AGRICULTURE & FOOD

Genome wide chromatin accessibility in the developing ovary of the Black Tiger Prawn

James Kijas[1], Moira Menzies[1], Amin Mohamed[1], Sean McWilliam[1], Jake Goodall[1], Russell McCulloch[1], Roger Huerlimann[2,3], Dean Jerry[2,3], Greg Coman[4] and Nick Wade[1,2]

[1]CSIRO Agriculture & Food, Queensland Bioscience Precinct, 306 Carmody Rd., St. Lucia, Brisbane, QLD 4067, Australia.
[2]ARC Research Hub for Advanced Prawn Breeding, Townsville, QLD 4811, Australia.
[3]College of Science and Engineering, James Cook University, Townsville, QLD 4811, Australia.
[4]CSIRO Agriculture & Food, Bribie Island Research Centre, 144 North St, Woorim, QLD 4507, Australia.

JAMES COOK UNIVERSITY AUSTRALIA

Black Tiger prawn (*Penaeus monodon*) is a major aquaculture production species farmed both locally in Queensland and interest is returning in it's production globally. Breeding programs are relatively new in comparison to terrestrial livestock species, and there is a lag in the development of the genomic tools needed to enhance applied breeding and understand the biology of key traits. Draft reference genome assemblies are now emerging, prompting us to begin generating the datasets needed to construct first pass functional annotation. We elected to use ATAC-seq [1] to collect chromatin accessibility data and to map active promoters and other elements of the gene regulatory machinery.

Study Design Focussed on Ovarian Development

We collected ovaries at three stages of development and generated ATAC-seq and RNA-seq data as described in Figure 1. This study design generates data to initiate functional annotation and pinpoint the genes and gene networks driving ovarian maturation.

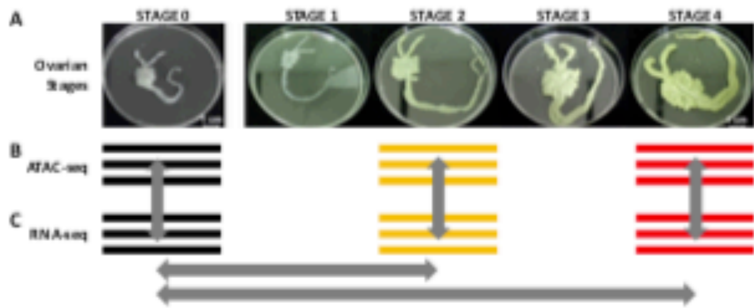

Figure 1: **Study Design. (A)** Prawn ovarian developmental stages are classified using a numerical system where non-eyestalk ablated control animals are stage 0, and stages 1 – 4 indicate increasing development. **(B)** ATAC-seq and RNA-seq **(C)** data was generated using three biological replicates for stage 0, 2 and 4 ovaries. Comparison of stage 2 and 4 datasets with control (stage 0) will be used to search for differential transcriptional and chromatin landscape signals (grey horizontal arrows). Similarly, co-analysis of the ATAC-seq and RNA-seq data will be used to characterise the control of gene expression (grey vertical arrows).

Table 1 shows summary statistics at the conclusion of the ATAC-seq data generation phase of the project. The FRiP scores and total number of chromatin accessibility peaks per sample indicate a dataset of moderate to high quality. Further, peak overlap analysis indicates higher rates of peak sharing within stage compared to between stages.

Sample	Read Count	FRiP%	Peaks
Stage0Rep1	366,854,307	41.8	20,410
Stage0Rep2	527,821,473	46.0	15,202
Stage0Rep3	503,343,465	32.3	11,805
Stage2Rep1	802,828,019	26.1	17,310
Stage2Rep2	566,162,819	31.5	21,402
Stage2Rep3	517,968,173	35.7	19,287
Stage4Rep1	330,988,116	41.7	25,384
Stage4Rep2	612,882,451	37.4	17,533
Stage4Rep3	503,472,847	51.0	27,984

Table 1: **ATAC-seq Data Summary.** Sample names indicate the stage and replicate number. The total number of reads from ATAC-seq is given. Following peak calling using Genrich [3], the fraction of reads in peaks is given as FRiP% along with the total number of peaks for each sample.

ATAC-seq TSS Enrichment Analysis

High quality ATAC-seq datasets show very strong read enrichment at the transcription start sites (TSS) of transcribed genes. We used TSS defined using two approaches to evaluate our ATAC-seq data (Figure 2). Using TSS derived by mapping a *de novo* transcriptome [4] to the available reference assembly [5], we demonstrate moderate levels of read enrichment. This suggests generation of the first known chromatin accessibility dataset from a crustacean has been successful.

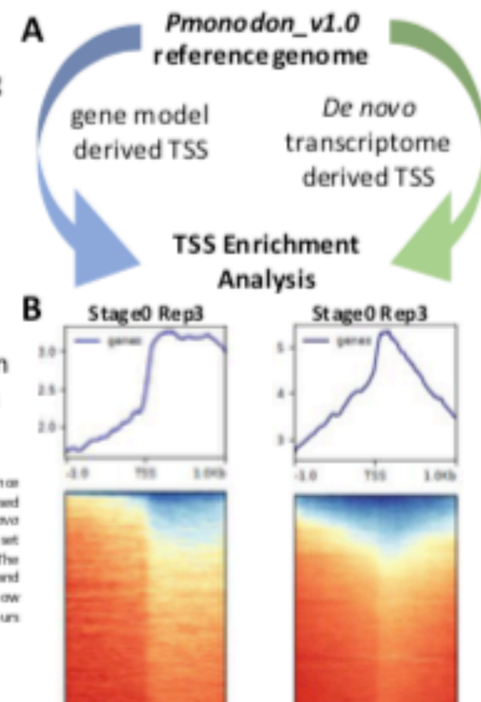

Figure 2: **TSS Enrichment Analysis. (A)** We used the draft *Pmonodon_v1.0* reference assembly (currently unpublished [5]) and defined TSS using two approaches. First, we used the associated gene models to define TSS (blue arrow). Secondly, we exploited a *de novo* multi-tissue transcriptome [4]. This was mapped back to the genome to define a second set of TSS (green arrow). Both TSS sets were then used to evaluate read enrichment **(B)**. The line diagram shows the distribution of genes as a function of bp distance from TSS and should take a sharp bell shape for high quality gene models. The associated heatmaps show the extent of read enrichment immediately surrounding the TSS. Increasingly cool colours (trending from red to blue) indicate increasingly strong enrichment.

Conclusions

1. We have generated data suitable for functional annotation of the Black Tiger Prawn.
2. Preliminary assessment of the ATAC-seq data shows it has moderate to high quality.
3. Our future goal is to better understand the genes driving maturation and their regulatory elements.

As Australia's national science agency and innovation catalyst, CSIRO is solving the greatest challenges through innovative science and technology.

CSIRO. Unlocking a better future for everyone.

FOR FURTHER INFORMATION
Dr James Kijas
E James.Kijas@csiro.au
T +61 7 3214 2442

REFERENCES
[1] Buenrostro et al. (2013) *Nat Methods* 10: 12138.
[2] Deenam et al. (2018) *Repro Res* 156: 527-544
[3] Genrich available at https://github.com/jsh58/Genrich
[4] Huerlimann et al (2018) *Sci Reports* 8: 13553.
[5] https://www.jcu.edu.au/prawn-breeding-hub.

ACKNOWLEDGEMENTS
We thank the ARC Research Hub for Advanced Prawn Breeding for providing prepublication access to the *Pmonodon_v1.0* reference genome assembly. This work was conducted using CSIRO strategic funds.