**Notes on RNA-seq gene expression analyses from Wasp Antennal Drumming Project**

**Hypothesis**: Antennal drumming and restricted access to nutrition influence development of worker-like phenotype in larvae.

**Prediction**: By simulating drumming and restricting colonies’ access to forage, larvae that would normally develop into new queens will develop worker-like phenotypes (including gene expression). **We have lipid data that support this hypothesis! So what about gene expression?**

**Replicates:** 30 samples, 5 groups (6 samples per group)

\*Note: odd numbered samples (-1, -3, -5) were run in Lane 1; even numbered samples (-2, -4, -6) were run in Lane 2.

**Groups**

* **Foundress-Reared (samples “F1” thru “F6”), OR “WORKER DESTINED”:** 
  + All individuals in this group were collected at the beginning of the year, when only the foundress tended the larvae. **We expect these larvae to develop into workers, naturally.**
* **Worker-Reared (samples labelled “DR”, “DU”, “NR”, and “NU”), OR “QUEEN DESTINED”:** 
  + All individuals in this group were collected later in the summer, when larvae would normally be developing into new queens. We expect these larvae to vary in phenotype based on the different treatments applied.
  + **Worker-reared Treatment Groups** 
    - “DR”: Larvae experienced artificial drumming (D) on the nest, and the colony experienced restricted foraging (R)
    - “DU”: Larvae experienced D, the colony experienced unrestricted foraging (U)
    - “NR”: Larvae experienced no artificial drumming (N), the colony experienced R
    - “NU”: Larvae experienced N, the colony experienced U. **These individuals are unmanipulated and should develop into new queens.**

|  |  |  |
| --- | --- | --- |
|  | **Drumming Treatment (D)** | **No Drumming Treatment (N)** |
| **Restricted Nutrition (R)** | **DR**  **\*\* Predict most WORKER-like gene expression** | **NR**  **\*\* Predict some WORKER-like gene expression**  **(nutrition-related genes)** |
| **Unrestricted Nutrition (U)** | **DU**  **\*\* Predict some WORKER-like gene expression**  **(drumming-related genes)** | **NU**  **\*\* Predict least WORKER-like gene expression** |

**Analyses Requested**

1. QC and filtering of raw data.
2. Read mapping to existing *P. fuscatus* transcriptome (from Toth lab) and expression counts for each sample and each transcript.
3. Can use existing best BLAST hits for annotation (from Toth lab). Facility also mentioned interest in trying out other annotation strategies.
4. Differential expression analysis.
   1. We would like the data analyzed in two ways:
      1. With only group as a factor. Data: from all 5 groups. Asking: which transcripts are differentially expressed in each of the 5 groups relative to each other? Result: DET lists for each of the 5 groups.
      2. 2x2 factorial design. Data: from 4 worker-reared groups only. Asking: How does drumming, nutrition and the interaction of these factors influence gene expression in developing larvae? Result: DET lists for nutrition and drumming factors.
5. After generating initial DET lists. **(Note: we would like to stop and meet and discuss results at this point. We may want to take it from here on our own).**

Comparisons of DET lists and expression patterns from each group. The main question we will aim to ask is: does the DR group have the most WORKER-like gene expression; that is, gene expression that is most like the “F” group?

**1)** Clustering analysis—what is clustering pattern of gene expression for the 5 groups?

Heatmaps to visualize expression patterns and identify clusters of genes showing patterns of interest.

* + All genes
  + All differentially expressed genes
  + All caste related differentially expressed genes (e.g., “F” v “NU”)

1. Linear discriminant analysis—which groups show the most similar expression patterns? Must limit to the top 25 genes (# variables must be smaller than the # of samples)
2. PCA clustering based on gene expression space
   1. Use the three different gene lists here too

\*\*Potential conundrum: we noticed, after submitting our samples, that samples collected later in the season might be a better representation of the data. Therefore, looking for outliers will be important. I can either give you an a priori list of samples to look out for, or we can discuss that after the initial analysis.