Individual Locus Effects on Mortality Risk in the Endangered St. Lawrence Beluga Whale

Sofia Day, Supervised by: Dr. Tim Frasier

Introduction

SLE Belugas (*Delphinapterus* l*eucas)*
- Habitat: St. Lawrence Estuary and River (Figure 1)
- Listed as Endangered by COSEWIC in 2014[2]

(GREMM 2019)

Mortality
- Initial population declines were a result of hunting practices (ended in 1979) and high cancer prevalence was observed in the mid-late 20th century[2].
- Despite protection from hunting and reduced environmental contamination, the population has not yet recovered[2].

What underlying factors may be contributing to high mortality rates?
- One hypothesis: inbreeding and/or low genetic diversity
- Recent work in the Frasier Lab found no correlation between inbreeding coefficients and mortality risk, and no signs of recent inbreeding within the population[3].
- However, specific regions of the genome were homozygous for deleterious alleles in neonates, particularly in regions associated with processing environmental contaminants[3].

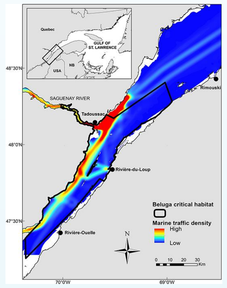

Figure 1. SLE beluga habitat (Polar Research 2021).

Research question & hypothesis

- Are specific regions of the genome associated with different causes of mortality?
- Aim to identify allelic and genotypic differences across the genome for individuals that died of different causes.

Methods

Data collection and sample selection
- Completed by previous collaborators
- Phenotype mortality data obtained from necropsies (1983-2022)
- Genomic data from preserved blood and tissue DNA extracts, sequenced using ddRADseq protocols

Data filtering
- Missing genotype calls, Hardy-Weinberg equilibrium, minor allele frequency, and relatedness
- Remaining dataset: 143 individuals and 15,452 single nucleotide polymorphisms (SNPs)

Analysis and visualization

1) Genome-wide association study (GWAS) analyses → logistic regressions using a GLM within PLINK
- Four causes of death: infectious disease, dystocia, neonate mortality, and neoplasia
- Two analyses for each death: allelic comparisons and genotypic comparisons

2) Examine the data set for epistatic interactions
- module clustering using the Weighted Gene Co-expression Network Analysis (WGCNA) R package and GWAS test results

Results

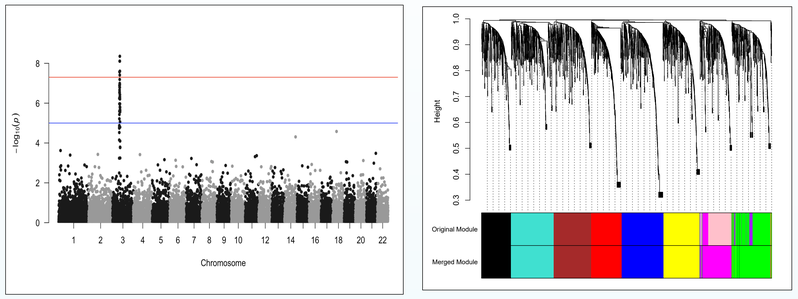

Figure 2. Example of a Manhattan plot (left - R graph gallery 2025) and a gene clustering dendrogram (right - Orton et al. *in review*), demonstrating ideal results with identifiable allelic associations to a trait and unique gene modules.

GWAS analyses – Manhattan plot

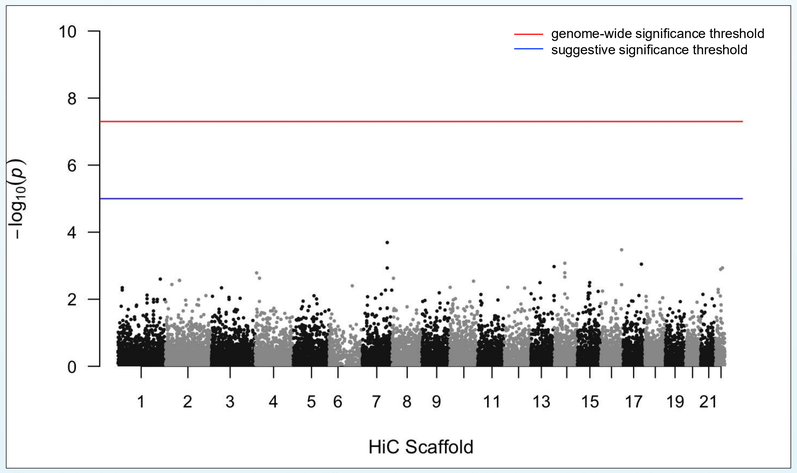

Figure 3. Logistic regression comparing allelic differences between individuals that died due to infectious disease (N = 42) and all other samples (N = 101).

Epistatic interactions – Dendrogram

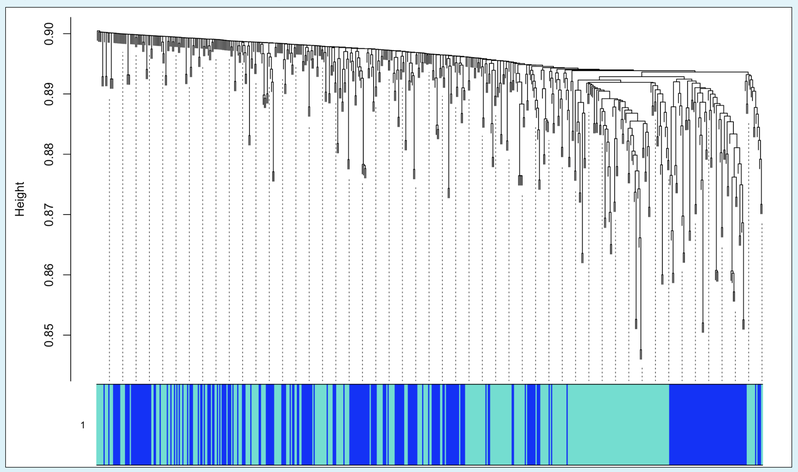

Figure 4. Clustering of epistatic gene modules associated with infection-related mortality, calculated using p-values from prior GWAS tests.

Discussion

1. GWAS analyses revealed no significant allelic or genotypic association to any of the four mortality categories (all p-values > $1x10^{-5}$) (Figure 3). Results suggest that the characteristics at any specific locus are not correlated with mortality risk.

2. Further analyses aiming to identify epistatic interactions revealed no significant gene module clustering (Figure 4), suggesting either a relatively homogeneous dataset or independent gene effects, with no epistatic interactions associated with the analyzed causes of death.

3. Although ddRADseq has proven to be a viable sequencing method for association analyses, other literature debates the use of this genomic sequencing approach given its reduced representation of the genome, conceivably resulting in allelic dropout, false positives, or limited power to detect effects[1].

   → Given recent findings of cluster associations in neonates using whole-genome data[3], associations that were not detected in these analyses, further work on this topic would likely require or should heavily consider whole-genome sequencing approaches.

(COSEWIC 2014)

(Lisa Barry)

Conclusion

- Further research is needed to better understand mortality rates in the Endangered St. Lawrence beluga whale population. Particularly as environmental conditions shift, genomics are becoming increasingly valuable in marine conservation efforts.

- While this work provided no evidence of locus-specific or epistatic effects associated with mortality risk from the analyzed causes, the use of ddRADseq methods should be re-examined, and future studies may consider a whole-genome sequencing approach.

Acknowledgments

- Dr. Tim Frasier, for supervision and support throughout my thesis.
- Lab graduate students Audrey and Ricky, for help with coding and data sorting.
- NSERC, for funding via a Discovery Grant to Tim Frasier.
- DFO and GREMM, for field and necropsy data.

References

(1) Atsawawaranunt K et al. 2026. Missing or mis-telling the story? Trade-offs for restriction-site associated compared to whole genome sequencing. Mol Ecol. 35(2):e70211. https://doi.org/10.1111/mec.70211

(2) COSEWIC. 2014. COSEWIC assessment and status report on the beluga whale *Delphinapterus leucas*, St. Lawrence Estuary population, in Canada. Ottawa: Committee on the Status of Endangered Wildlife in Canada. xii + 64 pp.

(3) Orton R, Mueller C, Garroway C, Ferguson SH, Michaud R, Frasier T (*in review*). Homozygosity accumulated through genetic drift and not through recent inbreeding underpins high mutation load in an endangered beluga whale population. Mol Ecol.