****

**­­­­­­­­­­­­­­­­**

 **UNDERGRADUATE SUMMER VACATION SCHOLARSHIP AWARDS – FINAL SUMMARY REPORT FORM 2020/21**

***NB: This whole report will be posted on the Society’s website therefore authors should NOT include sensitive material or data that they do not want disclosed at this time.***

**Name of student:**

Jie Yue Chan

**Name of supervisor(s):**

Professor Abigail Tucker

**Project Title: (no more than 220 characters)**

Fate decisions in the mammalian dental lamina

**Project aims: (no more than 700 words)**

Overall Aim: to understand the molecular mechanisms that drive whether a tooth forms or not from the dental lamina during development

Aim 1: To identify key molecular players in dental lamina fate decisions using an RNAseq dataset from the minipig

Aim 2: To test whether the identified genes show differential gene expression in a range of mammalian embryos with different tooth replacement strategies.

Methods

We have previously generated an RNAseq database created from dissected dental lamina from minipig embryos. Minipigs, as humans are diphyodont with two sets of teeth. Here we have collected tissue from the dental lamina between the 3rd and 4th premolars, which later regresses and does not form a tooth, and the dental lamina adjacent to the deciduous 3rd premolar, which goes on to form a replacement tooth (Popa et al., 2019). The RNA has been checked for quality, with sequencing performed by the Oxford Genomic Centre. The FastQ sequence were aligned to the minipig (Sus scrofa) genome and imported in to Degust (degust.erc.monash.edu) for differential gene expression visualization.

Bioinformatics were performed on this dataset and highlighted genes will be verified by performing immunofluorescence on a variety of mammalian species.

Mammalian embryos have all been collected and stored in Ethanol ready for wax embedding and sectioning.

These include:

Bat (diphyodont) (samples collected from Carolia colony at London zoo)

Human (diphyodont) (samples from Human Developmental Biology Resource)

Mouse (monophyodont) (samples bred in house at KCL)

Opossum (mixed monophyodont and diphyodont) (samples from colony at the Crick Institute)

Gene/protein expression will be imaged by confocal and apoptome microscopes.

Tucker, A.S., & Fraser, G.J. (2014). Evolution and developmental diversity of tooth regeneration. Sem. Cell Dev. Biol. 25-26, 71-80. PMID: 24406627

Popa, E.M., Buchtova, A., Tucker, A.S. (2019). Revitalising the rudimentary replacement dentition in the mouse. Development Feb 8; 146 (3). PMID:30658984

**Project Outcomes and Experience Gained by the Student (no more than 700 words)**

Throughout my project I learnt how to use bioinformatics tools to compare large RNA datasets by using the programming software R and the DESeq2 package provided by Bioconductor. I learnt lab techniques such as wax slide creation and immunohistochemistry and practiced on sample mouse slides. The bioinformatics took quite a time to learn so I concentrated on this and understood how to check whether highlighted genes were good candidates.

The following is the workflow that I have undertaken and experience gained. The R script used to generate list of differentially expressed genes was ran in R Studio.

1. Data input

The RNAseq database that was previously generated had been checked for quality, and the FASTQ files which contain the reads have been aligned to the reference genome of the minipig to generate BAM files for each tissue sample. The number of reads of different genes are counted using the function “featureCounts” from the “Rsubread” package.

The BAM files are linked to the correct version of GTF files (which are used to hold gene structure information specific to the species being investigated) using the argument “annot.ext”, and “isGTFAnnotationFIle” is set to “TRUE”. The output is a count matrix which are combined into a single count table using function “cbind”.

To run the DESeq analysis, the function “DESeqDataSetFromMatrix” is used from the “DESeq” package. Read counts are stored in argument “countData”, with a column metadata table of sample information and the design formula stored in the argument “colData”. The “design” argument takes data from the count table and column metadata to inform DESeq2 how to process the samples, which in my project was the deep compared to the interdental lamina to find genes responsible for tooth formation. Care is taken to ensure that the columns in count table and count metadata are in the same order.

2. Running the DESeq2 pipeline

The “DESeq” function from the package “DESeq2” ran on the dataset consists of 3 steps: estimate size factors (normalization), estimate dispersions, and apply Wald test statistics. The differences across the columns in the count table are mainly due to different library sizes and not gene expression changes, therefore normalization is required to correct size differences to allow the RNAseq data to be comparable. The dispersion is a parameter for the negative binomial distribution which is used in RNAseq data to see if they are differentially expressed. The gene-wise dispersions are estimated and shrunk to improve the accuracy of dispersion to model the counts. Finally, the Wald test is performed.

The results are called using the “results” function in the DESeq2 package. The “contrast” argument is used to choose a reference level for the factors, and the first factor specified is used as a control for DESeq2 to make the comparisons. The “alpha” argument specifies the FDR (false discovery rate) cutoff for the adjusted p value – this corrects for the multiple testing problem due to the high number of genes tested and reduces the number of false positives.

3. Result presentation

Weakly expressed genes are high in shot noise instead of biological noise and can be filtered out independently to improve the multiple testing adjustment.

The results table only uses gene IDs from Ensembl and is hard to understand, therefore gene names and its descriptions were added via the functions “useMart” –> “getBM” –> “match” in the “biomaRt” package. The final results table is saved as a CSV file via function “write.csv”.

Plots such as MA plots and PCA plots can be used to show differentially expressed genes.

Personal experience gained

I practised independent learning to familiarise myself with R and to understand the script and its individual components using various online resources and courses.

As I was working in silico in a mostly wet lab team, I gained unique insight of working in an interdisciplinary team as I had to communicate and collaborate between different labs to obtain the support I needed.

I also had experience running immunos in other tooth replacement models. I learnt techniques for microtoming and creating wax slides in a snake model and ran Sox9 immunos.



Figure 1: Sox9 immuno of salivary gland in snake model

**Please state which Society Winter or Summer Meeting the student is intending to present his/her poster at:**

Summer

**Proposed Poster Submission Details (within 12 months of the completion of the project) for an AS Winter/ Summer Meeting – (no more than 300 words)**

Proposed abstract to be submitted:

Understanding fate decisions is an essential goal in developmental biology with important ramifications for tissue regeneration and repair and our understanding of disease. Every cell in a body receives and provides cues to its neighbours allowing initially similar cells to embark on unique pathways. In this project we have investigated the decisions that determine whether a replacement tooth forms or not by comparison of the development of the dental lamina (tooth forming) and interdental lamina (non-tooth forming) in the diphyodont (two generations of teeth) minipig.

The deep dental lamina of the embryonic minipig was compared to the interdental lamina in the same embryo using RNAseq. The resulting data was analysed using DESeq2, and the results of the analysis showed a total of 10 differentially expressed genes including 2 undefined genes. The gene expressions of the output were validated using publicly available datasets on Eurexpress, Allen Brain and Genepaint. One candidate showed some expression in the tip of the epithelium of the molar placode. Molars are formed by adding teeth of the same family that buds off the dental lamina via serial addition. This mode of formation is similar to the pattern observed during replacement tooth formation, therefore, this gene was a good candidate gene for playing a role in tooth replacement. The expression pattern of this gene was therefore followed in the mouse, which does not replace its teeth, and opossum, which selectively replaces its teeth. The results highlight a selection of genes that may play key roles in determining whether a tooth is replaced or not.

**Brief Resume of your Project’s outcomes**: **(no more than 200-250 words)**.

*The title of your project and a brief 200-250 word description of the proposed/completed project. The description should include sufficient detail to be of general interest to a broad readership including scientists and non-specialists. Please also try to include 1-2 graphical images (minimum 75dpi). NB: Authors should NOT include sensitive material or data that they do not want disclosed at this time.*

Fate decisions in the mammalian dental lamina

Understanding fate decisions is an essential goal in developmental biology with important ramifications for tissue regeneration and repair and our understanding of disease. Every cell in a body receives and provides cues to its neighbours allowing initially similar cells to embark on unique pathways. In this project we have investigated the decisions that determine whether a replacement tooth forms or not by comparison of the development of the dental lamina (tooth forming) and interdental lamina (non-tooth forming) in the diphyodont (two generations of teeth) minipig.

The deep dental lamina of the embryonic minipig was compared to the interdental lamina in the same embryo using RNAseq. The resulting data was analysed using DESeq2, and the results of the analysis showed a total of 10 differentially expressed genes including 2 undefined genes. The gene expressions of the output were validated using publicly available datasets on Eurexpress, Allen Brain and Genepaint. One candidate showed some expression in the tip of the epithelium of the molar placode. Molars are formed by adding teeth of the same family that buds off the dental lamina via serial addition. This mode of formation is similar to the pattern observed during replacement tooth formation, therefore, this gene was a good candidate gene for playing a role in tooth replacement.

**Other comments: (no more than 300 words)**

M. I. Love, W. Huber, S. Anders: Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. bioRxiv (2014). doi:10.1101/002832

|  |
| --- |
| **Data Protection/GDPR**: I consent to the data included in this submission being collected, processed and stored by the Anatomical Society. Answer YES or NO in the Box below |
| Yes |
| **Graphical Images**: If you include graphical images you must obtain consent from people appearing in any photos and confirm that you have consent. A consent statement from you must accompany each report if relevant. A short narrative should accompany the image. Answer N/A not applicable, YES or NO in the box below |
| N/A no people in images |
| **Copyright**: If you submit images you must either own the copyright to the image or have gained the explicit permission of the copyright holder for the image to be submitted as part of the report for upload to the Society’s website, Newsletter, social media and so forth. A copyright statement must accompany each report if relevant. Answer N/A not applicable, YES or NO in the box below |
| Yes I own copyright |

 *Signature of student.........................Jie Yue Chan....................Date…30/09/2021……..*

 *Signature of supervisor………Abigail Tucker …..... Date…30/09/2021…*

END OF FORM

----------------------------------------------------------------------------------------------------------------------------------------

*File: USVRS 202021 Report Tucker Chan – no sigs*