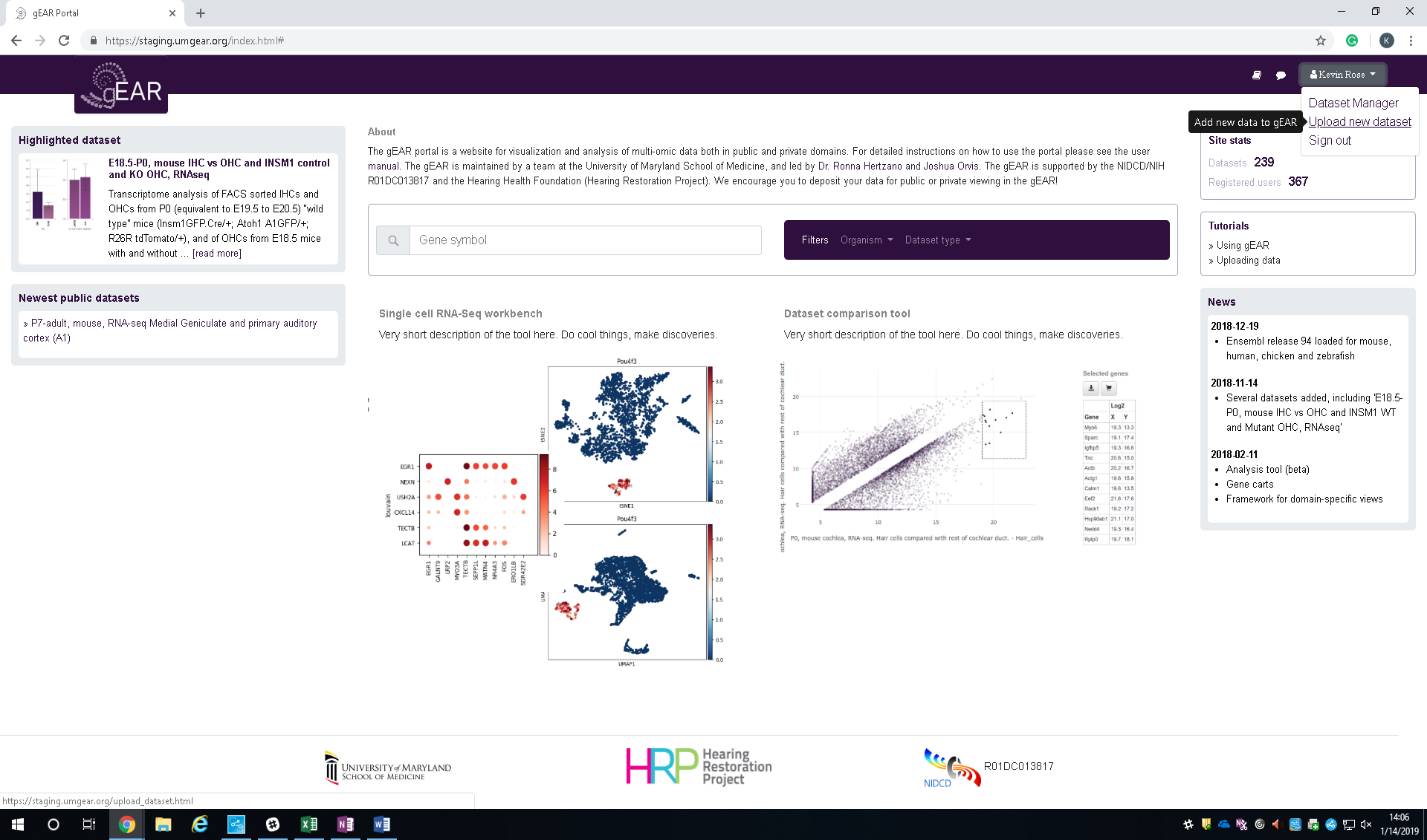
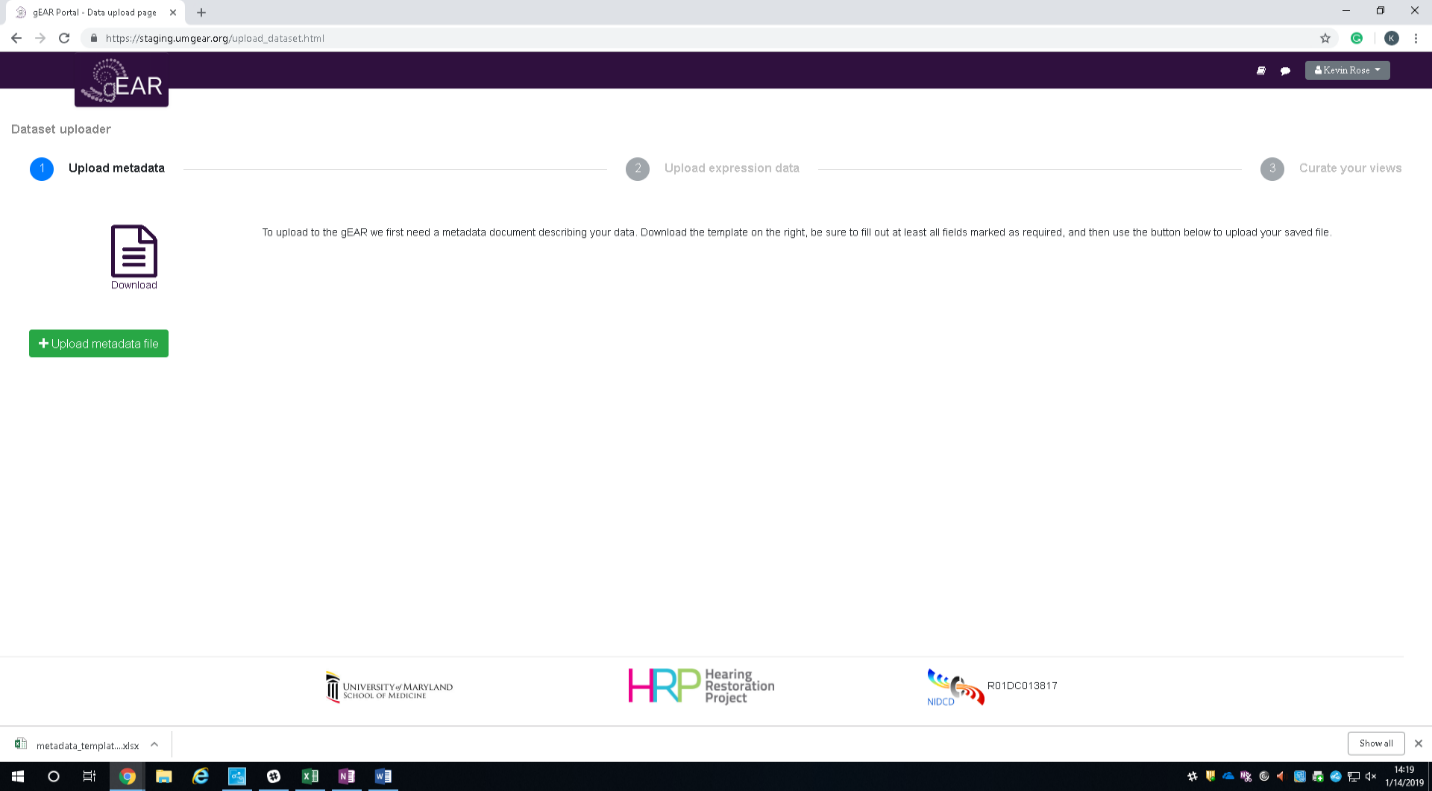
**gEAR Manual – Uploader:**

To be able to use the gEAR’s[**gEAR Uploader**](https://umgear.org/upload_dataset.html), you must be logged in. You can access the dataset uploader from the gEAR’s[**homepage**](https://umgear.org/)using the drop-down menu under your account name. The gEAR will prompt you through the upload process. However, for a detailed overview of each step (3 steps in total), please follow this guide. If you need any help, please[**contact us**](https://umgear.org/contact.html).



**Step 1 - Uploading Metadata:** Metadata is data [information] about the dataset to be uploaded. The gEAR makes this very easy with a simply to use excel sheet. To begin, click on the document labeled download which will download the metadata template excel sheet.

Required Fields (Colored):

**Title –** A concise title of the dataset (ideally less than 16 words). This typically includes the dataset type (e.g., microarrary, RNA-seq, scRNA-seq), the organism being studied, and experimental conditions (e.g, mutant vs wild type, treatment etc). Example – ‘Bulk RNA-seq analysis of neonatal mouse cochlear hair cells’.

**Summary –** A concise narrative of the experiments performed which describes how the data were produced. Example – ‘This study examined transcripts that are enriched in neonatal mouse cochlear hair cells. Hair cells were purified by FACS sorting for GFP fluorescence from the cochleae of transgenic mice in which the endogenous Atoh1 gene was fused with GFP’.

**Dataset\_type –** What type of expression data is being used? Examples – scRNA-seq, bulk RNA-seq, etc.

**Annotation\_source –** Reference annotation database. Example - Ensembl

**Annotation\_release\_number –** Release number of reference annotation. Example - 93

**Contact\_email –** Email address of the main person to contact about this dataset.

**Contact\_institute –** Name of the institution which the main person to contact is affiliated with.

**Contact\_name –** Main person to contact about this dataset (typically this is the principal investigator).

**Sample\_taxid –** NCBI numeric taxonomy ID assigned to organism. Taxid can be found [here](https://www.ncbi.nlm.nih.gov/Taxonomy/TaxIdentifier/tax_identifier.cgi).

Additional Fields:

**Geo\_accession –** GEO series ID (GSExxxx). If this ID is given, the gEAR uploader will automatically populate rows 8-18 in the metadata template. They include: contact\_email, contact\_institute, contact\_name, sample\_taxid, sample\_organism, platform\_id, instrument\_model, library\_selection, library\_source, library\_strategy, pubmed\_id.

**Sample\_organism –** Organism that the data were generated from. Example – Mus Musculus

**Platform\_id –** GEO platform ID (GPLxxxx). GEO [Homepage](https://www.ncbi.nlm.nih.gov/geo/).

**Instrument\_model –** Type of technology used to produce data. Example – Illumina HiSeq 2000

**Library\_selection –** The method used to select and/or enrich the material being sequenced. Example – transcriptomic

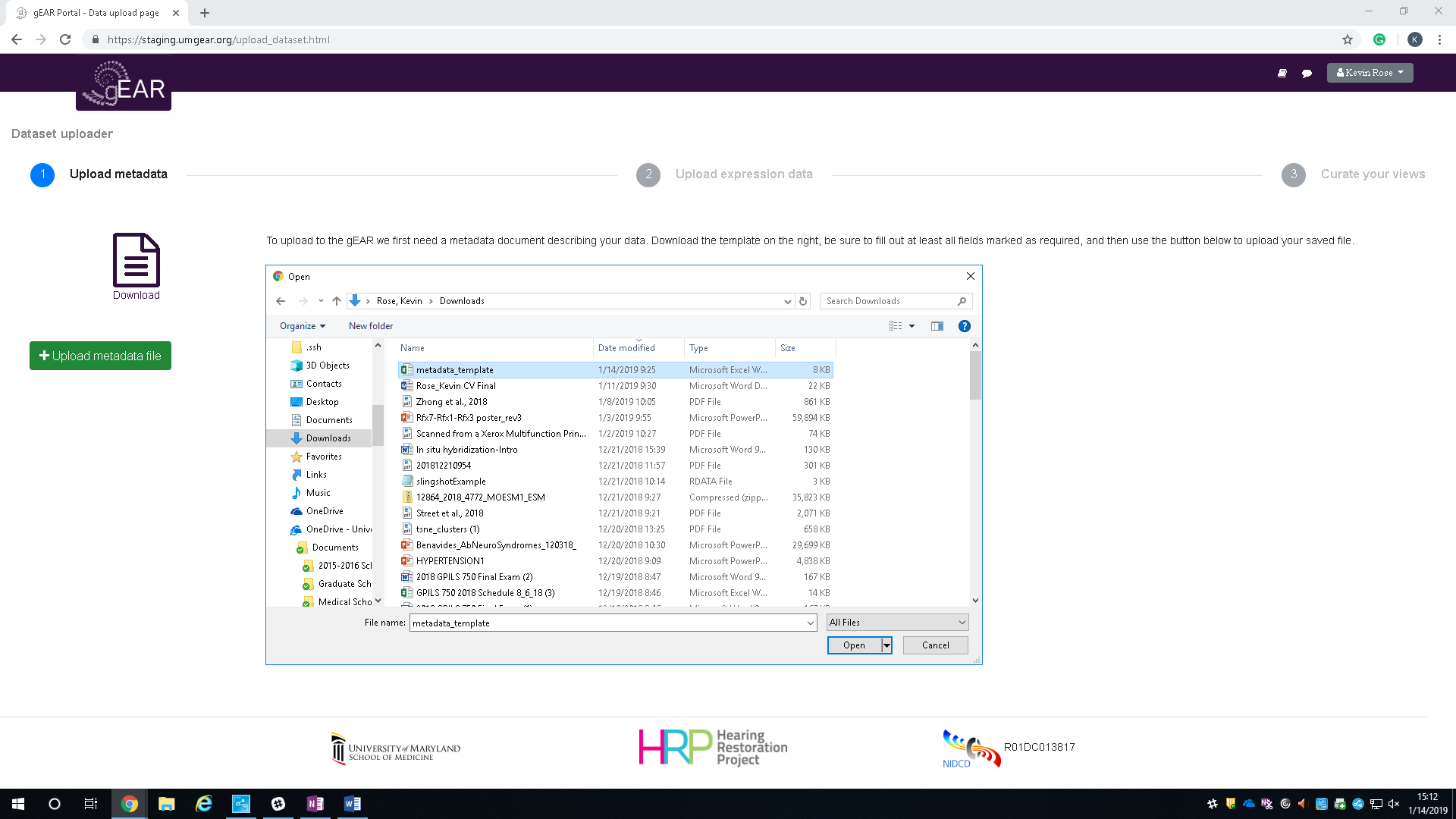
**Library\_source –** The type of source material that is being used. Example – cDNA

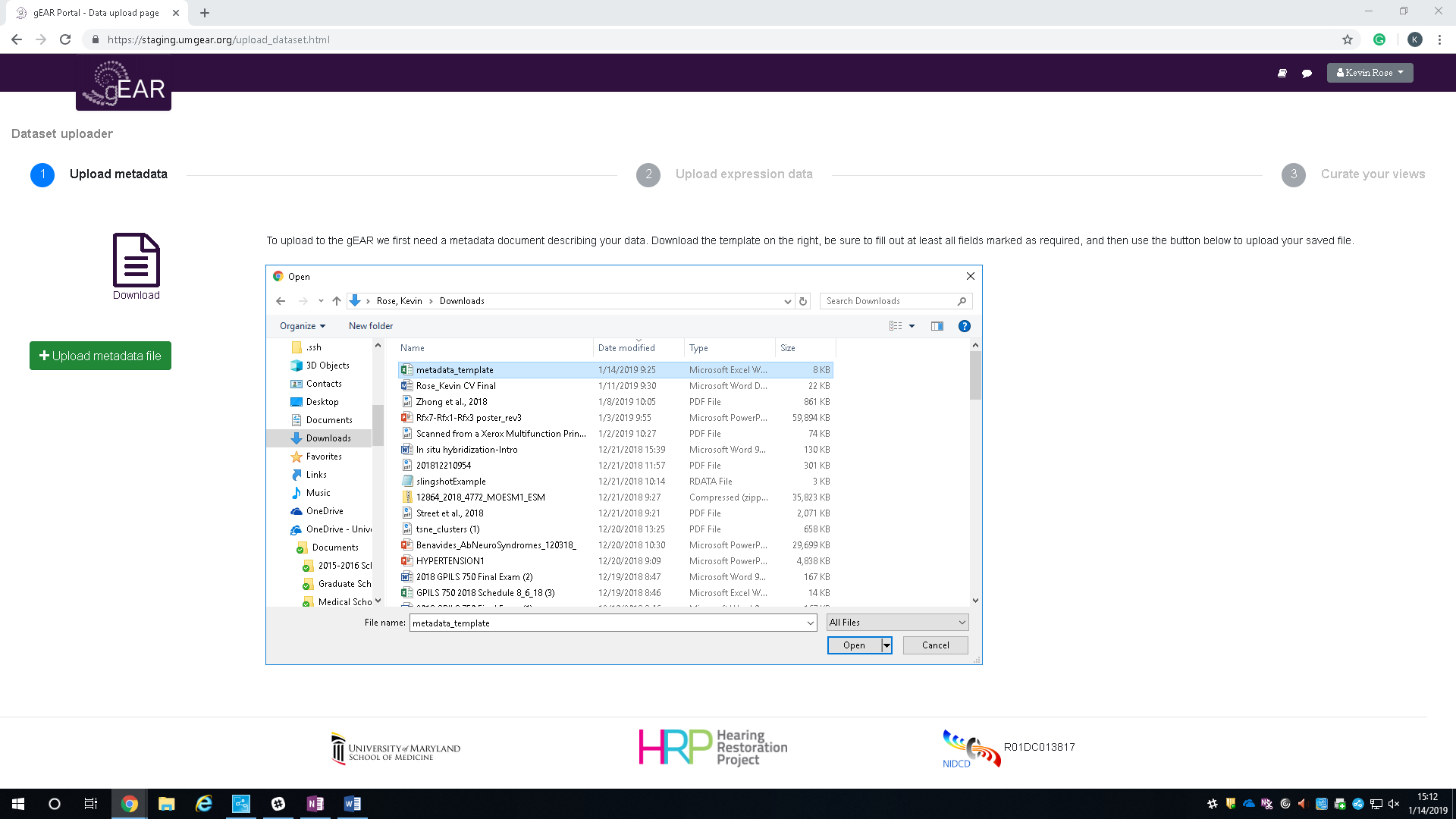
**Library\_strategy –** The sequencing technique intended for the library. Example – RNA-Seq

**Pubmed\_id –** Numeric PubMed ID.

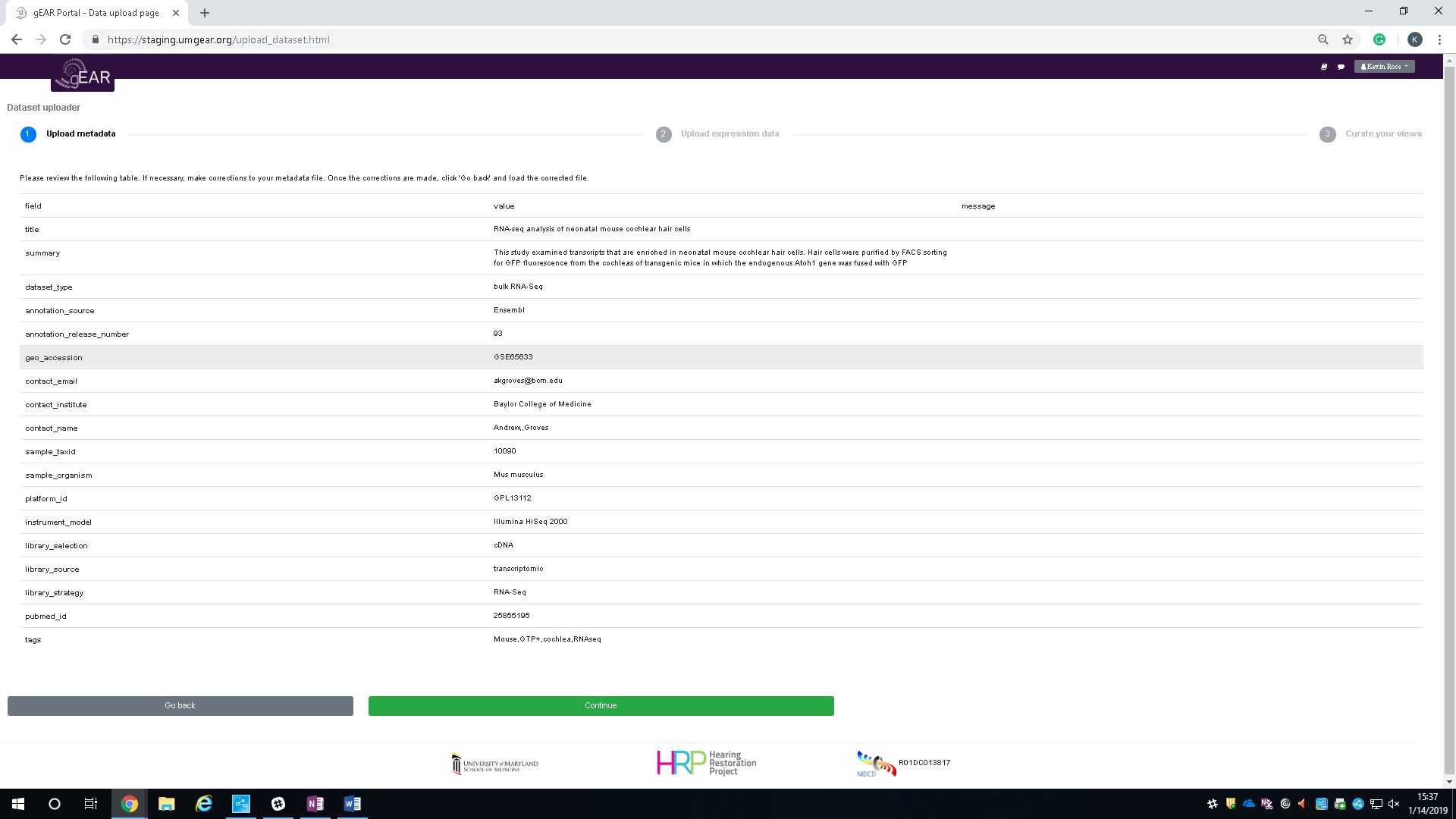
**Tags –** These are essentially keywords. Topics, categories, or subject-matter the dataset encompasses. Example – ‘Mouse, GTP+, cochlea’.

After the metadata template is complete, save it and upload it into gEAR. This is done by clicking on the **upload metadata file** button. Then, find the metadata template in your files and click open.





After uploading the metadata template, the gEAR will show you a table with all the information you just placed into the template. If all values are correct, you can proceed to the next step by clicking continue. If something is not correct, the gEAR will show an error message at the top of the webpage. The metadata template needs to be edited and uploaded again if this happens.

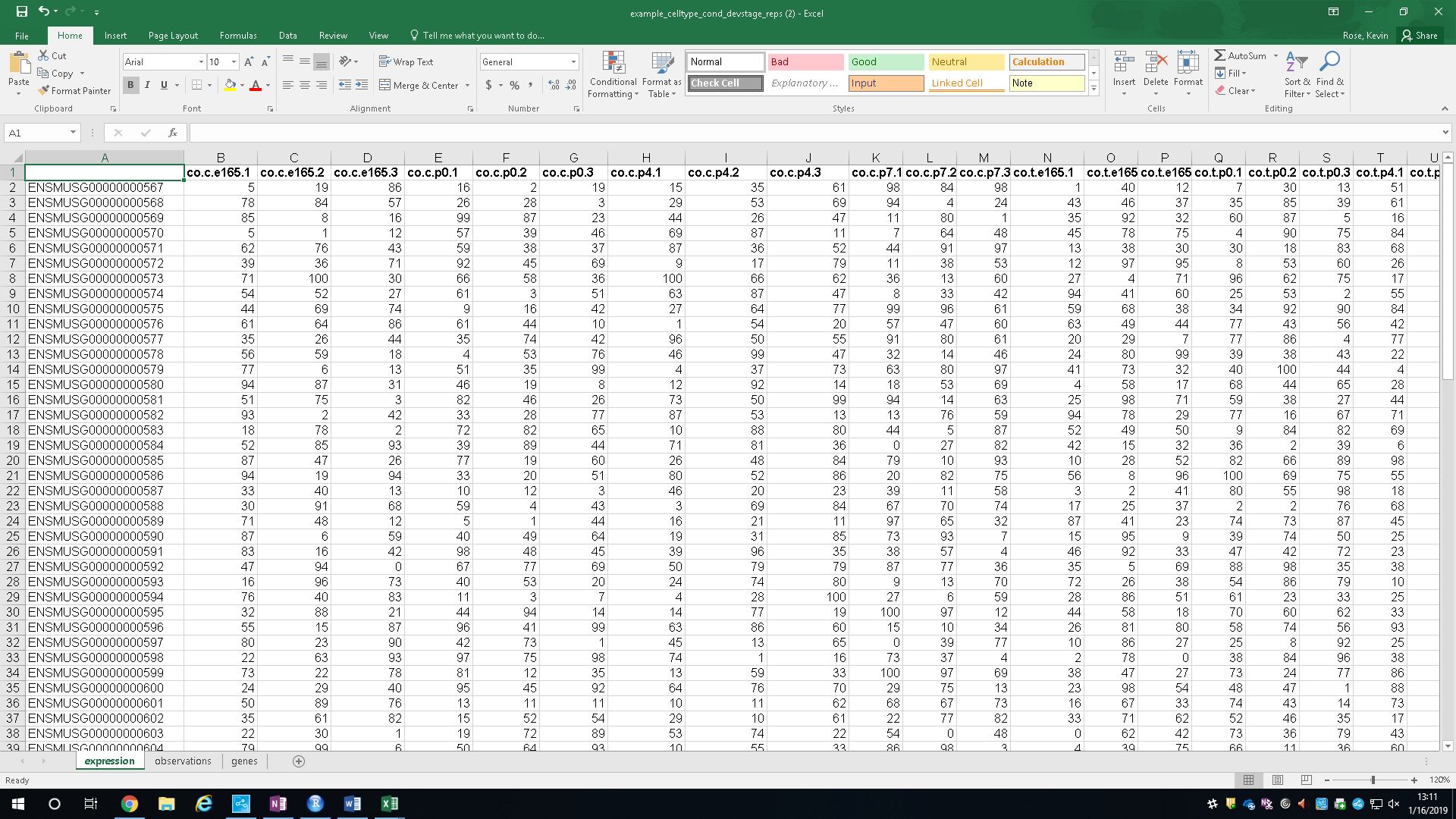


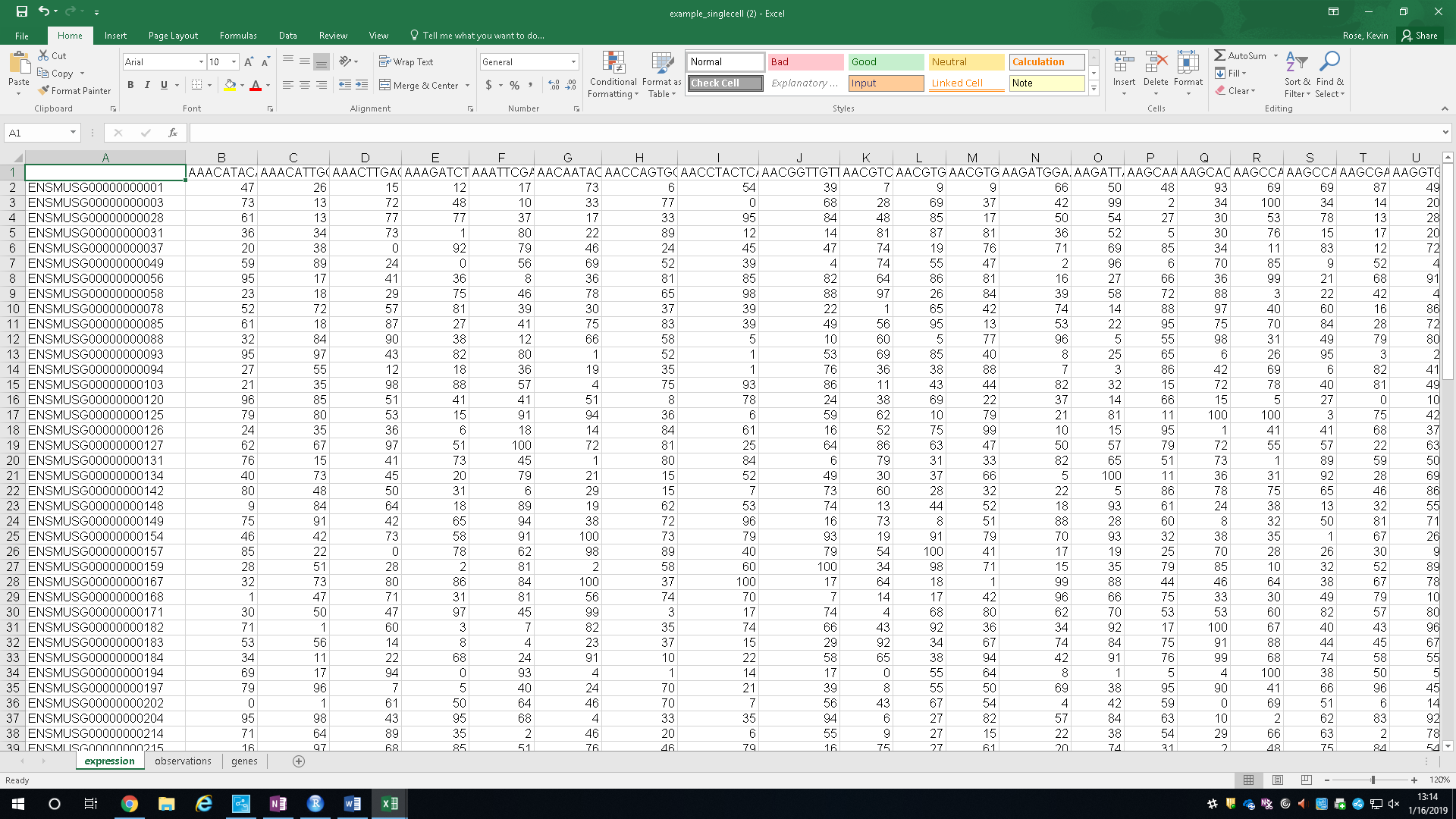
**Step 2 – Uploading Expression Data:**

Expression data is all the information from your experiment including reads, genes, and observations. However, the gEAR needs this information to be formatted in a excel sheet with **three tabs** labeled observations, expression, and genes.

**(1) Expression Datasheet -** Columns contain samples (scRNA-seq = single cells, bulk RNA-seq = tissue) and rows are genes and their expression values.

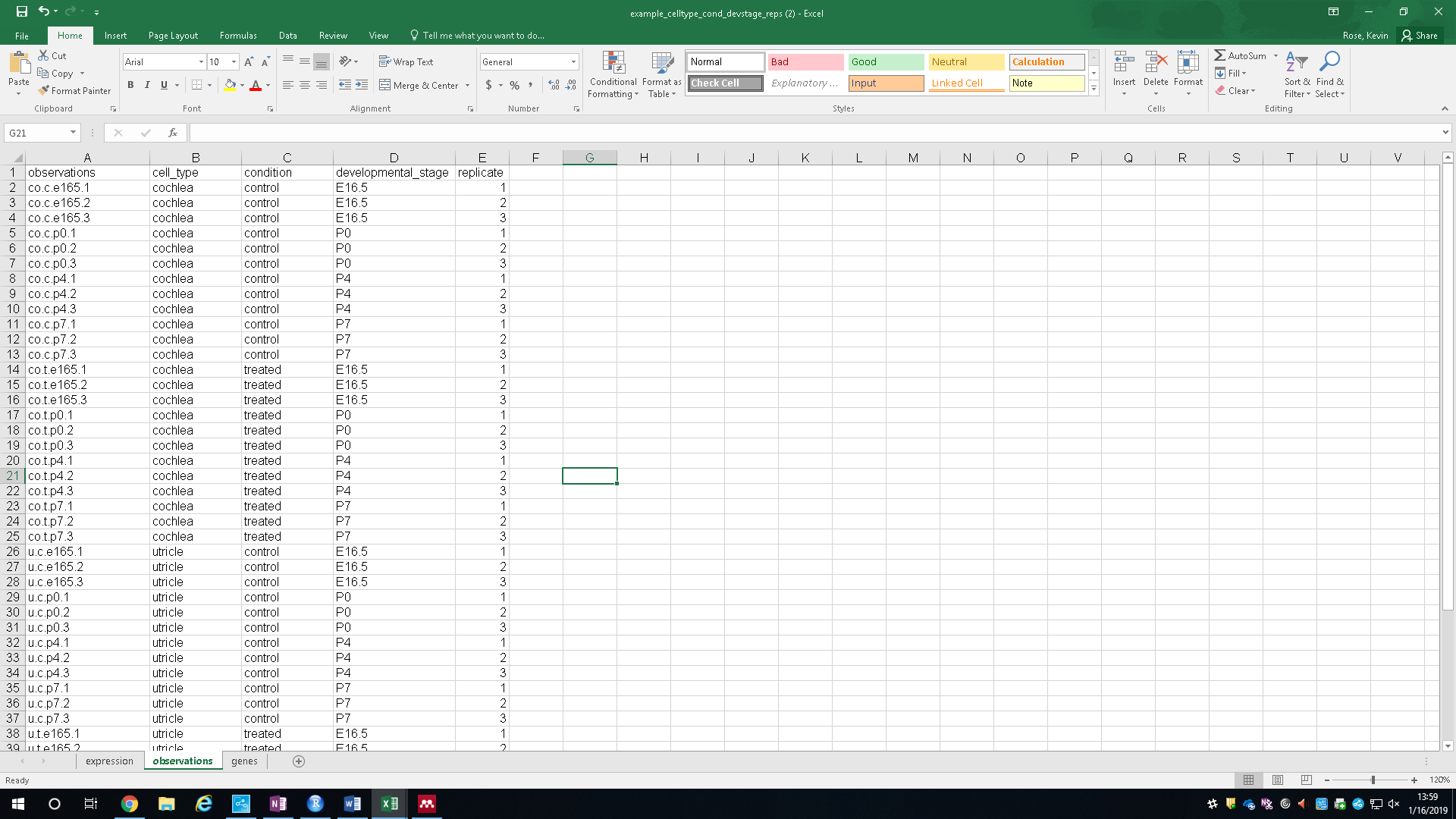
Bulk RNA-seq – Each column is a sample and each row is a gene labeled by its Ensembl ID.



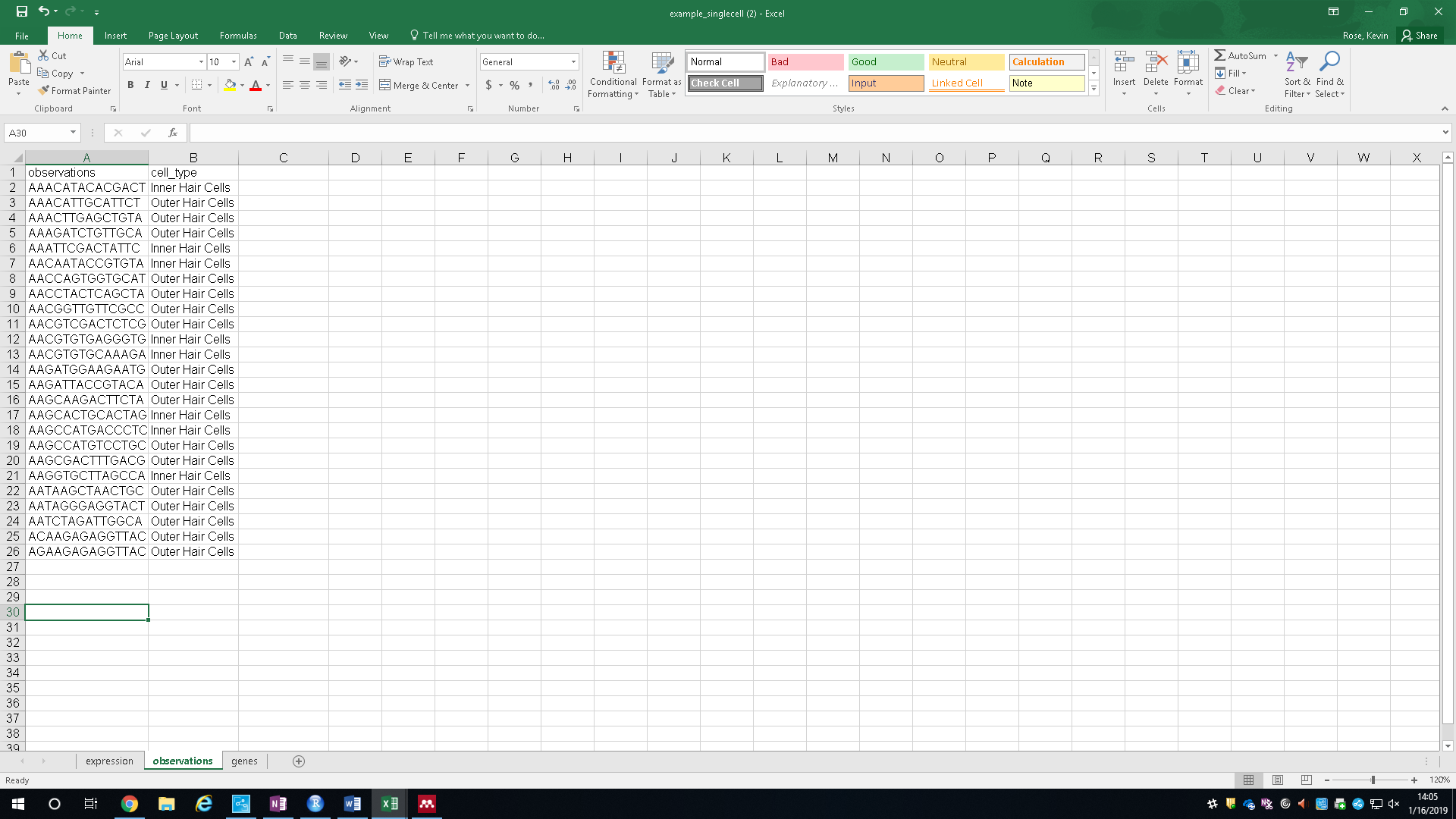
ScRNA-seq - Each column is a single cell (barcode) and each row is a gene labeled by its Ensembl ID.

**(2) Observations Datasheet -** This is the place to define the identity of each sample is from the expression datasheet (this file is essentially the column metadata). There are 5 columns that need to be filled: Observations – this is the sample name and has to match the column header in the expression datasheet.   
Cell type – can be tissue (e.g., brain, heart) or cell type within a tissue (e.g., neuron, astrocyte).   
Condition – treatment, control etc.   
Developmental stage – this is age (e.g., E16.5, P0).   
Replicate – this lets the gEAR know that the data are one of several. It is important for having the gEAR calculate standard errors at a later stage. If you need to add additional information, please contact us.

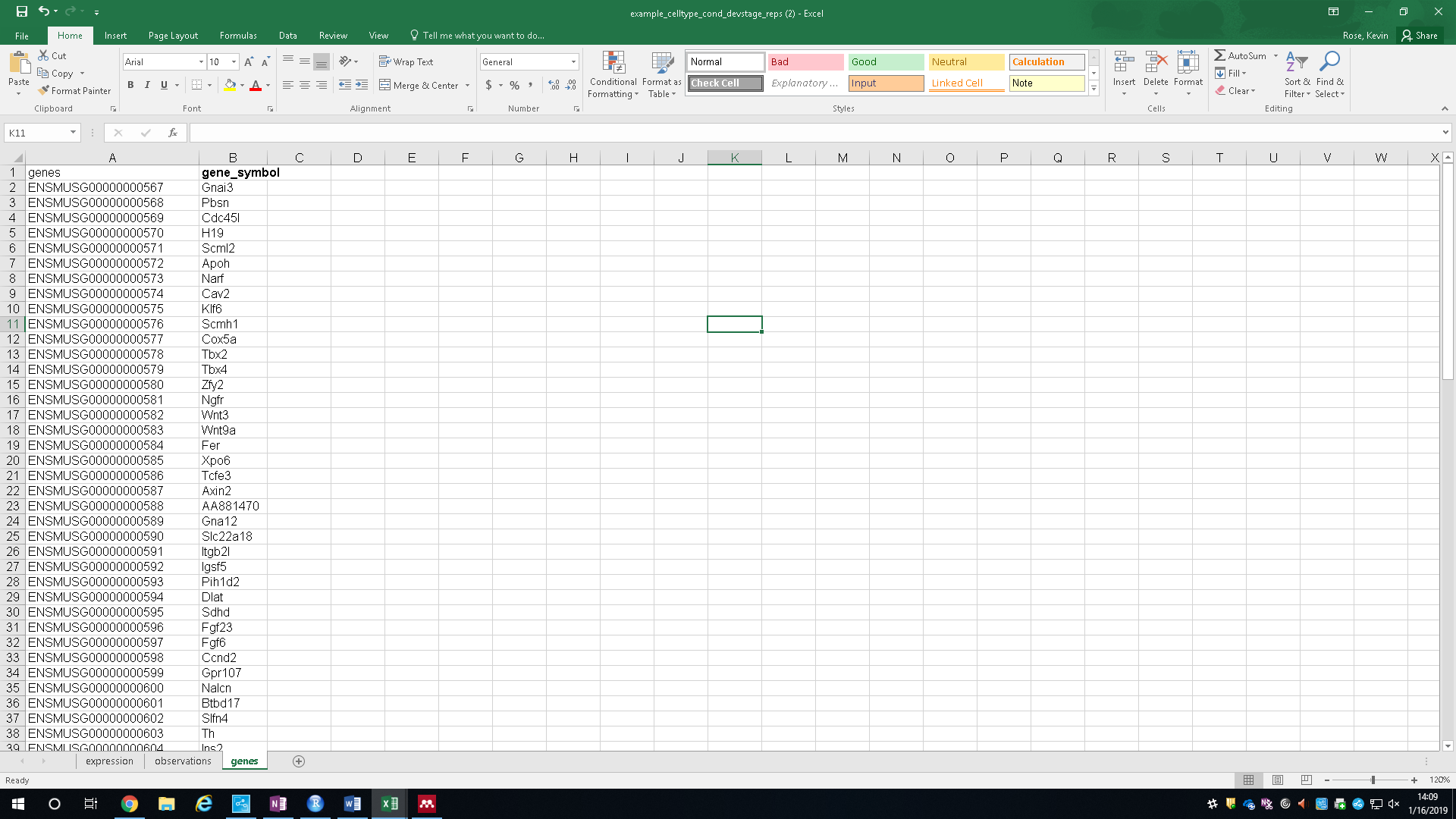
Bulk RNA-seq – observations are the names of the samples (same as columns in expression datasheet) and the cell type, condition, developmental stage and replicate for each observation follows.



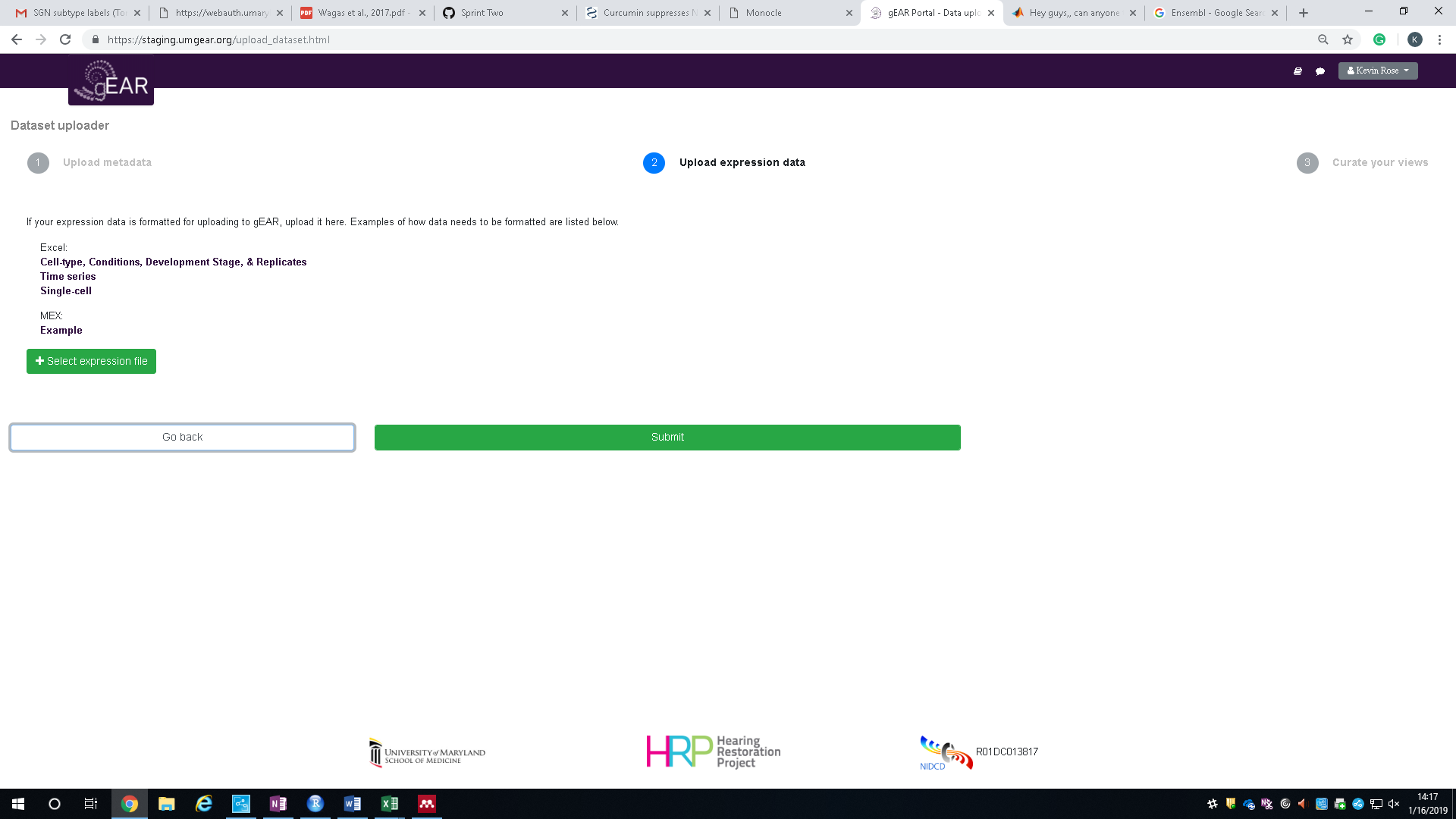
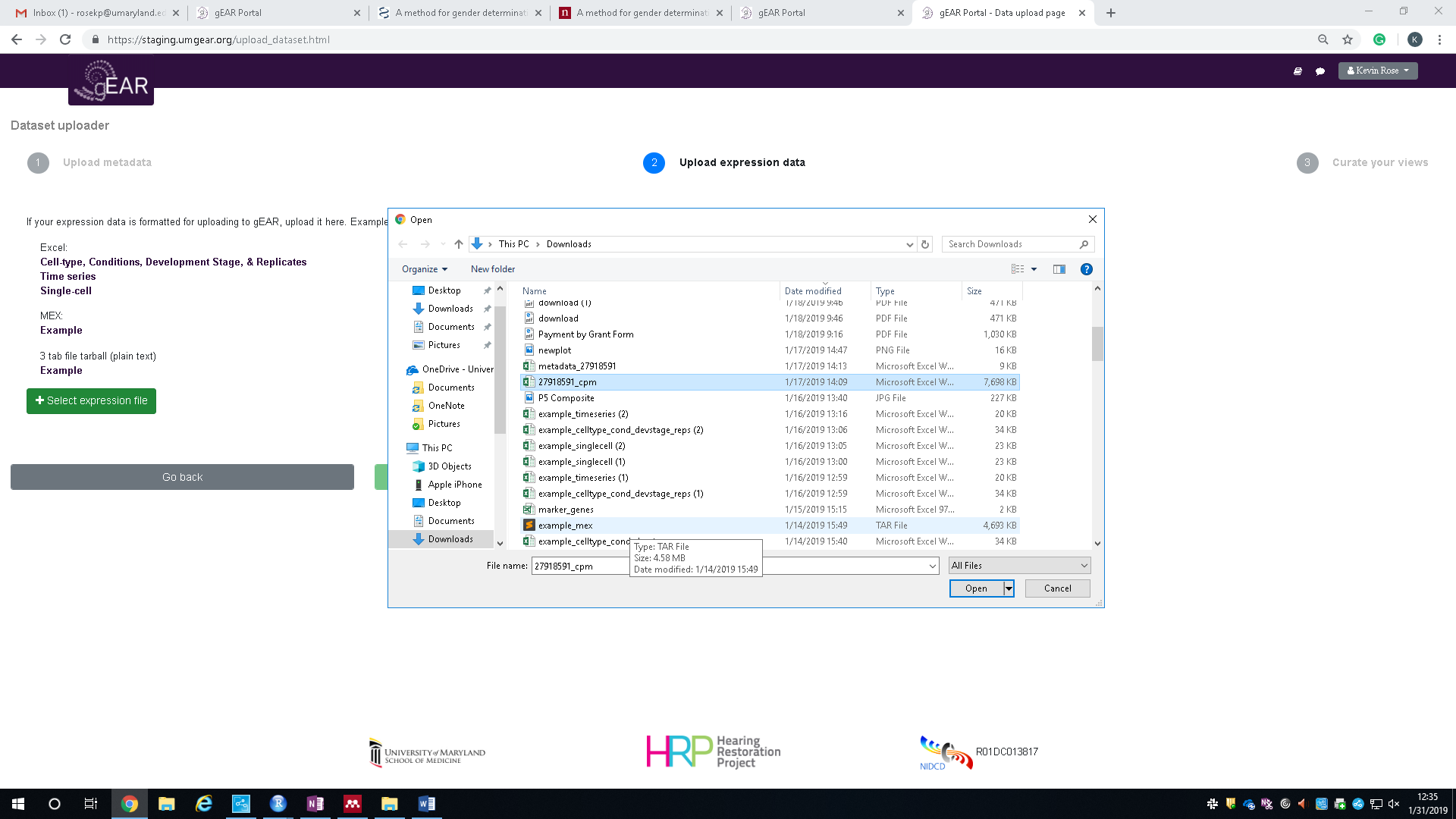
ScRNA-seq – observations are the barcodes for each cell (same as columns in expression datasheet) and the cell type follows.



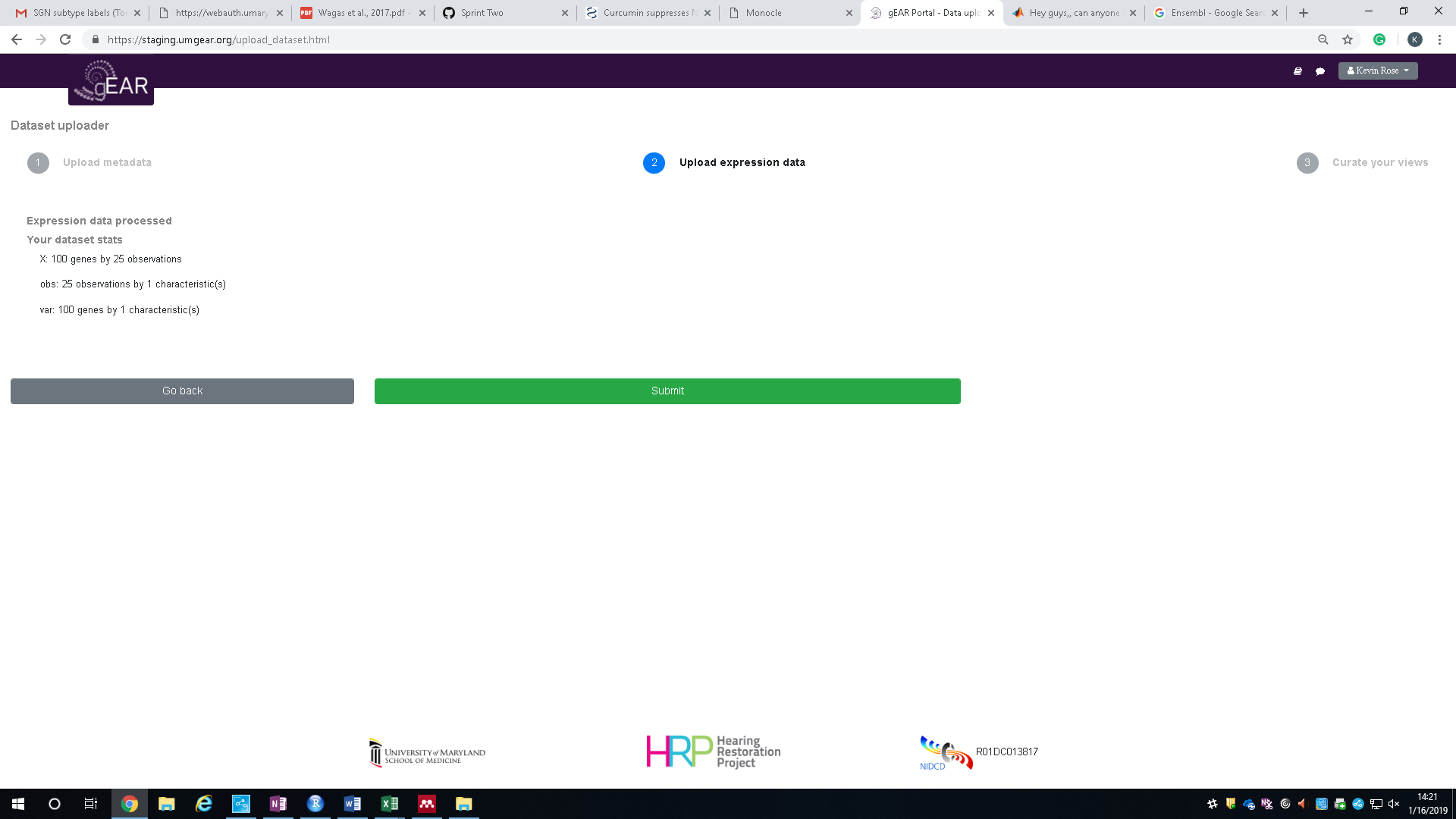
**(3) Genes:** This sheet is the row metadata sheet for the expression datasheet and consists of Gene ID”s (Ensembl) and their associated gene symbols (short gene names).



After the excel sheet is complete, you can upload it into the gEAR. Click on select expression file and open your datasheet. This step can take 30-60 seconds, depending on how large the dataset is.

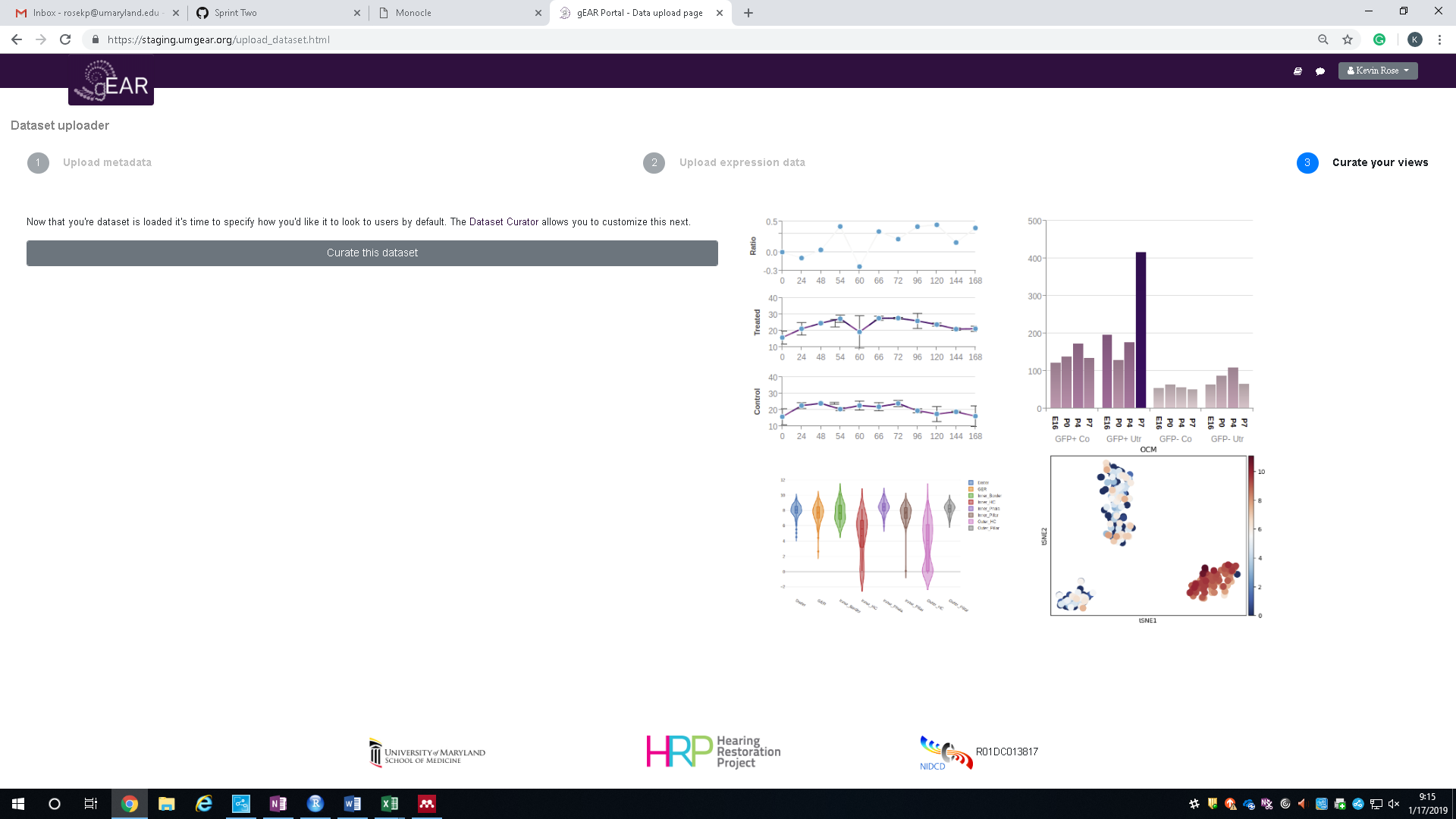


The gEAR will then confirm that your expression file is correct and inform you if there is any problems. Confirm that number of observations and genes are in your dataset and then click submit.

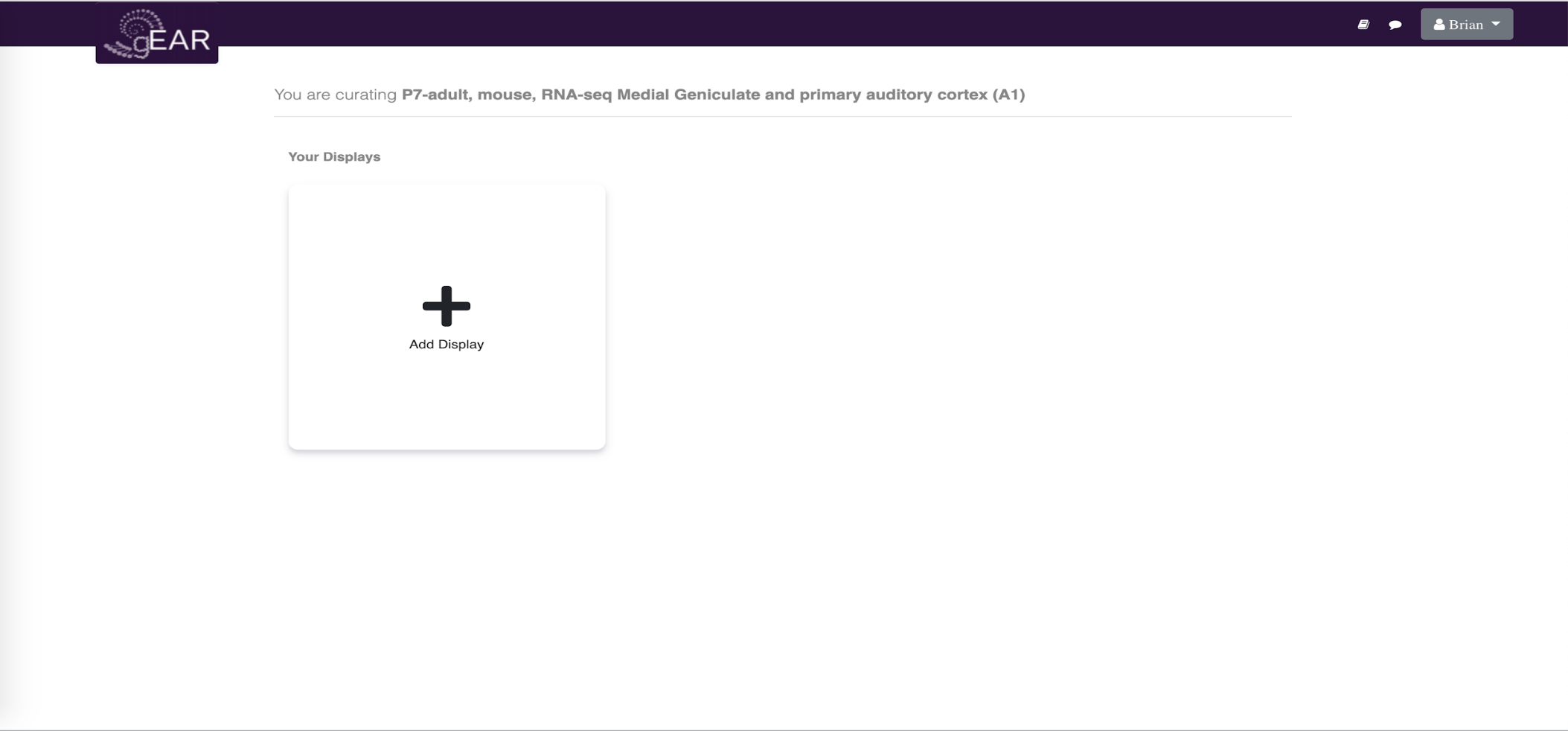


**Step 3 – Curating your dataset:**

Your dataset is now uploaded into the gEAR! The final step is to curate the dataset which means - specify how you would like it to look when you or other users browse your dataset using the gEAR (can have more than one display). To open the curator, click on curate this dataset.

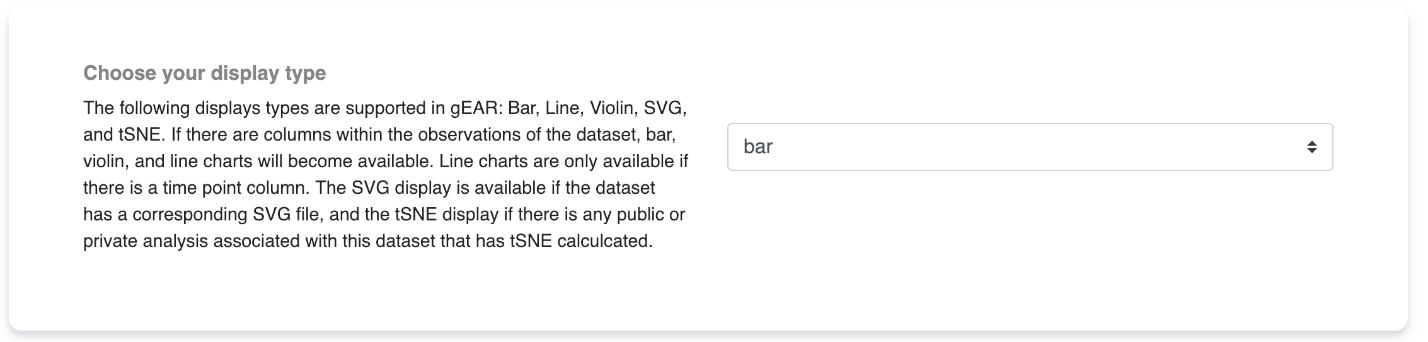


To make your first display, click on the **Add Display** panel.

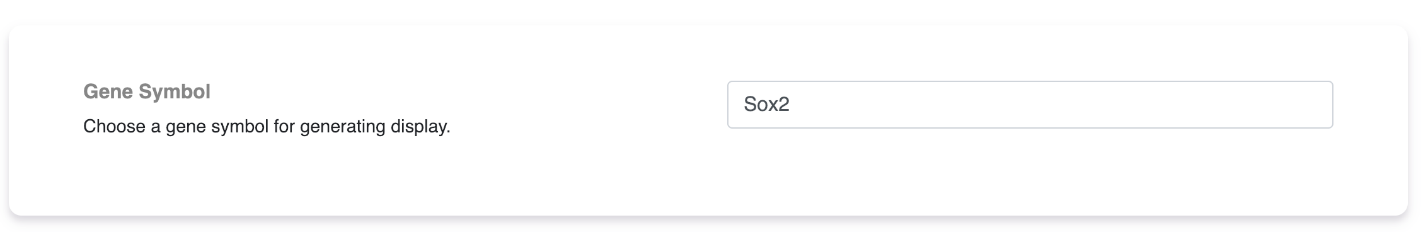


After clicking **Add Display**, you will begin the process of curating your dataset. This process includes choosing your display type, gene symbol to preview display, grouping, ordering, coloring, naming, and lastly, saving. Our example below will demonstrate curating a bar display.

**Choosing display type:**

Currently the gEAR supports bar graphs, violin plots and line graphs (SVG and tSNE support is database specific). You can change the plot type by using the drop-down menu. Display types are determined based on the columns that exist in your dataset. For example, line plots are only available if your dataset contains time\_point.

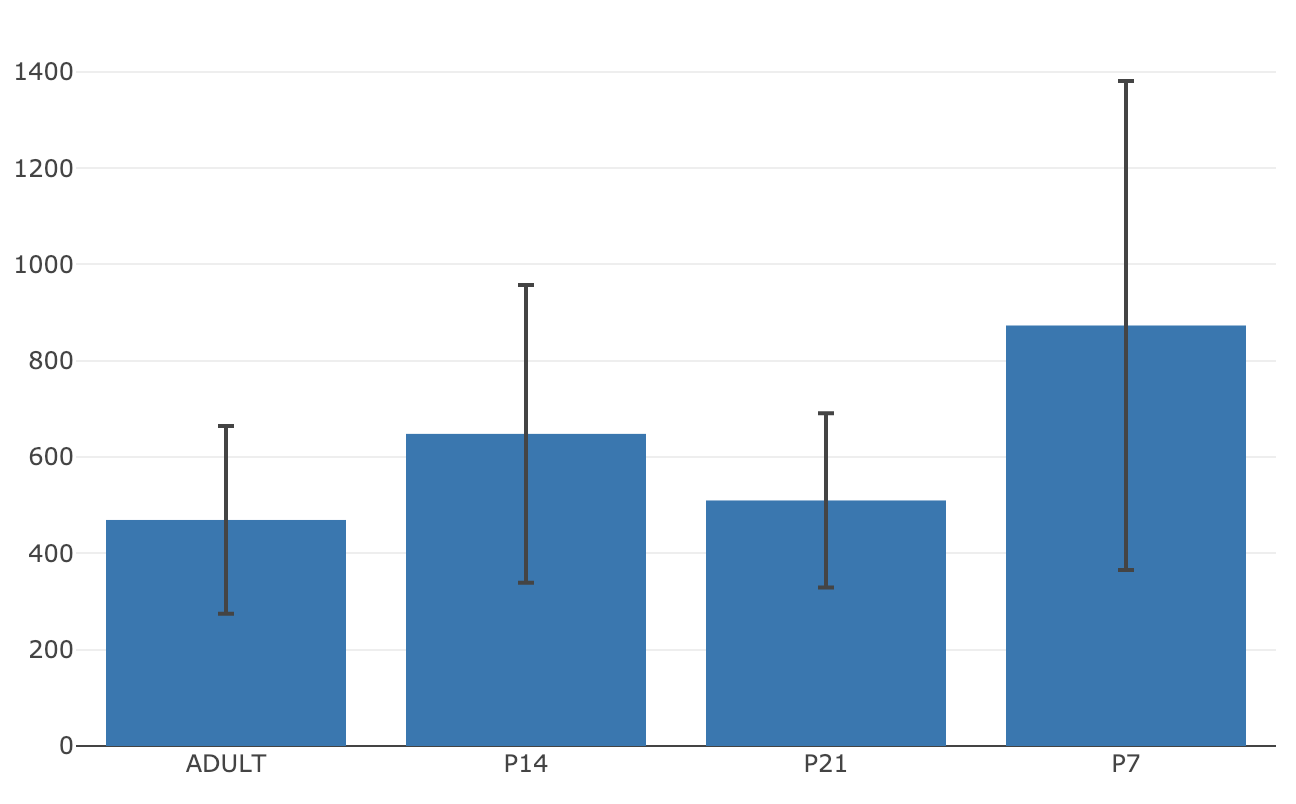
For the sake of testing out what the display will look at, enter a gene symbol that is in the dataset. Here we will be using Sox2.

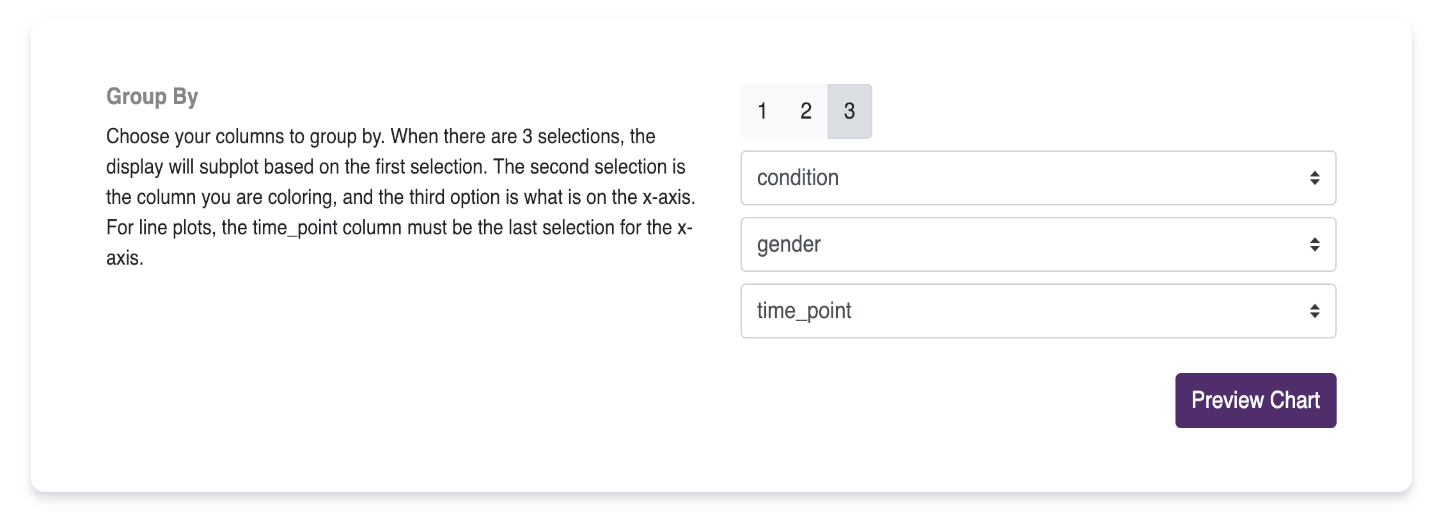
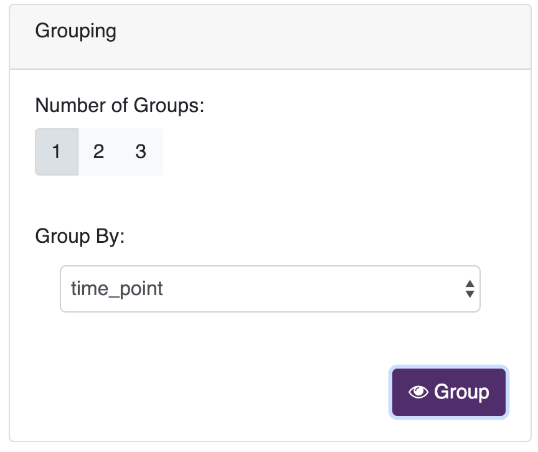


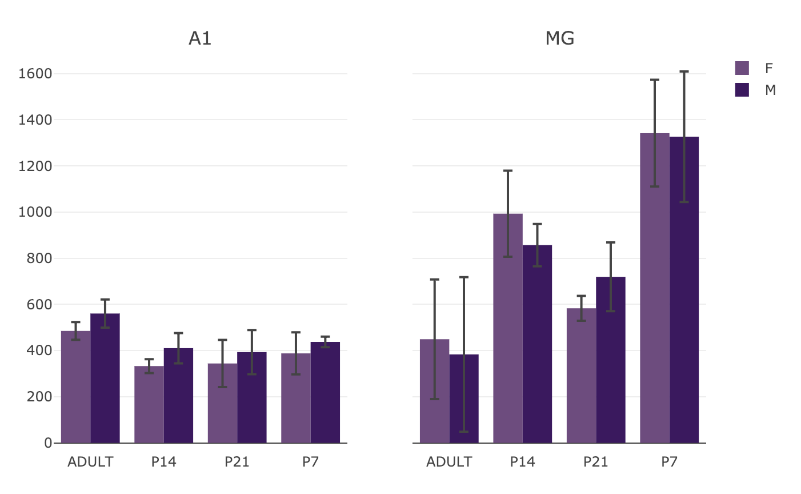
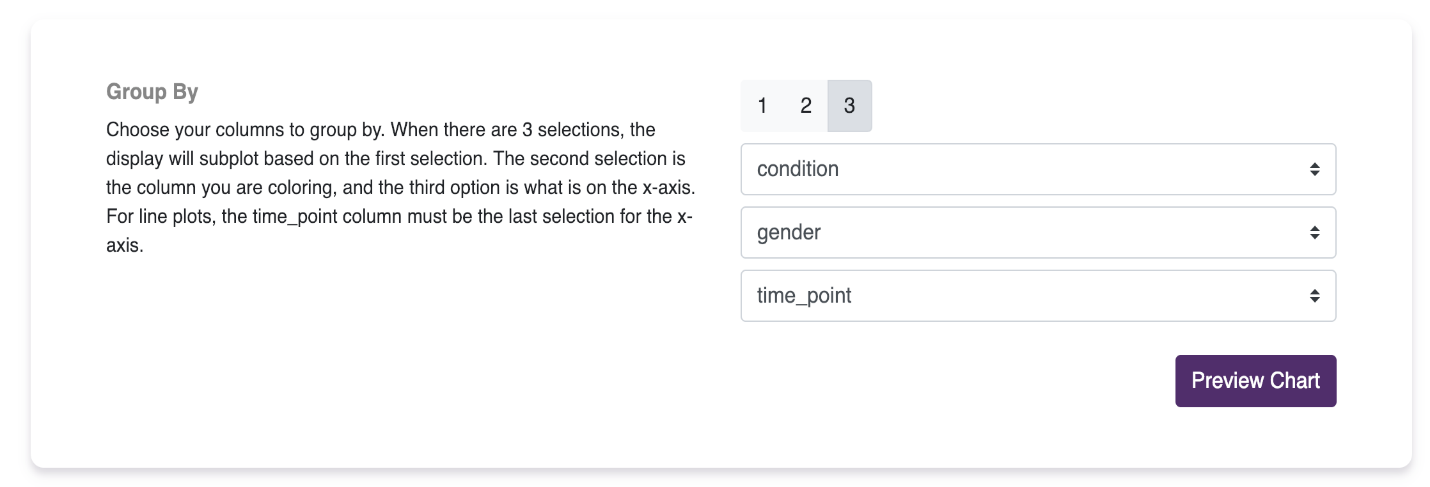
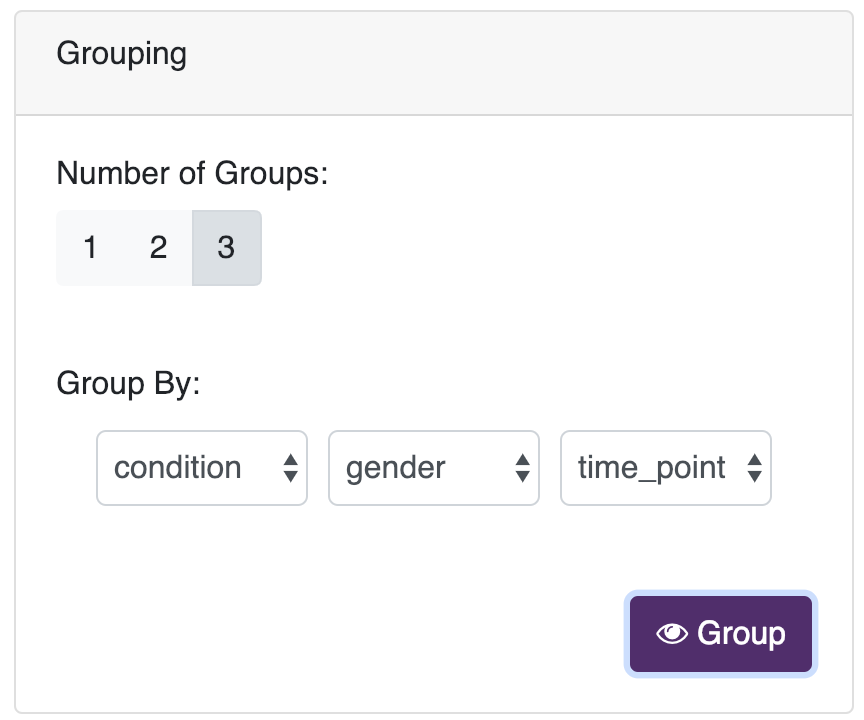
**Grouping:**

This display option lets you change the number of groups to display on your graph and how to group them. For example, the following is a dataset with samples obtained from two brain regions, at four time points and from both male and female animals, separately.

To group just by time point, for example, we can select 1 for the number of groups and time point for the group by option. This will aggregate the data and display time\_point across the x-axis. Click the **preview chart button** to preview the display.

****

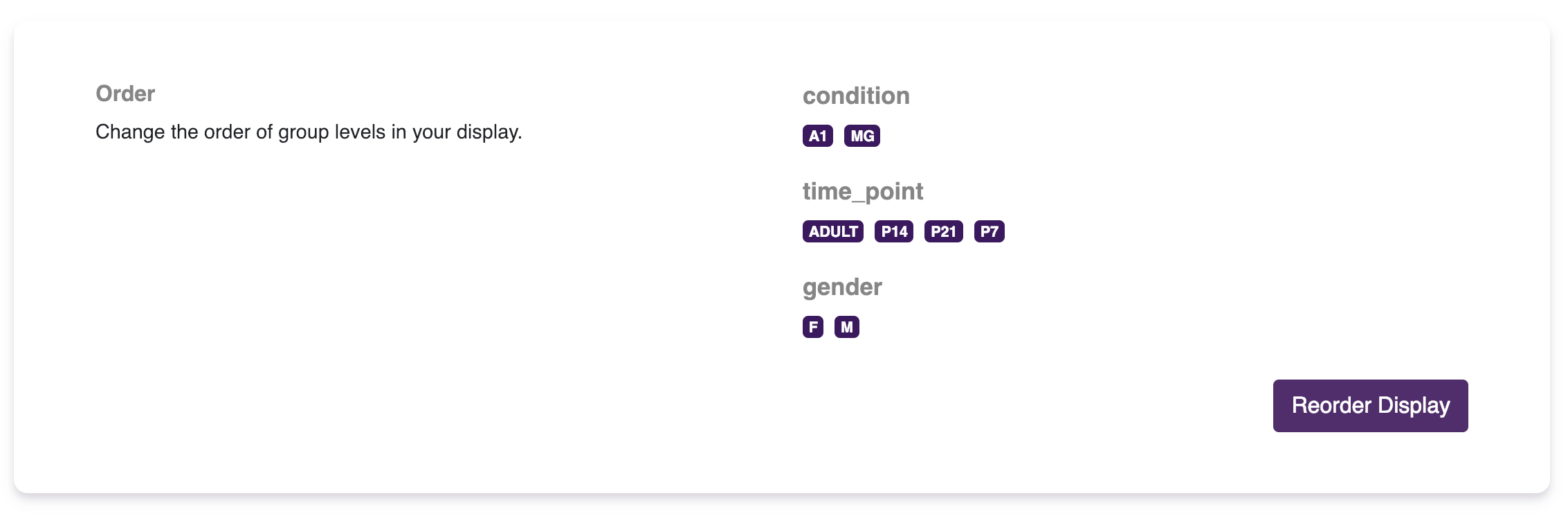


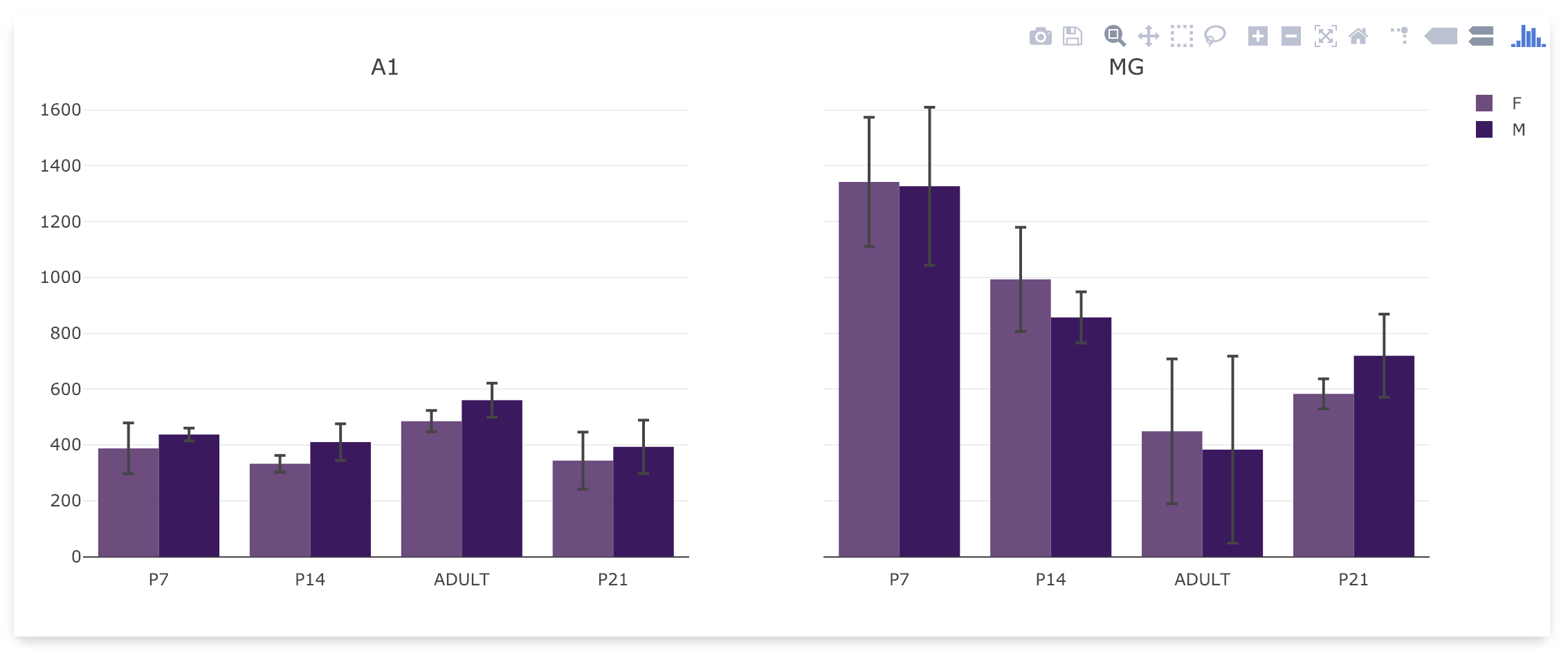
********However, if you want to show condition, gender, and time\_point you can select 3 groups and place these variables into the group by drop-down menus.

**Ordering:**

Sometimes, a dataset’s time\_point column represents ordinal categorical values, however, gEAR may not know how to order these values. The ordering panel allows you to adjust this order, including the order of other columns such as condition and gender. To change the order, simply click and drag the purple boxes into the order you want them.

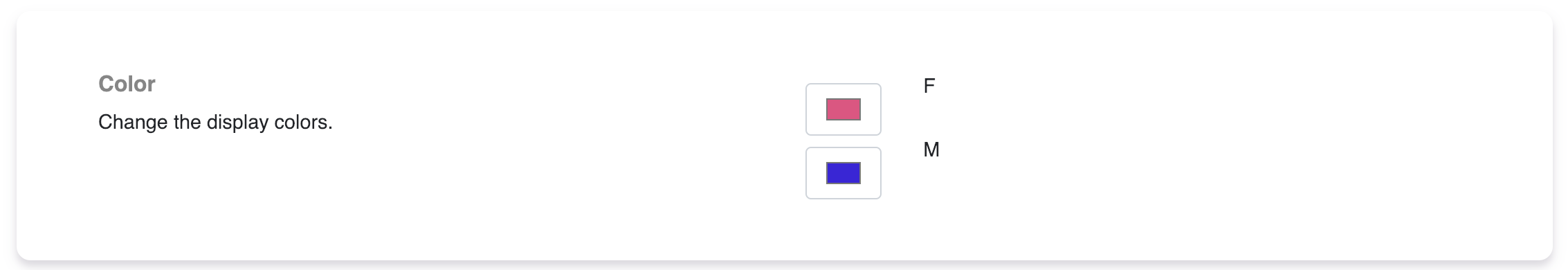
For example, the time points above are not in the correct order (ADULT, P14, P21, P7). We can fix this by simply dragging the labels to the correct order and clicking Reorder.

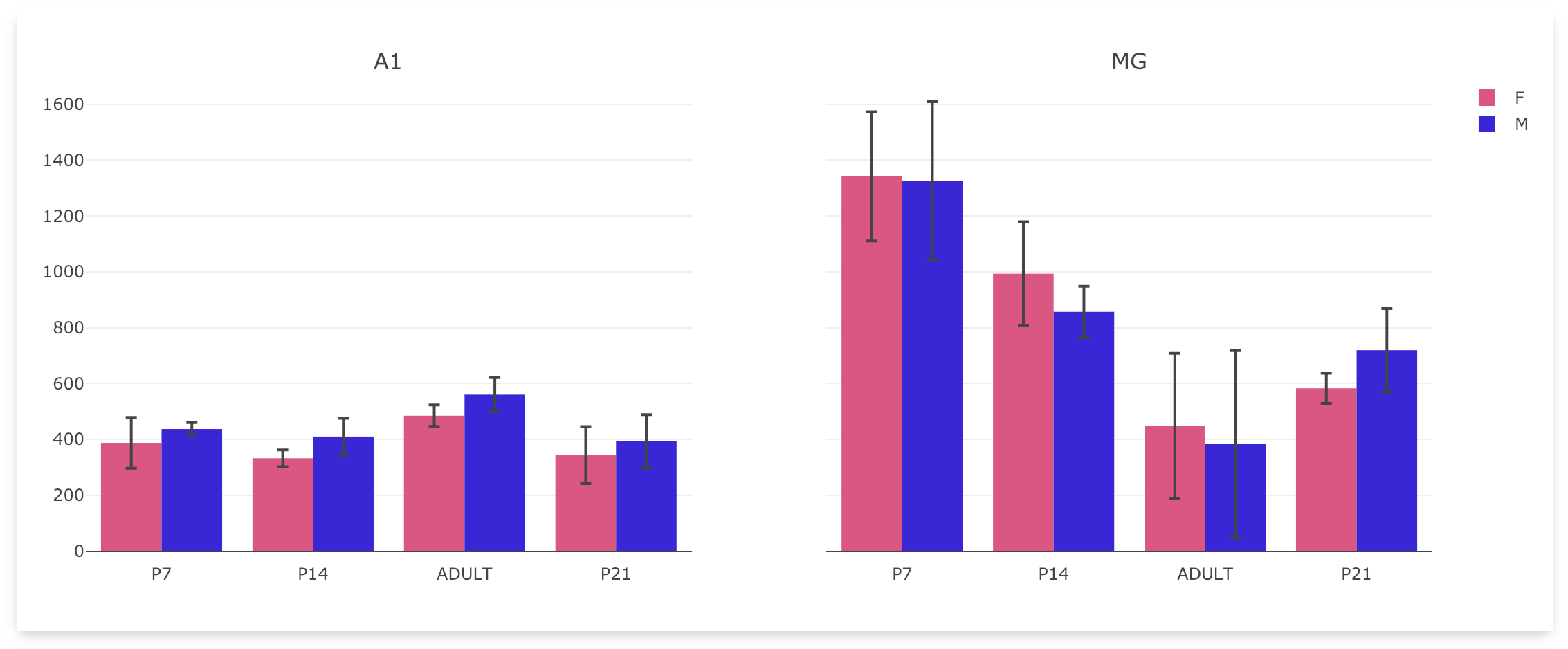




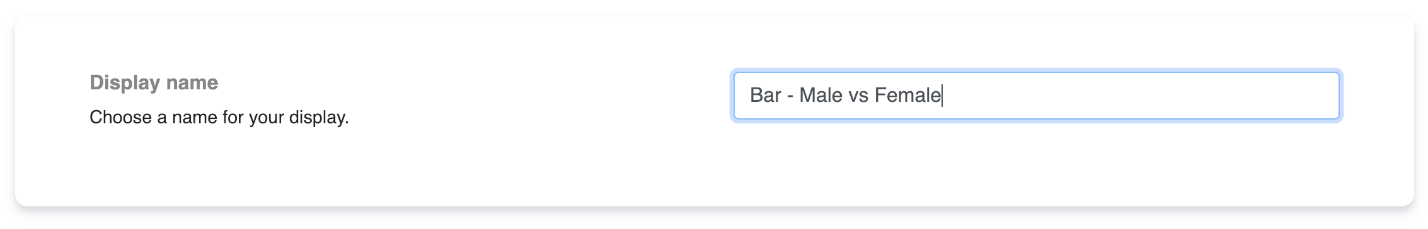
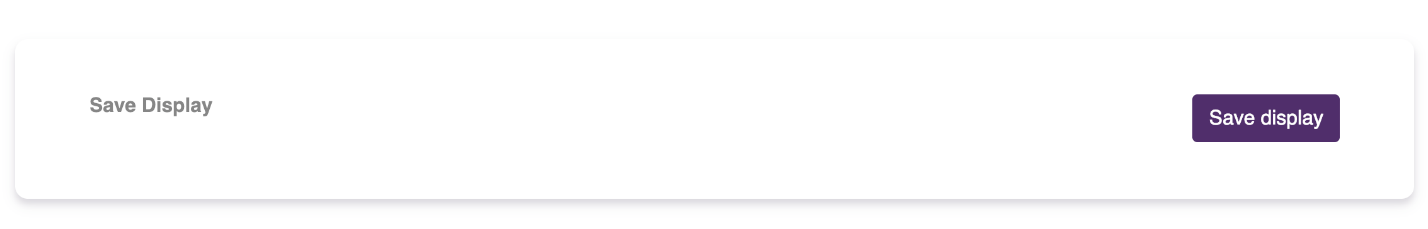
**Coloring:**

Allows you to change the colors of the different conditions on your graph. Click on the condition that you want to change colors of and select the color you want to change it to. Here we changed females to pink and males to blue.

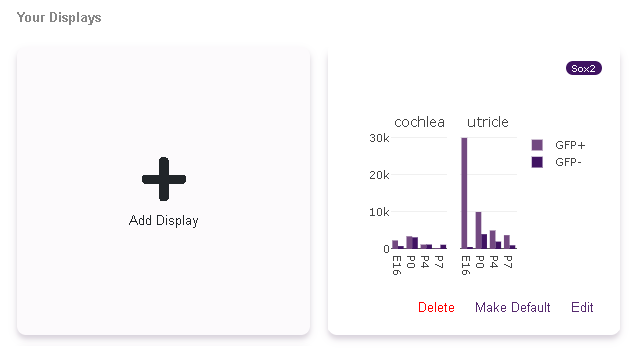




**Naming:**

****This will change the name of this curation of the dataset. Once you are happy with how the data is presented, change the name of your display to something representative of the visualization and press save. This will save your display and it will be accessible in the “Your Curations” panel.

****Once the display is saved it will show up in your display list on the curator. At any time you can delete this display or edit it. You can also make it the default display when searching a gene in this dataset. If you save more than one display, you will be able to toggle between displays.



****