**Data Wrangling**

**Introduction:** Prenatal exposure to excess glucocorticoids programs hypothalamic-pituitary-adrenal (HPA) function, increasing risk for psychiatric disease. The hypothalamic paraventricular nucleus (PVN) is centrally implicated in programming stress responsivity. The objective of this study was to understand how synthetic glucocorticoids influence gene expression in the PVN, and to further elucidate how this transcriptional signature relates to the phenotypic profile observed in the exposed animals.

**Methods:** mRNA library preparation was performed using Illumina TruSeq V2 mRNA enrichment using standard protocols. High-throughput sequencing were performed on an Illumina HiSeq 2500 sequencing system using standard run, following the protocol recommended by Illumina for sequencing mRNA samples. Sequencing was done for each biological replicates at 1 × 51 bp by the Donnelly Centre for Cellular and Biomolecular Research. Sequence reads were aligned to the *Cavia porcellus* reference genome (cavPor3.83) with Tuxedo Suit tools [1] accessed through The Galaxy Project [2]. Sequencing depth for RNA-seq samples averaged 45 million reads per biological sample with >85% overall alignment rate. Subsequent analyses were performed in R (version 3.2.3). Gene read counts were determined with Genomic Alignments (version 1.6.3) as described by the authors [3]. Outliers were removed using Cook’s distance with default cutoffs [4], and data with from genes with multiple readouts from the same gene name were summed and recast using melt and cast functions in R.

1. Trapnell, C., et al., *Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks.* Nat Protoc, 2012. **7**(3): p. 562-78.

2. Afgan, E., et al., *The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update.* Nucleic Acids Research, 2016. **44**(W1): p. W3-W10.

3. Lawrence, M., et al., *Software for Computing and Annotating Genomic Ranges.* PLoS Comput Biol, 2013. **9**(8): p. e1003118.

4. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.