10 down-regulated genes. Write this down as an additional list in the spreadsheet where you wrote down the top 10 up-regulated genes.

### Your Lab Hand-In So Far

You should now have:

1. **Figure 1:** A figure comparing APOE expression in the microglia from normal and AD donors (Q8)
2. **Figure 2:** A multipanel figure comparing the top two up-regulated genes (MT-ND3 and MT-ATP6) between normal and AD donors (Q10, Q12)
3. A short paragraph summarizing your analysis of the up-regulated genes (Q11, Q13)
4. A multipanel figure comparing the top two down-regulated genes (LINC02712 and 3X3CR1) between normal and AD donors (Q15)
5. A short paragraph summarizing your analysis of the down-regulated genes (Q16)
6. A spreadsheet with 2 labelled lists: the top 10 most up-regulated genes and the top 10 most down-regulated genes. (Q14, Q17)

## Thinking Critically About Confounds

This is a correlational study – donors were not assigned to have AD (thank goodness!). But that means we have to be especially worried about confounds – other ways in which these participants/cells might differ beyond AD status.

For example, in the control group, there are 5 female and 4 male donors (55% female). But in the AD group there are 46 female and 29 male donors (61% female). That means that the AD group is not really the same as the control group – females are over-represented. And that means that what we think is an AD difference might actually just be a sex difference. Uh oh!

We can check this worry – we can look at if our top genes that differ by AD seem to differ by sex as well.

### Examine if expression of MT-ND3 differs in the microglia of control males and females:

- In the disease drop down, check normal only
- Color code by MT-ND3
- Then in the sex drop-down, check males only and then females only, again screen shotting each group on their own and making a figure for side-by-side comparison.

**Q18: Within these normal donors, does it look like females express more MT-ND3 in their microglia? Could this be a confound? Does this complicate our interpretation that MT-ND3 is increased in AD? How?**

By the way: The SEA-AD researchers were aware of this issue, and in the paper they published they took additional steps to try to control for the imbalance of genders – science is hard!)

### Wrapping up: Some things to think about

- We saw that the comparison of Alzheimer’s Disease and normal cells was confounded by sex (the AD group had a lot higher proportion of females than the control group). Think about it: How do you think the original researchers of the SEA-AD project dealt with this issue?
- The SEA-AD project sequenced over 1.2 million cells. But these came from only 84 individual donors. What, really, is the sample size? Normally when we have millions of data points we can feel pretty confident about results generalizing, but should we be more cautious here? Why?

Chen, Shuo, Yuzhou Chang, Liangping Li, Diana Acosta, Yang Li, Qi Guo, Cankun Wang, et al. 2022. “Spatially Resolved Transcriptomics Reveals Genes Associated with the Vulnerability of Middle Temporal Gyrus in Alzheimer’s Disease.” *Acta Neuropathologica Communications* 10 (1): 188. <https://doi.org/10.1186/s40478-022-01494-6>.

Gabitto, Mariano I., Kyle J. Travaglini, Victoria M. Rachleff, Eitan S. Kaplan, Brian Long, Jeanelle Ariza, Yi Ding, et al. 2024. “Integrated Multimodal Cell Atlas of Alzheimer’s Disease.” *Nature Neuroscience* 27 (12): 2366–83. <https://doi.org/10.1038/s41593-024-01774-5>.

Keren-Shaul, Hadas, Amit Spinrad, Assaf Weiner, Orit Matcovitch-Natan, Raz Dvir-Szternfeld, Tyler K. Ulland, Eyal David, et al. 2017. “A Unique Microglia Type Associated with Restricting Development of Alzheimer’s Disease.” *Cell* 169 (7): 1276–1290.e17. <https://doi.org/10.1016/j.cell.2017.05.018>.

Papeo, Liuba, Beatrice Agostini, and Angelika Lingnau. 2019. “The Large-Scale Organization of Gestures and Words in the Middle Temporal Gyrus.” *Journal of Neuroscience* 39 (30): 5966–74. <https://doi.org/10.1523/JNEUROSCI.2668-18.2019>.