070 STAT RNAseq project

**Bioinformatic Methods**

All bioinformatic analyses and commands are provided in the GitHub repository: https://github.com/theo-allnutt-bioinformatics/070\_stat\_RNAseq. RNAseq reads were trimmed to remove Illumina adaptor sequences and low quality ends using Trimmomatic v0.36 (1). Trimmed reads were mapped to the coding sequences (cds) of the *Danio rerio* reference genome: ftp://ftp.ncbi.nlm.nih.gov/genomes/refseq/vertebrate\_other/Danio\_rerio/reference/GCF\_000002035.6\_GRCz11/ using BBmap v38.33 (2). Mapped read counts per sample were tabulated using a custom Python script and statistically analysed for differential expression using EdgeR (3) and ALDEX2 (4).

**References**

1. Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170.

2. https://jgi.doe.gov/data-and-tools/bbtools/bb-tools-user-guide/bbmap-guide/

3. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2009;26(1):139-40.

4. Fernandes, AD, Macklaim, JM, Linn, TG, Reid, G, Gloor, GB (2013). “ANOVA-Like Differential Gene Expression Analysis of Single-Organism and Meta-RNA-Seq.” PLoS ONE, 2013, volume 8, issue 7, e67019.