1. **Title**

Genetic diversity of interferon genes in platypus (Ornithorhynchus anatinus) populations across Australia

1. **Running head**

Genetic Diversity in the Platypus

1. **Student name**

Amy Maria Garber

1. **Abstract**

The platypus is one of Australia’s unique mammals with distinctive features, as well as being one of five remaining egg-laying species. It is endemic to eastern Australia and currently faces population concerns. Previous studies have demonstrated that large dams limit gene flow which as a result, reduce genetic variation. This can impact their survival and reproductive fitness. The interferon gene, essential for immune response, plays a crucial role in maintaining health and defence against diseases. This study aims to characterise genetic variation of the interferon genes across the platypus range, allowing for insights into adaptive potential and may guide conservation efforts. This research utilises whole genome sequencing data processed in previous studies from multiple databases and utilises a variety of analytical methods. Data analysis for this project involves the identification of interferon genes on specific chromosomes, followed by statistical tests to evaluate heterozygosity differences and estimate other genetic diversity indexes. The results have implications for understanding the platypus genetic variation and may guide conservation strategies.

1. **Keywords**

Platypus; IFN / ’Interferon gene’, ‘Genetic diversity’

**Progress report**

1. **Introduction**

The platypus (*Ornithorhynchus anatinus*) in the subclass Monotremata, is the only living species within the family Ornithorhynchidae (Bino et al., 2019; Grant & Fanning, 2008; O'Brien, 2008; Warren et al., 2008). They are regarded as one of the world’s most evolutionary distinct mammals and is one of five extant species of egg-laying mammals (Bino et al., 2019; Grant & Fanning, 2008; O'Brien, 2008; Warren et al., 2008). Plarypuses have a number of unique features, including venomous spurs in males, multiple sex chromosomes (n=10), a biofluorescent coat and electroreception units in their bills whiach are used to hunt freshwater macroinvertebrates (Martin et al., 2018; Jose L. Mijangos et al., 2022). The platypus is endemic to eastern mainland Australia, Tasmania and King Island, although there is a no native (i.e. introduced) small population on Kangaroo Island in South Australia (Bino et al., 2019; Warren et al., 2008). Currently, the platypus is listed as ‘Near Threatened’ by the International Union for Conservation of Nature (IUCN) with a decreasing population trend (Woinarski & Burbidge, 2016), and listed as ‘endangered’ in South Australia. However, the platypus is not currently listed as a threatened species under Australia’s Environmental Protection and Biodiversity Conservation Act 1999 (Hawke et al., 2020). Due to the unique nature of this species, there is need for national level risk assessment for the platypus, allowing for understanding of the conservation status, and to be able to prioritise their conservation efforts (Bino et al., 2020).

Previous studies looking into the platypus genome have found that there is four to 20 fold increase in genetic variation across platypus populations when compared to within populations (Jose L. Mijangos et al., 2022). In another study, results indicated no recent gene flow had occurred between locations (Martin et al., 2018). This study provided evidence indicating bottlenecking and long-term population declines (Martin et al., 2018). Hence, it is demonstrated that there is limited platypus dispersal between populations, with dams acting as major barriers (Martin et al., 2018; Jose L. Mijangos et al., 2022). Long term effects of this isolation of populations has had significant impact on gene flow and has reduced local populations sizes (Jose L. Mijangos et al., 2022). Genetic variation is essential for fitness and survival of species (Jose L. Mijangos et al., 2022). The reduced genetic variation, as well as reduced populations sizes, may lead to lower survival rates of individuals and reduced reproductive output, due to inbreeding depression or catastrophic events (Jose L. Mijangos et al., 2022).

Interferons (IFN) has been found to play a vital role in the innate immune system of vertebrates, because they IFN genes can be divided into three groups, type-I, type II and type III IFNs (Hughes, 1995). Within these divisions, there are many subgroups of IFNs present in mammals (Hughes, 1995).

Paragraph about how genetic diversity can be measured, see:

Box 3 in <https://doi.org/10.1016/j.tree.2017.09.012>

It is expected that, as found in other genetic studies (references), there will be higher variations of genetic diversity across populations than within. Results from this study provide insights about the adaptive potential of populations which can be used to guide future conservation efforts. My findings might help to develop targeted actions, such as...

1. **Material and Methods**

**Data collection**

In this study, I used two different whole genome sequencing (WGS) datasets. The first dataset, hereafter the “dam dataset”, comprises 26 platypus individuals collected from Tenterfield Creek (n = 11), below the dam (n = 8) and in Severn River above the dam (n = 7) (Jose L. Mijangos et al., 2022). The second dataset, here after the “whole range dataset”, was collected across the whole species range in eastern mainland Australia and Tasmania, with a total of 57 individuals (Martin et al., 2018).

**Laboratory work**

Summarise the lab methods from Martin 2018 and Mijangos 2022.

The dam dataset…

The whole range dataset…

**Bioinformatic pipeline**

The study made use of refreshed and updated reference genomes (Zhou et al., 2021), as well as original programs for read mapping using NextGenMap (Sedlazeck et al., 2013), genotype calling making us of Octopus (Cooke et al., 2021) and

The location of interferon genes were retrieved from…

Interferon genes analised in this study were…

Create a table with information of the IFN genes like location, function, etc

Identification of genes was completed using a variety of methods, including PCA analysis (Price et al., 2006), identification of FST outliers (Brauer et al., 2018), comparative genomics making use of the program Genespace ("GENESPACE: syntenic pan-genome annotations for eukaryotes," 2022) and genotype-environment associated studies (Brauer et al., 2018).

Following the study, the data was transferred to and stores in the Research Data Store (RDS) service at the University of Sydney. Platypus data is stored in Gadi, a supercomputer run by the National Computational Infrastructure (NCL).

Additional data available to use includes 214 platypus samples collected from nine different rivers across four nations in south-east Australia.

Furthermore, the 57 platypus samples collected from the whole species range in eastern mainland Australia and Tasmania have been sequenced in two different formats. The first is whole genome sequencing, and the second using methods described above (NextGenMap, Octopus), and have been mapped to the reference genome produced by Zhou *et al.* (2021).

**Data analysis**

We estimated genetic diversity in the IFN genes using three different methods, namely allelic richness (q = 0), Shannon information (q = 1), and heterozygosity (q = 2). These three measures whose contrasting properties provide a rich summary of diversity .

Estimates of genetic diversity were performed using the function gl.report.diversity from the R package dartR (Jose Luis Mijangos et al., 2022).

To tested whether genetic diversity estimates were significantly different between populations using re-randomization as implemented in the dartR function gl.test.heterozygosity.

1. **Progress**

Overall, I have achieved the objectives I set for myself up until this point. My goal for the first semester was to have my data analysis completed by the end of the break between semesters. This semester has been focussed on learning how to use the R Studio program required to analyse the genetic data as well as background research on the platypus and its genetics. I have researched and read through relevant literature available to further deepen my understanding of not just the platypus, but genetic studies that have previously been completed. Additionally, I have made a start on the statistical analysis using R Studio. I have downloaded all the necessary extensions and ensured they are functional. I have identified chromosomes that the interferon gene is present on within the platypus genome. The chromosome X1 has bene uploaded into Studio R and data analysis has begun.

My research project to date has consisted of regular meetings with my supervisor, Jaime Gongora, as well as weekly meetings with Jose Luis Mijangos who has worked with Jaime on previous platypus genetic research. My initial draft of the introduction and methods have been sent to both Jaime and Luis for review and feedback. I, with the help of my supervisors, have set up a clear plan and timeline for the rest of the year. This will help to keep me on track with my research project.

1. **Plan**

* To have read over feedback and comments provided on my draft introduction and methods by both Jaime and Luis, and make appropriate adjustments by 30/06/2023.
* Use the first few weeks of the holidays to complete statistical analysis, and prepare questions to ask for my next meeting with Jaime and Luis. I plan to have statistical analysis finalised by 23/07/2023.
* I plan to have a skeleton version of my manuscript completed by this date as well (23/07/2023).
* My final draft of my discussion and results will be completed by 31/07/2023.
* This ensures I have sufficient time to complete the draft manuscript submission by the due date (29/09/2023).
* Finally, my final manuscript will be completed by the due date (13/10/2023).

1. **Problems**

The biggest challenges to date have been becoming re-familiar with R studio. Although I have used the program before during my previous studies, this was a few years ago and I had to re-teach myself how to use the different screens and commands. Learning to use the dartR program has presented as an additional challenge. At the start of the project, I thought I had a deep enough understanding of R Studio, however as I have never used the dartR package before, this presented itself as an additional hurdle to overcome.

Another challenge I have found is ensuring I allocate sufficient time to work on the professionally focussed project during semester, while juggling all of my other subjects, assignments and examinations. As our third year has been the busiest and most challenging so far, I have found myself placing the project aside to focus on more urgent studies. I have managed to overcome this, however, by ensuring I have weekly / bi-weekly meetings with Luis, and by ensuring I have kept Jaime up to date on the work being completed.

1. **Assessment**

Across the rest of the year I will strive to continue to meet the objectives I have set out for myself. My ‘Progress’ and ‘Plan’ across time has been demonstrated by the use of a timeline image, as in Figure 1. This timeline has been agreed upon by myself, Jaime and Luis to ensure I remain on track. If this timeline is not possible, I will notify my advisor promptly to seek advice on the best course of action to get myself back on schedule. Ensuring these steps are taken will assist me in submitting a satisfactory manuscript by the due date next semester.

1. **Budget and request for funding**

At this time, there is no budget nor request for funding for my project. This is due to the fact that the data has been previously collected and processed.

1. **Tables**

NA

1. **Figures**

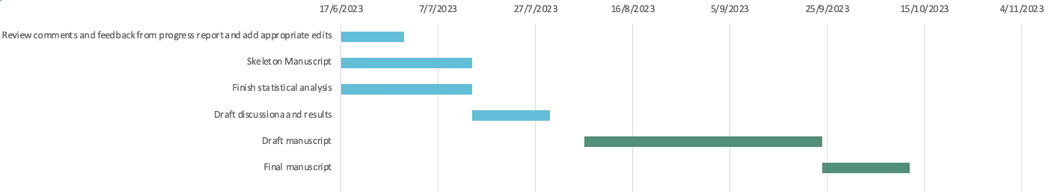


Figure 1. Timeline figure showing my planned progress on my professionally focussed research project across time. On the x-axis are the steps necessary to completing my manuscript, plotted against time (dd/mm/yyyy) on the Y-axis.

**References**

Bino, G., Kingsford, R. T., Archer, M., Connolly, J. H., Day, J., Dias, K., Goldney, D., Gongora, J., Grant, T., Griffiths, J., Hawke, T., Klamt, M., Lunney, D., Mijangos, L., Munks, S., Sherwin, W., Serena, M., Temple-Smith, P., Thomas, J., . . . Whittington, C. (2019). The platypus: evolutionary history, biology, and an uncertain future. *Journal of Mammalogy*, *100*(2), 308-327. <https://doi.org/10.1093/jmammal/gyz058>

Bino, G., Kingsford, R. T., & Wintle, B. A. (2020). A stitch in time – Synergistic impacts to platypus metapopulation extinction risk. *Biological Conservation*, *242*, 108399. <https://doi.org/10.1016/j.biocon.2019.108399>

Brauer, C. J., Unmack, P. J., Smith, S., Bernatchez, L., & Beheregaray, L. B. (2018). On the roles of landscape heterogeneity and environmental variation in determining population genomic structure in a dendritic system. *Molecular ecology*, *27*(17), 3484-3497. <https://doi.org/10.1111/mec.14808>

Cooke, D. P., Wedge, D. C., & Lunter, G. (2021). A unified haplotype-based method for accurate and comprehensive variant calling. *Nature biotechnology*, *39*(7), 885-892. <https://doi.org/10.1038/s41587-021-00861-3>

GENESPACE: syntenic pan-genome annotations for eukaryotes. (2022). 1917.

Grant, T., & Fanning, I. b. D. (2008). *Platypus*. CSIRO Publishing.

Harrison, G. A., McNicol, K. A., & Deane, E. M. (2004). Type I interferon genes from the egg‐laying mammal, Tachyglossus aculeatus (short‐beaked echidna). *Immunology and cell biology*, *82*(2), 112-118. <https://doi.org/10.1046/j.0818-9641.2004.01230.x>

Harrison, G. A., Young, L. J., Watson, C. M., Miska, K. B., Miller, R. D., & Deane, E. M. (2003). A survey of type I interferons from a marsupial and monotreme: implications for the evolution of the type I interferon gene family in mammals. *Cytokine (Philadelphia, Pa.)*, *21*(3), 105-119. <https://doi.org/10.1016/S1043-4666(03)00029-2>

Hawke, T., Bino, G., & Kingsford, R. (2020). *A national assessment of the conservation status of the platypus*. Australian Conservation Foundation. <https://trove.nla.gov.au/work/253398774>

Hughes, A. L. (1995). The evolution of the type I interferon gene family in mammals. *Journal of Molecular Evolution*, *41*(5), 539-548. <https://doi.org/10.1007/BF00175811>

Martin, H. C., Batty, E. M., Hussin, J., Westall, P., Daish, T., Kolomyjec, S., Piazza, P., Bowden, R., Hawkins, M., Grant, T., Moritz, C., Grutzner, F., Gongora, J., & Donnelly, P. (2018). Insights into Platypus Population Structure and History from Whole-Genome Sequencing. *Molecular biology and evolution*, *35*(5), 1238-1252. <https://doi.org/10.1093/molbev/msy041>

Mijangos, J. L., Bino, G., Hawke, T., Kolomyjec, S. H., Kingsford, R. T., Sidhu, H., Grant, T., Day, J., Dias, K. N., Gongora, J., & Sherwin, W. B. (2022). Fragmentation by major dams and implications for the future viability of platypus populations. *Communications biology*, *5*(1), 1127-1127. <https://doi.org/10.1038/s42003-022-04038-9>

Mijangos, J. L., Gruber, B., Berry, O., Pacioni, C., & Georges, A. (2022). dartR v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods in ecology and evolution*, *13*(10), 2150-2158. <https://doi.org/10.1111/2041-210X.13918>

O'Brien, S. J. (2008). The Platypus Genome Unraveled. *Cell*, *133*(6), 953-955. <https://doi.org/10.1016/j.cell.2008.05.038>

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics*, *38*(8), 904-909. <https://doi.org/10.1038/ng1847>

Sedlazeck, F. J., Rescheneder, P., & von Haeseler, A. (2013). NextGenMap: fast and accurate read mapping in highly polymorphic genomes. *Bioinformatics*, *29*(21), 2790-2791. <https://doi.org/10.1093/bioinformatics/btt468>

Sherwin, W. B., Chao, A., Jost, L., & Smouse, P. E. (2021). Information theory broadens the spectrum of molecular ecology and evolution: (Trends in Ecology and Evolution 32:12, p:948–963, 2017). *Trends in ecology & evolution (Amsterdam)*, *36*(10), 955-956. <https://doi.org/10.1016/j.tree.2021.07.005>

Warren, W. C., Hillier, L. W., Yang, S.-P., Heger, A., Locke, D. P., Miethke, P., Waters, P. D., Veyrunes, F., Fulton, L., Fulton, B., Graves, T., Wallis, J., Marshall Graves, J. A., Puente, X. S., Lopez-Otin, C., Ordonez, G. R., Eichler, E. E., Lin, C., Ze, C., . . . Chinwalla, A. T. (2008). Genome analysis of the platypus reveals unique signatures of evolution. *Nature*, *453*(7192), 175-183. <https://doi.org/10.1038/nature06936>

Woinarski, J., & Burbidge, A. (2016). *Ornithorhynchus anatinus*. The IUCN Red List of Threatened Species 2016: e.T40488A21964009. Retrieved 20/05/2023 from <https://www.iucnredlist.org/species/40488/21964009>

Zhou, Y., Shearwin-Whyatt, L., Li, J., Song, Z., Hayakawa, T., Stevens, D., Fenelon, J. C., Peel, E., Cheng, Y., Pajpach, F., Bradley, N., Suzuki, H., Nikaido, M., Damas, J., Daish, T., Perry, T., Zhu, Z., Geng, Y., Rhie, A., . . . Zhang, G. (2021). Platypus and echidna genomes reveal mammalian biology and evolution. *Nature (London)*, *592*(7856), 756-762. <https://doi.org/10.1038/s41586-020-03039-0>