Exploring Gene Expression with Bioinformatics Using GLDS-288

Curricular Unit
Teacher Materials

OVERVIEW

This series of lessons is designed to introduce students to mechanisms and importance of gene expression in biological systems. Students investigate physical changes observed in mice during spaceflight through transcriptome analysis. Emphasis is made on the difference between genomic changes (mutations) and gene expression on protein synthesis and overall function. After analyzing data using the GeneLab Galaxy platform, students create a hypothesis as to which pathways are being impacted by differential gene expression and will understand how transcription factors regulate transcription.

The module was designed with AP and IB biology in mind, but the lessons would be appropriate for an advanced honors biology course with some small modifications. Students should have learned the process of protein synthesis before starting this lesson as it assumes students are familiar with transcription and translation and the relationship between DNA and protein.

CONTENT OBJECTIVES

- Students will be able to explain how differential gene expression changes the phenotype of an organism.
- Students will be able to analyze gene expression data to create a hypothesis that explains the altered phenotype seen during spaceflight.
- Students will be able to describe how transcription factors regulate gene expression.

PACING AND SCHEDULING

This lesson was designed for four 90 min blocks (1 day Engage, 1 day Explore, 1 day Explore & Explain and Extend), equivalent to five 50 min periods (1 day Engage, 2.5 days Explore, 1 day Explain, 1 day Extend). **Note:** This activity is flexible and there are many sections noted in the Teacher's Notes that could be shortened or removed depending on your time constraints and the scope of your course.

TEACHING METHODS

- Teachers and students will need to access the Galaxy platform using an internet connection with a web browser such as Chrome
- A set of slides accompanies this curriculum unit
- Lessons as described are designed with students working in collaborative groups

CURRICULAR CONTENT

Unit Overview:

The unit will analyzing data from GLDS-288 in the GeneLab Data Repository. This experiment was done on mice and the spleen tissues were analyzed through RNA sequencing. The full dataset for this experiment is available, but in this unit, students will be using only the sequencing data from the ground control samples and the microgravity samples from the ISS.

ENGAGE-

Students are introduced to spaceflight factors and engage in discussions about why these factors make spaceflight difficult and the effects on biological processes. The GLDS-288 experiment is introduced, and students do research into the metadata available in the GeneLab Data Repository. They will identify independent and dependent variables and controls in the experiment to understand where the data they are analyzing originated from. Lastly, students consider why model organisms like mice are used in experiments and get a short introduction to the function of the spleen.

EXPLORE-

After getting familiar with the experimental design, students will get an overview of RNA sequencing and how RNA splicing creates problems for data analysis. Students will be working in the Galaxy platform to create and analyze volcano plots and over-expressed GO term plots. Their data analysis will lead to several pathways that are differentially expressed. Students conduct research to create a final hypothesis that explains how the differential expression of these pathways lead to the phenotypic changes observed during spaceflight.

EXPLAIN-

An excerpt from the principal investigator's (Horie et al, 2019) publication is included for students to compare their hypothesis and findings to the researcher's findings. This excerpt specifically identifies changes in transcription factors and causing pathways involved red blood cell function. This is used as an opportunity for instruction on the function of transcription factors and how they interact with promoter and operator sequences as a mechanism for gene regulation. Students revisit their hypothesis from earlier to add in, or clarify how transcription factors may be altering the pathways they identified.

EXTEND-

Finally, students will apply their understanding of gene expression and the role of transcription factors to the *lac* operon. This operon is commonly included in AP biology curriculum and this module provides a unique introduction to this topic.

Background Information:

Modified from GL4HS Manual (Blaber, 2021)

Genome (Genomics) – the total DNA of a cell or organism. This includes DNA that encodes for genes, non-coding DNA and genetic material contained outside of the nucleus. Genes can be studied individually through techniques such as PCR and quantitative PCR, or they can be studied genome wide e.g. microarrays and sequencing techniques.

Transcriptome (transcriptomics) – the study of the RNA molecules within an organism, including messenger RNA (mRNA), ribosomal RNA (rRNA), transfer RNA (tRNA) and all other non-coding RNAs. Generally, transcriptomics focuses on studying the mRNA within a cell which is the template for protein synthesis. The transcriptome is studied through expression-based techniques, such as, microarrays or sequencing techniques, e.g. RNA-Seq.

RNAseq-- RNA-seq is a powerful tool that enable us to analyze the expression profile of an entire organism's genome. RNA-seq is more accurate than the previous microarray technology which was based on the principle that two DNA strands with complimentary sequence pair with each other by forming hydrogen bonds and complementary nucleotide base pairs (hybridization). RNA-seq, however, fragments RNA into small pieces, which are then converted to complimentary DNA (cDNA) and introduced to a flow cell whereby the actual sequence is determined using the addition of fluorescent tags. This results in millions of reads that must then be aligned and assembled to determine the sequence of the transcriptome and determine expression levels for each gene.

Galaxy Platform-- Galaxy is a powerful open source computational analysis program that was developed by the Nekrutenko lab in the Center for Comparative Genomics and Bioinformatics at Penn State, the Taylor lab at Johns Hopkins University, and the Goecks Lab at Oregon Health & Science University, along with contributions from the community. The program provides access to hundreds of genomic analysis tools that can be used to analyze microarrays, next gen sequencing, proteomics, metabolomics, imaging, and many more assays.

RNA STAR-- **STAR** (Spliced Transcript Alignment to a Reference) is a super-fast spliced alignment program. It also has several other advantages, including being able to perform an unbiased search for splice junctions without needing any prior information on location, sequence signals, or intron length. It is also capable of aligning a read with multiple splice junctions, indels and mismatches and those with low-quality ends. **STAR** can find the spliced junctions *de novo*, but they can also be supplied to the program when building the reference index. **STAR** finds splice junctions by using an approach based on 'maximum mappable length'. **STAR** splits a read into pieces (50 bp by default) and finds the best portion that can be mapped per piece. The longest matching sequences are called the maximum mappable prefixes (MMPs). The different parts of the read that are mapped separately are called seeds. The first part of the read that was mapped is called seed 1. It then searches again to find the next longest sequence to that exactly matches the reference genome (MMP) and this will be seed 2. By sequentially searching for the unmapped portions of reads, this makes the algorithm very efficient. It does this by using an uncompressed suffix array (SA) to search for the MMPs which allows for quick searching against large genomes. If there is not an exact matching sequence for each part due to mismatches or indels, then **STAR** will extend the previous MMP. If this still does not give good alignment, then the poor-quality end sequence or adapter sequence will be soft clipped

goseq -- Gene Ontology analysis highlights over- or under-represented biological processes based on your gene list. Gene Ontology (http://geneontology.org/) maintains a controlled hierarchical vocabulary of terms and definitions to describe molecular functions, biological processes, and cellular components. To perform this analysis, we will use **goseq** to test for over-represented gene categories.

In additional to GO analysis, this tool also corrects potential length bias in differentially expressed genes.

Handouts with Teacher's Notes:

EXPLORE

Space is an extremely harsh environment. NASA has been studying the effects of spaceflight on humans since 1961 when the human space program was started. After all this research, there is a lot that we still don't know about how space impacts biological processes or how these impacts are happening. As space biologists, we work on determining the mechanisms behind **phenotypic changes** we see during spaceflight. This is often due to changes in **gene expression**.

Teacher's Note- Use Slide 2 to introduce spaceflight factors and as you talk through these, have students begin brainstorming the impacts. Slide 3 has more specific ways that the factors impact humans. Students can add this to their tables after their own ideas. A discussion about what students know/think happens to humans in space is a great way to get them thinking about spaceflight factors.

Spaceflight Factors: Slides 2 & 3

There are five main spaceflight factors that are the stressors that cause the biological changes we see during spaceflight and are the biggest hurdle for us to overcome before considering long term space missions.

	Radiation	Isolation	Distance from Earth	Altered Gravity Fields	Hostile Environment
	Impacts on cell cycle- cancer DNA damage	Behavior Sleep impacts (no 12hr cycle on ISS)	Implications for medical care Communication	Muscle/bone loss Fluid shifts	Gas exchange Environmental conditions

Teacher's Note- *This section is NOT included in the Student Guide, but if time allows, having a discussion about model organisms may be useful for your students.*

We use model organisms, like mice, often during spaceflight experiments. Think about the positives and negatives of using mice as model organisms and what other examples of model organisms could be.

Pros for Using Mice	Cons for Using Mice	Other Examples of Model Organisms
Small- space is expensive on ISS Short life span Mammals Lots of similarities to humans	They aren't humans! Size creates differences in dosing and metabolism Hard to compare aging in mice to humans	C. elegans (nematodes) Drosophila Zebrafish E. coli

Teacher's Notes: Slide 5 shows students how to navigate to the tab in GeneLab, Use Slide 6 to recap when reviewing design Slide 5 & 6

The Goal: Use the GeneLab platform to analyze data to determine possible biological pathways contributing to **differential gene expression** in space in the spleens of mice.

Background:

The data we are going to analyze in GeneLab originated from a spaceflight experiment that flew on the Mouse Habitat 1 mission which launched with the JAXA to the ISS. To better understand the transcriptomic data, we need to get more information on how the experiment was setup. Visit this link for information on this experiment (GLDS-288).

1. Locate the Treatment Protocol information. Using this information, identify what independent variable(s) were investigated.

Spaceflight and artificial gravity

2. Given the independent variable(s), identify the treatment groups that were used.

Earth control ISS microgravity

ISS Artificial gravity (centrifuge) **There are pictures of the mouse centrifuges used in the JAXA mission page in extra resources**

3. Browse through the remainder of the Protocol section. What is the dependent variable in this experiment?

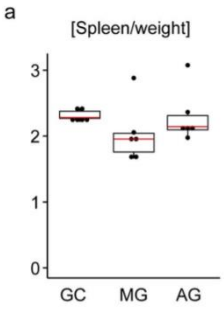
Gene expression (RNAseq data)

4. Identify at least 3 controls that were used in the experiment and explain the importance of each. Males, age, diet etc

Now that we understand more about the experimental design, we need to familiarize ourselves with the organ of interest: the spleen. After watching this video, answer the following question and do a little more research on structure of this organ. Video on Slide 6

1. What are the two main functions of the spleen? Filtering blood for old red blood cells and producing/housing immune cells

The figure to the right shows a graph of the weight of the mice from each of the experimental groups in our study. GC stands for Ground Control, AG is Artificial Gravity and MG is Microgravity. **Identify trends in the data and create a working hypothesis about why you think this might be happening.**



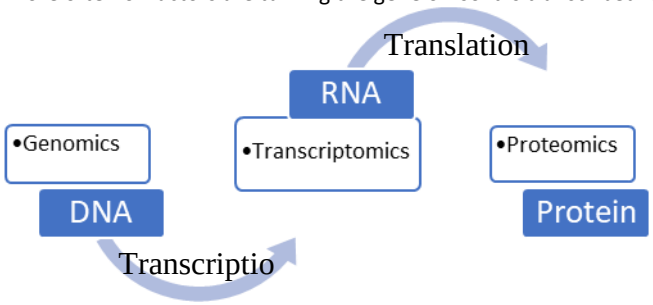
Guide students to think about why the spleen loses weight in space. Guide students to think about what is in the spleen- (blood) and how that could connect with weight loss. It is generally thought that the spleen loses weight due to lower red blood cell production and lower functioning.

EXPLORE

Teacher's Notes- Review background information on RNAseq. An in-depth explanation is not necessary for complete understanding of the data. Emphasis should be on the fact that these changes are NOT due to mutations—they are gene expression. **An optional extension activity here could be to have them find an academic journal article that uses either genomics or transcriptomics and share the goals of the research to the class to compare and contrast when each -omic is most appropriate.**

RNA Sequencing Slide 8

Often, when we refer to sequencing, we are talking about genetic sequencing—or determining the order of nucleotides in DNA. This is a helpful technique when you are interested in mutations or differences between individuals or determining the alleles present in certain conditions. However, in different environmental conditions, like spaceflight, we see changes in the phenotypes that are often not a result of mutations in DNA, but instead the result of genes being turned on or off due to the difference in the environment. This is **differential gene expression**; the DNA itself is not changed but other factors are either turning the gene on so it is transcribed more often or factors are turning the gene off so it is transcribed less often.



In order to study gene expression, information from mRNA is going to be most helpful. In order to collect this information, we use RNA sequencing. Using RNA sequencing to survey genes that are actively transcribed is called **transcriptomics**.

RNA sequencing is a complex process, but in effect, the mRNA transcripts from a cell, tissue or organism is isolated from the cells, transformed into complementary DNA (cDNA) which can then be sequenced, then each piece of cDNA is compared to a known genome to determine which genes the original mRNA transcript was from in a process called **alignment**. In this way, we can use DNA sequencing methods to determine which genes were turned off or on which then gives us important information about the ways biological processes are changes in response to the environmental factors.

1. Explain the differences between genomics and transcriptomics. *Hint: think of the types of data that is collected in each.*
2. Think of two situations where genomics would be the most advantageous.
3. Think of two situations where transcriptomics would be the most advantageous.

Teacher's Notes- Depending on your classroom setting, RNA STAR could be removed from the activity. It is a great opportunity to talk about RNA processing, but if that is not part of your curriculum or you covered it elsewhere, the lesson can be completed without it.

RNA Processing and RNAStar Slide 9 & 10

Much of the process of RNA sequencing happens after the sequencing has been completed. Much of this is done using bioinformatic algorithms that help show trends in the data, run alignments, and create visualizations that scientists need to make sense of all this information. RNAStar is one of the alignment algorithms that GeneLab uses on the Galaxy platform. What

is particularly helpful about RNAStar is that it is **splice-aware**. To understand what this means and why it is so helpful, let's think about mRNA processing.

In eukaryotic genomes, one gene does not equal one protein. In other words, we can arrange the pieces of a gene in different ways through **RNA splicing** to create different transcripts which will then make different proteins.
Pre-mRNA is what the gene directly reads. It contains **exons** and **introns**. Exons are the portions that are expressed in the final protein. Introns are interrupting pieces that are removed. To create different protein products, different introns are removed or left in to create new transcripts.

1. How does RNA splicing complicate the process of alignment?

It cannot look for the complete transcript- instead it has to find pieces spaced out by introns.

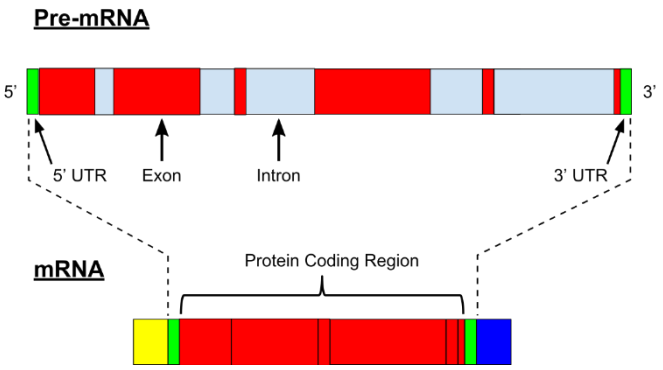
2. RNAStar is a splice-aware algorithm. What do you think this means?

It "knows" that transcripts will not be found as entire segments in the genomes.

3. If we searched a genome for an entire mRNA transcript would we find a match?

No

4. How does RNA STAR accommodate splicing when it runs alignments?



USED WITH CREATIVE COMMONS LICENSE FROM HTTPS://COMMONS.WIKIMEDIA.ORG/WIKI/FILE:PRE-MRNA.SVG

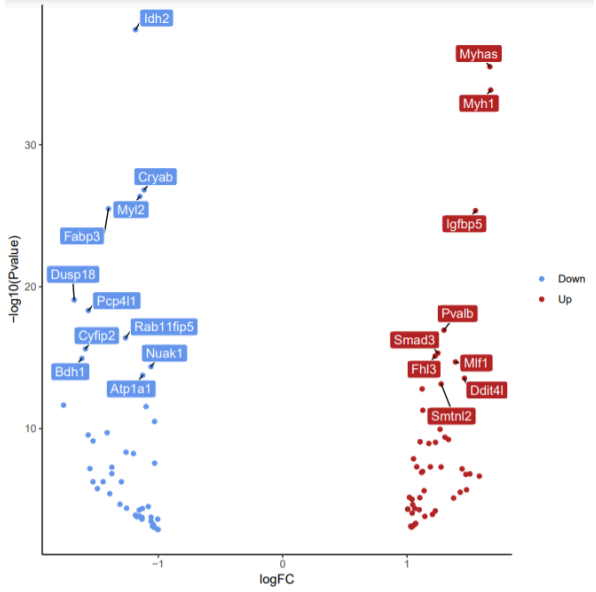
It breaks the transcripts into pieces and then finds matches for individual chunks. It then ensures that those pieces are in linear stretches (i.e. not on different chromosomes) to determine that they do come from the same gene.

Bioinformatics Analysis:

We will use a dataset from the GLDS-288 experiment that has already been run through alignment and the genes present in the flight group have been compared to the ground control group. We will not consider the artificial gravity group for our analysis. This data has been **normalized** to help us ignore genes that were not differentially expressed between the two groups as well as genes that showed expression changes, but were not significant enough for us to consider. The first visualization we are going to create is a volcano plot.

Teacher's Notes- if you leave out RNA STAR, quickly explain that alignment is how we determine which genes the RNA transcripts we collected are from.

An example of what we might see is on the next page.



Slide 12

- Each point represents one gene.
- The x-axis is showing the **fold change** which is how many times more or less the gene was expressed during spaceflight.
- The y-axis is showing statistical significance, or how confident we are that these genes changed because of the environment (space).
- Red points are genes that are **upregulated**.
- Blue points are genes that are **downregulated.**
- The labels are the 20 most significantly changed gene IDs.

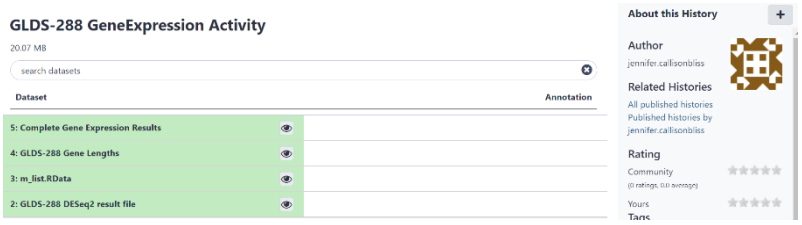
Teacher's Guide to Galaxy- Opening the Shared Data

1. Click on the link here to open the shared history with all necessary files for this activity.
2. The link should take you to Galaxy where you will need to sign in using Google. It is imperative you log in with a handy Gmail account so that you can save your progress.
3. Once you log in, you want to navigate to the **GLDS-288 GeneExpressionActivity** history (you may need to click on the link above again after you log in)

4. You should see the page below once you are logged in and in the history.

5. Click on the Plus Sign on the top right to add the files to your history.

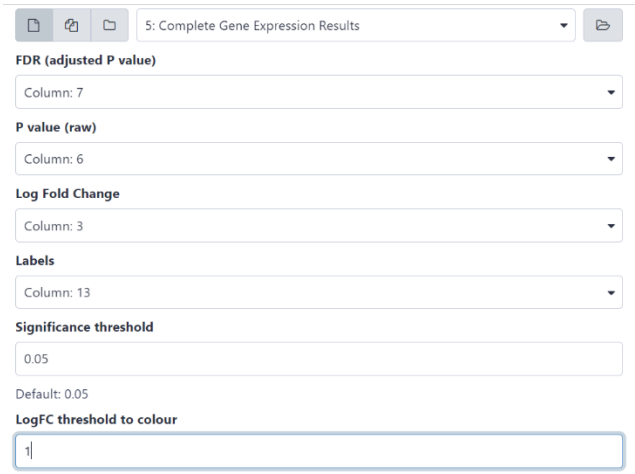
6. Rename the history to **GLDS-288 Analysis** and click Import



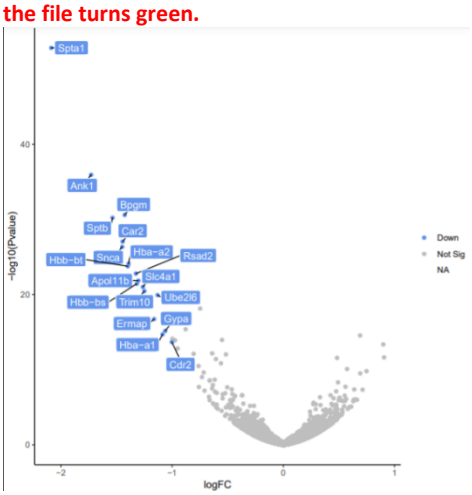
Teacher's Guide to Galaxy- Volcano Plots

1. Click on the Analyze Data tab to go to the main window. **Ensure that the history you just added shows up on the right.**
2. Scroll to RNAseq in the tools on the left > volcano plot
3. Click the drop down to choose **Complete Gene Expression Results**
4. *FDR (adjusted P value):* Column 7
5. *P value (raw):* Column 6
6. *Log Fold Change:* Column 3
7. *Labels:* Column 13
8. *Significance threshold:* Enter 0.05
9. *LogFC threshold to color:* Enter 1
10. *Points to label:* Significant
11. *Only label top-most significant:* Enter 20
12. All other parameters to be left on default settings.
13. Click Execute

Your parameter fields should look like the image to the right:



The volcano plot student should end up will look like the image below.
To see the plot, click the "eye" button on the file in your History when the file turns green.



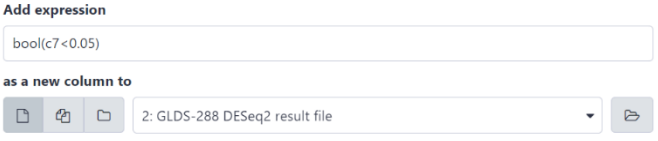
Have a discussion about the gray points on this plot— the example plot used a filtered dataset where genes with small changes and low significance were removed. This is a good chance to talk about what genes to select for research and why we are not worried about the genes in gray (they are not changed much between ground and spaceflight and they have high p-values).
Students will do research on 3 genes using GeneCards. Slide 13 has an easy guide to GeneCards for students to find information. It would likely be helpful to walk through the database on one gene together as a class. **Allow time for students to share their research with another group so they have information about more than 3 of these genes.** A class discussion of patterns or connected genes would like help push students in the right direction in the next steps. Genes involved in red blood cell synthesis or differentiation may be of particular interest, but allow students to find their own connections as well.

<u>Gene Ontology Analysis:</u>
Gene Ontology is a database of gene functions and when we run data through goseq it decreases the complexity of the data so we can analyze it easier. Instead of looking up every individual gene that is differentially expressed, goseq runs the genes through Gene Ontology which determines if several genes function in similar biological pathways to better show us trends in expression. When we run our analysis, we will get a plot showing several significant pathways, how much each pathways' expression was changed in flight versus ground control and how statistically significant the changes were.
We have to first alter our files to run them in goseq: Slides 14-16

Teacher's Guide to Galaxy- goseq

Adding a column based on statistical significance (P value)

1. Click on Text Manipulation > Compute

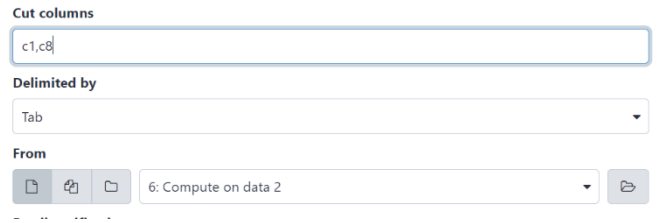


2. *Add expression:* bool(c7<0.05)
3. Use the dropdown menu to choose the **GLDS-288 DESeq2 Results File** (DO NOT USE THE FILTERED FILE WE USED FOR THE VOLCANO PLOT)
4. Leave all other settings and click Execute

This will create a file with True or False added as a final column to our DESEq2 results. True will indicate that the expression change was statistically significant.

Cut out all columns of the results file we edited EXCEPT for the gene identifier and True/False

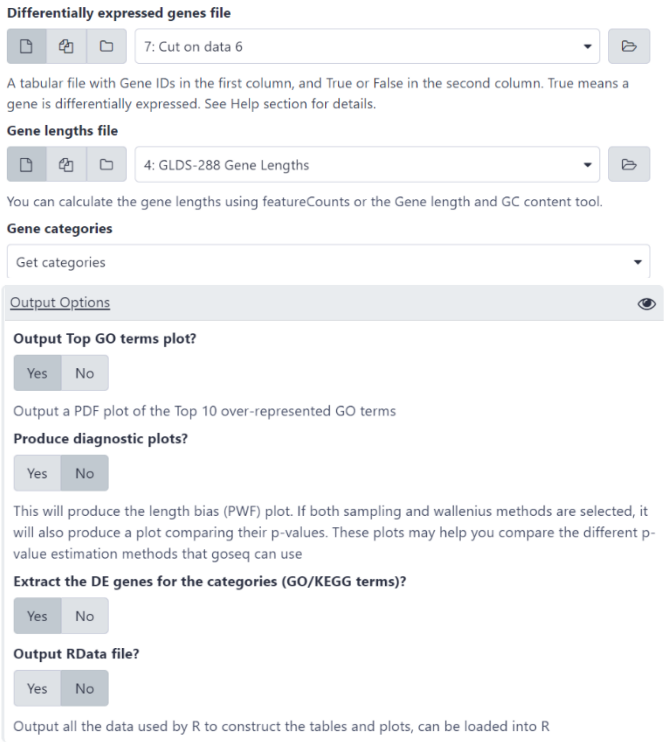
1. Find Cut under the same menu as Compute
2. ***Cut columns:*** c1,c8 from the **Compute file** we just generated
3. Execute
4. After the file is green in your history, Click the pencil icon to change the name of the file to **Genes and T/F**



Running goseq

This algorithm compares all the genes that were up/downregulated and runs them through the Gene Ontology (GO) database. This database puts genes into larger categories based on their functions and the pathways they are involved in.

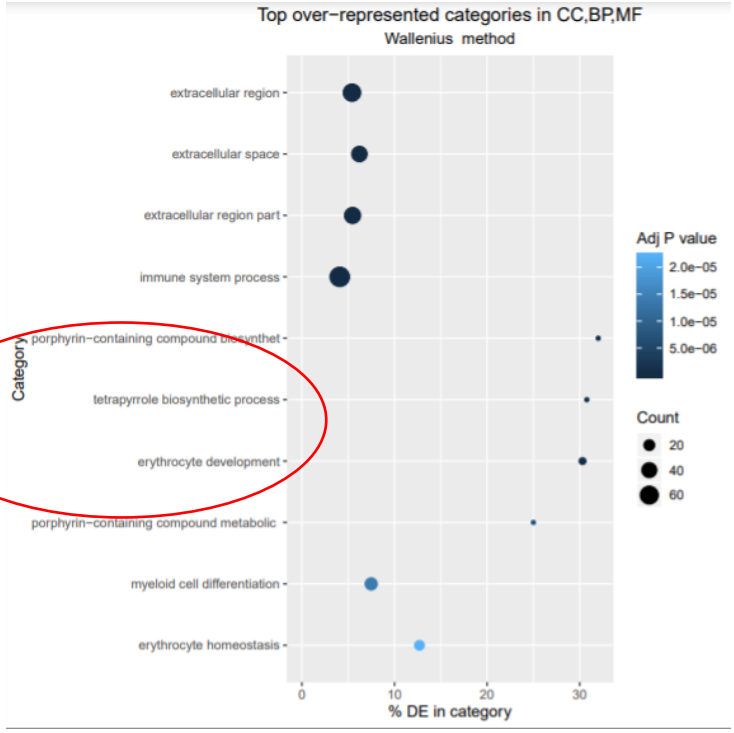
1. Search goseq in the toolbar menu
2. Choose the **Genes and T/F file** for the first file
3. ***Gene Lengths:*** Select the **GLDS-288 Gene lengths file**
4. ***Select a Genome:*** Mouse
5. Find ***Output Options***
   - o ***Output Top GO Terms Plot?****:* Yes
   - o ***Extract the DE genes?*** Yes
6. Execute



To the right is what parameter fields should look like (the Cut on data 6 is the **Genes and T/F file**).

Goseq will produce three files-

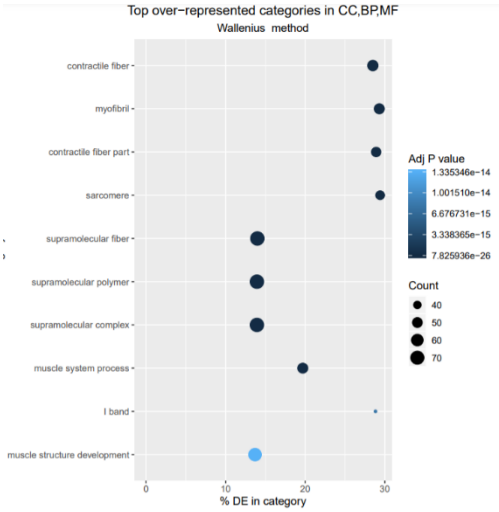
Students will only use the 2nd file- Overrepresented GO terms plot. The 1st file (Ranked category list) is nice because it allows you to see more pathways than the plot shows—scroll to column 6 to see GO terms.



To the left is the GO Terms plot the students should generate.

If students pick these pathways encourage them to research what the extracellular matrix is and how it relates to the immune system

These pathways are all related to red blood cells and heme synthesis. **Note:** not many genes are changed in these pathways, but they are extremely differentially expressed.



Slide 17- Example GO Plot to the right

- Each dot represents ALL of the genes in one GO pathway
- Each pathway is labeled on the left
- The color of the dot is the P value (statistical significance)
- The darker the color, the MORE significant the changes are
- The x-axis how much change in gene expression there is
- The size of the points represents how many genes were altered in that pathway

1. Are there pathways that seem related to each other?
2. Discuss some of the tradeoffs in choosing which pathways to research further—should you choose pathways where lots of genes are interacting, higher DE changes or choose based on the P values?

Slide 17, 18, 19- again, it would be really helpful to go through the database with the students for a pathway. This database is **NOT** as user-friendly as other websites so students may need additional guidance on making meaning from these pathways. Spend time on the information on Slide 18 with the Inferred Tree View—again, this can be difficult for students to understand. Encourage students to do a Google search for more clarification if they need it. **For struggling students, point them towards immune system development or erythrocyte development.**

Choose 2 pathways to research using geneontology.org and complete the table to determine whether these pathways give you clues as to how space is impacting the spleen.

GO Process	Description	Notes/Other Related Processes

Share your finding with another group to determine if there are overlaps in processes. Pick one GO term. **Why do you think this process is affected by spaceflight?**

Push students in these initial explanations to make connections between how the genes were changed (downregulation) and how that would impact the pathway it's involved in and **why** that may be happening in space.

EXPLAIN

Now that you have created hypothesis for how the spleen is being affected by spaceflight, read the excerpts of this paper to see how Horie et al (2019) analyzed this data and the conclusions they came to. Slide 20

Our data further suggest that spaceflight causes a reduction in the expression level of genes related to erythrocytes in the spleen. Spaceflight reportedly caused a reduction of the red cell mass in astronauts[39], which was proposed to be due to the suppression of erythropoiesis. In addition, a reduction in the number of erythroid cells in the spleen of rats after 22 day spaceflight was reported[23]. Notably, the results of colony formation assays suggest that erythropoiesis is reduced in the bone marrow of flight mice[40]. As extramedullary haematopoiesis occurs in the spleens of mice[41], the mechanisms controlling the extramedullary erythropoiesis may be impaired in mice experiencing spaceflight.

Overall, our data suggest that relatively long-term spaceflight down-regulates the expression of genes related to erythrocytes in the spleen. This down-regulation is likely due to the reduction of transcription factors GATA-1 and Tal1, which control the expression of these genes. Detailed investigation of the possible association between the down-regulation of these gene and the development of anaemia during space flight should be addressed in future studies.

1. What are erythrocytes and how are they related to the spleen's function?
   Red blood cells, spleen is designed to clean blood cells and has a role in blood cell formation
2. What is erythropoiesis?

Making red blood cells

3. Explain what is happening to the spleen in terms of **differential gene expression**. (What genes are being turned on/off and how is that impacting the spleen?) Answers should vary.

4. Based on this excerpt, what is a transcription factor? How do they contribute to gene expression? Answers should vary.

Slides 21 & 22-

Transcription factors is a vague term used to describe a lot of different proteins. Define the three main types of transcription factors below.

Activators- Bind to promoters and must be present in order for RNA polymerase to bind

Enhancers- Cause genes to be transcribed at a higher level than with an activator alone (not all genes have enhancers)

Repressors- Bind to DNA to stop transcription completely (often bind to operator sequences)

Return to the hypothesis you created in the Explore section. Modify your explanation to include how transcription factors may be changing your pathway.

EXTEND

Slide 23

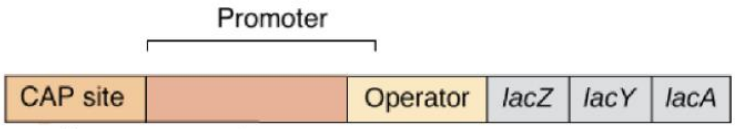
Optional: If you do not need to cover operons, this may be eliminated.

Let's apply what we learned about gene expression in the spleen to gene expression in bacteria by looking at the *lac* operon. This is a tightly controlled operon in bacteria because it determines what kind of nutrient they can break down—lactose or glucose. It is important to only express genes when the environment warrants it, otherwise the bacteria are wasting energy creating enzymes they don't need.

Based on the description of the parts of the *lac* operon and our understanding of gene expression, determine how the transcription factors interact for the scenarios below.

- The operon has 3 genes, all are necessary to metabolize lactose
- CAP is an **activator**
- *lac* **repressor** binds to the operator
- *lac* repressor is a lactose sensor
- CAP senses low glucose concentration

The *lac* operon:



USED WITH CREATIVE COMMONS LICENSE FROM https://www.khanacademy.org/science/ap-biology/gene-expression-and-regulation/regulation-of-gene-expression-and-cell-specialization/a/the-lac-operon

Scenario 1:

Bacteria have access to a high glucose/no lactose media.

Transcription Factors Bound	RNA Polymerase Active?	*lac* operon expressed?
lac bound b/c no lactose CAP unbound because lots of glucose	No	*No*

Scenario 2:

Bacteria have no glucose available and high lactose.

Transcription Factors Bound	RNA Polymerase Active?	*lac* operon expressed?
lac unbound CAP bound	Yes	*Yes*

Scenario 3:

Bacteria are in a media of equal parts glucose and lactose. They preferentially break down glucose and are now running low.

Transcription Factors Bound	RNA Polymerase Active?	*lac* operon expressed?
lac unbound b/c lactose is present CAP bound because low glucose	Yes	***Low expression***

Scenario 4:

Bacteria have run out of both glucose and lactose in the media.

Transcription Factors Bound	RNA Polymerase Active?	*lac* operon expressed?
Lac bound b/c no lactose CAP bound b/c no lactose	No	*No*

STANDARDS ALIGNMENT

NGSS AP Biology Science Practices

- Visual Representations 2.C- Explain how biological concepts or processes represented visually relate to larger biological principles, concepts, processes, or theories.
- Questions and Methods 3.C- Identify experimental procedures that are aligned to the question, including a. Identifying dependent and independent variables. b. Identifying appropriate controls. c. Justifying appropriate controls.
- Representing and Describing Data 4.B- Describe data from a table or graph, including a. Identifying specific data points. b. Describing trends and/or patterns in the data. c. Describing relationships between variables.
- Argumentation 6.A- Make a scientific claim
- Argumentation 6.B- Support a claim with evidence from biological principles, concepts, processes, and/or data.
- Argumentation 6.C- Provide reasoning to justify a claim by connecting evidence to biological theories.

AP Biology Standards

- IST-1.N.6 c and d- In eukaryotic cells the mRNA transcript undergoes a series of enzyme-regulated modifications— c. Excision of introns and splicing and retention of exons. d. Excision of introns and splicing and retention of exons can generate different versions of the resulting mRNA molecule; this is known as alternative splicing.
- IST-2.A.1- Regulatory sequences are stretches of DNA that interact with regulatory proteins to control transcription.
- IST-2.A.3 b- The phenotype of a cell or organism is determined by the combination of genes that are expressed and the levels at which they are expressed— Induction of transcription factors during development results in sequential gene expression
- IST-2.B.1- Both prokaryotes and eukaryotes have groups of genes that are coordinately regulated— a. In prokaryotes, groups of genes called operons are transcribed in a single mRNA molecule. The lac operon is an example of an inducible system. b. In eukaryotes, groups of genes may be influenced by the same transcription factors to coordinately regulate expression.
- IST-2.C.1- Promoters are DNA sequences upstream of the transcription start site where RNA polymerase and transcription factors bind to initiate transcription.
- IST-2.C.2- Negative regulatory molecules inhibit gene expression by binding to DNA and blocking transcription.
- IST-2.D.1- Gene regulation results in differential gene expression and influences cell products and function.

REFERENCES

GLDS-288 Dataset: Kenta H. "Transcriptome analysis of murine spleen in space", GeneLab, Version 5, http://doi.org/10.26030/9ntw-qy67

GLDS-288 Publication: Horie K, Sasanuma H, Kudo T, Fujita SI, Miyauchi M, Miyao T, Seki T, Akiyama N, Takakura Y, Shimbo M, Jeon H, Shirakawa M, Shiba D, Yoshida N, Muratani M, Takahashi S, Akiyama T. Down-regulation of GATA1-dependent erythrocyte-related genes in the spleens of mice exposed to a space travel. Sci Rep. 2019 May 21;9(1):7654. doi: 10.1038/s41598-019-44067-9. PMID: 31114014; PMCID: PMC6529412.

GL4HS Manual: GeneLab for High School Bioinformatics Manual. Blaber, Elizabeth. 2021.

AUTHOR
Jennifer Callison-Bliss, Wheeler High School (Marietta, Georgia)
Edited by GL4HS Staff