Supplementary Material 2 for Meuler M, Casimo K (2024) A Case for the Use of Open Data as a Tool to Incorporate Socioscientific Topics into Neuroscience EducationJ Undergrad Neurosci Educ 22(2):A90-A98.

Lesson 1: Brain Donation and Bioethics

Learning Objectives:

- Students will be able to reflect on the importance of bioethics within biomedical research
- Students will be able to describe the process of both living and post-mortem brain donation, and how these types of donations provide different biological data
- Students will be able to appreciate and articulate why some people may choose to not donate their brain to science
- Students will be able articulate the importance of neurodiversity within brain science
- Students will be able to reflect on the ethical implications of policies of expressed vs. presumed consent
- Students will be able to articulate the importance of consent within biomedical research

Part 1: Introduction

Biomedical research is integral to society. Learning more about the biology of humans allows us to uncover the mysteries of diseases that have large impacts on individuals and communities, from neurological diseases like Alzheimer's disease to viral infections like COVID-19. Researching pathogens and diseases that impact our communities often requires the use of **human subjects** at some point in the research process.

Although these projects enroll humans as subjects to be studied, it is imperative that this research respects the autonomy and dignity of participants throughout the research process. While the field of modern biomedical research has numerous regulations in place to prevent ethical misconduct, the field has a long and storied history of exploiting communities of color for the purposes of biomedical research (Washington, 2006).

In this lesson, you will explore the role of bioethics within the context of anatomical donations. Recognizing that the field of neuroscience has a complex history of medical racism and exploitation, this lesson provides suggested readings and resources that explore the historical perspectives of biomedical research.

While biomedical research is a broad field, this lesson will specifically focus on biomedical research involving the human brain. By detailing the process of human brain donation and its role within biomedical research, this lesson aims to foster a critical reflection on why someone may or may not choose to donate their brain to science. After exploring the factors impacting a person's decision to donate their brain, the lesson concludes by asking students to engage in a class debate on two distinct policies of consent that are used in various nations. This lesson is particularly reading heavy, though it is intended to provide you with a foundational understanding of the role brain donation plays within science. At the end of this lesson, we hope you gain an understanding of the importance of bioethics and how the field of modern biomedical sciences works to ensure that we do not repeat the injustices of the past. Human volunteers in science are integral to the field itself. Without the generous acts of individuals who choose to donate their brains to science, much of the research conducted on the human brain would not be possible.

Reference: Washington, H.A. (2006). Medical apartheid: the dark history of medical experimentation on Black Americans from colonial times to the present. New York: Doubleday.

Feeling Safe and Comfortable Participating in Research

The field of neuroscience has the unique yet challenging task of studying the most complex piece of organized matter in the known universe. Attempting to uncover the mysteries of the brain requires a deep ethical consideration of how to study something as personal and unique as the human brain.

Unfortunately, the role of ethics and morals within scientific research has not always been respected. The field of science has a long and storied history of violating the rights and dignity of individuals who did not properly consent to participating in scientific research. Historically, the individuals who bore the majority of harm inflicted by science were more often than not from communities of color and/or individuals who identified as women. Despite this history of scientific harm, communities of color continue to be disproportionately impacted by issues involving bioethics.

Understanding the history of biomedical ethics and the disproportionate impact nonconsensual research has had on communities of color is imperative in order to ensure that modern biomedical research does not repeat the harms of the past. These biomedical harms are relevant to both the field of neuroscience and other fields of science more broadly. We would need several lessons in order to adequately cover the history of injustices inflicted by science, as the history of science is an entire field in and of itself! Although we do not have time in this lesson to dive deeply into the history of biomedical research, there are several great resources you can reference to learn more about this subject:

Suggested readings:

- *The Immortal Life of Henrietta Lacks* by Rebecca Skloot
- *Medical Apartheid: The Dark History of Medical Experimentation on Black Americans from Colonial Times to the Present* by Harriet A. Washington
- *The Mismeasure of Man* by Stephen Jay Gould
- *Superior* by Angela Saini
- *Inferior* by Angela Saini

Understanding the history of science is crucial to ensure that the field of modern biomedical research does not repeat the harms of the past. In this lesson, we will explore how consent and medical ethics both play crucial roles in how modern neuroscience ensures that its research is conducted in an ethical manner. Studying the human brain is an exciting yet challenging endeavor, and thus, bioethics plays a key role in guiding this crucial and important research.

How do we study the human brain?

Given that research on the human brain requires samples of human brain tissue and/or human brains to study, several ethical questions arise. **How** do we study the human brain, and **whose** brains do we study?

While some scientists use model organisms, such as mice, to study the brain, these models by themselves are not sufficient for fully understanding the human brain. Instead of relying solely on mice, neuroscientists seek out opportunities to ethically study the human brain. Often, these researchers rely on individuals who are willing to donate their brain to science following their death. In order to explore the process of brain donation from a bioethical and sociological perspective, we will use the Allen Institute for Brain Science as a case study.

The **Allen Institute for Brain Science** is a division of the Allen Institute, a nonprofit scientific research institute that was founded in 2003. The mission of the Allen Institute for Brain Science is to focus on defining and understanding cell types of the mammalian brain to ultimately better understand brain development, evolution, and disease. The Allen Institute for Brain Science is working towards creating a complete list of the different "parts" – cell types – in the brain. In addition to creating a list of the brain's parts and cell types, the Allen Institute also aims to determine how these cell types connect and function in the brain and what changes happen in cells in the aging brain and in neurodegenerative diseases such as Alzheimer's disease (AD). The Allen Institute for Brain Science conducts basic research on both healthy and diseased brains. This type of basic research means that scientists study healthy and diseased brains in order to gain a better understanding of the progression of Alzheimer's disease. The Allen Institute does not conduct clinical trials on AD or other types of applied biomedical research. In order to carry out this research and achieve this mission, the Allen Institute relies heavily on the generous act of individuals in the community who donate their brains to science.

In order to explore the importance of brain donation, this lesson will address four main questions:

1. *What is brain donation and how does it differ from other types of biomedical donations, such as organ donation?*

2. *How do beliefs concerning brain donation differ on the basis of an individual's cultural background?*

3. *Why is there no such thing as a "normal" brain?*

4. *What is the role of consent within brain donation?*

In many ways science helps individuals, whether that is through developing treatments for disease or providing insights into how the world works. However, equally important to how science helps people is to think about how people can help science. Brain donation is one small example of how science cannot be done without communal support and involvement. We will explore this concept in the activities that follow!

Part 2: What does it mean to donate your brain to science?

Brain donation is the process of granting permission for your brain and/or brain tissue to be used for the purposes of scientific research. Although the brain is indeed an organ, brain donation is a different process than **organ donation**. While being an organ donor means that your organs will be given to other people in need of transplants to keep them alive, being a brain donor means your brain will only be used for research purposes. The process of brain donation is also different than **whole-body donation** to science. Whole-body donation is a process where an individual consents to their body being used for biomedical research on their other organs or for their body to be given to a medical school for instructional purposes. Thus, there are three distinct processes for choosing to consent to brain donation, organ donation, or whole-body donation. These processes are also mutually exclusive, meaning that donors can only choose one type of donation. For example, if you elect to donate your brain to science, you are ineligible to be an organ donor.

The research done at the **Allen Institute for Brain Science** involves two different types of generous brain donations from individuals:

1. Post-mortem brain donation

2. Living brain tissue donation

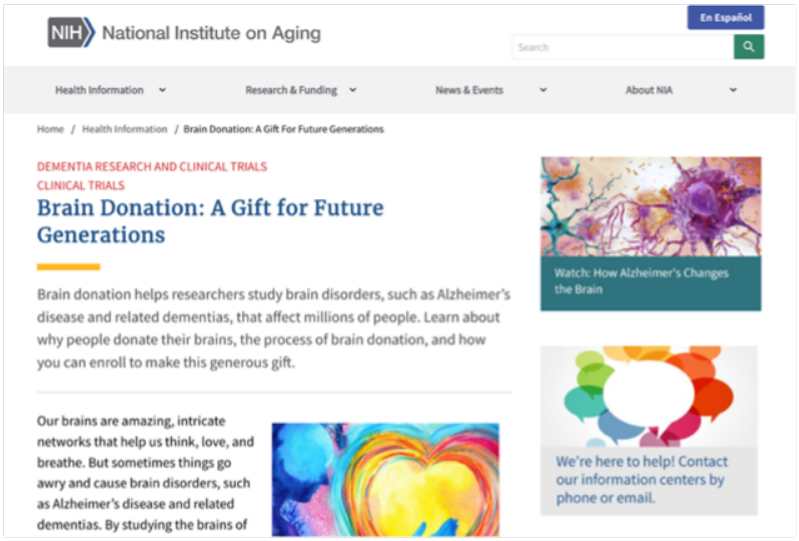
Post-Mortem Brain Donation:

Post-mortem is a Latin phrase meaning "after death." Post-mortem brain donation is the process whereby an individual chooses to sign up for their brain to be donated to science once they have passed away. Post-mortem donation allows scientists to gather crucial data about the brain's activity after cell death, such as which genes it was expressing and in what quantities.

In order to gain a better sense of post-mortem brain donation and its importance to science and society at large, take 5 minutes to read the following article. After reading through the article, answer the reflective questions that follow.

Article: This article is linked here, or the full link is listed below.

https://www.nia.nih.gov/health/brain-donation-gift-future-generations



Reflective Questions:

- **Why do you think the National Institute on Aging refer to post-mortem brain donation as a "Gift for Future Generations?"**

- **Who does the National Institute on Aging consider to be "high priority" individuals for brain donation? Why are these individuals identified as especially important to add to the brain donor bank?**

Living brain donation:

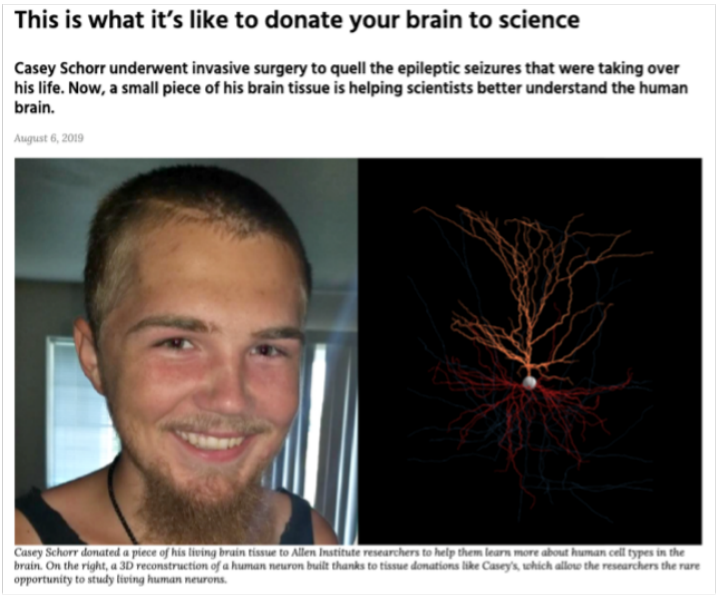
While post-mortem brain donation is available to everyone, everywhere, living brain donation is possible for a few individuals undergoing specific types of neurosurgical procedures in cities with living tissue donation programs. Living brain donation is a process where a small section of living brain tissue is taken from a living donor. How is this possible?

These donations only take place when individuals who are already scheduled for brain surgery opt in to provide a small piece of the brain tissue to researchers. This is only an option when the individual is already scheduled for a procedure where a portion of their brain tissue is being removed. For example, an individual who is having brain surgery for epilepsy to remove the site generating the seizures could choose to be a living brain donor. Because the surgery for epilepsy already requires removal of some healthy brain tissue that the surgeons must remove in order to reach the site generating the seizures, the individual undergoing the surgery could choose to donate this removed tissue to researchers. Individuals who are not already undergoing a surgery that requires removal of brain tissue are not eligible to be living brain donors.

Living brain tissue provides a unique opportunity for scientists to study the activity of brain cell's while the brain tissue is still alive. At the Allen Institute for Brain Science, researchers have been able to keep donated living brain tissue alive for up to 5 days.

10 minute activity:

In order to understand the process of living brain donation, read this short article written by the Allen Institute's Rachel Tompa, PhD, about the generous donation made by Casey Schorr. After reading the article, be sure to answer the reflective questions that follow.



This is what it's like to donate your brain to science

Casey Schorr underwent invasive surgery to quell the epileptic seizures that were taking over his life. Now, a small piece of his brain tissue is helping scientists better understand the human brain.

August 6, 2019

Casey Schorr donated a piece of his living brain tissue to Allen Institute researchers to help them learn more about human cell types in the brain. On the right, a 3D reconstruction of a human neuron built thanks to tissue donations like Casey's, which allow the researchers the rare opportunity to study living human neurons.

If the embedded link above does not work, here is the full link to the article:

https://alleninstitute.org/what-we-do/brain-science/news-press/articles/what-its-donate-your-brain-science

Reflective Questions:

- **What can scientists study with living brain tissue that they cannot study with post-mortem brain tissue?**

- **Think about Casey's story and his experience with living brain donation. What reasons did he list for why he donated his living brain tissue to science?**

- **In addition to those listed by Casey, can you think of any other reasons as to why someone may opt to donate part of their living brain to science?**

- **Why do you think someone would opt not to donate part of their living brain tissue to scientists?**

- **Spend 5 minutes and discuss your answers to the reflective questions above with a classmate, a friend, a family member, etc.**

Optional: For more information about living brain donation, read this additional article by Rodrigo Pérez Ortega about research being conducted at the Allen Institute that uses living human brain tissue donations: https://www.science.org/content/article/how-donate-piece-your-brain-science-while-you-re-still-alive

Now that we understand the difference between living and post-mortem brain donation, let's discuss some factors that may influence a person's decision to donate their brain to science.

Reasons why someone may choose to not donate their brain to science:

While donating your brain to science as either a living or post-mortem donor is incredibly helpful to science and the advancement of biomedical research, it is important to note that some people may choose not to participate in brain donation.

Although there are a variety of reasons as to why someone may opt not to donate their brain to science, cultural and/or religious beliefs and practices are major factors that may dissuade someone from being a brain donor. Reasons as to why someone may object to donating their brain to science could include beliefs that one's body must remain intact for proper burial and/or travel to an afterlife or cultural norms that dissuade individuals from discussing death-related topics. Recalling what we learned at the start of this lesson, it is also important to recognize that historical injustices within the field of medicine against communities of color can also play a role in why an individual may not feel comfortable donating their brain to science. No matter what a person cites as a reason that they would opt not to be a brain donor, their decision should be respected. Scientists greatly appreciate the donations of those who choose to participate, but we do not want anyone to feel pressured to make a decision about donation unless they are comfortable doing so.

What types of brains does science "want" to study?

A common misconception about brain donation is that science only needs to study brains that have a specific characteristic. For example, some people tend to think that if scientists are studying dementia, then those scientists only want people who have dementia to donate their brains. This is not true! Science benefits when people with ALL types of brains sign up for brain donation.

But what does it mean to say that science benefits from the donation of *all types of brains*?

No such thing as a "normal brain:"

Word choice is incredibly important within science. For this reason, it is important to clarify what we mean when we discuss what a "healthy" brain is. A healthy brain refers strictly to disease pathology. This means a healthy brain is one that does not display symptoms or biological signs that are characteristic of known neurological diseases. There are many different types of neurological diseases, such as Alzheimer's disease, Parkinson's disease, and Huntington's disease.

"Normal" is a word that should be avoided when referencing human brains and/or bodies. Specifically in terms of the brain, a "healthy" brain is not synonymous with a "normal" brain. In fact, many scientists support the theory that there is no such thing as a "normal" brain. Instead, many scientists have called for a widespread recognition of neurodiversity and its role within the field of neuroscience.

In order to explore the danger of using the word "normal" to reference brains, please take 5 minutes to read this short article by Thomas Armstrong in the *AMA Journal of Ethics*.

If the embedded link does not work, the full link is listed here:

https://journalofethics.ama-assn.org/article/myth-normal-brain-embracing-neurodiversity/2015-04

doi: 10.1001/journalofethics.2015.17.4.msoc1-1504

Reflective Questions

- **If someone asked you to define what a "normal" brain was, what would you say? Do you think it is possible for someone to have a normal brain?**

- **Why would it be beneficial to have someone who identifies as neurodivergent consent to donate their brain to science?**

As explained by Armstrong, neurodiversity is a key part of neuroscience. One aspect of neurodiversity is the idea of neurodivergence. Neurodivergent is a term that was coined in the 1990s by Judy Singer. Nicole Baumer and Julia Frueh of Harvard Health Publishing define **neurodiversity** as "the idea that people experience and interact with the world around them in different ways; there is no one 'right' way of thinking, learning, and behaving, and differences are not viewed as deficits." This understanding of neurodiversity highlights the importance of having a diverse range of individuals donate their brains. There is no such thing as a "normal" brain. Instead, scientists will often differentiate brains based on two distinct criteria:

• To explain whether or not a brain displays signs of known diseases, scientists will describe a brain as either healthy or diseased

• To explain an individual's brain function and behavior, scientists will often use terms such as "**neurotypical**" or "**neurodivergent**"

Any combination of these two criteria is possible. For example, it is possible for someone to have BOTH a healthy and neurodivergent brain. Given the vast amount of neurodiversity that exists among the population, science benefits when anyone and everyone decides to donate their brain!

Reference: *https://www.health.harvard.edu/blog/what-is-neurodiversity-202111232645*

Want to learn more about neurodiversity? Here are a list of other resources you can explore outside of this lesson:

• **Article:** *What is neurodiversity*? by Nicole Baumer and Julia Frueh https://www.health.harvard.edu/blog/what-is-neurodiversity-202111232645

• **Book:** *Out of My Mind* by Sharon Draper

• **Book:** *Furiously Happy* by Jenny Lawson

• **Podcast:** *Being Seen*

Part 3: Understanding the Role of Consent in Brain Donation

At the start of this lesson, you read about the importance of ethics within the context of biomedical research. Discussions of ethics within research also require a consideration of how **consent** will be obtained from study participants. Obtaining consent is essential to ensure that participants are aware of what their brain tissue will be used for, the purpose of the study to which they are donating their brain, and other details about the research itself. The ethical guidelines surrounding consent and brain donation differ based on what state and/or country you reside in. Comparing these policies to one another provides valuable insight into the field of bioethics.

Washington State:

The Allen Institute is located in the state of Washington. In Washington, donating your brain to science falls under the Uniform Anatomical Gift Act: https://app.leg.wa.gov/rcw/default.aspx?cite=68.64. This act specifies who is authorized to provide consent for brain donation, which can vary depending on whether or not the donor is still alive or if the donor has passed away.

United States:

The United States operates from an "**expressed consent**" policy for organ and brain donations. This means that people are assumed to NOT be an organ or a brain donor until a person explicitly signs up and consents to be one. For organ donation, you can often sign up for this when getting a driver's license in the United States. The policy of expressed consent is also practiced in countries such as Denmark, Brazil, Canada, and the United Kingdom. For brain donation, there are several different ways through the NIH NeuroBioBank to sign up to be a brain donor. The Allen Institute for Brain Science recommends https://braindonorproject.org/ for inquiring about the process of donating your brain based on which state you reside in.

Globally:

Some countries, such as Spain, Belgium, France, Norway, and Singapore, operate from a "**presumed consent**" policy, also known as an "opt-out" policy. Instead of needing to sign up to donate your organs, you are assumed to consent to donating your organs unless you explicitly opt-out of doing so.

A policy of presumed consent requires individuals to remove themselves from the list of eligible donors, whereas a policy of expressed consent requires individuals to add themselves to the list of eligible donors.

Reference: https://journalofethics.ama-assn.org/article/presumed-vs-expressed-consent-us-and-internationally/2005-09

Activity: Bioethics Policy Evaluation

Now that you have heard about the different bioethical policies of expressed vs. presumed consent, it is now time for you to evaluate the benefits and drawbacks of these two approaches. In addition to the information provided in this lesson, feel free to find outside resources that further explain potential pros/cons of each of these policies.

After researching these two policies further, please fill out the following table:

	A policy of presumed consent	A policy of expressed consent
Pros		
Cons		

Post-Debate Reflection:

1. After learning about both expressed and presumed consent, which policy do you think is best practice from a bioethical perspective? Be sure to explain your answer in detail below.

2. List at least two reasons why consent is imperative within the context of biomedical research:

3. After what you have learned in today's lesson, do you think you would donate your brain to science? Why or why not?

Choosing to donate or not donate are both completely valid positions and there is no one "right" answer. Each individual has the right to decide whether or not brain donation is something that aligns with their beliefs.

Conclusion:

Understanding the role of consent within biomedical research is a crucial part of science. Given the history of injustices perpetuated by scientific research against individuals, particularly people of color and/or individuals who identify as women, it is imperative to reflect upon the importance of ethics within research. We hope you walk away from this lesson with an appreciation for the field of bioethics and the bioethical considerations that go into the work done at the Allen Institute for Brain Science.

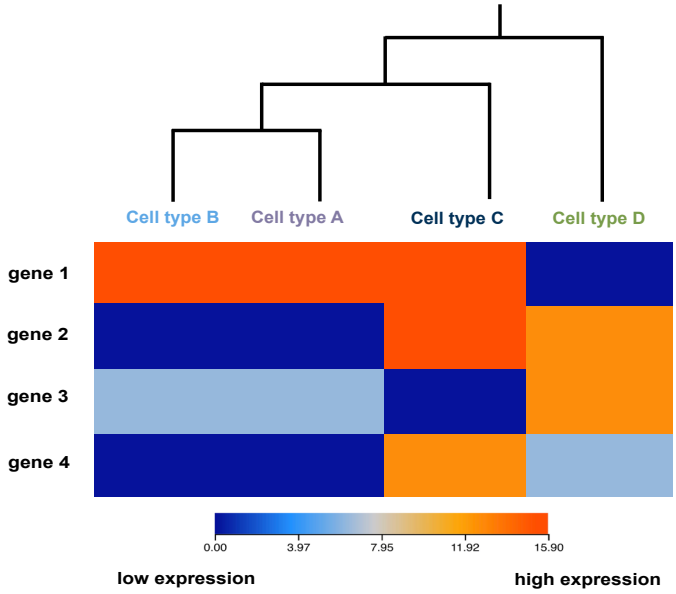
If you continue on to complete the following three lessons in the Cell Types, Health, and Disease: An Interdisciplinary Exploration of Alzheimer's Disease lessons, you will have the chance to explore the open data from the Allen Institute that was obtained from 84 donors who consented to post-mortem brain donations. Some were healthy and some of them had Alzheimer's Disease. As you will see, the generous donation from these 84 individuals has allowed researchers at the Allen Institute to gather a large quantity of highly detailed data on the healthy human brain and pathological changes that may be associated with Alzheimer's disease. Without the 84 individuals who consented to donate their brain, none of this research would be possible.

If you develop your own lesson plan using Allen Institute resources, we invite you to share your experience with us at communications@alleninstitute.org. Teachers are also encouraged to publish original lessons using our open data, tools, and other resources, and to share those lessons with us.

Supplementary Material 2 for Meuler M, Casimo K (2024) A Case for the Use of Open Data as a Tool to Incorporate Socioscientific Topics into Neuroscience EducationJ Undergrad Neurosci Educ 22(2):A90-A98.

ALLEN INSTITUTE

Lesson 2: The Importance of Basic Research in Brain Science



Learning Objectives:

- Students will be able to articulate what transcriptomic data is and how it is gathered
- Students will be able to articulate how dendrograms and heatmaps can be used in conjunction with transcriptomic data in order to further differentiate cell types from one another
- Students will be able to apply basic principles of interpreting data visualization to complex transcriptome datasets
- Students will explore the nuanced differences between basic and applied research
- Students will be able to defend the importance of both basic and applied research within the field of biomedical science

Introduction:

If you completed Lesson 1, you learned about the process through which an individual can donate their brain to science. Keeping what you learned about brain donation in mind, you will now explore what type of basic research underlies the field of neuroscience.

The Allen Institute, located in Seattle, Washington, is a nonprofit scientific research center that specifically prioritizes basic research. **Basic research**, also known as **foundational research**, is curiosity-driven research that aims to further our understanding of a particular phenomenon. For example, studying the healthy human brain is one type of basic research, where scientists further our understanding of the brain itself, rather than focusing on treating disease states or other concrete applications. A "healthy" brain is a term that refers strictly to a brain that is not known to have a neurological disease of some kind. "Healthy" is not synonymous with "normal." In fact, given the wide array of neurodiversity that exists within the population, "normal" is not a term used in science to describe brains. For more information about neurodiversity in brain science, see *Lesson 1: Brain Donation and Bioethics*.

While studying a healthy human brain is considered a type of basic research, the field of neuroscience also conducts a significant amount of applied research. Applied research specifically seeks to understand options for treatment and/or a cure for a particular disease. Both basic and applied research are necessary in order to further our scientific understanding and improve human health. Basic science helps provide a foundation of knowledge for later translational research and is crucial for the field of biomedical science as a whole. Studying a healthy, neurotypical brain allows us to understand what may change later during the onset of disease or the development of neuroatypical characteristics. In addition to these practical applications, basic research is also helpful for gaining a foundational understanding of the healthy human brain to expand our general knowledge of the field of neuroscience.

While both basic and applied research are integral to the field, science is often limited by time, money, and resources. **Given these restrictions, how do scientists decide whether to conduct basic or applied research?**

In this lesson, we will use current research from the Allen Institute for Brain Science as a specific case study in the process of conducting basic scientific research. Founded in 2003 by Paul G. Allen, the Allen Institute for Brain Science is a division of the larger Allen Institute.

The Allen Institute's mission is to conduct big, open, and team science. **Big science** means conducting large-scale, basic research projects that aim to uncover the mysteries of biological systems such as the brain, the immune system, and the cell. **Open science** means that much of the data generated by the Allen Institute is made openly available to the public. Lastly, **team science** means that each of the Allen Institute's projects relies heavily on scientific collaboration and large teams of people, versus a more traditional lab model where a Principal Investigator (or PI) leads a group of people who may not have the chance to collaborate with other labs.

In this lesson, you will have the chance to explore some open data available on the Allen Institute for Brain Science's website. In order to further understand the importance of basic scientific research, you will have the chance to interpret the transcriptomic data of the healthy human brain. By looking at dendrograms and heatmaps, you will strengthen your data analytic skills while also exploring why it is so important to further our understanding of the healthy human brain and its composite parts!

Activity 1: What would you fund?

Before we dive into looking at the actual Allen Institute for Brain Science's data, we want to do a quick exercise that prompts you to consider what type of research, you, yourself, would choose to fund. The type of research scientists can pursue is often restricted by limited resources and opportunities for funding. Given the limited availability of funding, scientists must communicate why their research is important and the impact it would have on society. This can pose a challenge for scientists pursuing basic research projects, since funding agencies typically are looking for research that will have specific implications for disease. This activity will provide you with the opportunity to explore how scientists in both applied and basic research can advocate for the importance of their research.

Scenario: You are given 10 million dollars to invest. Two teams of scientists approach you with a grant proposal. The grant proposals are written documents that outline these groups of scientists' plan for what they will study, how they will study it, and WHY you should choose to invest your money to help them study it.

Proposal 1: The scientists in group 1 plan to study a possible treatment for Alzheimer's disease (AD). AD is a neurological disease that results in a decline in memory skills and cognitive functioning and is estimated to impact 24 million people globally. They are interested in testing whether a drug that is thought to break down phosphorylated tau protein works to treat AD. Scientists hypothesize that phosphorylated tau protein form neurofibrillary "tangles" that accumulate with neurons and disrupt normal functioning.

Reference: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3405821/#:~:text=In%20the%20US%2C%20approximately%205.5,as%20high%20as%2024%20million

Proposal 2: The scientists in group 2 plan to study the varying cell types of the human brain. These scientists define cells as different "types" when they have different structure/forms (morphologies), gene expression, and electrophysiological activity. These scientists will study brain tissue donated from healthy, neurotypical individuals who had no known neurological diseases or conditions. The output of this research will be descriptions and classifications of different cell types within the human brain. This research hopes to further science's foundational knowledge of the structure and function of the human brain.

Activity: Given these descriptions of the two proposed projects, fill out the table below.

	Potential benefits	Potential limitations/drawbacks
Proposal 1: Researching a treatment for Alzheimer's Disease		
Proposal 2: Researching the cell types of the human brain		

Final Decision:

Now it's time for you to make your investment decision. How do you plan on allocating your money? Is one group getting all 10 million dollars, are you splitting it evenly, or are you distributing it in a particular manner between the two groups? Be as specific as you can in explaining how you will allocate your money and WHY you made the decision you did: (1 paragraph)

This activity was meant to encourage you to think about the benefits and/or drawbacks of both basic and applied research. There is no one right answer as to how money should be allocated between the two, as they both have their own unique role in the field of science.

Basic Research in Practice: Studying a Healthy Human Brain

Now that you have had the chance to explore the importance of basic research, let's dive into basic research in practice by looking at some of the open data made available from the Allen Institute for Brain Science. In particular, we will be looking at the data the Allen Institute for Brain Science has collected from the **Middle Temporal Gyrus (MTG)** of five healthy, neurotypical donor brains. While the Institute has collected a diverse range of data from each of these donor brains, this lesson will specifically be looking at the **transcriptomic data** collected from the brain cells in each donor's MTG.

In order to understand what transcriptomic data is, let's first establish the difference between a **genome** and a **transcriptome**.

Genome vs. Transcriptome:

When you hear the term "genome," what do you think of?

A genome is a complete library of the genes/genetic material present within somatic cells. Somatic cells are the cells within the body that are not sperm or egg cells. The National Human Genome Research Institute defines a genome as "a fancy word for all your DNA."

Think about your own genome and the somatic cells within your body. Somatic cells could be anything from the cells that make up your skin to the cells that make up your brain.

Knowledge Check

- **Do your skin cells carry a different genome (set of genes) than your brain cells? In other words, do skin cells have different genes than brain cells?**

- **But if skin cells and brain cells contain the same genes, what makes one a skin cell and the other a brain cell?**

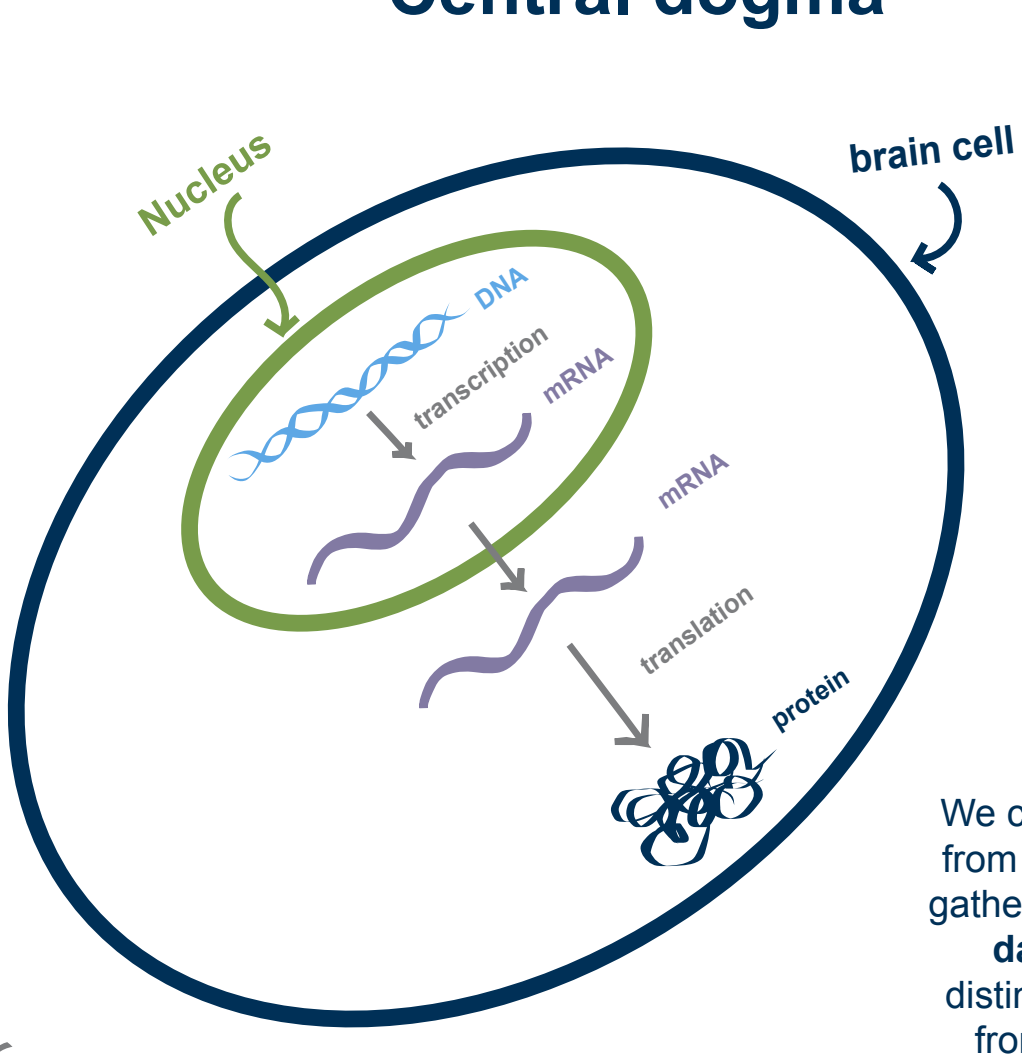
Even though your skin cells and brain cells contain the same genes, this does not mean that they **express** all of the same genes.

How can we measure gene expression?

In order to answer this question, scientists gather **transcriptomic data**.

What is transcriptomic data?

One method scientists use to differentiate cell types from one another is to collect transcriptomic data. This transcriptomic data identifies **which genes a cell is transcribing into RNA transcripts** and in **what quantities**. If a cell, and more specifically, that cell's nucleus, contains a specific RNA transcript, this indicates that the cell is expressing the specific gene associated with that RNA. By (1) isolating nuclei, (2) sequencing the mRNA transcripts found within the nuclei, and (3) counting those transcripts, we can tell **which genes** the cell is expressing and **how much** these cells are expressing these genes. The figure below provides a quick reminder of what transcription is within the context of the central dogma of biology:

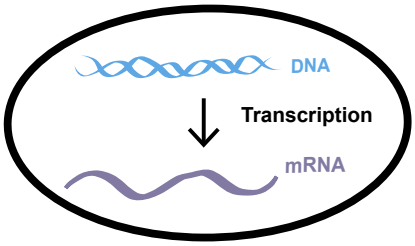


We can isolate nuclei from brain tissue and gather **transcriptomic data** to help us distinguish cell types from one another

Transcriptomic data is growing increasingly popular within the field of neuroscience. Before you can analyze transcriptomic data, it is important for you to understand both what it is and how it is gathered. The figure below details how scientists gather transcriptomic data and use it to study different types of cells, also called "cell types," within the brain.

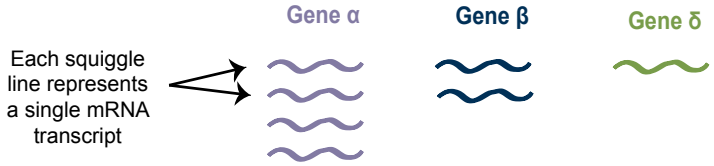
Transcriptomic data is a key tool used to construct a **transcriptome** for a particular cell. What is the difference between a genome and a transcriptome? While a genome is a complete catalog of all genes available for every cell in the body, a transcriptome is a catalog of the RNA transcripts found within a cell at a specific point in time. A transcriptome is often used as a method of measuring gene expression within a cell.

Transcriptomic data is gathered by following the steps outlined in the figure below:



Isolate the nuclei from the cells in the sample of brain tissue and extract the RNA found in each nucelus.

Sequence the mRNA transcripts found in each cell's nucleus in order to determine which genes each brain cell was expressing.



Count the number of mRNA transcripts found for each gene. This allows us to quantify how much each cell was expressing each gene.

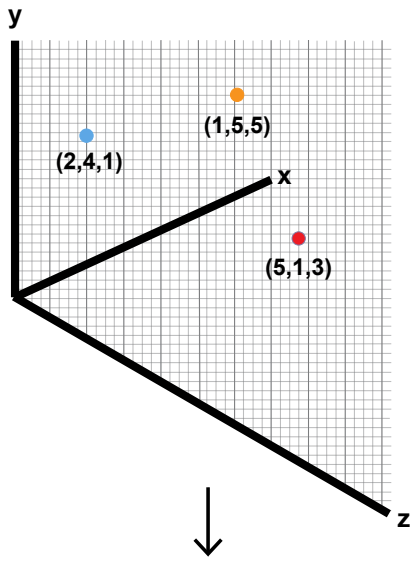
	Gene α	Gene β	Gene δ
# of mRNA transcripts found in Cell 1	4	2	1

Create a table that shows how much each cell was expressing each gene.

	Gene α	Gene β	Gene δ
# of mRNA transcripts found in Cell 1	2	4	1
# of mRNA transcripts found in Cell 2	1	5	5
# of mRNA transcripts found in Cell 3	5	1	3
repeat count for thousands of cells...	...	...	...

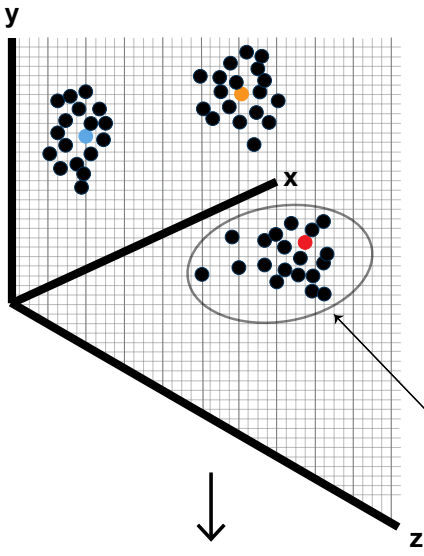
Repeat this process for THOUSANDS of cells. Remember, this means we are counting how much EACH cell was expressing EACH gene. If we wanted to create a table that listed the data in full, this data table would have thousands of rows.

x = gene α value
y = gene β value
z = gene δ value



If we wanted to create a graph that plotted the inital data for cell 1, cell 2, and cell 3 and their relative amount of expression of gene alpha, gene beta, and gene delta, we would need a 3D graph like the one on the left.

x = gene α value
y = gene β value
z = gene δ value



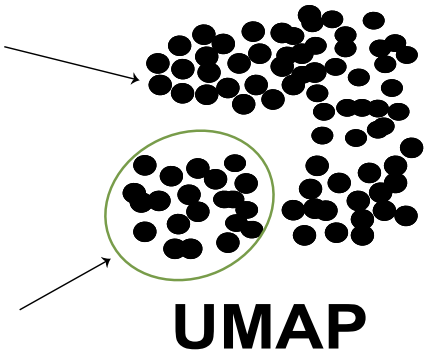
We can repeat this process for the thousands of cells that were collected from the brain tissue sample. Notice that the cells begin to cluster based on how similar their gene expression for gene alpha, gene beta, and gene delta is to one another. These clusters help us identify which cells may be more similar and/or dissimilar to one another!

when we plot the gene expression data for more cells, we notice that cell 3 (red dot) clusters next to these other cells from the sample

	Gene α	Gene β	Gene δ	repeat for thousands of genes...
# of mRNA transcripts found in Cell 1	2	4	1	...
# of mRNA transcripts found in Cell 2	1	5	5	...
# of mRNA transcripts found in Cell 3	5	1	3	...
repeat count for thousands of cells...	...	...	...	...

In addition to collecting data on gene expression for thousands of cells, scientists will add another layer of complexity by measuring the gene expression of these thousands of cells for THOUSANDS of genes. A table displaying this data would have thousands of rows and thousands of columns. Since the graph would now have much more than just 3 dimensions, we will need a special type of tool to graphically represent this data in a way that humans can visualize.

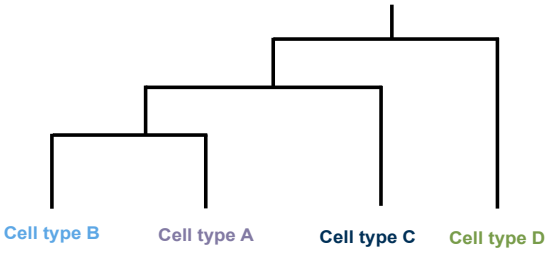
Each dot represents a single nucleus isolated from a single brain cell



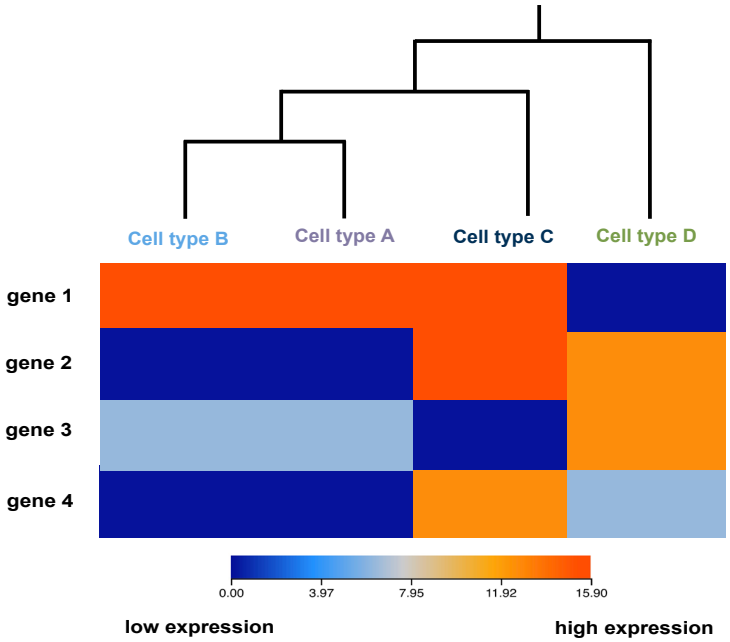
Identify clusters in the data--these clusters represent cells that are more like each other than they are like any other cells

UMAP

In order to plot this many-dimensional graph in a way humans can visualize, we use a dimensionality reduction tool, such as a UMAP, to plot it in a 2D space. Dimensionality reduction is a technique that helps represent many-dimensional data in just two or three dimensions.



Organize the clusters identified in the UMAP to construct a dendrogram that displays hierarchical relationships between the clusters based on each cell type's similarity and dissimilarity of gene expression.



Use a heatmap below the dendrogram to compare the level of gene expression between each cell type for specific genes of interest.

As we established above, understanding **which genes a cell expresses** and in **what quantities** are two key pieces of information that help us distinguish different types of cells from one another. While transcriptomics can be used in a variety of scientific studies to understand organisms and/or cells, this lesson will focus on its applications within the field of brain science.

If two cells show a difference in which genes they express and/or if they express genes in different quantities, scientists can use this as evidence to support the hypothesis that these cells may be different **types** of cells. Organizing cells into different cell types is an essential part of understanding the brain. There are other ways to define cell types – historically, the most common was based on their shape – but in this lesson we will be working only with cell types defined based on the transcriptome. The driving motivation behind research to discover these different **cell types** is that in order to understand the whole (the brain), we have to first understand its parts.

In this lesson, you will have the opportunity to analyze transcriptomic data from healthy human brains. In this part of the lesson, we will use the Human Brain Atlas from the Allen Institute for Brain Science to explore how scientists can use gene expression to create a "map" of cell types in the healthy, neurotypical brain by identifying different "types" of cells in the brain.

Before we can look at this data, we will need to know how to read both a dendrogram and a heatmap, which are two ways of visualizing data. We will start by learning about dendrograms.

Activity 2: Dendrograms

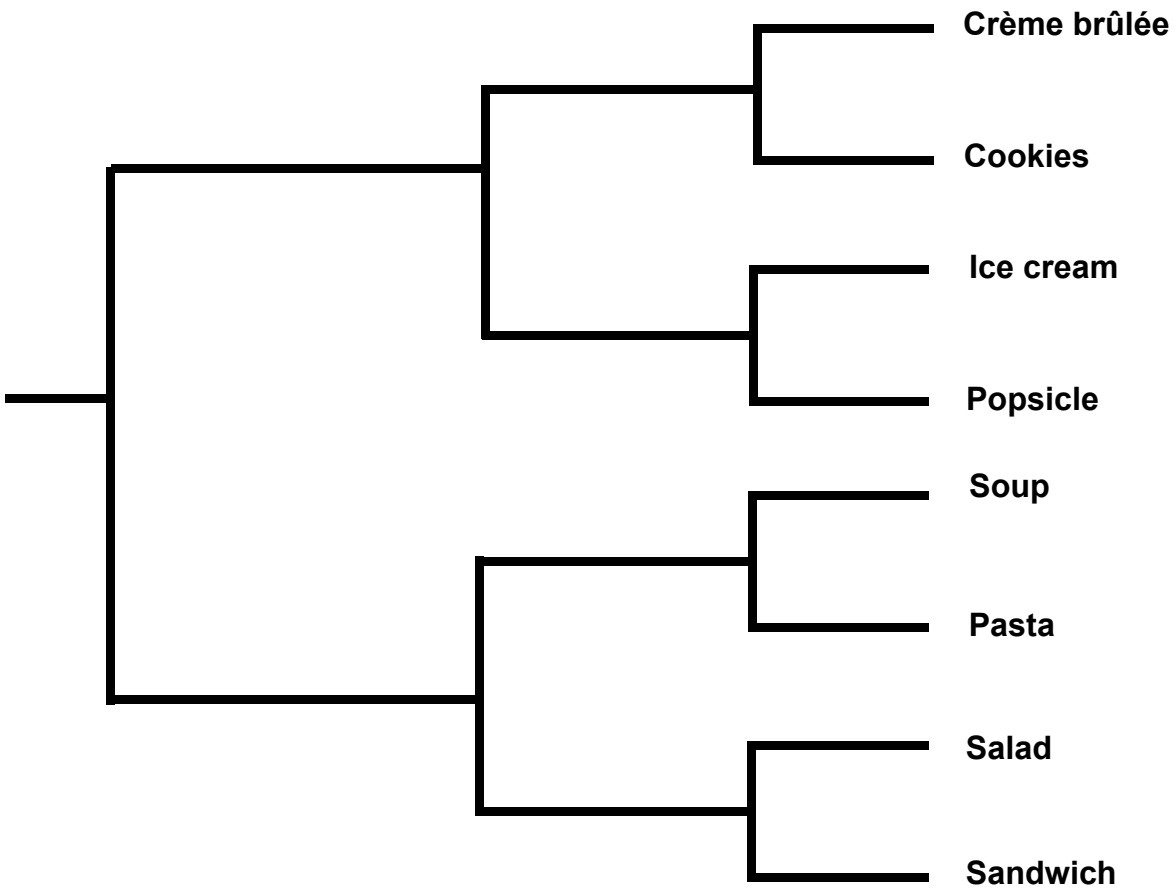
What is a dendrogram?

A dendrogram is one type of visual aid used to show relatedness between different items of interest. It is often used to cluster items in a hierarchal manner, grouping items that are more similar to one another closer together on the "tree" and items that are more dissimilar further apart. One type of dendrogram is a "phylogeny," which is an evolutionary tree that shows relatedness between species based on common ancestry within the context of evolutionary biology.

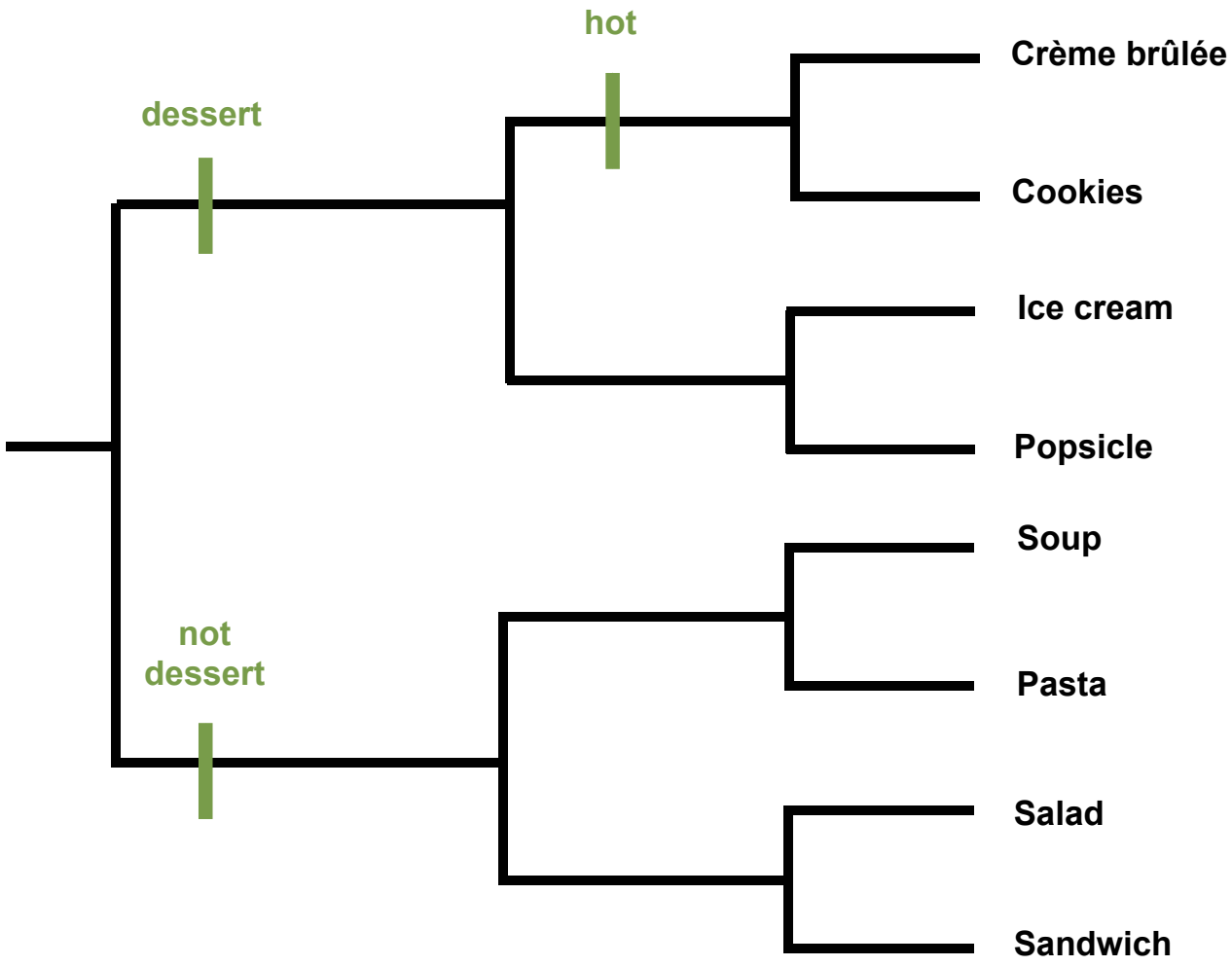
Dendrograms can be used to categorize almost anything. For example, imagine that we wanted to organize the following foods based on how similar/dissimilar they are to one another:

- Crème brûlée
- Ice cream
- Pasta
- Popsicle
- Salad
- Cookies
- Sandwich
- Soup

Step 1: Look at the dendrogram below. This dendrogram compares these foods based on their level of similarity to one another.



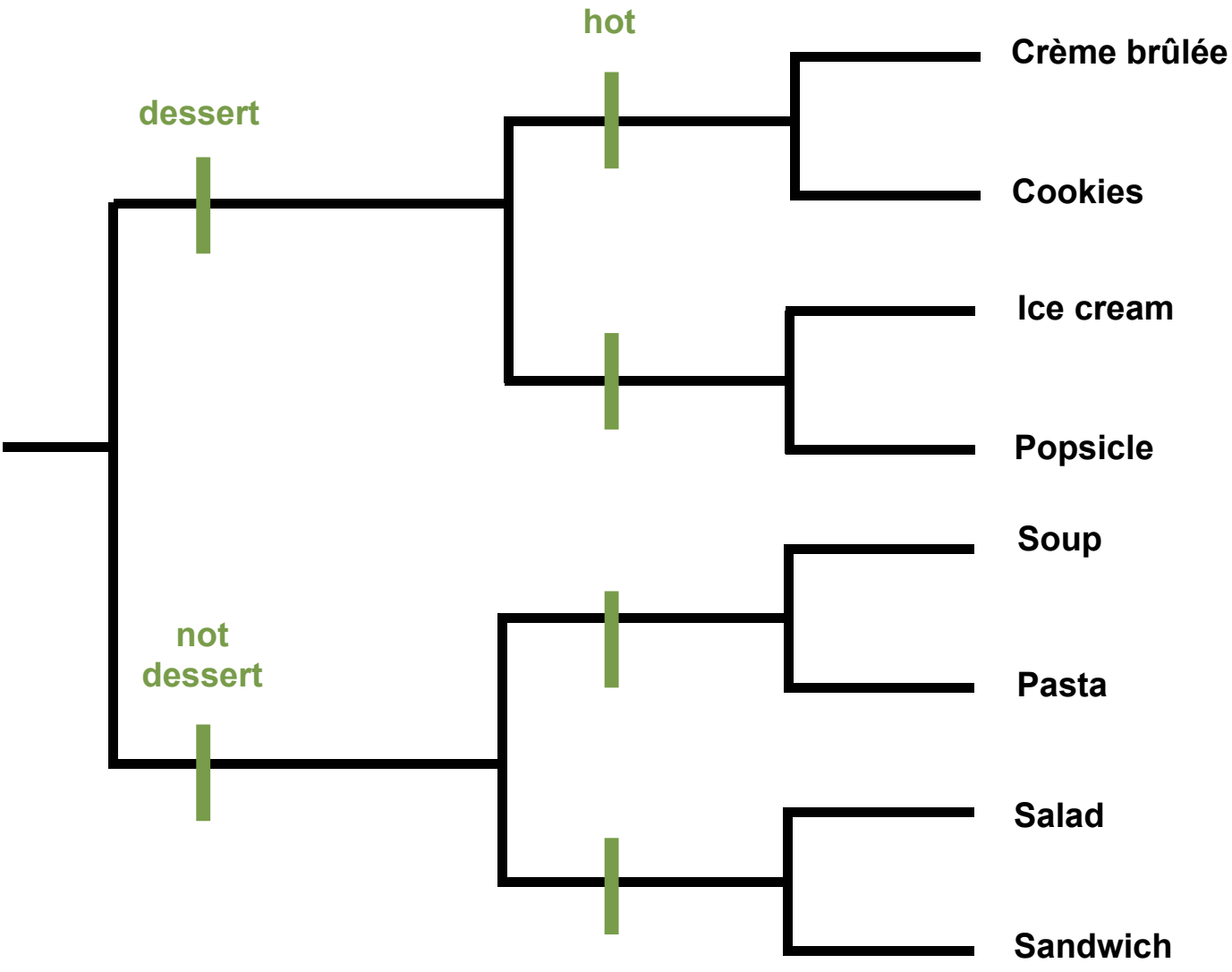
Food items that are grouped closer together on the tree are more similar to one another, while types of food that are grouped further apart are more dissimilar. This dendrogram can be enhanced by adding labels that describe what characteristics of these food items distinguish them from one another:



Step 2: Notice the green dashes on the dendrogram. These green dashes represent specific characteristics of the food. For example, the first green dash shows branching based on whether or not a food item is considered a dessert or not a dessert. If we follow the "dessert" branch, we see another branching event where the crème brûlée and cookies are on the branch labeled "hot," referencing the temperature at which they are usually served/eaten.

Step 3: We have already mapped the first few characteristics of the food items on the dendrogram. Now it is your turn to choose what other characteristics of the food items you can add to further distinguish them from one another.

Mark other characteristics that you could use to distinguish these foods from one another on the dendrogram. Three green dashes have been left blank for you to fill in on the dendrogram below:



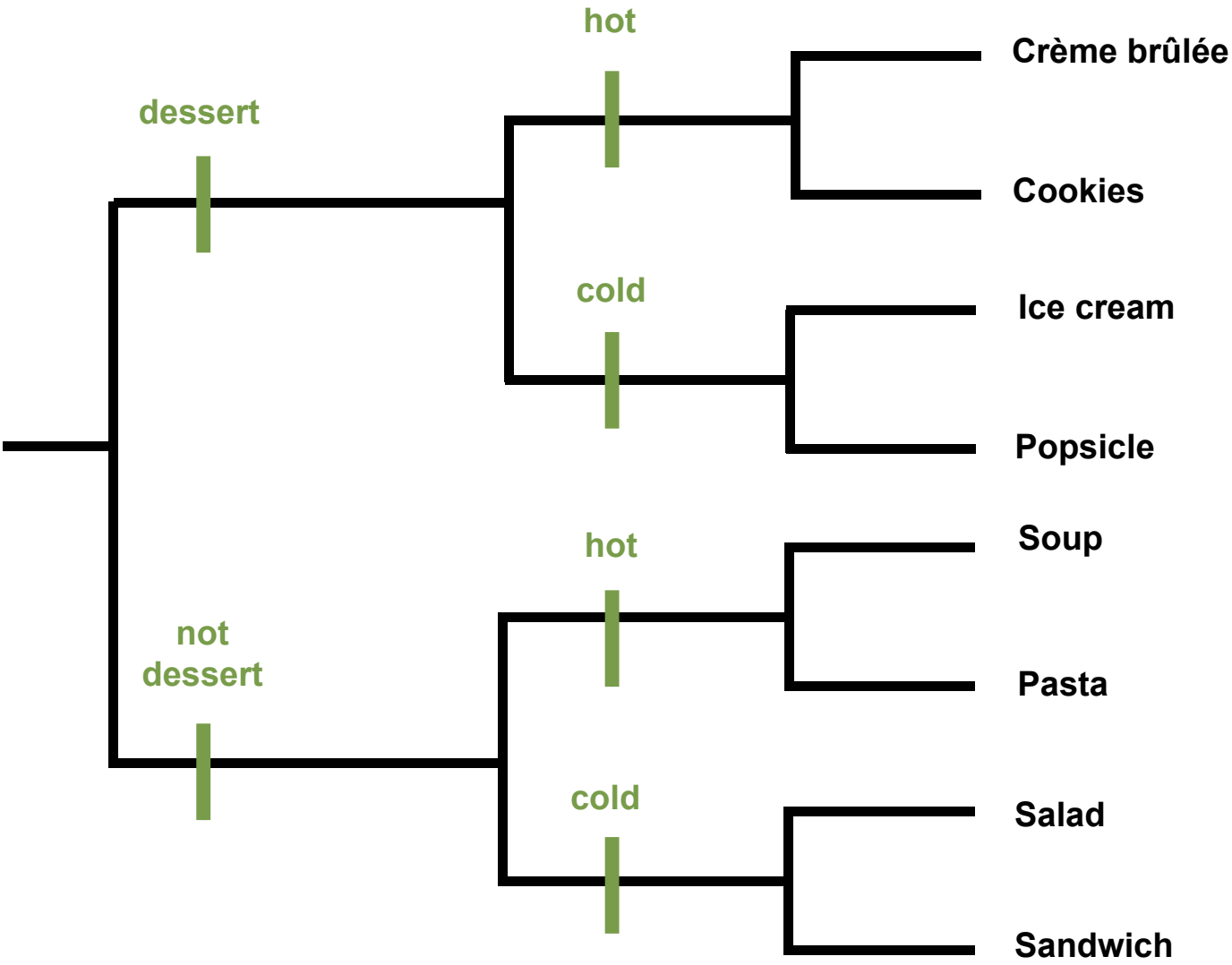
Knowledge Check

- **Based on the characteristics you marked on the tree, why do you think soup and pasta are grouped closest together on the dendrogram compared to the other food items?**

- **Can you think of other characteristics other than temperature/type of food that you could use to distinguish these foods from one another?**

Check Your Work:

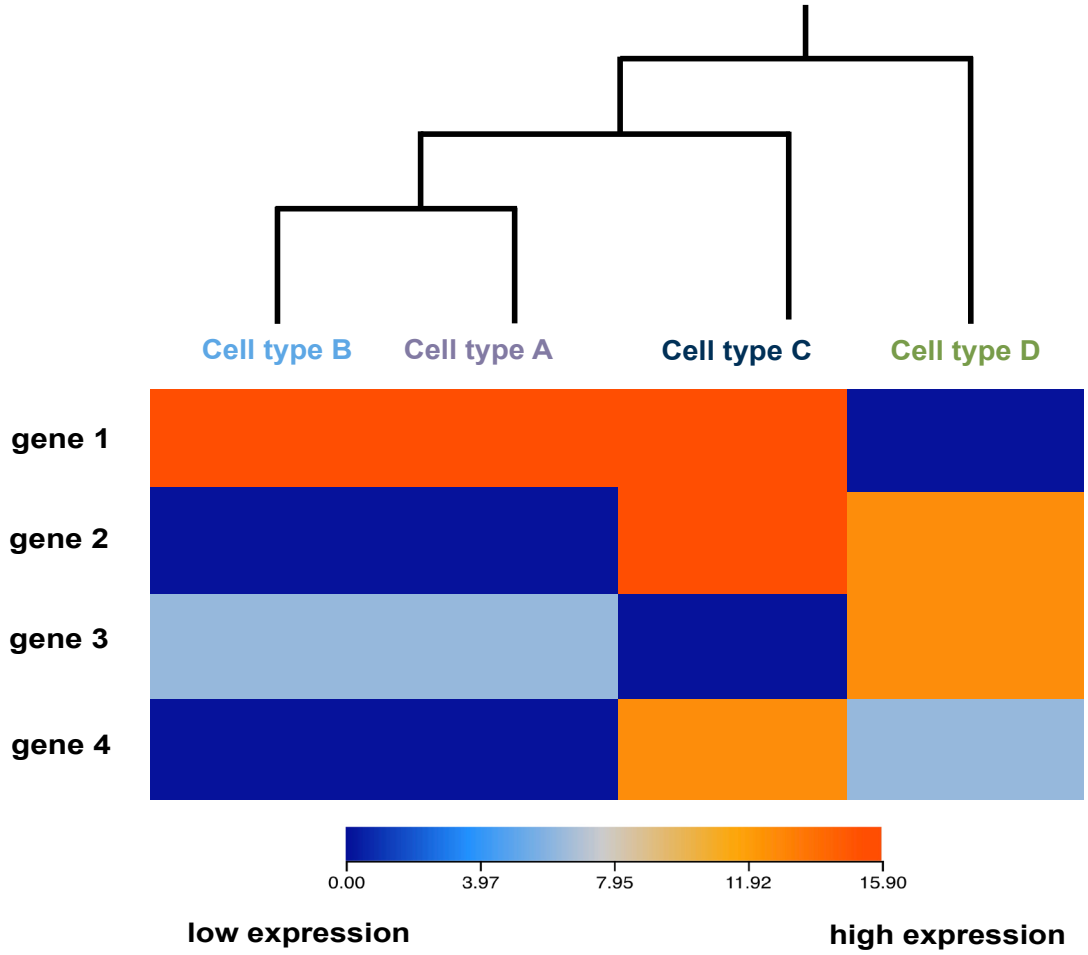
Pictured below is the dendrogram we created together and the characteristics we used to label the dendrogram. You may have chosen other food characteristics that could also be correct!



Activity 3: Heatmaps

Now that you are familiar with what dendrograms are and how they organize data in a hierarchal manner, we will build on the idea of dendrograms by adding another type of data visualization to our analysis. In this section, we will focus on interpreting heatmaps. Heatmaps are helpful ways to represent data as colors. Because heatmaps use color instead of numbers, this often makes heatmaps easier to interpret at a glance than simply a table of data.

Example heatmap:



Notice that this figure contains both a dendrogram (shown above) AND a heatmap. The dendrogram above shows 4 cell types. The dendrogram shows which cell types, based on a transcriptomic analysis of their entire genome, are most similar to one another. The heatmap below uses colors to show the relative level of expression for genes 1, 2, 3, and 4 within each cell type. This is the same gene expression data that was used to define the cell types using the clustering method mentioned above, which we will explore more in lesson 4. The gene expression data is now organized and visualized by cell type classification.

Knowledge Check

To practice interpretating a dendrogram and a heatmap in combination with one another, answer the practice questions below using the example heatmap provided on the previous page:

- **Is gene 1 highly expressed across all four cell types? How can you use the heatmap to answer this question?**

- **Based on the data about relative gene expression provided within the heatmap, why do you think cell type A and cell type B are grouped so closely together on the dendrogram?**

Activity 4: Analyzing Basic Research Data

The Allen Institute for Brain Science has constructed an extremely detailed dendrogram of the different cell types of the human brain. The dendrogram featured below was constructed using transcriptomic data taken from donor brain tissue. As you learned in lesson 1, brain donation plays a significant role within biomedical research. Although this dendrogram looks significantly more complicated, remember what you learned in activity 1 and 2 about how they are constructed! The dendrogram featured below uses the same concept as you saw before, except now, a significantly higher number of cells and a higher number of genes are being compared.

To view the dendrogram, click here.

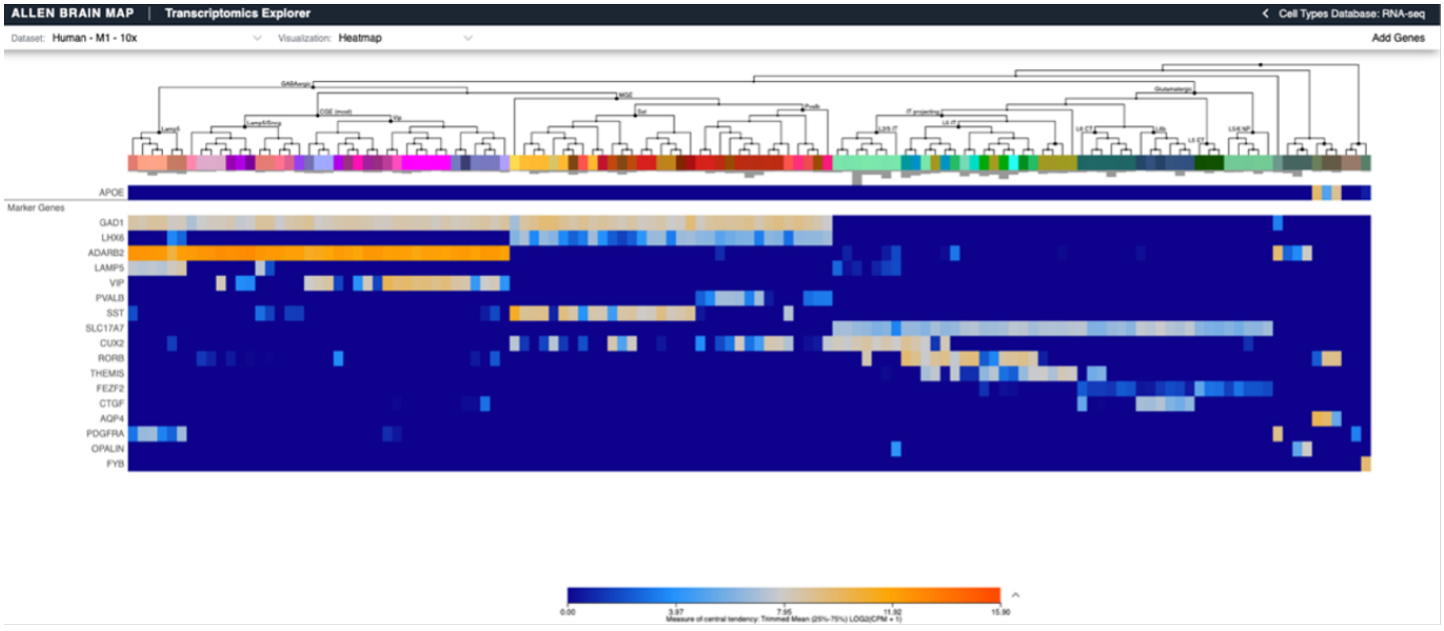
The full link is also included here: https://celltypes.brain-map.org/rnaseq/human_m1_10x?selectedVisualization=Heatmap&colorByFeature=Cell+Type&colorByFeatureValue=GAD1

The dendrogram can also be accessed by going to:

1) SEA-AD.org

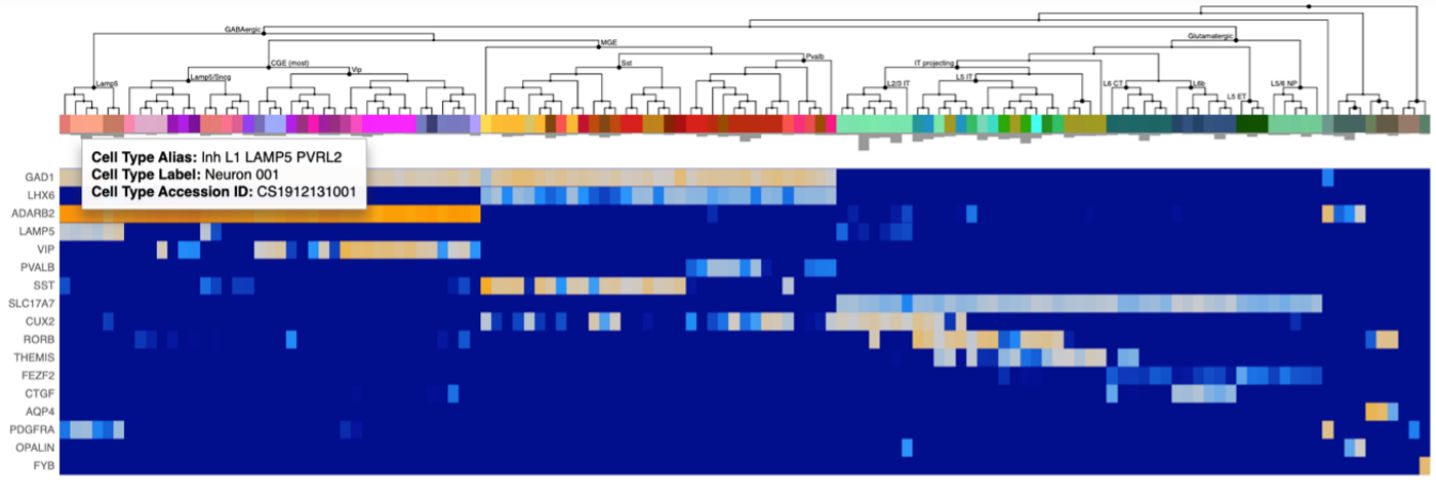
2) On the Seattle Alzheimer's Disease Brain Cell Atlas page, scroll until you find the "cell types" section. Under "Cell Types" click on the "Transcriptomics Explorer (Reference MTG) link.

Your screen should look like this:



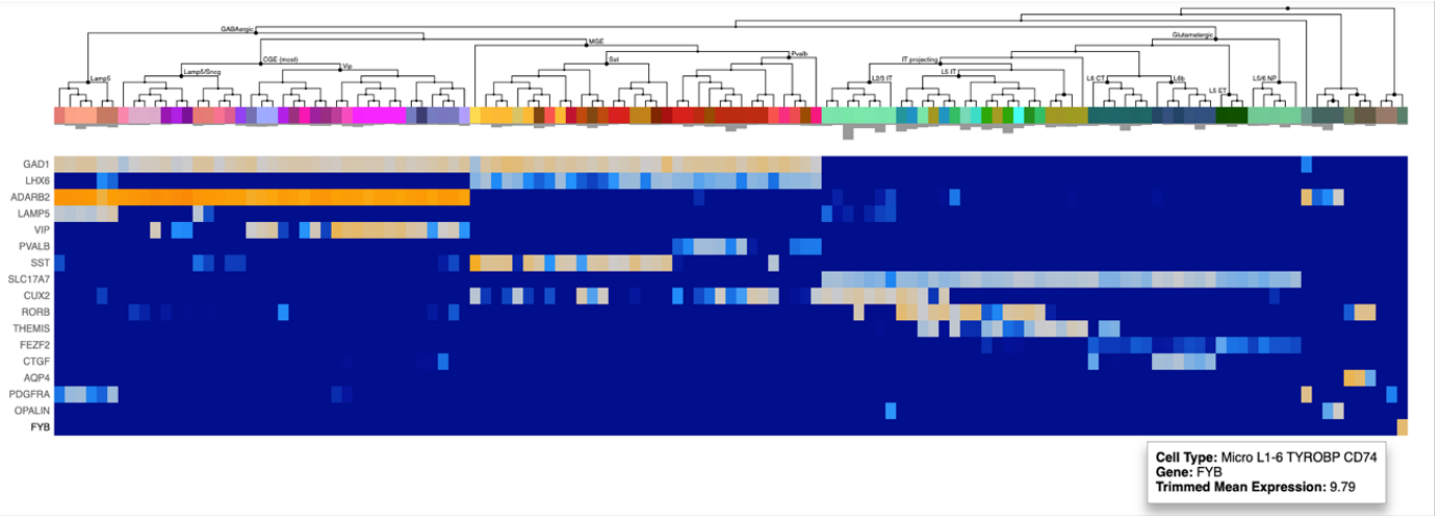
This dendrogram has several interactive elements that you can use to better analyze the data.

Step 1: Hover your mouse over the colors at the end of the **dendrogram**. Notice that hovering over these colors allows you to tell which **cell type** the color represents. For example, when you hover your mouse over the first color on the far left of the dendrogram, we are told this peach color represents the cell type labeled "Neuron 001."

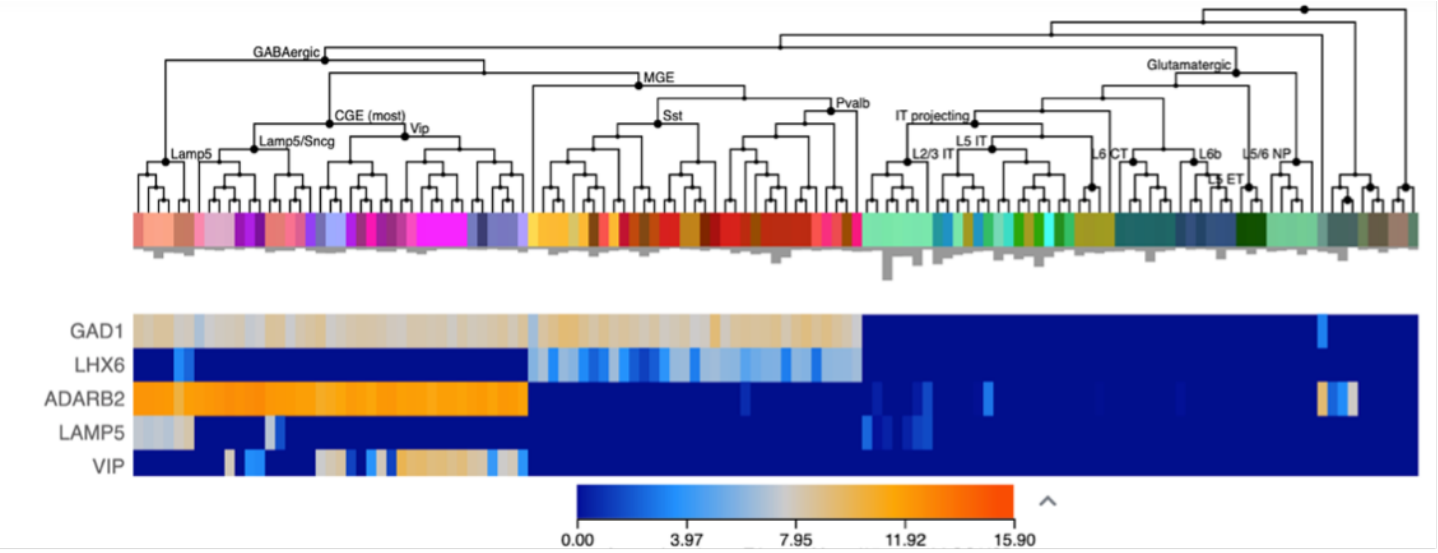


Step 2: Look at the list of genes on the far left. Next to the list of genes is a heatmap. This heatmap uses the exact same scale and color gradient as we saw in Activity 2!

Step 3: Move your cursor to hover over the bottom right hand corner of the heatmap (the orange square). Notice that hovering over this square allows you to see which cell type the box in the square map represents, which gene it is displaying level of expression for, and the "mean expression" of that gene for that specific cell type.



Step 4: Look at the branching of the dendrogram



Notice on this dendrogram we have two main branches. One branch is labeled glutamatergic (excitatory) neurons and one is labeled GABAergic (inhibitory) neurons.

In the heatmap, you can see the level to which each **cell type** expresses the **specific genes listed**.

For example, if we look at the very top row of the table, we can see the extent to which each of the **cell types express GAD1**. Looking at the GAD1 gene expression, it appears that only the GABAergic neurons appear to express GAD1. We can see the level of that expression by using the key at the bottom of the page that shows the gradient of color and the relative level of expression associated with each color.

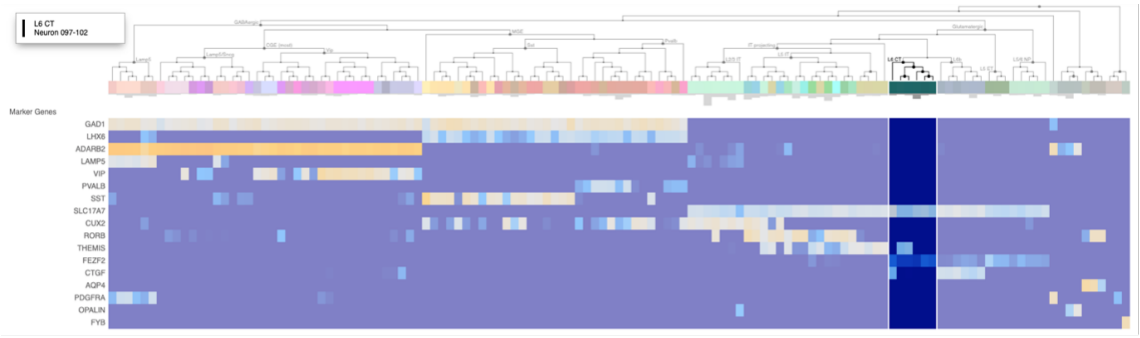
Step 5: Click on the node on the dendrogram labeled "**L6 CT**." Notice that clicking on this node allows you to look more clearly at only the gene expression for the cells that fall under the "L6 CT" branch on the dendrogram.

In order to practice interpreting a dendrogram and a heatmap, answer the following practice questions.

Practice Questions:

- Looking at the dendrogram, it appears like GAD1 is expressed across all types of GABAergic brain cells. According to the heatmap below the dendrogram, which gene seems to be expressed only amongst Glutamatergic brain cells?

- Click on the node on the dendrogram labeled "L6 CT." Notice that clicking on this node allows you to look more clearly at only the gene expression for the cells that fall under the "L6 CT" branch on the dendrogram. You should see something like this:



- Which two genes in this heatmap do ALL the L6 CT cells tend to express to at least a small extent?

- **Which two cell types appear to express a small amount of the THEMIS gene? Instead of the long name of the cell type, feel free to answer using the neuron number listed below the cell type's name.**

- **Why is it helpful for scientists to find genes that are only expressed in a small number of cell types?**

- **This dendrogram uses only transcriptomic data to organize cells into different cell "types." Aside from transcriptomic data, what other type of data do you think would be helpful to compare cells to one another and classify different cell types?**

- **4. The transcriptomic data used to create this dendrogram and heatmap was gathered from brain tissue donated by individuals who were considered healthy, neurotypical donors. A healthy donor is someone considered to have no known neurological diseases. In other words, these donors were not known to have neurological conditions such as dementia, Alzheimer's disease, etc. A neurotypical donor is someone who is considered to have the "standard" brain functioning and processing. Why do you think it is important for scientists to study these types of healthy, neurotypical brains in addition to studying brains with neurological conditions?**

Conclusion:

Throughout the course of this lesson, you have explored why it is important to study both healthy and diseased brains, and why basic research, like that conducted at the Allen Institute, is so integral to the field of science as a whole. In order to explore the power of basic research, you had access to open transcriptomic data available to the public from the Allen Institute for Brain Science.

If you continue on to complete lessons 3 and 4, you will have the opportunity to dive into the basics of disease research by looking at the data from the Seattle Alzheimer's Disease Brain Cell Atlas from the Allen Institute for Brain Science and its research partners. While this lesson challenged you to practice your data analytic skills for dendrograms and heatmaps, lessons 3 and 4 will introduce you to different methods, such as neuropathology image analysis and UMAP interpretation.

If you develop your own lesson plan using Allen Institute resources, we invite you to share your experience with us at communications@alleninstitute.org. Teachers are also encouraged to publish original lessons using our open data, tools, and other re-sources, and to share those lessons with us.

Supplementary Material 2 for Meuler M, Casimo K (2024) A Case for the Use of Open Data as a Tool to Incorporate Socioscientific Topics into Neuroscience EducationJ Undergrad Neurosci Educ 22(2):A90-A98.

ALLEN INSTITUTE

Lesson 3: Societal and Biological Perspectives on Alzheimer's Disease

Learning Objectives:

- Students will be able to evaluate what research questions can and cannot be answered via certain datasets based on the demographic composition of the study cohort
- Students will be able to evaluate the known risk factors for Alzheimer's disease
- Students will be able to analyze immunolabeled brain tissue for known biomarkers of Alzheimer's disease pathology
- Students will be able to assess what steps the NIH has taken to increase diversity in biomedical research
- Students will be able to propose possible ways the diversity of biomedical research cohorts can be improved in future research endeavors

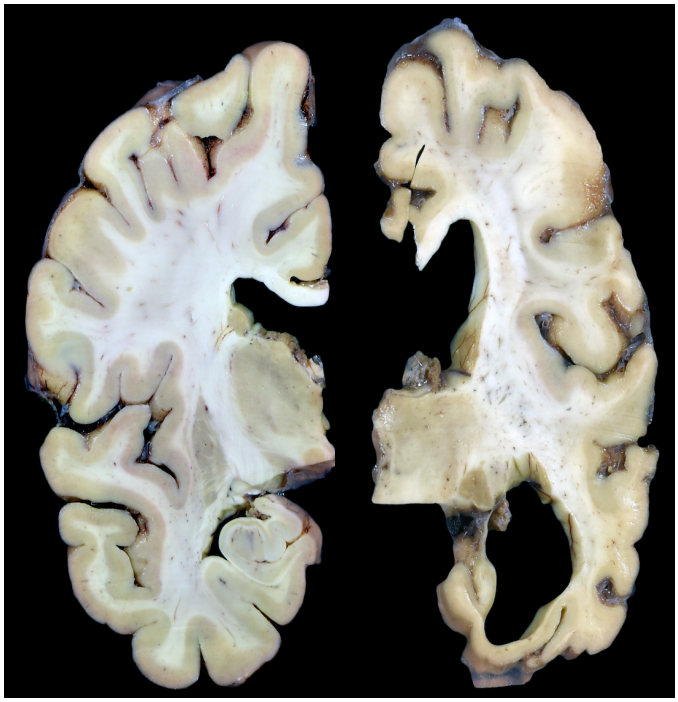


Image from UW Medicine
Left: healthy brain
Right: brain with AD pathology

Introduction

In lesson 1, you learned about the process of brain donation and how biomedical research can use donated brain tissue to perform critical analyses of disease pathology. In this lesson, we will extend our knowledge of brain donation from basic science looking at healthy brains to the basic science of disease. The brain samples donated by these 84 individuals continue to provide crucial data about the early pathogenesis of Alzheimer's disease (AD) for the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD).

In this lesson, you will use the open data provided by the Allen Institute's SEA-AD study to conduct a qualitative neuropathology image analysis. Prior to looking at the images taken of each donor's brain tissue, you will get to know the demographics of the 84 donors who generously donated their post-mortem brain tissue for the purposes of this study. This lesson consists of three brief activities, which are summarized below:

Activity 1: Biomedical Research and Demographics

- First, you will read a brief article that overviews some history of biomedical research and how it continues to face challenges in recruiting diverse study participants that are representative of the general population.

Activity 2: Exploring the Donor Index

- Next, you will have the opportunity to do a deep dive into the SEA-AD donor index, which provides key demographic information about the 84 donors who were a part of the project.

Activity 3: Neuropathology Image Analysis

- Lastly, you will have the chance to do a side-by-side comparison of images taken of each donor's brain tissue samples and look for specific biomarkers that are suspected to play a role in AD pathology.

At the end of this lesson, you will have developed a deeper understanding of how scientists use brain tissue samples to study the pathology of AD. We hope that you walk away from this lesson recognizing that while an important part of biomedical research is studying **how** a disease impacts people, it is equally important to investigate **who** the disease impacts.

What is Alzheimer's disease (AD)?

AD is a specific type of dementia that affects an individual's behavior, thinking, and memory. AD and dementia are not synonymous:

- **Dementia** is an umbrella term used to describe a class of symptoms that include a decline in memory, reasoning skills, and other cognitive functions. There are several different types of dementia, including vascular dementia, dementia with Lewy bodies, Parkinson's disease dementia, and several others.

- **AD** is a specific disease that accounts for 60-80% of dementia cases. All AD causes dementia, but not all dementia cases are caused by AD.

Since the first known case of AD was documented in the early 1900s, a considerable amount of research has been conducted in an attempt to understand the pathology of the disease. **Pathology** is a term used to describe the general cause and effects of a specific disease.

What is the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD)?

The SEA-AD project at the Allen Institute for Brain Science is a collective effort from scientists to gain a deep molecular and cellular understanding of the early pathogenesis of AD. The data collected within this study are derived from a full spectrum of 84 older adult donors, also referred to as "aged donors." This cohort of 84 donors includes both healthy controls and those with high AD pathology and cognitive dementia symptoms. In addition to gathering clinical and demographic information from each patient, Allen Institute scientists also gather transcriptomic and pathogenic data from each donor's brain tissue.

Data and specimens were obtained from the Adult Changes in Thought (ACT) Study from Kaiser Permanente Washington Health Research Institute (KPWHRI), and the University of Washington Alzheimer's Disease Research Center (ARDC). The ACT study from Kaiser Permanente specifically follows initially healthy donors starting at 65 years of age and through the rest of their lifespan. This type of longitudinal data allows scientists to gather crucial medical and demographic information about each donor over their lifespan and at their time of death.

The SEA-AD project is one type of basic biomedical research that seeks to understand the early pathogenesis of AD. It is also important to investigate **who** the disease impacts. Historically, biomedical research has struggled to recruit diverse cohorts of study participants. In order to understand why biomedical research has historically studied cohorts of predominately white, cisgender males, we will read the following article by Oh et al. (2015).

Activity 1: Biomedical Research and Demographics

Instructions:

- Read through "Diversity in Clinical and Biomedical Research: A Promise Yet to Be Fulfilled" by Oh et al. (2015) doi: 10.1371/journal.pmed.1001918.
- This is roughly a 10 minute read. The article can be accessed here or by clicking: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4679830/
- After you have finished reading through the article, answer the reflective questions listed below.

Reflective Questions

1. What was the 1993 National Institutes of Health (NIH) Revitalization Act?

2. Since the passage of the NIH Revitalization Act in 1993, what percentage of cancer clinical trials funded by the National Cancer Institute included enough minority participants to meet the NIH's own criteria and goals?

3. What barriers are in place that may discourage some people of color (POC) from participating in biomedical research?

4. What solutions does the article propose that could help to improve the diversity of biomedical research cohorts?

5. What solutions for improving the diversity of biomedical research cohorts can you think of that the article did not mention?

Biomedical Research on Alzheimer's Disease:

As you read in Oh et al.'s 2015 paper, understanding who is impacted by a disease is an integral part of biomedical research. While the field of AD research still has work to do to ensure that it recruits diverse cohorts of study participants, AD researchers have worked hard to identify several possible demographic and/or socioeconomic factors that may have an association with AD. Although there are several suspected demographic factors that impact one's risk for developing AD, this lesson will focus on the following three:

1. Age
2. Race and ethnicity
3. Biological sex

1. **Age:**

Individuals who are 65 or older are at the greatest risk for developing AD. The risk for developing AD doubles every five years after age 65. Age is currently the greatest known risk factor for developing AD.

AD is considered to be younger-onset/early-onset when it occurs in people younger than 65 years of age. Early-onset AD is far less common than AD in people 65 or older. For more information on early-onset AD, please visit:

https://www.alz.org/alzheimers-dementia/what-is-alzheimers/younger-early-onset

2. **Race and ethnicity:**

Both race and ethnicity have been identified as possible risk factors for developing AD. Research has shown that older individuals who identify as Black are roughly twice as likely to develop AD compared to older individuals who identify as White. Additionally, older individuals who identify as Hispanic are roughly 1.5x more likely to develop AD compared to older individuals who identify as White. Research on the racial disparities of AD suggest that the difference in prevalence could be due to socioeconomic factors and not biological factors. These socioeconomic factors could include access to healthcare, access to education, levels of stress experienced by an individual, etc.

Although race and ethnicity are frequently used as synonyms, they carry distinct meanings. The U.S. Census Bureau defines **race** as a person's self-identification with one or more social groups, which can include White, Black or African American, Asian, American Indian, Alaska Native, Native Hawaiian, and/or Other Pacific Islander. Federal statistical standards conceptualize a person's **ethnicity** into one of two categories: Hispanic (or Latino/a/x) or Not Hispanic (Latino/a/x). If a person is Hispanic/Latino, they can self-report/identify as any race.

For more information about race/ethnicity and Alzheimer's disease, please visit:
https://aaic.alz.org/downloads2020/2020_Race_and_Ethnicity_Fact_Sheet.pdf

3. **Biological sex:**

Of the more than 6 million Americans currently living with AD, **roughly two-thirds are female**. Age is the greatest known risk factor for developing AD, and because females tend to live longer than males on average, some scientists believe that the different rates of AD between people of different sexes may be attributable to this difference in average lifespan. However, recent research suggests that the female genome may contain certain **genetic factors** that raise their risk for AD.

Reference:

www.cnn.com/2022/06/30/health/female-alzheimer-gene-discovered-wellness-scn/index.html

Distinguishing between sex and gender:

While sex and gender are frequently used as synonyms, they have distinct meanings. **Sex** is a biological classification that is based on an individual's reproductive organs and sex chromosomes. While XX is typically used as the marker for biological females and XY as the marker for biological males, there are also individuals who are **intersex** and carry other sets of sex chromosomes and/or reproductive organs.

While sex is based on a person's biological characteristics, **gender** is defined as "set of social, psychological, or emotional traits, often influenced by societal expectations that classify an individual as either feminine or masculine." Cisgender individuals identify with the gender they were assigned at birth, while transgender individuals identify as a different gender than assigned at birth. Individuals also can identify as nonbinary or genderfluid. These are by no means the only options for an individual's gender identity, and one's sense of gender identity can be fluid over time.

For more information about gender identity and the difference between sex and gender, visit https://medicine.yale.edu/whr/about/mission/definitions/.

While biological sex and genetics may influence a person's genetic risk for developing AD, a person's **gender** identity may influence the financial or emotional burden that a person experiences due to AD. A majority of AD and dementia caregivers are women, and many women who have taken on caregiving roles for individuals with AD face financial burdens. To learn more about how gender can impact the social and/or financial barriers women face because of AD, please visit https://www.alz.org/alzheimers-dementia/what-is-alzheimers/women-and-alzheimer-s.

The importance of demographics in biomedical research:

Because age, race/ethnicity, and sex have all been identified as possible characteristics associated with AD, it is important that AD studies recruit a diverse range of participants from different age groups, different races/ethnicities, and different sexes. In this next section, you will have the chance to explore the demographic composition of the Allen Institute's SEA-AD donor cohort. Understanding who the donors/study participants are will inform the type of scientific questions we can ask about AD and its pathology.

Activity 2: Exploring the Donor Index

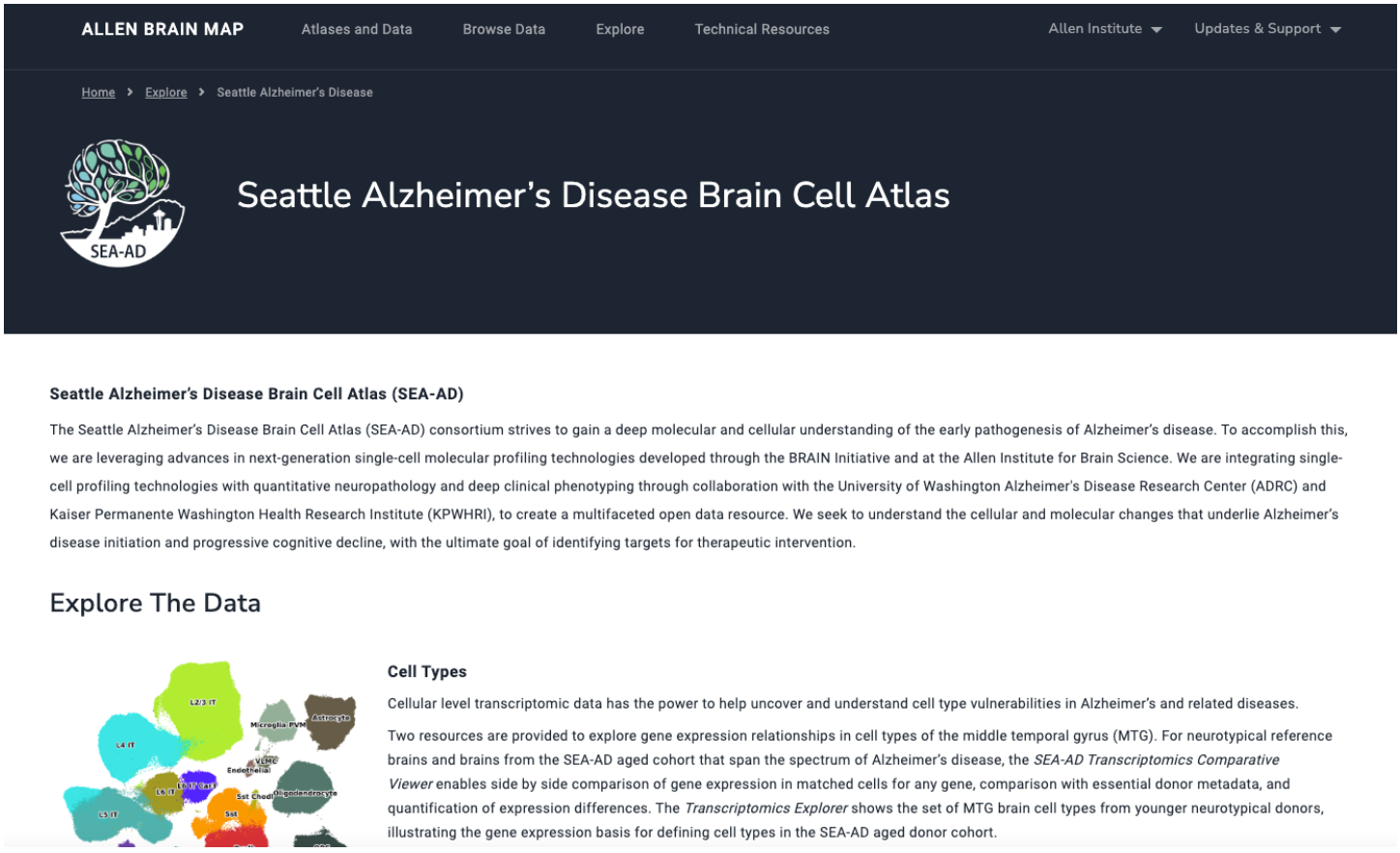
Who are the 84 donors of the SEA-AD study?

In order to explore who the 84 donors of the Allen Institute's SEA-AD project are, we need to first become familiar with how to filter through the data provided in the donor index.

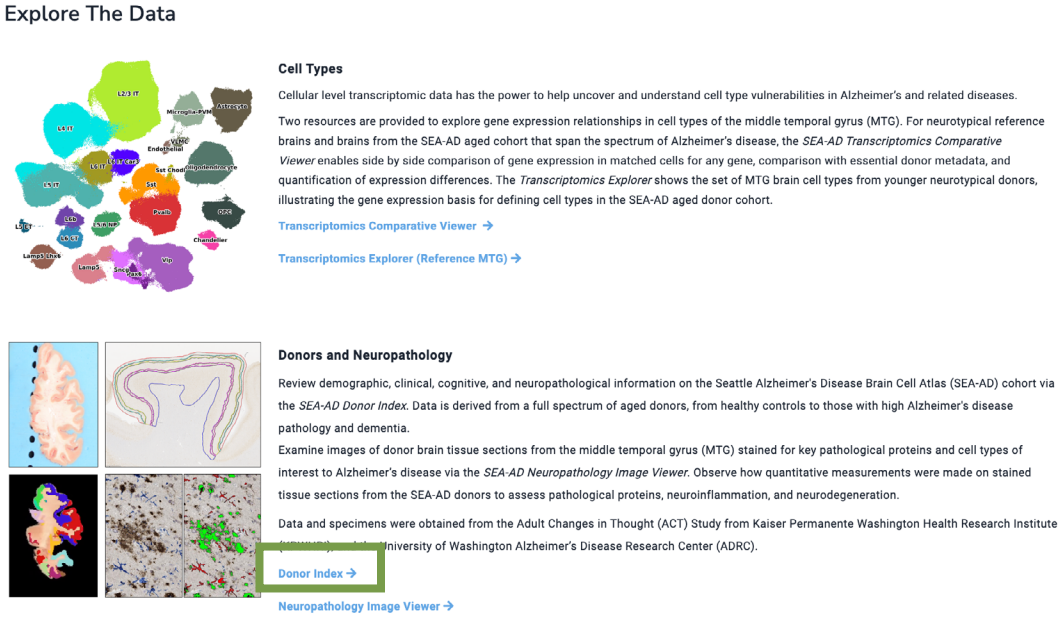
Tutorial: How to Navigate the SEA-AD Donor Index

Step 1: Go to the Seattle Alzheimer's Disease Brain Cell Atlas, linked here:
- https://portal.brain-map.org/explore/seattle-alzheimers-disease

Your screen should look like this:



Step 2: Scroll down this page until you see "Donors and Neuropathology." Under this section, you want to click on "donor index."



After opening the donor index, your screen should look like this:

ALLEN BRAIN MAP | SEA-AD - Metadata and Neuropath | Donors | Terms | Help

Filtered by None

Showing 84 Donors

Donor ID	Age at dea...	Sex	APOE4 sta...	Cognitive s...	Years of ed...	ADNC	Braak stage	Thal phase	CERAD sc...	Lewy body d...	Total micro...
H20.33.040	90+	Male	N	Dementia	13	Intermediate	Braak IV	Thal 4	Frequent	Not Identified ...	0
H20.33.036	90+	Female	N	No dementia	15	High	Braak V	Thal 5	Moderate	Not Identified ...	0
H21.33.019	75	Male	N	No dementia	15	Low	Braak 0	Thal 1	Sparse	Not Identified ...	1
H21.33.040	83	Male	Y	No dementia	17	High	Braak V	Thal 4	Frequent	Olfactory bulb...	3
H20.33.029	90+	Female	N	Dementia	13	High	Braak V	Thal 4	Moderate	Not Identified ...	0
H20.33.015	88	Male	N	Dementia	18	Intermediate	Braak V	Thal 3	Moderate	Not Identified ...	0
H21.33.012	90+	Female	N	Dementia	21	Intermediate	Braak IV	Thal 3	Sparse	Neocortical (D...	0
H21.33.003	78	Male	N	No dementia	16	Not AD	Braak 0	Thal 0	Absent	Not Identified ...	0
H20.33.002	90+	Female	N	No dementia	12	Not AD	Braak IV	Thal 0	Absent	Limbic (Transi...	0
H21.33.030	89	Male	Y	No dementia	17	Intermediate	Braak III	Thal 3	Moderate	Brainstem-pre...	3
H20.33.046	90+	Male	N	Dementia	21	High	Braak VI	Thal 5	Frequent	Not Identified ...	1
H20.33.031	87	Female	N	Dementia	12	High	Braak VI	Thal 4	Frequent	Not Identified ...	1
H20.33.032	90+	Male	N	No dementia	17	High	Braak V	Thal 5	Moderate	Not Identified ...	11
H21.33.027	90+	Male	Y	Dementia	18	High	Braak V	Thal 5	Moderate	Not Identified ...	0

Note in the top left corner, we are told that the data table is currently displaying data for **all 84 donors.**

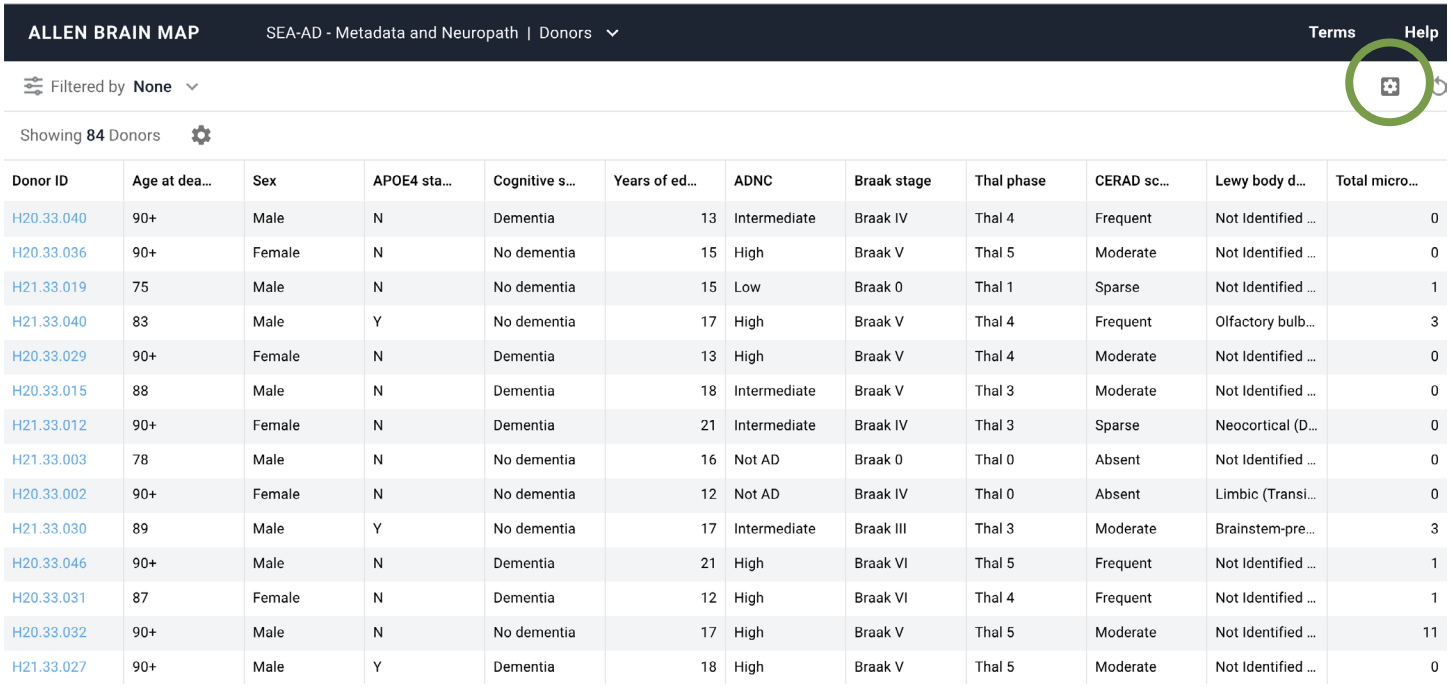
Step 3: Each column contains a different piece of data for each donor. While some are self-explanatory columns, others are less intuitive. Look at the following key for a brief explanation of what each column contains:

Title	Description
Donor ID	Donor Identification Number
Age at death	Age at death
Sex	Biological sex, defined by presence of Y chromosome
APOE4 status	Presence of at least one copy of the apolipoprotein (APOE) e4 allele
Cognitive status	Clinical diagnosis of dementia derived from DSM-IV
ADNC	Overall Alzheimer's disease neuropathologic change score
Thal phase	Extent of the anatomical distribution of amyloid beta plaque deposits
Braak stage	Extent of the anatomical distribution of neurofibrillary tangles (NFTs)
CERAD score	Semiquantitative neuritic plaque density
Lewy body disease pathology	Anatomical distribution of Lewy bodies
LATE-NC stage	Extent of Limbic-predominant Age-related TDP-43 Encephalopathy-Neuropathologic Change
Microinfarcts	Number of microscopic strokes identified in diagnostic screening sections
PMI	Post-mortem interval in hours (time from death to brain removal)
Race/ ethnicity	Self-reported race/ ethnicity
Years of education	Number of years of education starting in grade school
Age of dementia diagnosis	Age when dementia clinically diagnosed
Consensus clinical diagnosis	Clinical diagnosis determined by a consensus of providers who evaluated the individual
CASI score	Most recent Cognitive Abilities Screening Instrument assessment score
Interval from last CASI	Interval (in months) between last CASI assessment and death
MMSE score	Most recent Mini-Mental State Exam assessment score (here derived from CASI components)
Interval from last MMSE	Interval (in months) between last MMSE assessment and death
MoCA score	Most recent Montreal Cognitive Assessment score
Interval from last MoCA	Interval (in months) between last MoCA and death
Fresh brain weight	Brain weight (in grams) at time of brain removal
Brain pH	pH of intraventricular cerebrospinal fluid at time of brain removal
Overall CAA Score	Severity of Cerebral Amyloid Angiopathy
Atherosclerosis	Severity of plaque deposition in arteries
Arteriolosclerosis	Severity of thickening of arterioles
RIN	RNA integrity number of brain tissue sample

Step 4: The default setting is for the donor index to display the data for all 84 donors. However, you can use the filter tools to curate which donors you are evaluating.

There are two options to filter the donor data:

First, go to the top right-hand corner of the spreadsheet and click on the settings icon that looks like a gear. This icon is in the green circle in the figure below:

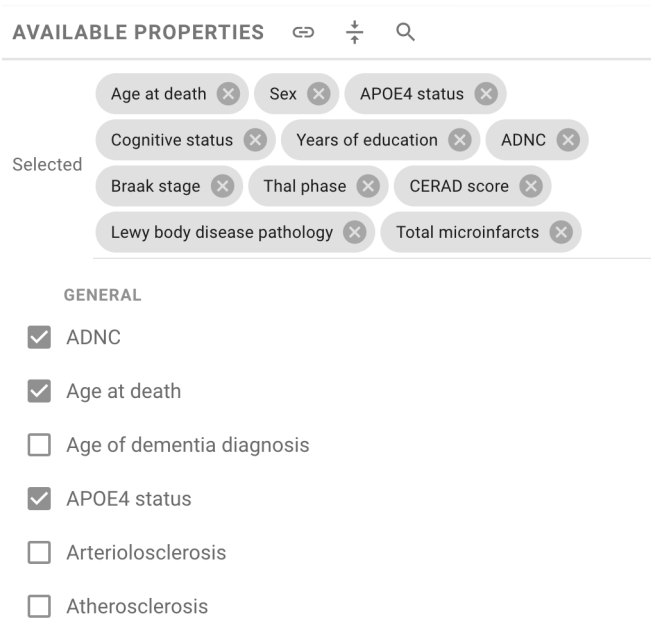


Donor ID	Age at dea...	Sex	APOE4 sta...	Cognitive s...	Years of ed...	ADNC	Braak stage	Thal phase	CERAD sc...	Lewy body d...	Total micro...
H20.33.040	90+	Male	N	Dementia	13	Intermediate	Braak IV	Thal 4	Frequent	Not Identified ...	0
H20.33.036	90+	Female	N	No dementia	15	High	Braak V	Thal 5	Moderate	Not Identified ...	0
H21.33.019	75	Male	N	No dementia	15	Low	Braak 0	Thal 1	Sparse	Not Identified ...	1
H21.33.040	83	Male	Y	No dementia	17	High	Braak V	Thal 4	Frequent	Olfactory bulb...	3
H20.33.029	90+	Female	N	Dementia	13	High	Braak V	Thal 4	Moderate	Not Identified ...	0
H20.33.015	88	Male	N	Dementia	18	Intermediate	Braak V	Thal 3	Moderate	Not Identified ...	0
H21.33.012	90+	Female	N	Dementia	21	Intermediate	Braak IV	Thal 3	Sparse	Neocortical (D...	0
H21.33.003	78	Male	N	No dementia	16	Not AD	Braak 0	Thal 0	Absent	Not Identified ...	0
H20.33.002	90+	Female	N	No dementia	12	Not AD	Braak IV	Thal 0	Absent	Limbic (Transi...	0
H21.33.030	89	Male	Y	No dementia	17	Intermediate	Braak III	Thal 3	Moderate	Brainstem-pre...	3
H20.33.046	90+	Male	N	Dementia	21	High	Braak VI	Thal 5	Frequent	Not Identified ...	1
H20.33.031	87	Female	N	Dementia	12	High	Braak VI	Thal 4	Frequent	Not Identified ...	1
H20.33.032	90+	Male	N	No dementia	17	High	Braak V	Thal 5	Moderate	Not Identified ...	11
H21.33.027	90+	Male	Y	Dementia	18	High	Braak V	Thal 5	Moderate	Not Identified ...	0

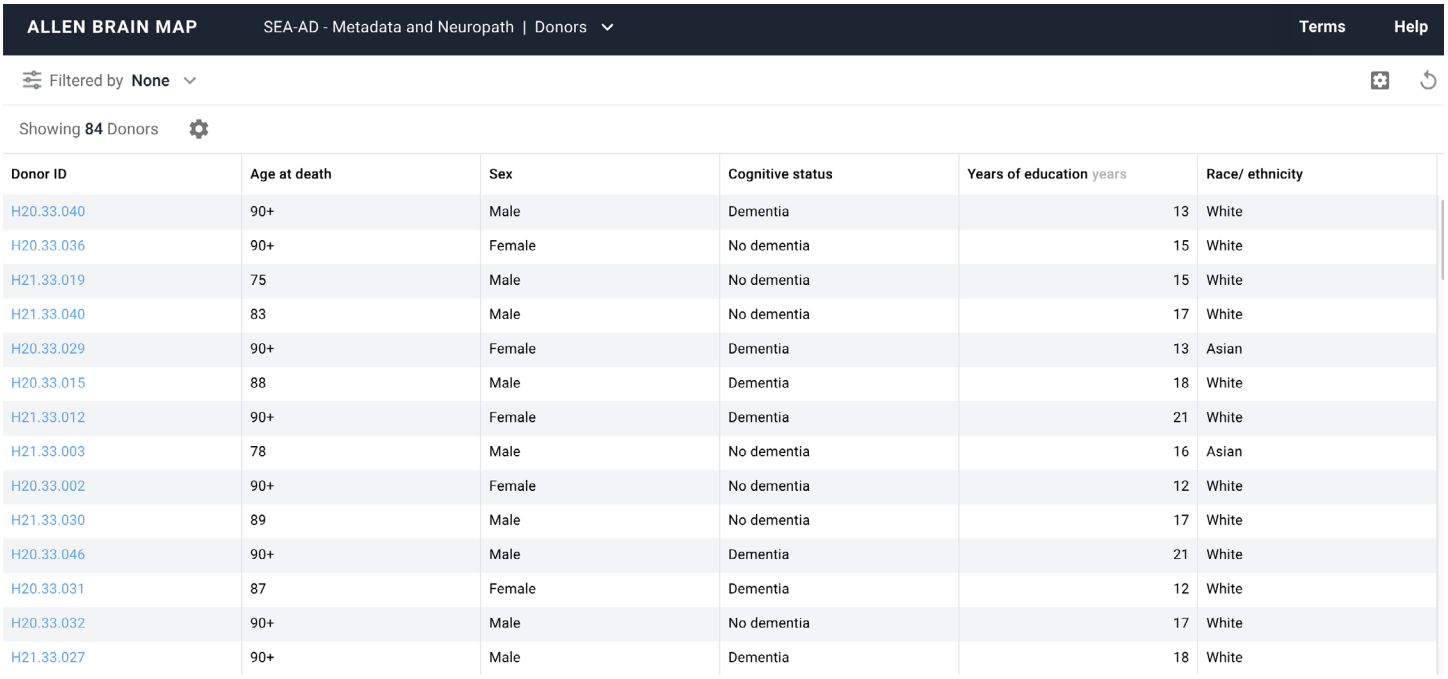
A drop-down menu should display in the top right corner of your screen.

Step 5: Using this drop-down menu, select the following criteria to filter the data by (note: you will also have de-select some of the characteristics that the donor index selects by default):

- Age at death
- Sex
- Cognitive status
- Years of education
- Race/ethnicity

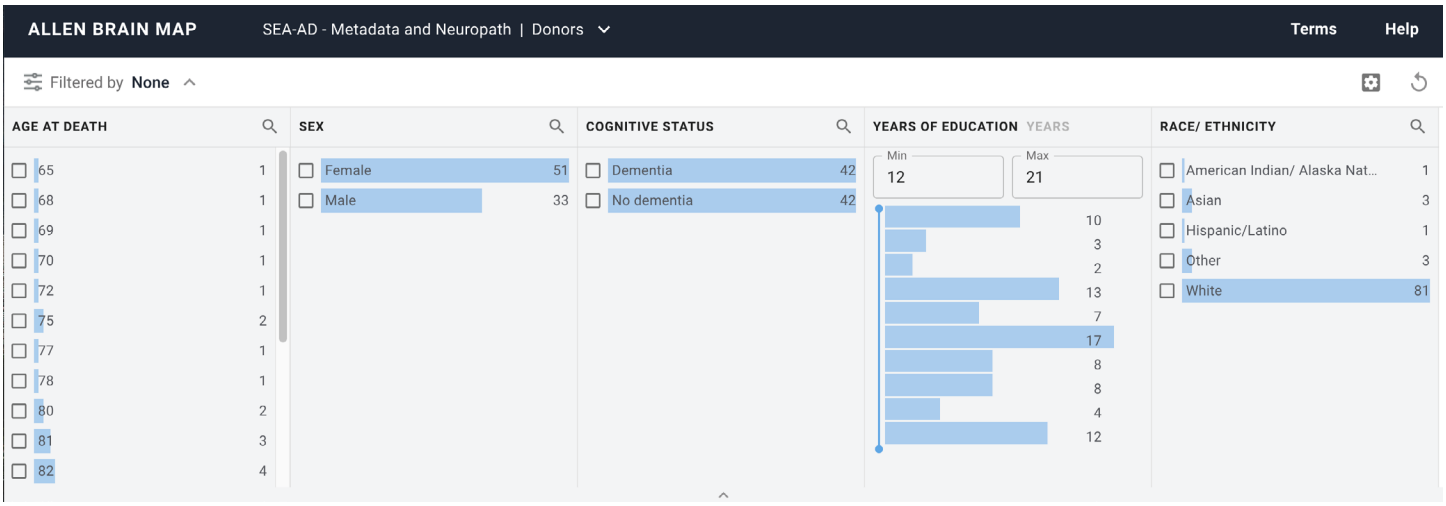


Notice that selecting these criteria in the "available properties" drop-down menu changes the table so that it looks like this:



Donor ID	Age at death	Sex	Cognitive status	Years of education years	Race/ ethnicity
H20.33.040	90+	Male	Dementia	13	White
H20.33.036	90+	Female	No dementia	15	White
H21.33.019	75	Male	No dementia	15	White
H21.33.040	83	Male	No dementia	17	White
H20.33.029	90+	Female	Dementia	13	Asian
H20.33.015	88	Male	Dementia	18	White
H21.33.012	90+	Female	Dementia	21	White
H21.33.003	78	Male	No dementia	16	Asian
H20.33.002	90+	Female	No dementia	12	White
H21.33.030	89	Male	No dementia	17	White
H20.33.046	90+	Male	Dementia	21	White
H20.33.031	87	Female	Dementia	12	White
H20.33.032	90+	Male	No dementia	17	White
H21.33.027	90+	Male	Dementia	18	White

Step 6: In addition to the "available properties" drop-down menu, there is also a way to directly filter the data by clicking on the arrow next to "Filtered by: None." After clicking on the arrow, your screen should look like this:



Notice that the options to filter the data match the one's you selected in the "available properties" box. The blue bars help to visualize the extent to which the donors fall into the selected categories. The numbers next to each trait tell you the exact number of donors that fall into each of the categories.

Knowledge Check

Now that you know how to filter the data, let's think back to the three demographic features we learned about earlier that possibly play a role in risk for developing AD:

1. Age
2. Race/ethnicity
3. Sex

1. Would this cohort of 84 donors allow us to study the possible association between race/ethnicity and AD? Why or why not?

2. Would this cohort of 84 donors allow us to study the possible association between sex and AD? Why or why not?

Understanding the strengths and limitations of your data:

The lack of racial and ethnic diversity of the 84 donors in this study is one of its limitations. Earlier in the lesson, you read Oh et al.'s (2015) article on the lack of diversity in biomedical research studies. The field of science continues to make efforts towards addressing this demographic gap in its research to ensure that our study populations are representative of society as a whole. Conclusions from this study with regard to race and ethnicity are limited, but we can explore impacts with regard to sex and gender. While everyone is encouraged to donate their brains to science, it is imperative that people feel comfortable doing so. Improving outreach and education efforts to diverse audiences would help the field of biomedical sciences deepen its foundational knowledge of AD pathology.

While this cohort of 84 donors does not allow us to explore questions about a possible association between AD and race/ethnicity, this cohort does allow us to ask several interesting and crucial questions about a possible association between a person's sex and their risk for developing AD.

In order to explore the possible association between sex and AD, we can use the collection of neuropathology image data available from the Allen Institute's SEA-AD study.

Activity 3: Neuropathology Image Analysis

As a part of the SEA-AD study, Allen Institute scientists imaged sections of donor brain tissue from the middle temporal gyrus (MTG) of all 84 donors. Before imaging these samples, the scientists applied a variety of stains that dye specific proteins and cell types of interest that are believed to play a role in AD pathology. But how do the scientists know what to look for in the brain tissue?

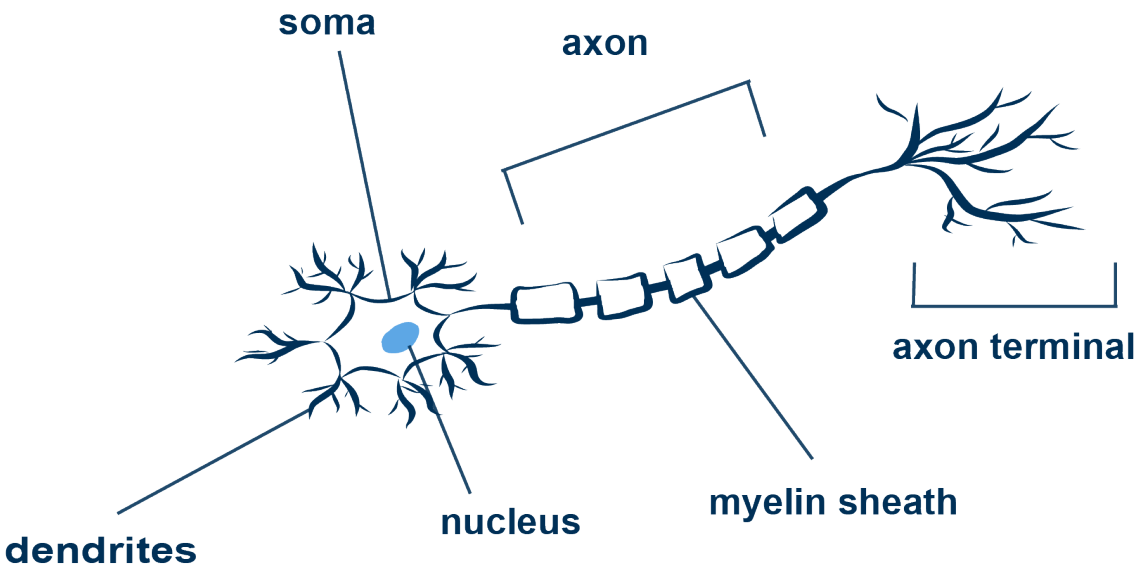
In order to look for pathological signs of AD in each donor's sample, scientists look for specific biomarkers. **Biomarkers** are biological signs of disease. Research on AD has identified several possible biomarkers for the disease. While several different possible biomarkers have been identified by AD researchers, this lesson will focus on two in particular: **beta-amyloid plaques** and **neurofibrillary (tau) tangles**.

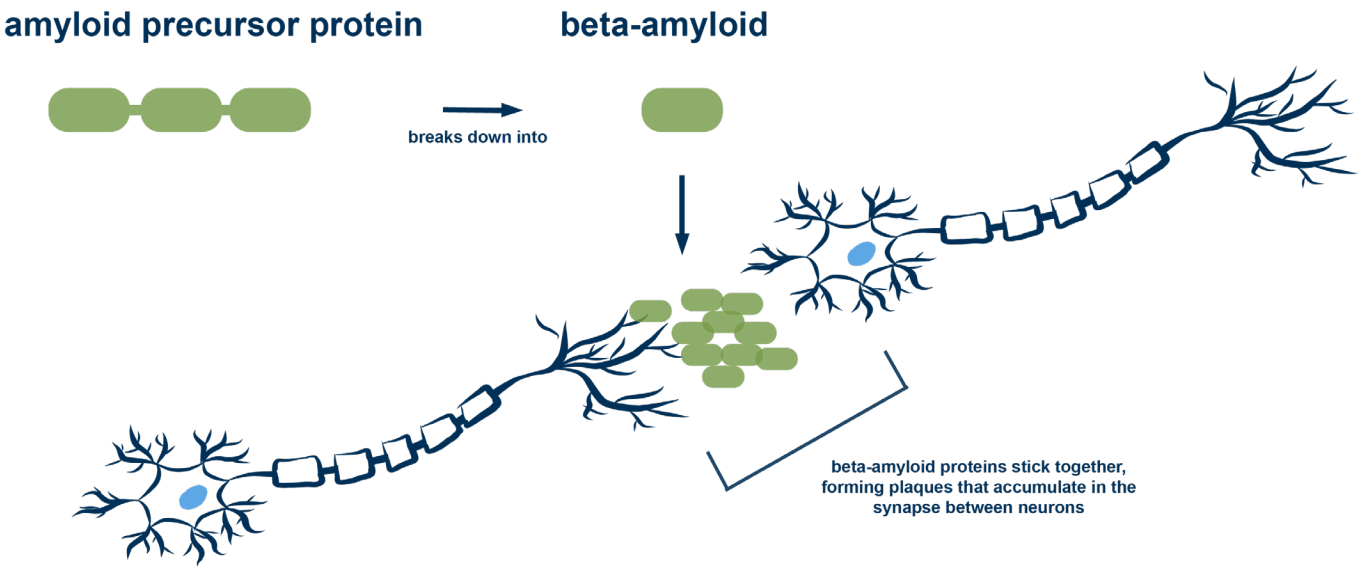
Before looking at the images using the Allen Institute's SEA-AD Neuropathology Image Viewer, we will explore what beta-amyloid plaques and neurofibrillary (tau) tangles are and the current theories surrounding how they play a potential role in AD pathology.

Note: The images below depict cartoon diagrams of neurons and non-neuronal cells. While these cartoons can be helpful to understand basic neural anatomy, they do not capture the full complexity and/or diversity of brain cells. For a deep dive into the structure and function of neurons, please check out the Neurons: Beyond the Textbook lesson located at https://alleninstitute.org/about/education-outreach/neurons-beyond-textbook/.

1. Beta-Amyloid Plaques

Research into AD biomarkers has identified beta-amyloid plaques as a possible biological hallmark of AD pathology. Amyloid precursor protein (APP) is a larger protein that is broken down into a smaller protein called beta-amyloid 42. In patients with AD, abnormally high levels of beta-amyloid appear to clump together to form plaques. These plaques can accumulate in the synapses between neurons and disrupt neuronal functioning.





Scientists measure the extent of the anatomical distribution of beta-amyloid plaque deposits by organizing pathology into different Thal phases. The **Thal phases** range from stage 0 (least severe extent of beta-amyloid plaque accumulation) to stage 5 (most severe beta-amyloid plaque accumulation).

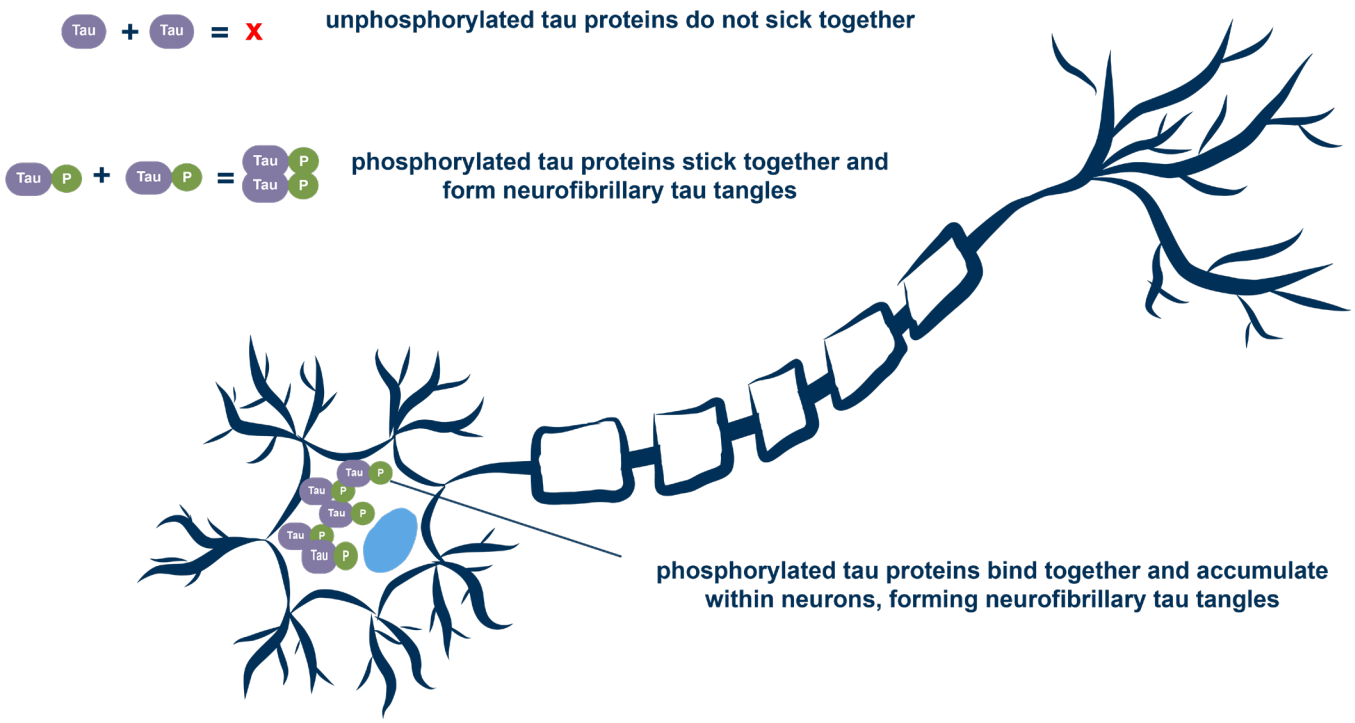
Reference: https://www.nia.nih.gov/health/what-happens-brain-alzheimers-disease

2. Neurofibrillary (Tau) Tangles:

An additional potential biomarker for AD is neurofibrillary (tau) tangles. While beta-amyloid plaques accumulate in between neurons, these tau tangles appear to accumulate at abnormally high levels inside neurons. Tau proteins in healthy neurons typically bind to microtubules. In patients with AD, these tau proteins undergo a chemical change when they are phosphorylated. This phosphorylation causes tau proteins to bind to each other rather than to microtubules, and these tau "tangles" then accumulate inside neurons. The accumulation of these tangles appears to disrupt the transport system within a neuron.

Scientists measure the extent of the anatomical distribution of neurofibrillary tangles by organizing pathologies into Braak stages. **Braak stages** range from stage 0 (least severe amount of neurofibrillary tangles) to stage 6 (most severe extent of neurofibrillary tangles). Roman numerals are frequently used to describe Braak stages. See below for how to interpret roman numerals for numbers 1-6:

- 1 = I
- 2 = II
- 3 = III
- 4 = IV
- 5 = V
- 6 = VI



Neuropathology Image Analysis:

Now that you are familiar with a few of the suspected biomarkers for AD, you can use this knowledge to look at actual neuropathology images from the donors in the Allen Institute's SEA-AD study. In particular, we will be looking at the presence/absence of beta-amyloid plaques within each tissue sample, as well as the relative distribution of these plaques between patient samples.

The neuropathology image viewer organizes the images by each donor's ID number. The Allen Institute uses a donor ID number rather than the donor's name in order to maintain the donors' anonymity and to respect their privacy.

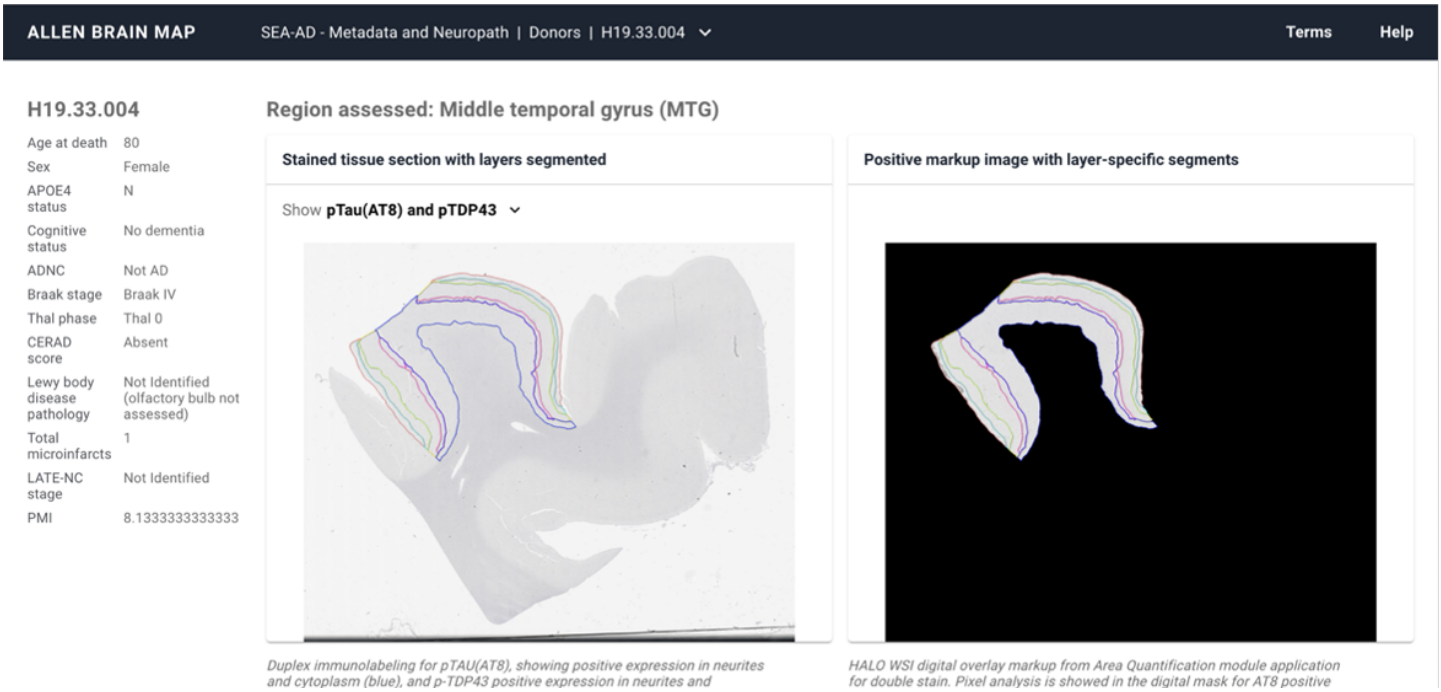
Donor H19.33.004:

We still start by looking at the neuropathology images of donor H19.33.004.

Step 1: Open up the neuropathology image viewer for donor H19.33.004:

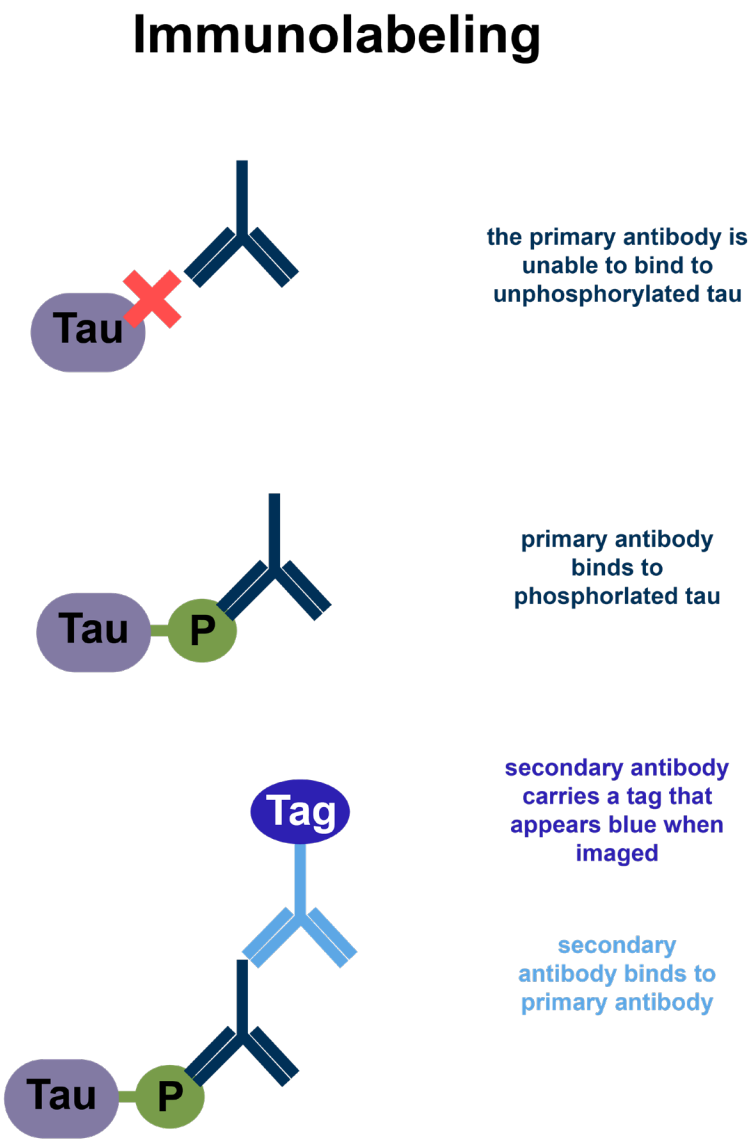
https://knowledge.brain-map.org/data/JGN327NUXRZSHEV88TN/donors/B2YX5RFBGNHG-F6R18GG

Your screen should look like this:

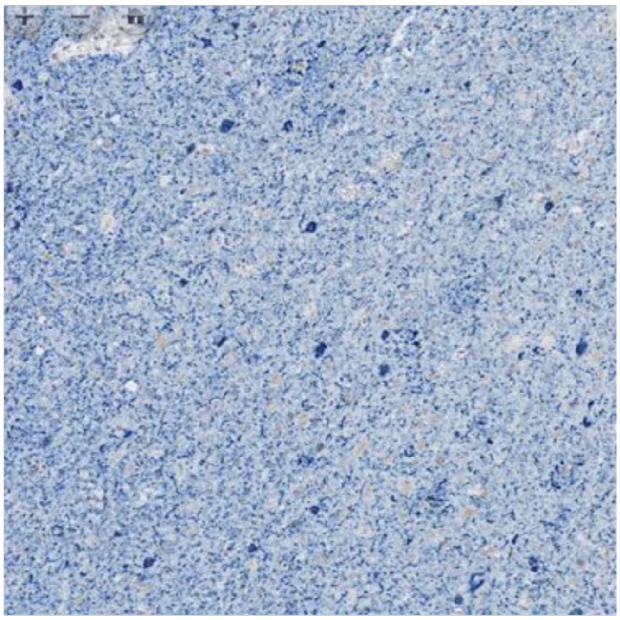


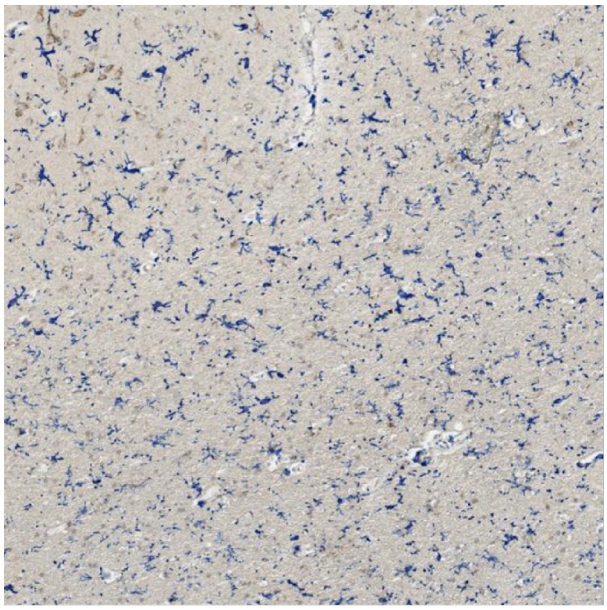
There are two different options for viewing this tissue sample. On the left, you can view the image under the "stained tissue section with layers segmented" option. The lefthand panel allows you to choose which immunolabeled tissue sample you are looking at.

Immunolabeling is the process where antibodies are used to visualize specific proteins. See the figure below to see the process of immunolabeling:



- **Which of the following images of immunolabeled brain tissue appears to have a larger amount of phosphorylated tau protein (tau tangles)? How do you know?**

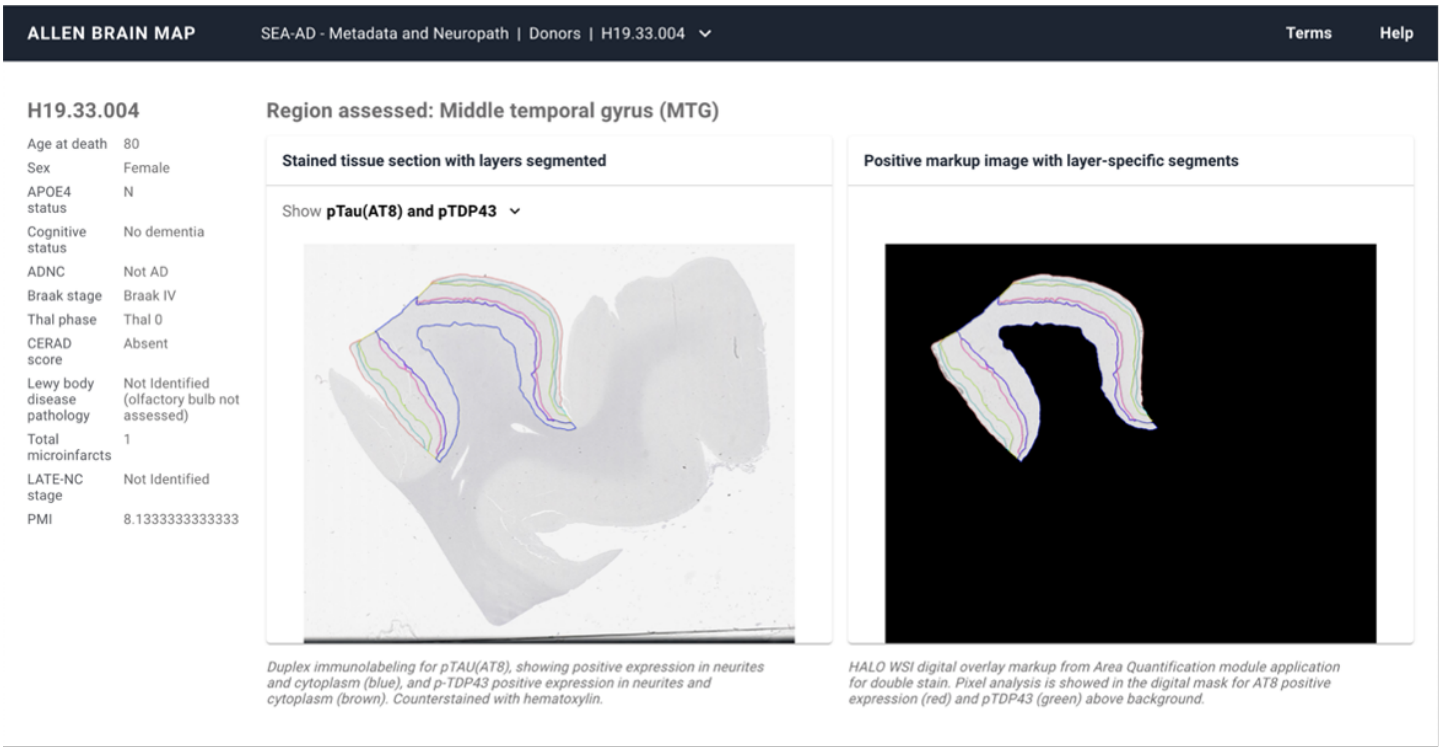




Please write your answer in the box below:

The default setting shows immunolabeling for "**Abeta(6E10) and IBA1**," which allows researchers to stain for beta-amyloid plaques. You can change this setting to select for stains for a variety of proteins. We will be changing the immunolabeling we are looking at in order to search for the presence or absence of neurofibrillary (tau) tangles.

Step 2: Click on the drop down menu and change the immunolabel from "Abeta(6E10) and IBA1" to "**pTau(AT8) and pTDP43**." Your screen should now look like this:



The **pTau(AT8)** and **pTDP43** are two different immunolabels that allow researchers to visualize the following proteins in tissue samples:

• **pTau(AT8):** AT8 stands for anti-phospho-tau and is a monoclonal antibody that binds to **phosphorylated tau protein**. Remember from before that phosphorylated tau protein tends to form neurofibrillary tangles that can accumulate within neurons. These tangles are suspected to play a role in AD pathology, and thus are often viewed as hallmark of AD.

• **pTDP43**: This stands for phosphorylated transactive response DNA-binding protein 43 (TDP43). This protein is suspected to form pathologic aggregates in AD
Reference: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3830649/

Recall from the immunolabeling process that the secondary antibody used is connected to a fluorescent tag that allows scientists to localize and visualize the proteins of interest. Allen Institute scientists used a tag for **phosphorylated tau tangles** that appears blue and a tag for **phosphorylated TDP43** that appears brown.

On the right, the "positive markup image with layer-specific segments" displays only part of the sample that has been labeled by layer of the cortex. In this image, phosphorylated tau proteins (tangles) appear red and phosphorylated TDP43 protein appears green. The colored lines on the image in blue, pink, green, and yellow are separating out the different layers of the cortex.

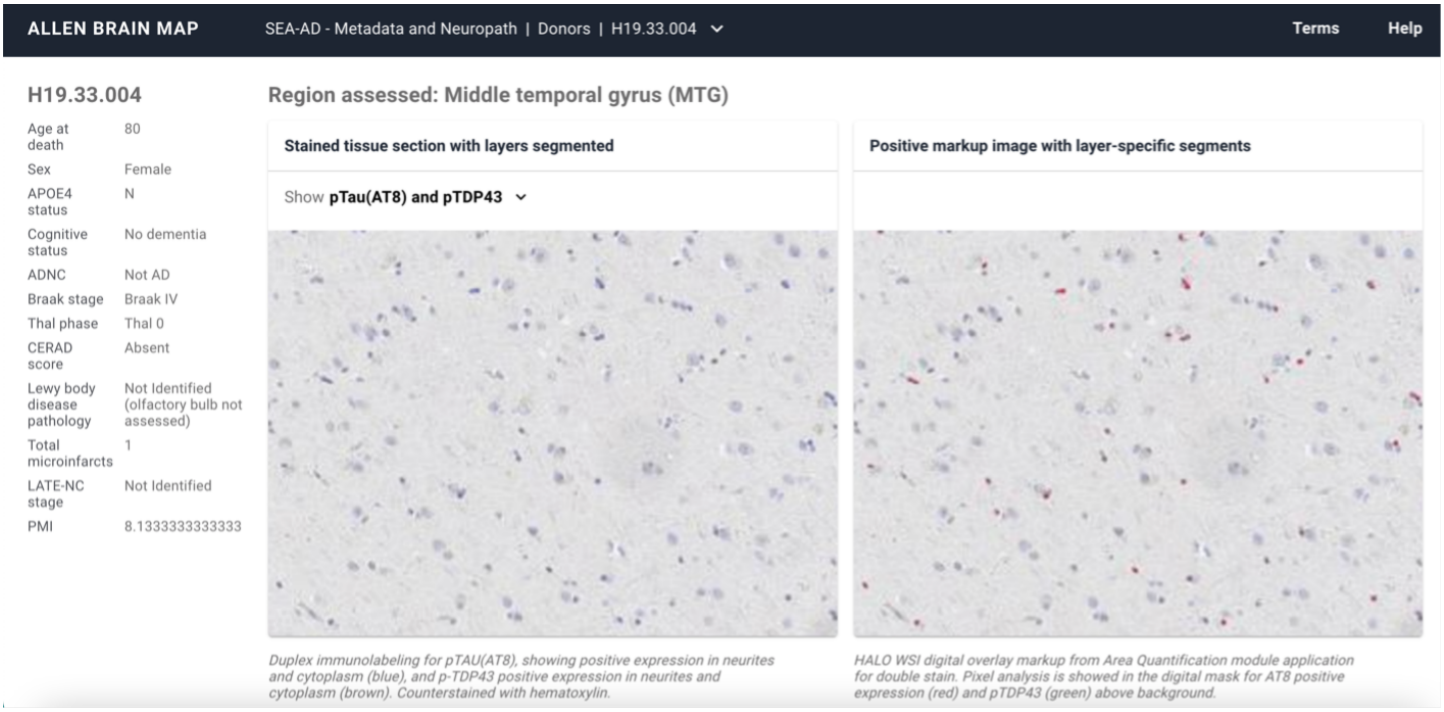
To get acquainted with the neuropathology image viewer, fill in the following table about donor H19.33.004:

	Donor H19.33.004
Sex	
Cognitive Status	
ADNC	
Braak Stage	

1. What does ADNC tell us about the donor? (Hint: if you are unfamiliar with what ADNC measures, please look back at the key featured earlier in the lesson, which is also linked here.)

2. What does the Braak stage tell us about the donor?

Look at the image on the left. Zoom in to any part of the sample. This is one example of what you see, but your image could look different based on where you zoom in:



1. Do you see any phosphorylated tau tangles in this sample?

2. Based on the characteristics of donor H19.33.004, did you expect to see a large or small number of phosphorylated tau tangles? Explain your answer.

Donor H20.33.031:

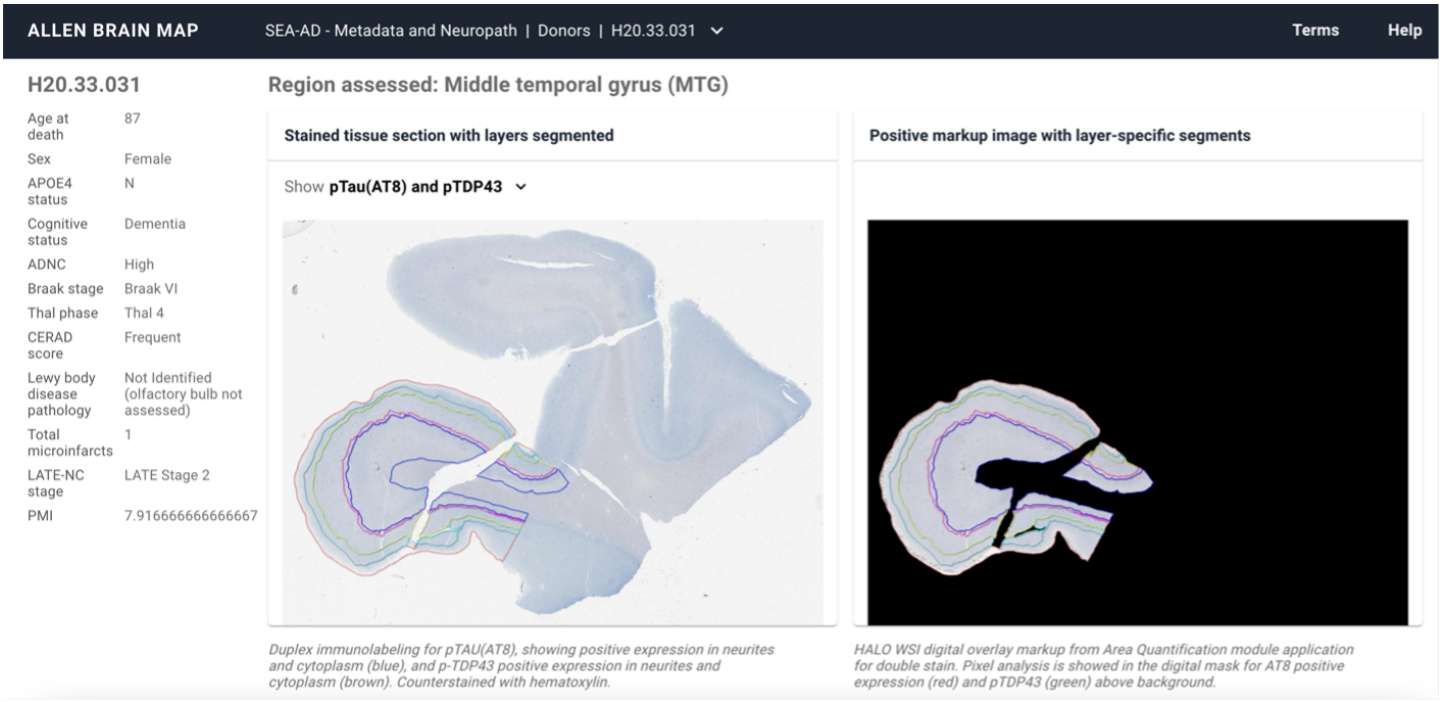
The neuropathology image viewer allows us to compare images between donors. In a **new tab**, go to:

https://knowledge.brain-map.org/data/JGN327NUXRZSHEV88TN/donors/A1NYU4F5DY621E8HO43

Note: opening this link in a new tab will allow you to do a side by side comparison of donor H19.33.004 from before and H20.33.031.

Change the immunolabel to **pTau(AT8) and pTDP43**.

Your screen should look like this:



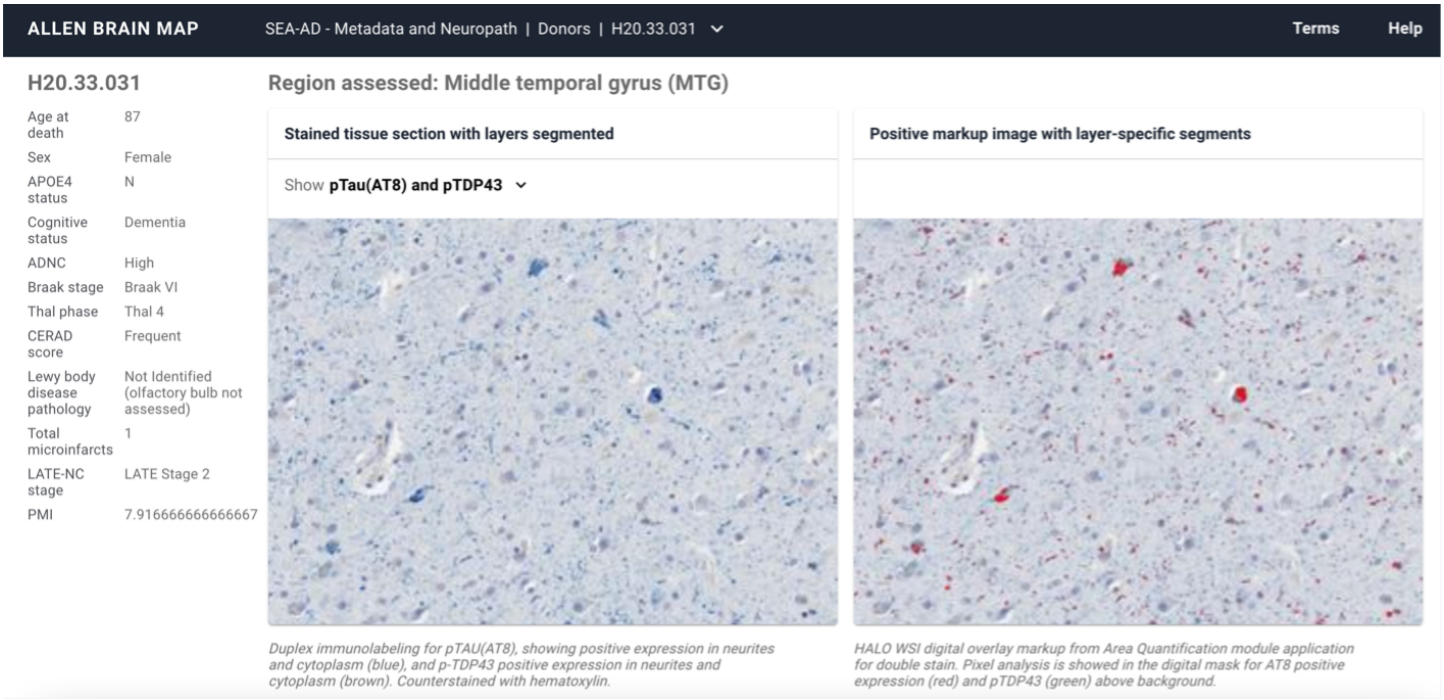
On the following page, fill in the table below to orient yourself to the characteristics of this donor:

	Donor H20.33.031
Sex	
Cognitive Status	
ADNC	
Braak Stage	

1.Notice that the stained tissue section with layers segmented (the image on the left) has a different overall shape than the tissue section from the previous donor (H19.33.004). Why do you think that is?

2. Looking only at the donor characteristics, what is different between donor H19.33.004 and H20.33.031?

Zoom in on the image on either the left or the right. This is one example of what you might see, although the image could differ based on where you chose to zoom in:



1. Do you see any tau tangles in this image?

2. Based on the characteristics of donor H19.33.004, did you expect to see a large or small number of tau tangles? Explain your answer.

3. Looking only at the neuropathology images, what is different between donor H19.33.004 and H20.33.031?

After comparing the neuropathology images between two female donors, let's do a comparison amongst male donors.

Donor H21.33.019 and Donor H20.33.046:

In a new tab, open: https://knowledge.brain-map.org/data/JGN327NUXRZSHEV88TN/donors/6JRB7FMYBVE5F3HDTM5

In ANOTHER new tab, open: https://knowledge.brain-map.org/data/JGN327NUXRZSHEV88TN/donors/HB9ENTM30SNBC4EP0PU

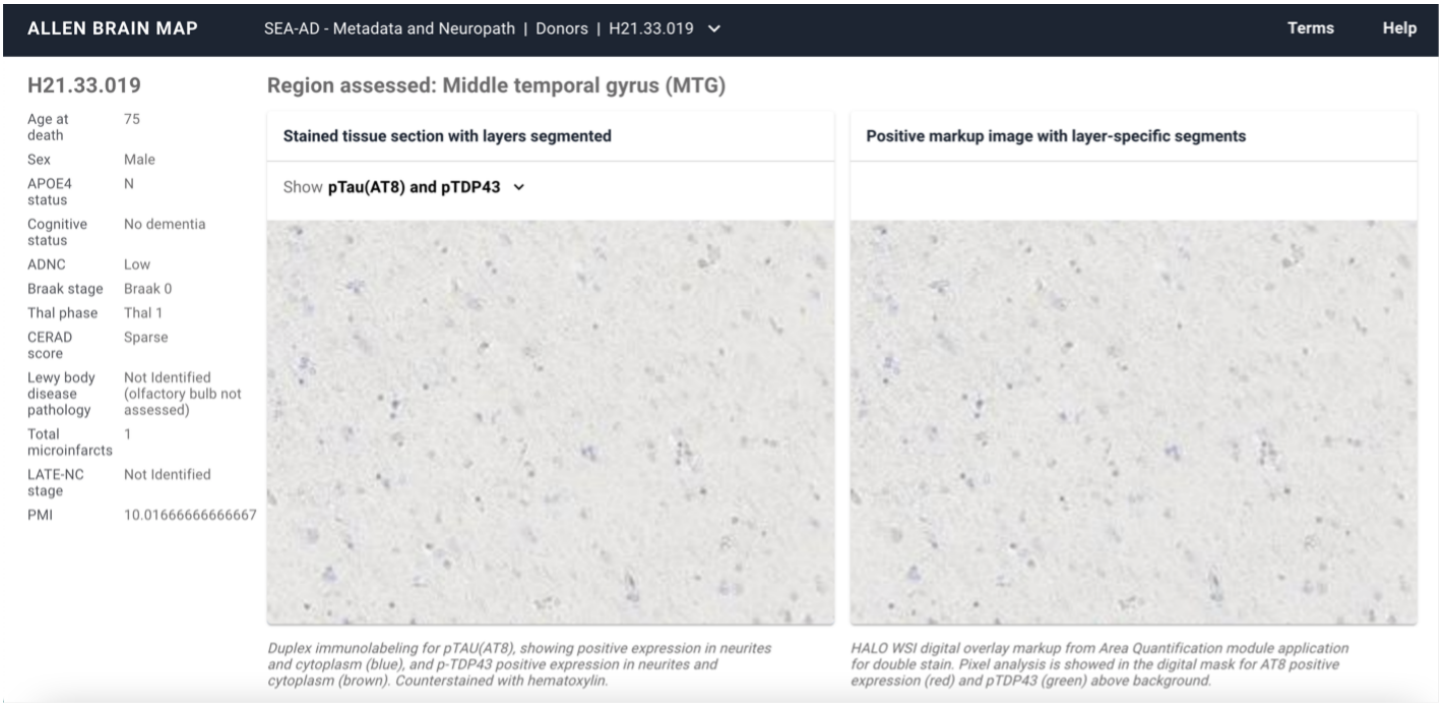
To clarify: in one tab you should have open the neuropathology images for donor H21.33.019, and in another tab you should have open the images for donor H20.33.046.

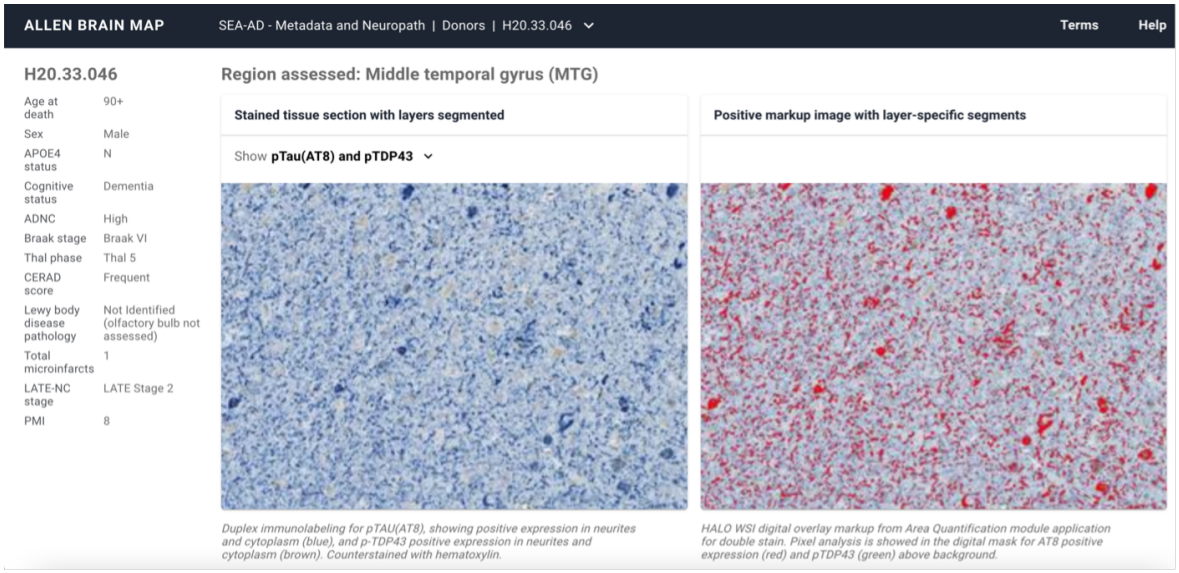
For both of these donors, be sure to change the immunolabel from the default to the pTau(AT8) and pTDP43 immunolabel.

Fill in the table below to do a side-by-side comparison of the two donors and their characteristics:

	Donor H21.33.019	Donor H.20.33.046
Sex		
Cognitive Status		
ADNC		
Braak Stage		

Zoom in on both of the donor images to compare:





1. What do you notice is similar between the donor images?

2. What do you notice is different between the donor images?

3. Did you expect donor H.21.33.019 and donor H.20.33.046 to have similar neuropathological images or different? Why or why not?

Qualitative vs. Quantitative Analyses:

Qualitative Analysis:

The comparative analysis you just performed was a qualitative assessment of Alzheimer's neuropathology. We were looking for the presence/absence of beta-amyloid plaques, but we were not attempting to quantify or measure the amount of these plaques.

Quantitative Analysis:

After scientists perform a qualitative analysis of neuropathology images, they often continue their analysis by quantitatively measuring the amount of biomarkers present in a sample.

Reflective Questions

1. After viewing the neuropathology images, how would you predict scientists go about quantitatively measuring biomarkers within a sample?

2. A common next step for scientists in a quantitative image analysis is to perform an expert annotation of the sample. What do you think it means for someone to expertly annotate a neuropathology image?

3. After looking at these neuropathology images, what other things would you want to explore in future qualitative or quantitative analyses of these images?

Conclusion:

Throughout the course of this lesson, you had the opportunity to explore AD pathology through qualitatively analyzing neuropathology images from the Allen Institute for Brain Sciences' SEA-AD project. We began the lesson by reflecting on the history of biomedical research and how the field of science continues to strive toward improving the diversity of its study cohorts. After exploring the demographic characteristics of the 84 donors included within the Allen Institute's SEA-AD study, you viewed the neuropathology images from a few of these donors.

In lesson 4, you will have the opportunity to explore AD pathology even further by analyzing transcriptomic data. Rather than view specific images of the brain tissue, you will instead analyze transcriptomic data from each donor's brain sample and explore how scientists can use data about gene expression to explore AD pathology.

For more information on the Allen Institute for Brain Science's SEA-AD project, please visit: https://portal.brain-map.org/explore/seattle-alzheimers-disease.

Supplementary Material 2 for Meuler M, Casimo K (2024) A Case for the Use of Open Data as a Tool to Incorporate Socioscientific Topics into Neuroscience EducationJ Undergrad Neurosci Educ 22(2):A90-A98.

ALLEN INSTITUTE

Lesson 4: Analyzing Transcriptomic Data to Explore Alzhiemer's Disease Pathology

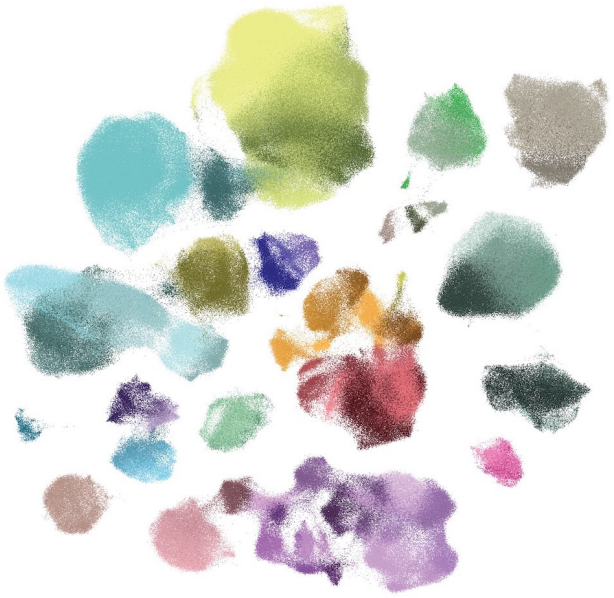


Image is from the Seattle Alzheimer's Disease Brain Cell Atlas consortium

Learning Objectives:

- Students will be able to articulate how transcriptomic data is collected and processed
- Students will be able to interpret transcriptomic data visualized within Uniform Manifold Approximation and Projections (UMAPs)
- Students will be able to navigate the Seattle Alzheimer's Disease Brain Cell Atlas (SEA-AD) database in order to interpret transcriptomic data
- Students will be able to filter data based on specific biomarker and/or demographic characteristics they are interested in exploring
- Students will be able to independently perform an analysis of transcriptomic data using the cellxgene interface
- Students will be able to compare gene expression between cell types using the cellxgene interface
- Students will be able to navigate the NIH gene database to explore the known functions of genes

Introduction

In this lesson, we will explore how scientists use information gathered about a cell's gene expression to explore the pathology of Alzheimer's Disease (AD). AD is a progressive neurological disorder that impacts a person's cognitive and memory abilities.

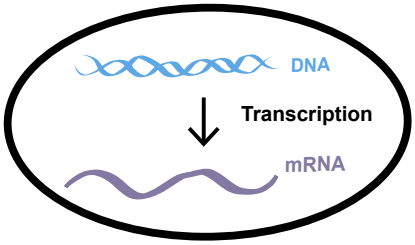
Although AD pathology can be studied in a variety of ways, scientists at the Allen Institute for Brain Science approach this research by paying extremely close attention to the cellular process of transcription analyzing which genes a cell is expressing and in what quantities.

Although transcriptomics can be used across a variety of fields, this type of data is becoming increasingly prominent within neuroscience. Neuroscientists at the Allen Institute who study AD pathology rely heavily on transcriptomic data to study brain tissue donated by individuals both with and without dementia after their deaths.

Review: Transcriptomic Data

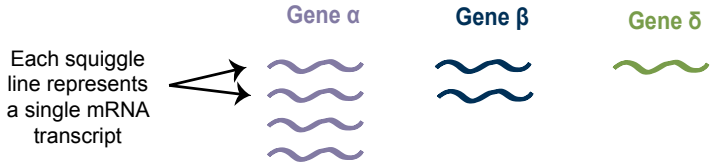
As you learned in lesson 2, transcriptomic data is a type of data that allows scientists to investigate which genes a cell is transcribing/expressing and in what quantities. If a cell, and more specifically, that cell's nucleus, contains a specific RNA transcript, this indicates that the cell is expressing the specific gene associated with that RNA. By (1) isolating nuclei, (2) sequencing the mRNA transcripts found within the nuclei, and (3) counting those transcripts, we can tell **which genes** the cell is expressing and **how much** these cells are expressing these genes. Repeating this process for thousands of cells from a sample of brain tissue allows researchers to find similarities and dissimilarities between cells on the basis of their gene expression. These patterns of similarity and dissimilarity are then used to classify certain cells as specific "types." The graphic below explains in detail how scientists gather and interpret transcriptomic data.

Carefully read through the graphic below to review how transcriptomic data is gathered and interpreted:



Isolate the nuclei from the cells in the sample of brain tissue and extract the RNA found in each nucleus.

Sequence the mRNA transcripts found in each cell's nucleus in order to determine which genes each brain cell was expressing.



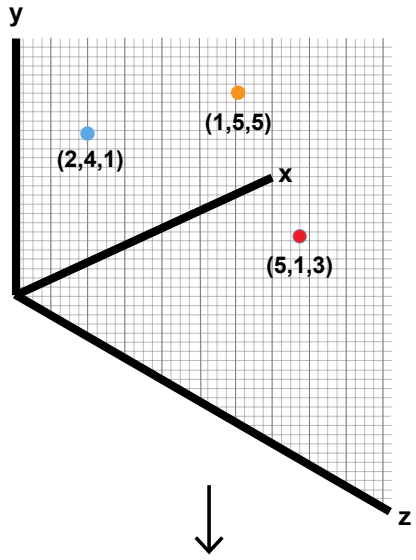
Count the number of mRNA transcripts found for each gene. This allows us to quantify how much each cell was expressing each gene.

	Gene α	Gene β	Gene δ
# of mRNA transcripts found in Cell 1	4	2	1

Create a table that shows how much each cell was expressing each gene.

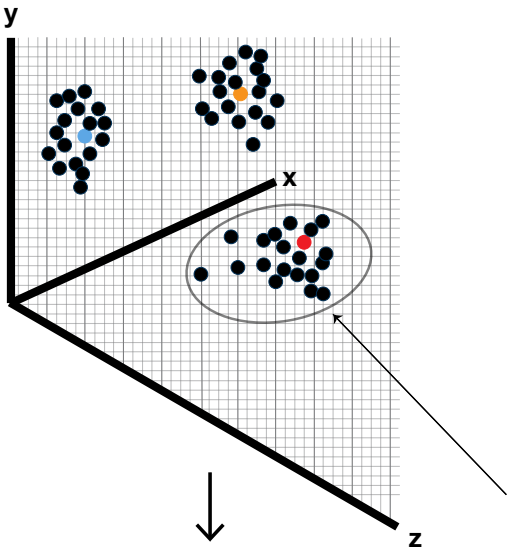
	Gene α	Gene β	Gene δ
# of mRNA transcripts found in Cell 1	2	4	1
# of mRNA transcripts found in Cell 2	1	5	5
# of mRNA transcripts found in Cell 3	5	1	3
repeat count for thousands of cells...	...	...	...

Repeat this process for THOUSANDS of cells. Remember, this means we are counting how much EACH cell was expressing EACH gene. If we wanted to create a table that listed the data in full, this data table would have thousands of rows.



x = gene α value
y = gene β value
z = gene δ value

If we wanted to create a graph that plotted the inital data for cell 1, cell 2, and cell 3 and their relative amount of expression of gene alpha, gene beta, and gene delta, we would need a 3D graph like the one on the left.



x = gene α value
y = gene β value
z = gene δ value

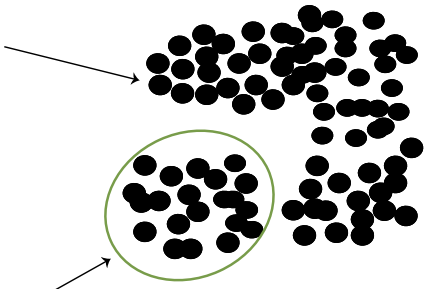
We can repeat this process for the thousands of cells that were collected from the brain tissue sample. Notice that the cells begin to cluster based on how similar their gene expression for gene alpha, gene beta, and gene delta is to one another. These clusters help us identify which cells may be more similar and/or dissimilar to one another!

when we plot the gene expression data for more cells, we notice that cell 3 (red dot) clusters next to these other cells from the sample

	Gene α	Gene β	Gene δ	repeat for thousands of genes...
# of mRNA transcripts found in Cell 1	2	4	1	...
# of mRNA transcripts found in Cell 2	1	5	5	...
# of mRNA transcripts found in Cell 3	5	1	3	...
repeat count for thousands of cells...	...	...	...	...

In addition to collecting data on gene expression for thousands of cells, scientists will add another layer of complexity by measuring the gene expression of these thousands of cells for THOUSANDS of genes. A table displaying this data would have thousands of rows and thousands of columns. Since the graph would now have much more than just 3 dimensions, we will need a special type of tool to graphically represent this data in a way that humans can visualize.

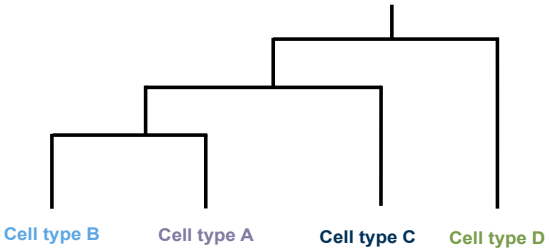
Each dot represents a single nucleus isolated from a single brain cell



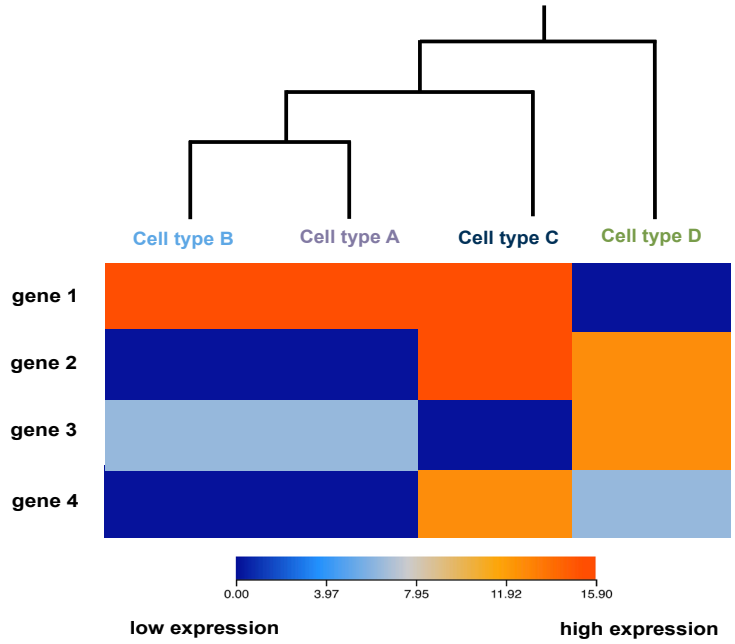
Identify clusters in the data--these clusters represent cells that are more like each other than they are like any other cells

UMAP

In order to plot this many-dimensional graph in a way humans can visualize, we use a dimensionality reduction tool, such as a UMAP, to plot it in a 2D space. Dimensionality reduction is a technique that helps represent many-dimensional data in just two or three dimensions.



Organize the clusters identified in the UMAP to construct a dendrogram that displays hierarchical relationships between the clusters based on each cell type's similarity and dissimilarity of gene expression.



Use a heatmap below the dendrogram to compare the level of gene expression between each cell type for specific genes of interest.

Exploring the different cell types of the brain helps us to understand how healthy brains function. This information is crucial to gather to understand what changes in the brain when certain neurological diseases arise. A common neurological disease that impacts memory capabilities and cognitive functioning is AD. Understanding the different cell types of the brain is crucial to AD research, as researchers can ask questions such as:

- *Do aged brains from people with dementia contain the same types of cells as brains from young adults?*
- *Do specific types of cells show changes in gene expression between a healthy brain and a brain with AD?*

Mapping the Brain:

Researchers at the Allen Institute are exploring these questions and more as a part of the Allen Institute for Brain Science's Seattle Alzheimer's Disease Brain Cell Atlas, which is abbreviated as SEA-AD. As a part of this project, scientists gather large amounts of transcriptomic data from post-mortem donated brain tissue from individuals within the Seattle area. This data is then uploaded into an online database called the SEA-AD Cell Atlas which is freely available to the public. The SEA-AD Cell Atlas consists of data from 84 older adults with and without AD. As of the writing of this lesson in 2022, the database features data from the middle temporal gyrus (MTG), and more data of the same types from other brain regions will be added in later updates. The Allen Institute for Brain Science has previously studied the MTG in depth in healthy donors, and it is affected relatively early in the progression of AD.

The SEA-AD project strives to gain a deep molecular and cellular understanding of the early pathogenesis of AD. The data collected within this study are derived from a full spectrum of aged donors, from healthy controls to those with high AD pathology and dementia. Data and specimens were obtained from the Adult Changes in Thought (ACT) Study from Kaiser Permanente Washington Health Research Institute (KPWHRI), and the University of Washington Alzheimer's Disease Research Center (ADRC). The ACT study is a study that follows initially healthy donors at age 65 for the rest of their lives. This study records a large range of medical information about each donor, each donor's demographic characteristics, and collects vital information about each donor while they are living and after they have passed.

In order to strengthen your understanding of AD, we will complete a guided tutorial of the SEA-AD database. This database is available online at: SEA-AD.org.

The SEA-AD database is vast and full of many different resources. Given the time constraints of today's lesson, we will focus on only one of the databases that contains the SEA-AD data: **CZ cellxgene**

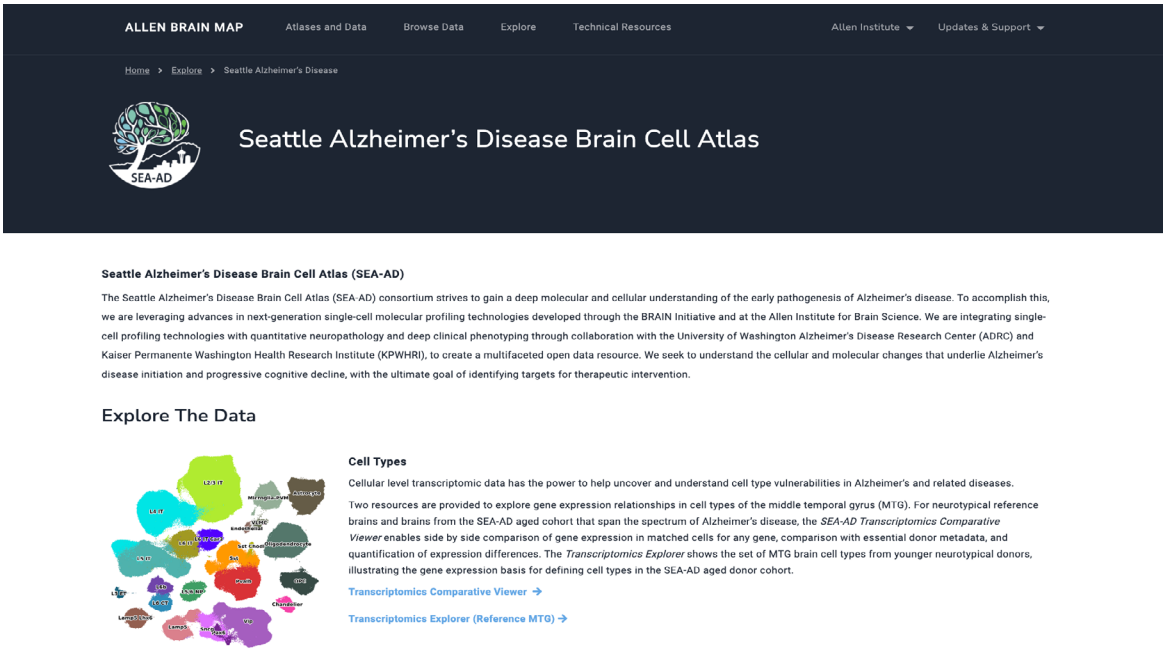
In order to orient you to the CZ cellxgene interface and how it can be used to explore questions related to Alzheimer's disease pathology, the following section provides you with a detailed tour of the database.

Database Tutorial:

In this section, we will explore how to use the ***cellxgene*** *interface. We will learn how to interpret UMAPs and how to filter the data to explore gene expression across cell types, donors, and genes of interest.*

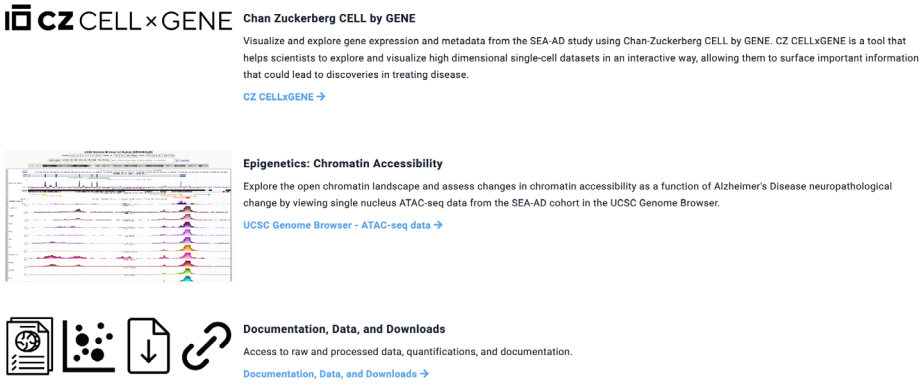
1. In order to access the Allen Institute for Brain Science's Seattle Alzheimer's Disease Brain Cell Atlas, go to SEA-AD.org.

Your web page should look like this:

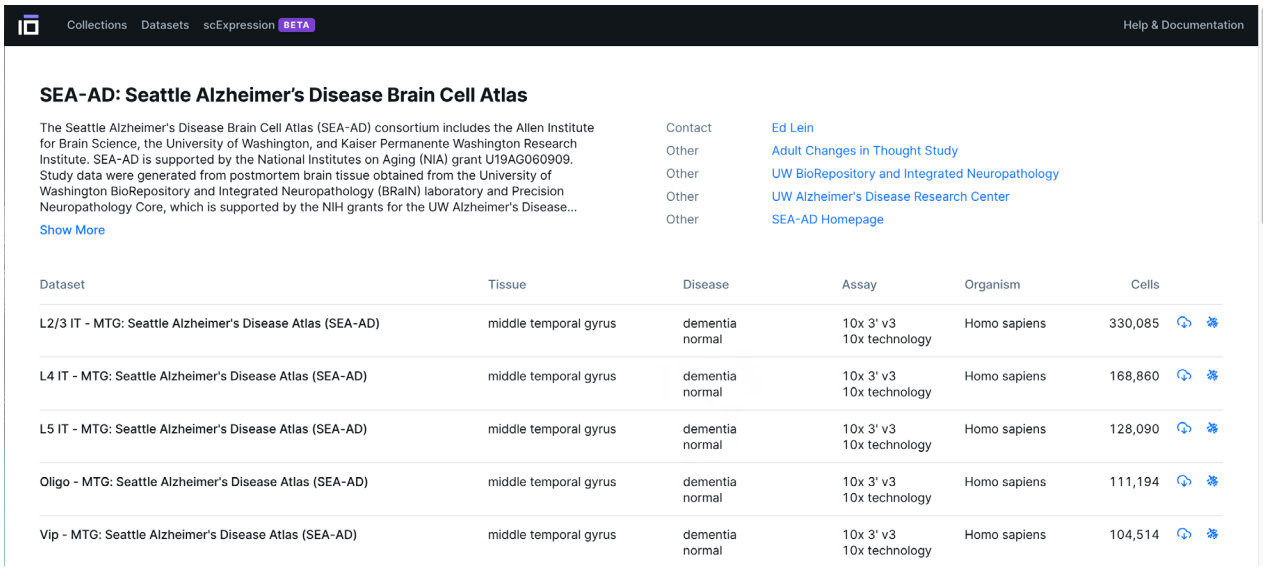


2. Scroll down until you see the **CZ cellxgene** link under the "Explore The Data" header.

3. Click on the "**CZ cellxgene**" link that is in blue.



After clicking the CZ cellxgene link, your page should look like this:



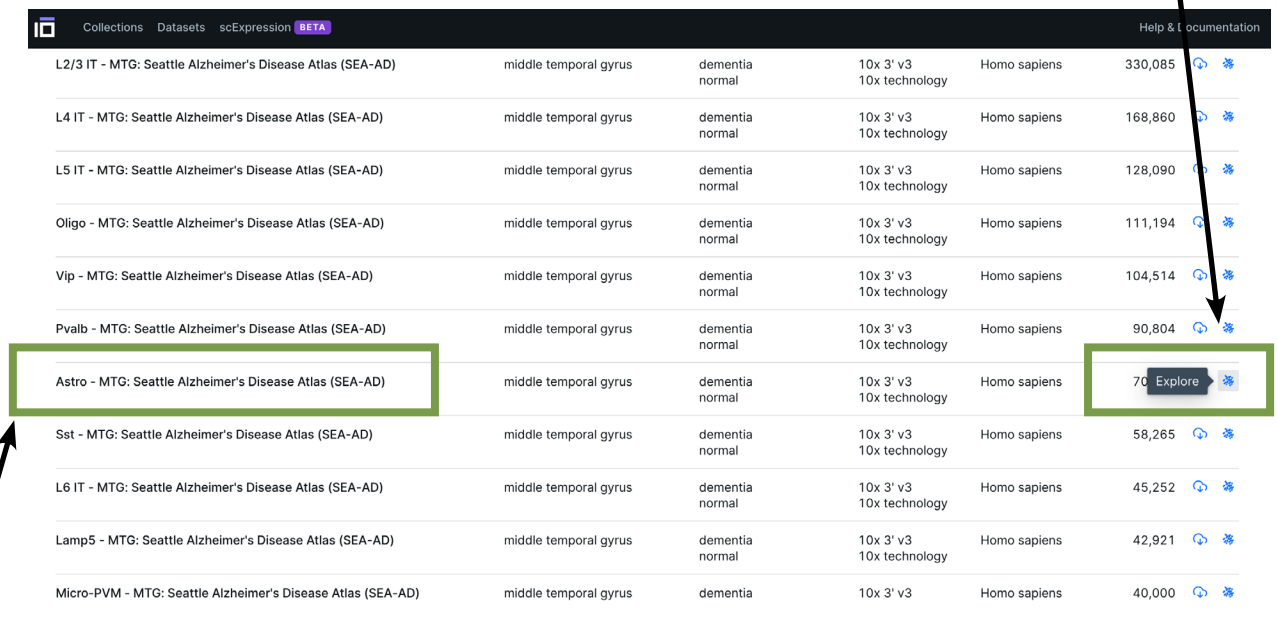
Each dataset linked on this page is for a different **cell type** within the brain. For example, the first dataset is titled "L2/3 IT – MTG: Seattle Alzheimer's Disease Atlas (SEA-AD)." This is the transcriptomic data from cells within the L2/3 layer of the middle temporal gyrus (MTG) within the human brain.

Cell types of interest: Astrocytes and Microglia

Although there are several different cell types listed within the cellxgene database, there are two particular cell types that are of interest to AD researchers. Scientists hypothesize that **astrocytes** (abbreviated as "Astro") and **microglia** (abbreviated as "micro") are two specific types of cells that may play a role in AD pathology. The SEA-AD database contains data for multiple types of brain cells, but for the purposes of this tutorial, we will look at the data for astrocytes and microglia.

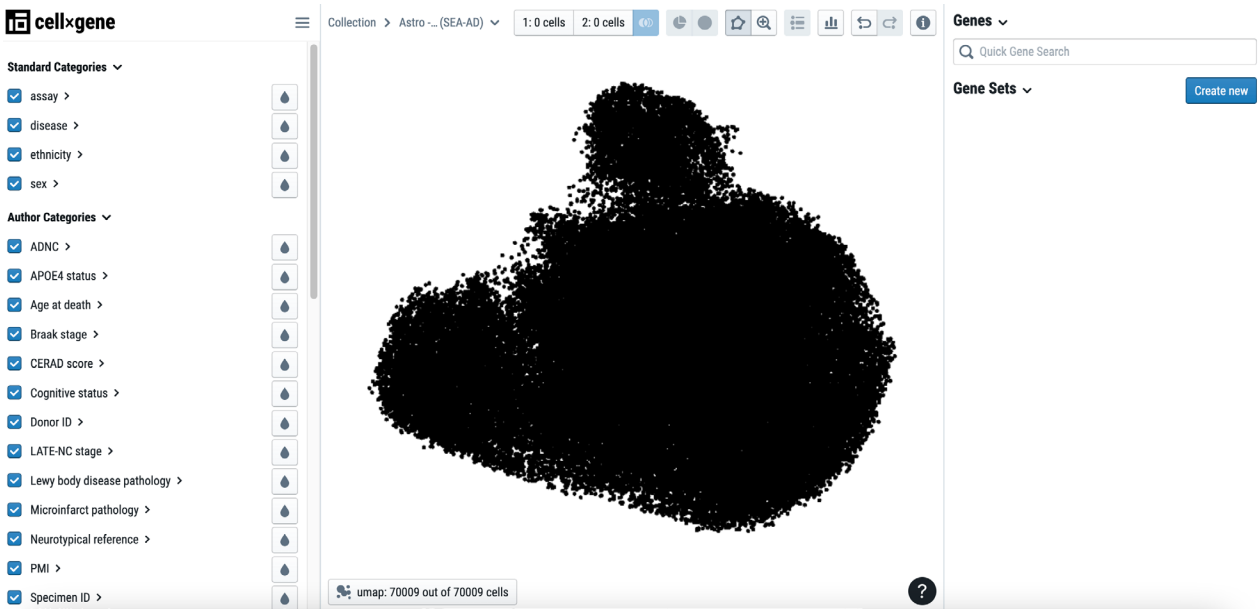
4. For the purposes of this tutorial, scroll down on the page until you find the dataset labeled “Astro-MTG: Seattle Alzheimer’s Disease Atlas (SEA-AD).” On the right-hand side, click on the “explore” icon.

the explore button can be found here next to the download icon



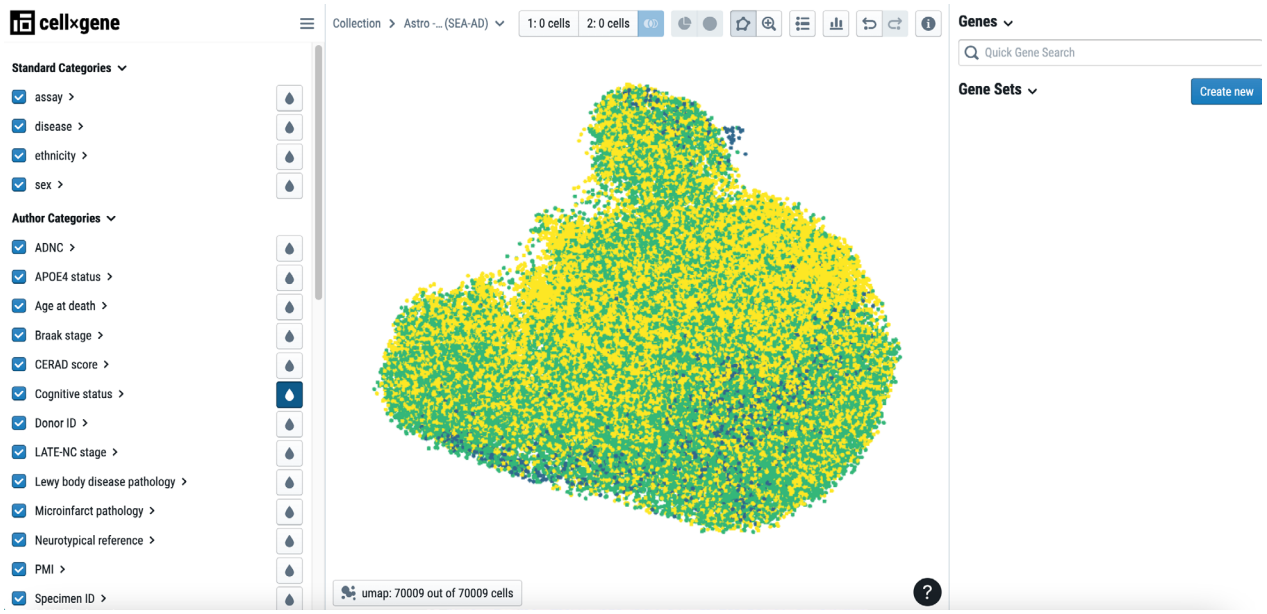
Scroll until you find the dataset labeled “Astro-MTG”

After clicking on “explore,” your screen should look like this:



In order to analyze this figure, we can apply filters. Select the paint drop icon next to each of the categories. For this tutorial, click on the paint drop icon next to “cognitive status.”

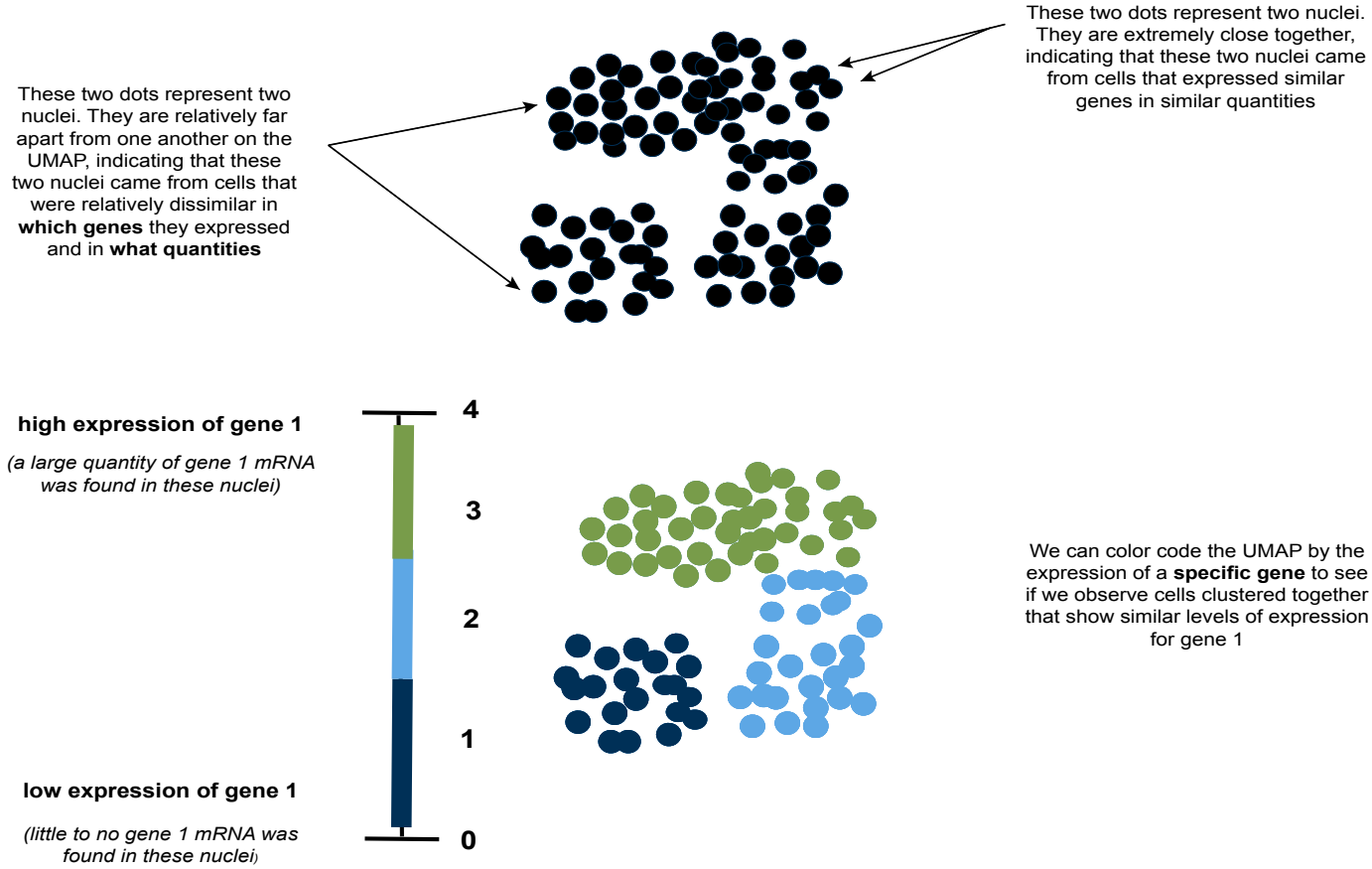
The color of the UMAP should change, and your screen should now look like this:



This is a special type of graph, called a **Uniform Manifold Approximation and Projection (UMAP)**. UMAPs are helpful ways of displaying many types of data, including transcriptomic data. In other words, these graphs help us to compare gene expression between cells. To review how UMAPs are generated and the type of data they display, refer to the transcriptomic data infographic included at the start of the lesson.

In order to learn how to analyze a UMAP, read through the figure listed below and discuss it with a neighbor/classmate/friend.

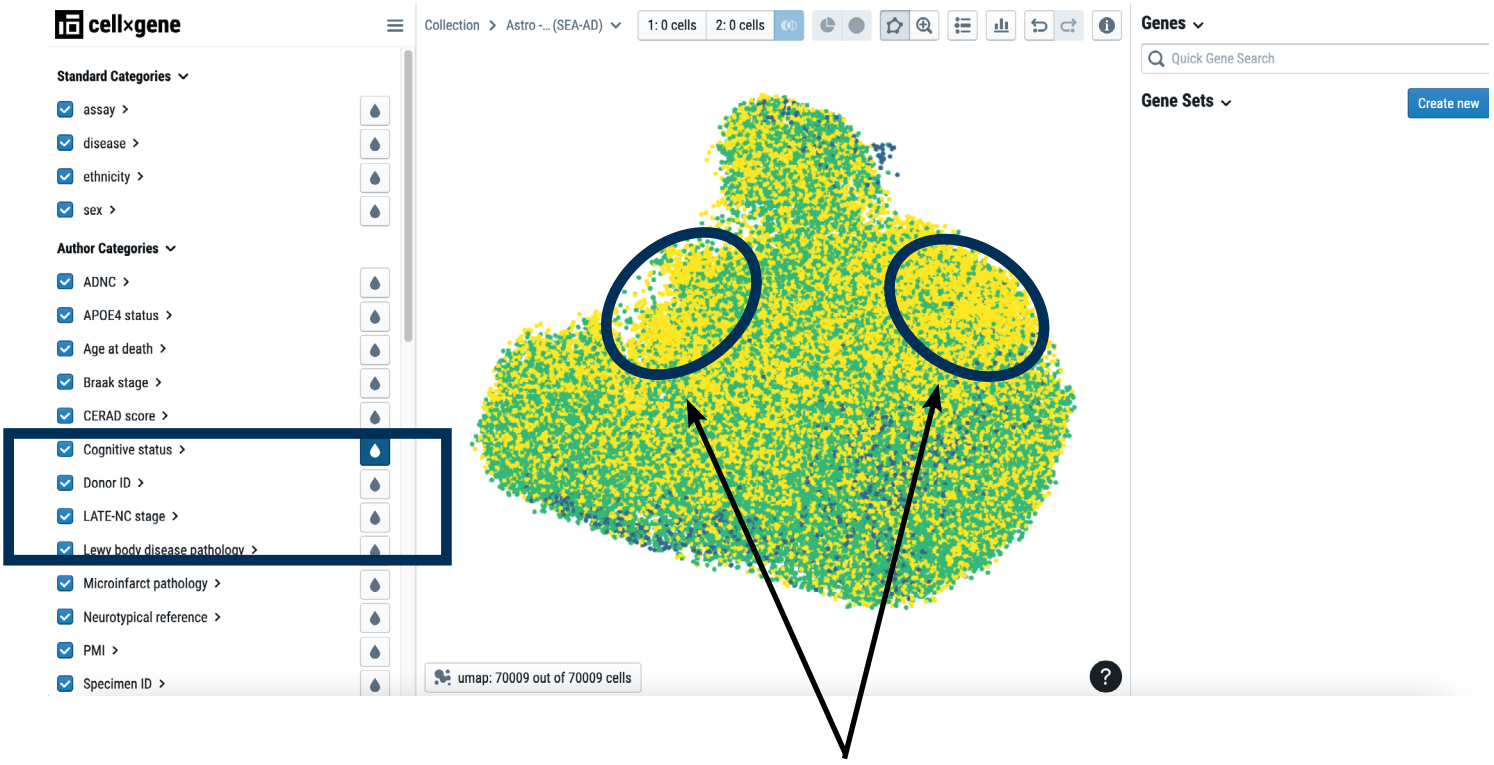
How do we analyze a UMAP?



Filtering by donor characteristics:

Dementia Status

Now that we know how UMAPs are generated and interpreted, let's return to the cellxgene data. By looking at the colors and clusters in the UMAP below, we can compare gene expression between cells taken from patients with dementia (yellow) and without (green).



Yellow dots represent nuclei from patients with dementia. Notice here we see these nuclei clustered together away from the green nuclei. The green nuclei represent nuclei isolated from brain tissue from donors who did not have dementia.

Knowledge Check

1. Does there appear to be a difference in gene expression between cells taken from patients with dementia compared to those without dementia? How can you tell?

2. Is this UMAP comparing gene expression generally, or is it looking at a specific gene?

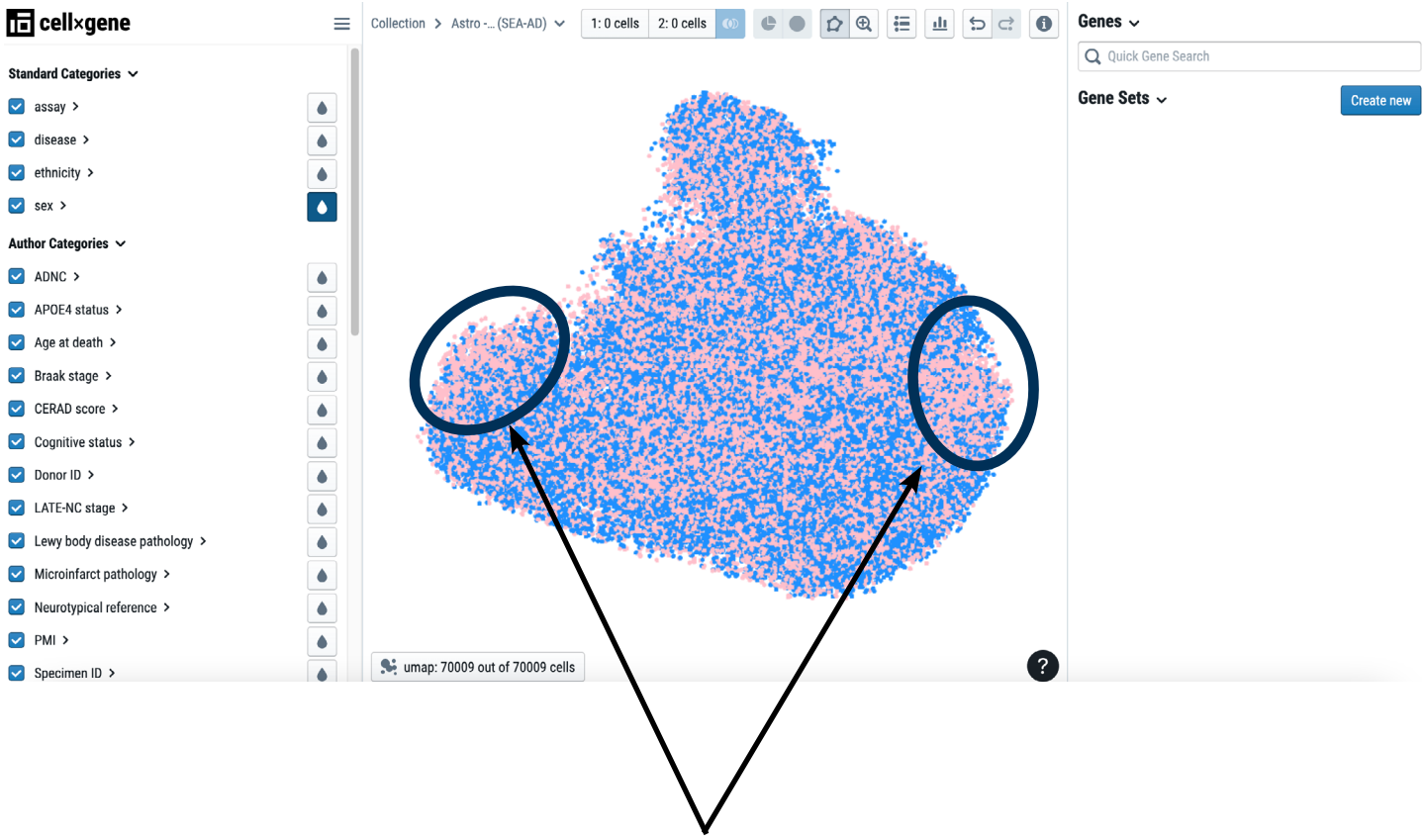
Biological Sex

In addition to filtering by dementia status, we can also compare gene expression of cells taken from biological males and biological females. As discussed in lesson 3, sex refers to a person's biological classification based on their reproductive organs and sex chromosomes, while gender is defined as "set of social, psychological, or emotional traits, often influenced by societal expectations that classify an individual as either feminine or masculine."

Reference: For more information about gender identity and the difference between sex and gender, visit https://medicine.yale.edu/whr/about/mission/definitions/.

Unclick the rain drop icon next to "cognitive status." Your screen should go back to a black UMAP.

Next, click on the raindrop icon next to "sex." Your screen should look like this:



Notice these clusters of pink on the UMAP

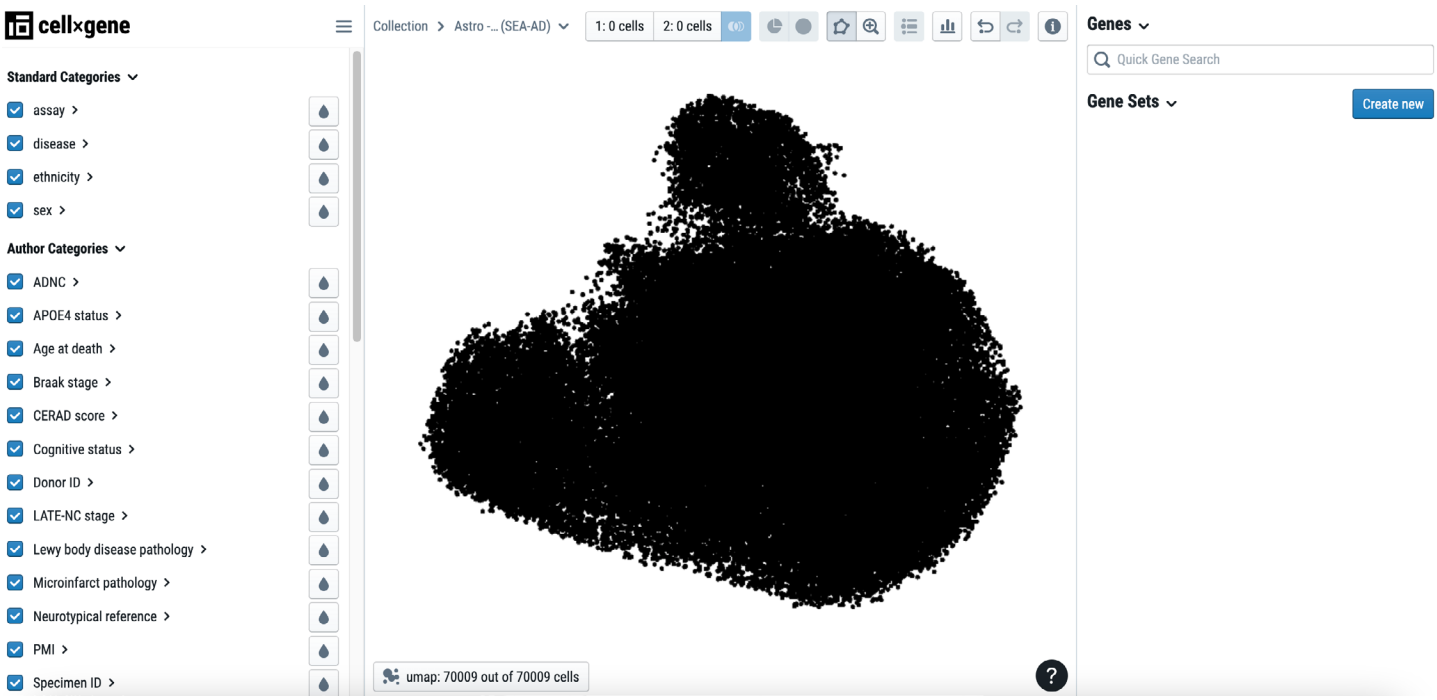
Knowledge Check

1. In your own words, describe what the pink clusters of dots represent within the UMAP on the previous page:

Filtering by gene:

In addition to filtering by one of the donor characteristics, such as cognitive status or donor sex, the cellxgene interface also allows us to compare gene expression between cells for **a specific gene**.

First, deselect the rain drop icon next to "sex." Once again, your screen should now show the black UMAP.



To filter by gene, go to the right hand column labeled "genes." For this tutorial, we will look at the gene labeled "APOE."

APOE is one of several genes that are suspected to play a role in Alzheimer's disease. Specific alleles of the APOE gene, such as APOE4, have been associated with increased risk of Alzheimer's disease. It is important to note that just because an individual has an allele of APOE associated with higher risk, this does not confirm that they will eventually develop AD. For the purposes of this lesson, we are looking at APOE expression broadly and not expression based on certain APOE alleles.

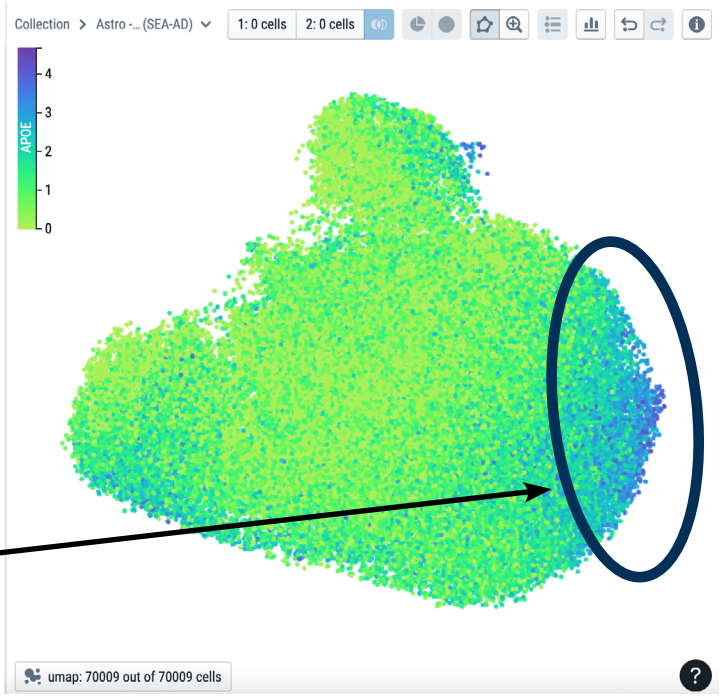
To filter the data by expression of the APOE gene, type in APOE into the "genes" search box.

Next, click the rain drop icon next to the APOE gene. Your screen should look like the figure below:

This key shows which color corresponds to high or low expression of APOE

Blue dots represent nuclei that had high levels of APOE expression.

In this UMAP, we see these blue dots clustered together on the right hand side.



Knowledge Check

1. What do the cluster of blue cells/nuclei at the right-hand of the UMAP tell us about the data?

Looking for Differential Gene Expression:

In addition to looking at the expression between cells of a single gene, the CZ cellxgene tool also allows us to identify genes that are differentially expressed between two categories. This is a particularly helpful feature of the cellxgene interface since it allows us to identify potential genes of interest that show different levels of expression based on specific characteristics like a donor's sex, age, etc.

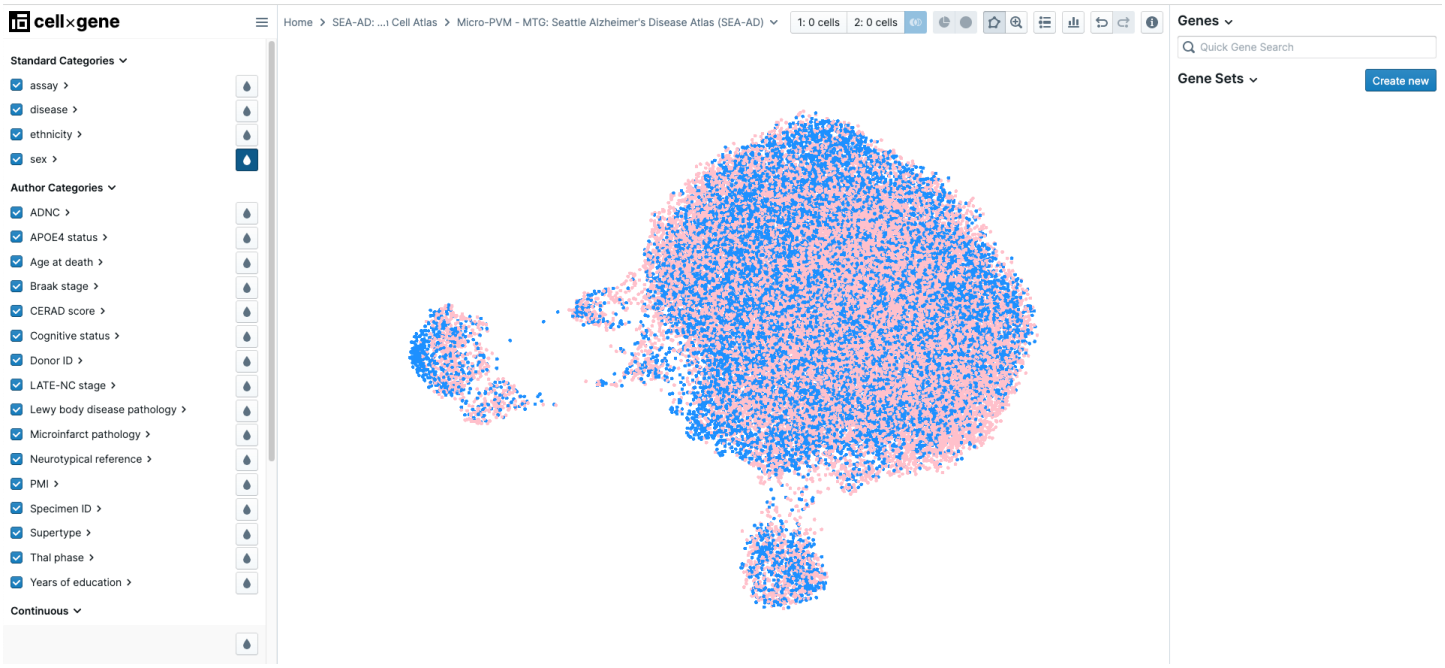
We will explore how to do this below:

1. We will start with a blank slate in cellxgene. Go to: https://cellxgene.cziscience.com/collections/1ca90a2d-2943-483d-b678-b809bf464c30

2. In the earlier part of the tutorial, we explored the data for astrocytes cells from the SEA-AD cohort. For this activity, we will switch to exploring the data for microglial cells. Scroll down and select the explore icon next to: "Micro-PVM – MTG: Seattle Alzheimer's Disease Atlas (SEA-AD)." If you are struggling to locate this, you can also access it directly from this link: https://cellxgene.cziscience.com/e/c76098ba-eed3-45b1-98f2-96fcac55ed18.cxg/

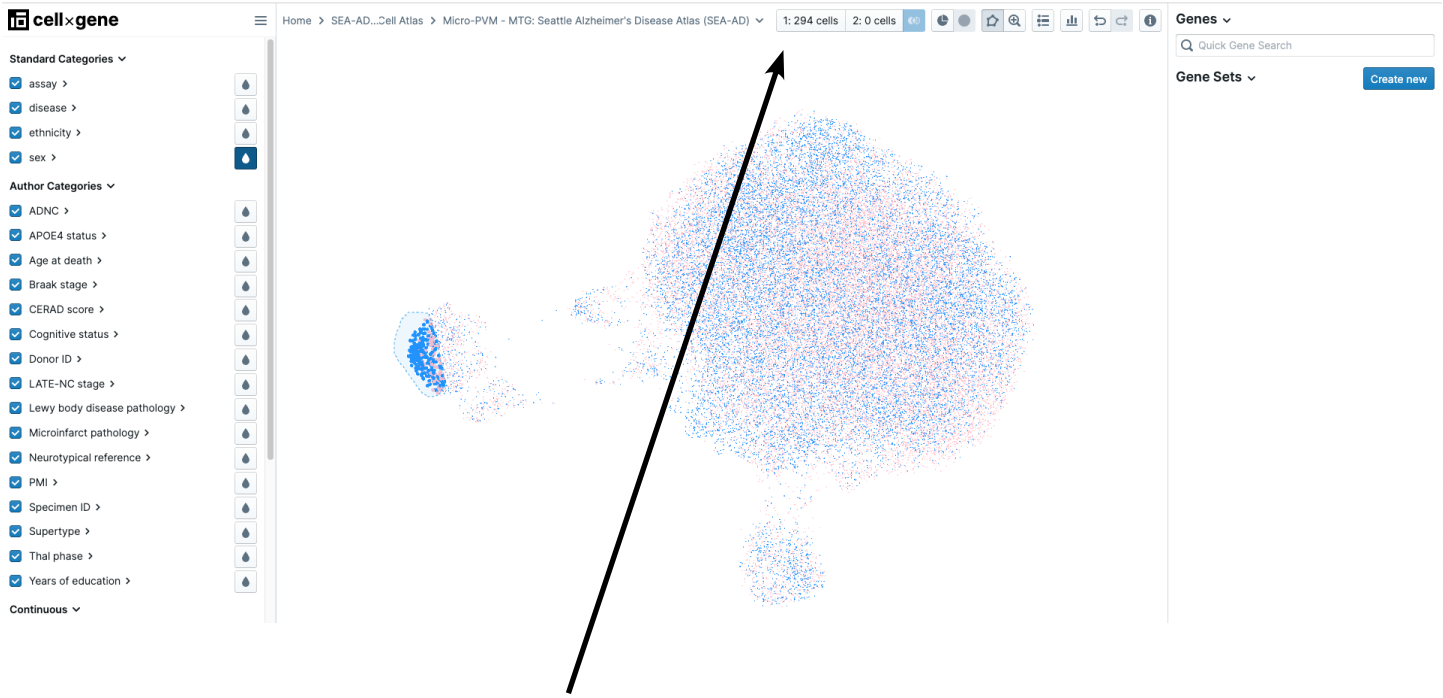
3. This page should open to display an all-black UMAP. We can filter this to explore specific donor criteria of our choosing.

4. Say we wanted to compare gene expression between the microglial cells in males and females. To filter by sex, select the paint drop icon next to "sex." Your screen should now look like this:



5. Notice at the top of the UMAP, there is a box labeled 1: 0 cells and 2: 0 cells. This is the highlight feature that enables you to highlight specific PARTS of the UMAP and compare gene expression between those clusters of cells.

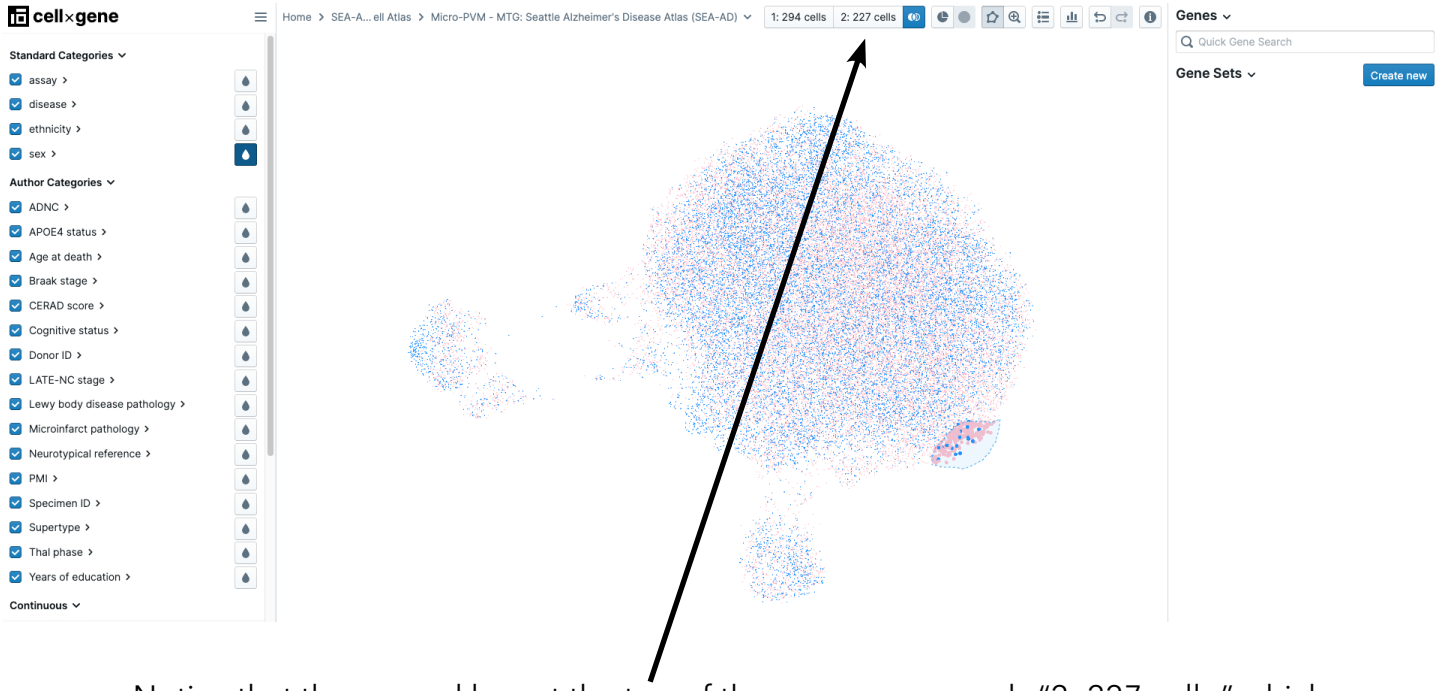
6. To start, hold down your cursor and drag over the cluster of blue dots on the lefthand side of the UMAP. After you have your cells selected, click on the box labeled "1:0 cells." This box should change and tell you how many cells you have highlighted. Your screen should look something like this:



Notice that the group 1 box now reads 1: 294 cells. This means that in the screenshot above, the user selected a total of 294 cells.

7. Now we need to select a second group of cells to compare with this first one! Click anywhere on the screen to deselect group 1. Next, repeat the same process and highlight the cluster of pink dots on the righthand side of the UMAP. Don't forget to click on the 2:0 cell box after you have highlighted your cells to save your selection.

Your screen should now look something like this:

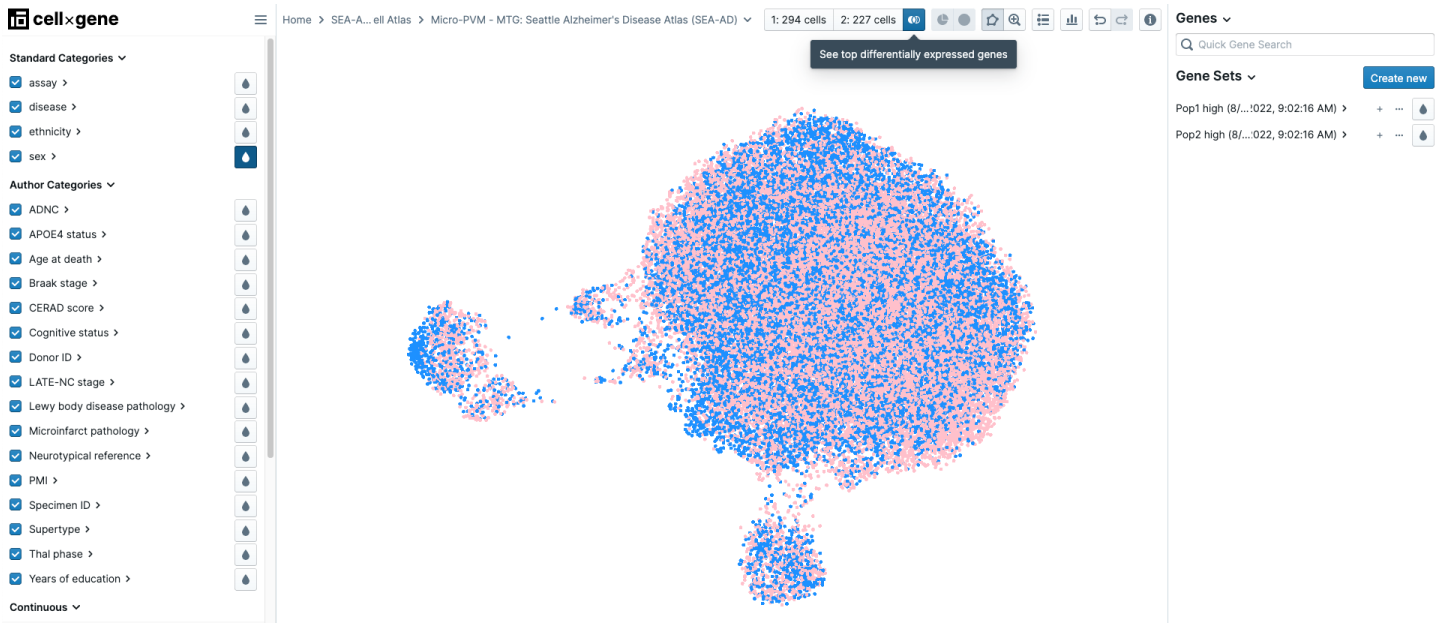


Notice that the second box at the top of the screen now reads "2: 227 cells," which tells us that this second selection contains 227 cells.

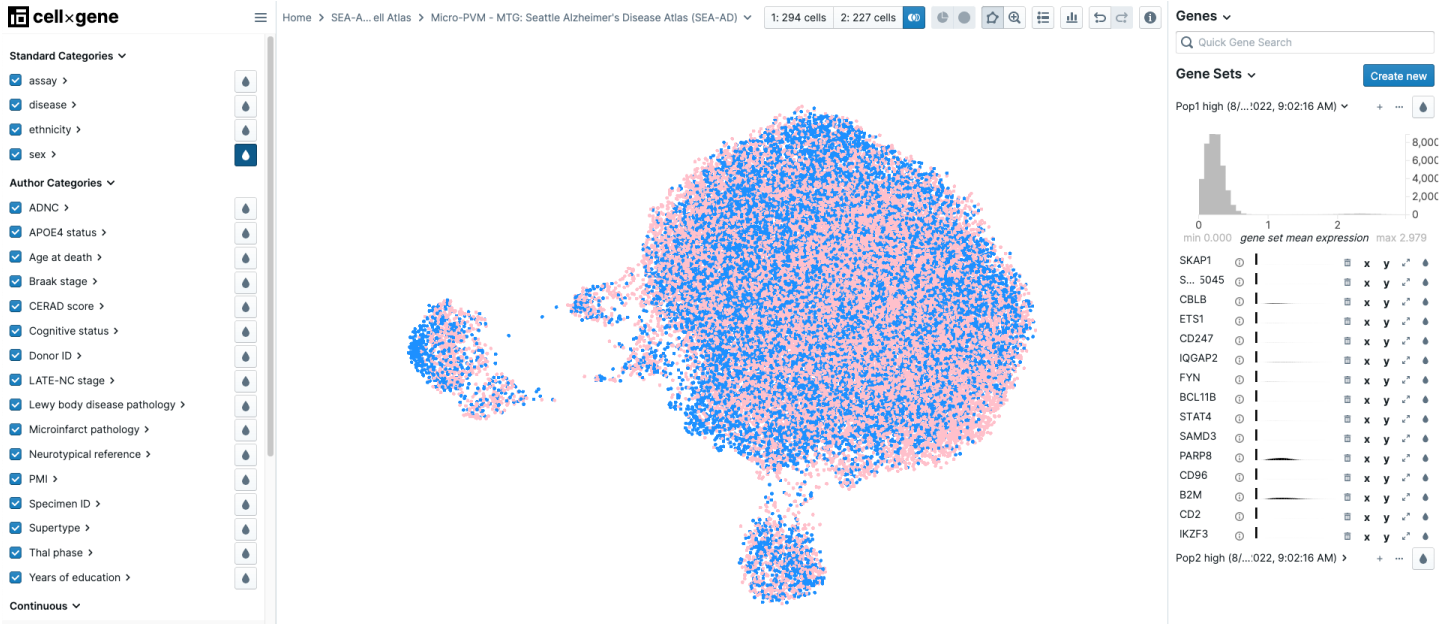
Note: It is likely that the numbers in your highlighted area may be slightly different than the ones given in this example depending on how many cells you highlighted and where on the UMAP you highlighted.

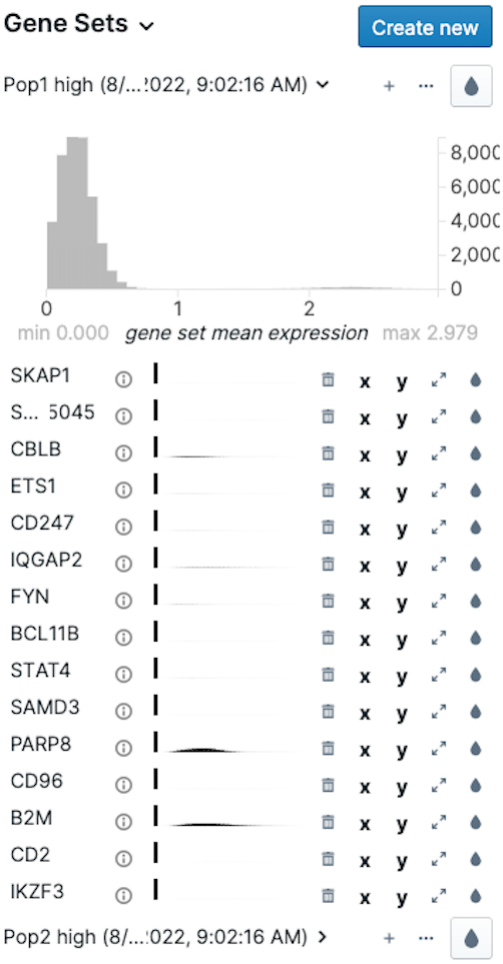
8. Since we have now highlighted two groups of cells, we can use cellxgene to find the top genes that are differentially expressed between these two groups of cells.

To do this, click on the blue icon next to the two boxes displaying the number of cells you highlighted in groups 1 and 2:



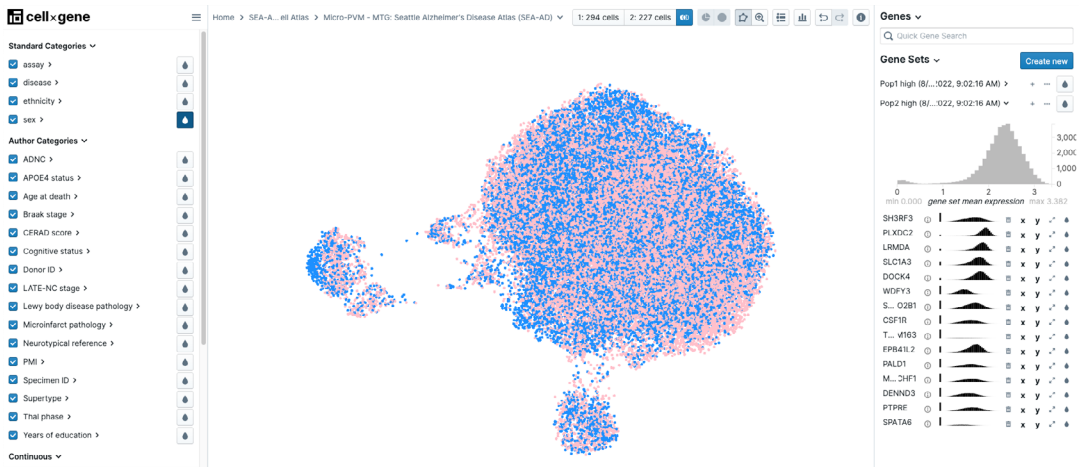
9. After clicking on this icon, cellxgene will begin drawing data on the top genes expressed in each group of cells. When the data loads, click on the "Pop1 high" bar under "Gene Sets." After clicking on "Pop1high," your screen should look something like this:





The "Gene Sets" column tells you which genes have the highest mean expression in population 1 **relative** to population 2. In the example column shown below, CBLB, PARP8, and B2M are the three genes that the first group of cells express the most.

10. Cellxgene allows us to compare the genes that are most highly expressed in population 1 relative to population 2, and vice versa. To look at the gene expression of group 2, click on the arrow next to "Pop2 high." Your screen should look something like this:



Knowledge Check

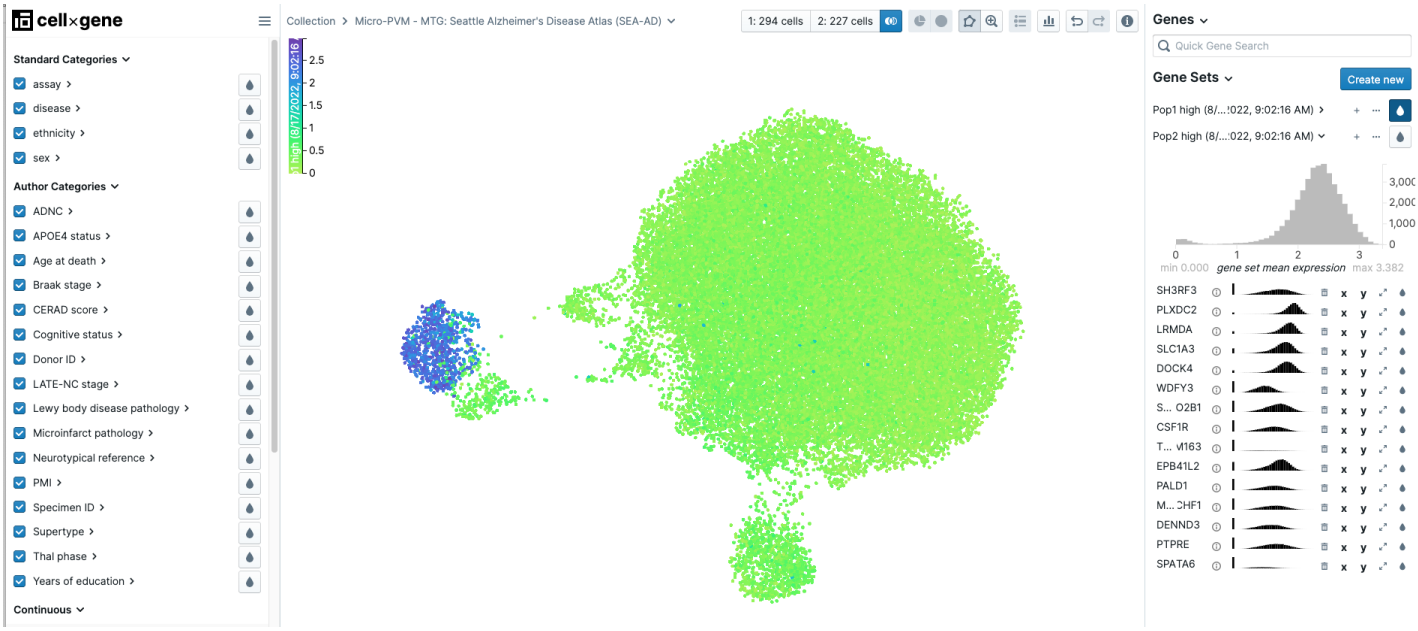
- **Look at the screenshots on the previous pages and think back to what you learned about how UMAPs are constructed. Based on the areas of the UMAP that were highlighted for population 1 and population 2, would you expect these two populations of cells to have similar gene expression? Why or why not?**

- **Looking at the screenshots on the previous pages, which genes appear to be expressed in the highest quantities in population 1 compared to population 2? How can you tell?**

- **Which seven genes appear to be expressed in the highest quantities by the group 2 of cells relative to group 1?**

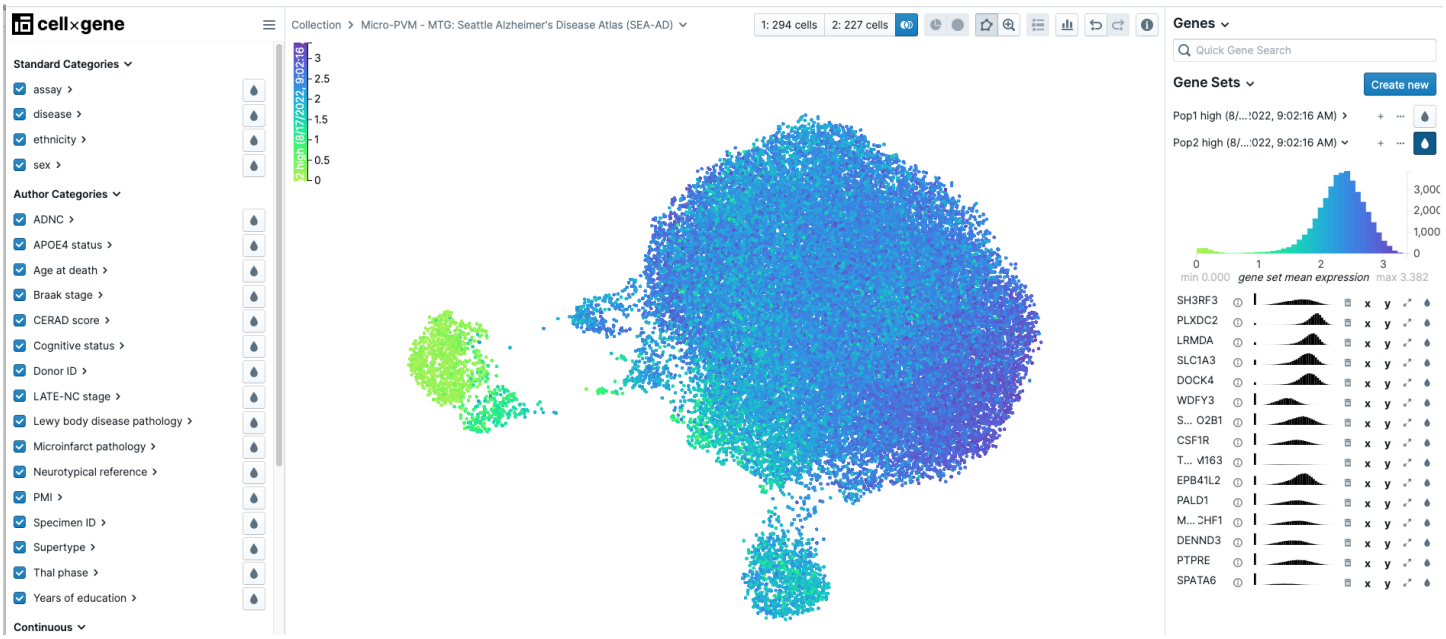
11. In addition to telling which genes are expressed at the highest levels in group 1 relative to group 2, we can also filter the UMAP to see which genes group 1 expressed in the highest quantities relative to group 2 by clicking on the rain drop icon next to “Pop1.”

To do this, click on the paint drop icon next to Pop1. Your screen should look like this:



12. You can also filter the UMAP by the genes that group 2 expressed in the highest quantities relative to group 1 by clicking on the rain drop icon next to "Pop2."

After clicking on the rain drop icon, your screen should look like this:



Now that you have an understanding of how to use the cellxgene interface to analyze UMAPs and identify genes of interest between specific cluster of nuclei/cells, you will now have the opportunity to design your own experiment using the cellxgene tool!

Your turn:

To help you design your own experiment using the cellxgene interface, we have created the following table for you to fill out. This table requires you to consider several questions concerning which cell type you will explore, which clusters of cells/nuclei you will compare, and what genes you will explore further.

To begin filling out the table, start by going to the list of datasets available through the Seattle Alzheimer's Disease Brain Cell Atlas: https://cellxgene.cziscience.com/collections/1ca90a2d-2943-483d-b678-b809bf464c30

Remember that each of the datasets included for this list is for a specific cell type. For example, the first file titled "**L2/3 IT – MTG: Seattle Alzheimer's Disease Atlas (SEA-AD)**" represents the data from the nuclei isolated from brain tissue from the L2/3 layer of the middle temporal gyrus (MTG) within the human brain.

Step 1: Pick which cell type you will explore

Example: "Oligo – MTG: Seattle Alzheimer's Disease Atlas"

Cell type database you will use in cellxgene: _____

Step 2: Pick which donor characteristic you will filter the UMAP by

Example: dementia status, sex, age at death, etc.

Donor characteristic you will filter by:_____

Step 3: Perform your own analysis

Follow the procedure provided earlier in the tutorial. Highlight 2 clusters of cells that you want to compare gene expression for and fill out the table below with the information provided for each of your clusters by cellxgene.

	How many cells did you select for this population?	What 3 genes did this population of cells express at the highest levels relative to the other population you selected?
Population 1 of cells		
Population 2 of cells		

Step 4: Explore functions of genes that are differnetially expressed

Now that you have identified which genes are differentially expressed between these two clusters in your UMAP, you will now have the chance to explore what these genes' known functions are. We can explore each gene's function by going to the NIH gene database!

Go to the NIH gene database: https://www.ncbi.nlm.nih.gov/gene

Using the NIH gene database, research the three genes that cellxgene identified your group 1 cells expressing in the highest quantity. In order to find a gene's function, all you do is copy over the name of the gene into the NIH database search bar and click search! Make sure you are looking at the gene expression in humans, which are listed as "homo sapiens" in the database. Fill out the table below with the information you find:

Population 1	Official full name of gene	According to the NIH gene database, what are the known function(s) of this gene?
Gene 1		
Gene 2		
Gene 3		

Step 5: Repeat for population 2

Repeat this process for the three genes that cellxgene identified being expressed the most in your group 2 cells:

Population 2	Official full name of gene	According to the NIH gene database, what are the known function(s) of this gene?
Gene 1		
Gene 2		
Gene 3		

Reflective Questions:

1. What demographic characteristic did you choose to explore and why?

2. You used the NIH gene database to explore the known function(s) of specific genes. Were the functions of the genes expressed at the highest level in your group 1 relative to group 2 similar to the function of the genes expressed at the highest level in your group 2 relative to group 1?

3. What questions about AD pathology do you believe scientists need to answer? How could these data help them?

Conclusion:

Throughout this lesson, you had the opportunity to learn about the value of transcriptomic data and how it can be used to study disease pathology. In addition to understanding how transcriptomic data is obtained, this lesson challenged you to critically analyze graphical representations of this data using the CZ cellxgene interface.

For those interested in learning more about the Allen Institute and its research into AD pathology, check out https://portal.brain-map.org/explore/seattle-alzheimers-disease

While this lesson focused on transcriptomic data analysis, additional lessons available at https://alleninstitute.org/about/education-outreach/teaching-materials/ cover other aspects of disease pathology. These other lessons include bioethical discussions of brain donation, activities that explore the importance of basic scientific research, and explorations of the biological and social hallmarks of AD.

If you develop your own lesson plan using Allen Institute resources, we invite you to share your experience with us at communications@alleninstitute.org. Teachers are also encouraged to publish original lessons using our open data, tools, and other resources, and to share those lessons with us.