An example of DNA methylation as a means to quantify stress in wildlife using killer whales

Carla A. Crossman1*, Lance G. Barrett-Lennard2 & Timothy R. Frasier1

The cumulative effects of non-lethal stressors on the health of biodiversity are a primary concern for conservation, yet difficulties remain regarding their quantification. In mammals, many stressors are processed through a common stress-response pathway, and therefore epigenetic changes in genes of this pathway may provide a powerful tool for quantifying cumulative effects. As a preliminary assessment of this approach, we investigated epigenetic manifestations of stress in two killer whale populations with different levels of exposure to anthropogenic stressors. We used bisulfite amplicon sequencing to compare patterns of DNA methylation at 25 CpG sites found in three genes involved in stress response and identified large differences in the level of methylation at two sites consistent with differential stress exposure between Northern and Southern Resident killer whale populations. DNA methylation patterns could therefore represent a useful method to assess the cumulative effects of non-lethal stressors in wildlife.

Although lethal threats such as directed hunting or incidental mortality (e.g., due to habitat loss or accidental mortality such as bycatch) represent the largest threats to global biodiversity1-3, there is increasing recognition that non-lethal threats are compromising the viability of many populations, particularly when experienced in combination and/or when they have synergistic effects4-6. For example, the global colony collapse of honey bees and the global declines in amphibian populations have both been attributed to the combined effects of multiple non-lethal threats7,8. As a result, there is a growing emphasis on the need to consider the cumulative effects of non-lethal threats in the management and conservation of biodiversity9-11.

Many non-lethal threats—including those from anthropogenic activities such as physical disturbance and contaminant exposure—are thought to increase stress on individual animals, with subsequent negative fitness consequences. Despite being well-documented in humans and mice, quantifying the negative fitness consequences of stress in wildlife populations remains a challenge. Stress hormones can be measured in blood, saliva, excrement and hair/feathers12,13; however, collection of these samples is difficult with many species and stress hormones often only capture a snapshot of an animal’s recent stress exposure. Quantifying the cumulative effects of multiple stressors poses an additional obstacle, particularly if their effects are synergistic, as meta-analyses often suggest14.

One approach that may be particularly informative for understanding the cumulative impacts of stressors is the study of methylation patterns15,16, and in particular that of genes involved in stress response. DNA methylation is an epigenetic modification to DNA that occurs primarily at 5'-CpG-3' (cytosine-guanine dinucleotide—CpG) sites17. CpG sites are not evenly distributed throughout the genome, but rather are concentrated in ‘CpG islands’ in the promoter region of many genes where they act as important regulators of gene expression18.

Stress-related changes in DNA methylation have been detected in genes in the hypothalamic pituitary adrenal (HPA) axis22-26—a main pathway regulating stress response in mammals27. The HPA axis is regulated by a suite of genes including the brain derived neurotrophic factor (BDNF) and the corticotropin releasing factor (CRF)28. CRF initiates the stress response pathway that culminates in glucocorticoids (stress hormones) being released29. The cellular responses to these hormones are mediated largely by the glucocorticoid receptor (NR3C1)30. The complexity of the stress response pathway relies on proper functioning of these and other genes and therefore altered methylation patterns could have significant downstream effects. For example, controlled laboratory experiments on mice and rats have identified changes in methylation patterns in response to stressors in genes in the

1 Biology Department, Saint Mary’s University, Halifax, NS, Canada. 2 Coastal Ocean Research Institute, Ocean Wise Conservation Association, Vancouver, BC, Canada. *email: crossman.ca@gmail.com
HPA axis\cite{22-26}, and similar results have been identified in human studies\cite{25,31}. Therefore, analyses of methylation patterns of stress-response genes have the potential to be particularly informative in assessing stress in wildlife populations because these patterns represent the cumulative impacts of the suite of stressors experienced by an individual, and therefore provide a method to quantify their cumulative and/or synergistic effects.

The Southern Resident population of killer whales (\textit{Orcinus orca}) in the Northeast Pacific is an endangered population (N = 74)\cite{32} whose main identified threats are the cumulative impacts of multiple non-lethal stressors (e.g., underwater noise, vessel disturbance, food availability and toxic contaminants)\cite{33}. To the north is the closely related and more robust population of Northern Resident killer whales (N ≈ 300)\cite{34}. Northern and Southern Resident killer whale populations have similar dietary preferences and inhabit adjacent, slightly overlapping ranges\cite{35,36}. The Southern Resident killer whale range has greater overlap with heavily inhabited coastlines and transited waterways compared to Northern Residents and as such the population experiences higher levels of exposure to anthropogenic stressors. We predict that the differential exposure to stressors will manifest themselves as differences in the amount of DNA methylation at CpG sites in genes involved in stress response.

Understanding the combined biological impacts of these non-lethal threats on individual fitness might therefore best be explored by examination of epigenetic differences in genes related to stress. As a step towards this goal, we identified the methylation patterns of stress-related genes in individual Southern Resident killer whales and compared them to those in Northern Resident individuals.

**Results**

Using amplicon sequencing of sodium bisulfite treated DNA extracted from biopsy samples of skin from free swimming killer whales, we identified percent methylation for CpG sites in the promoter region of three genes known to be key players in stress response: \textit{BDNF}, \textit{CRF} and \textit{NR3C1}; as well as \textit{ACTB} (β-actin)—a gene commonly used as a control as it regulates components of the cytoskeleton\cite{37} (Figure S1). We used a Bayesian regression model to tease apart the relative effects of individual, age, sex, population, and CpG site on methylation patterns (Fig. 1). Neither age nor sex had appreciable effects on methylation patterns, with the mean of the posterior distributions falling close to zero (Fig. 2a), confirming that methylation of the targeted CpG sites does not change with age or under the different hormonal exposures experienced across the two sexes. Thus, any differences in methylation patterns should be due to differences in the levels of external stressors. To compare the relative effects of being in a given population on the level of DNA methylation at a given site, we took the difference in the posterior probability distributions at each site between populations. If the population of origin had little to no impact on the amount of methylation, the distribution will centre around zero, whereas a positive distribution would indicate hypermethylation in Southern Residents compared to Northern Residents and a negative distribution would be indicative of hypomethylation. As expected, CpG sites within the promoter of the control gene (\textit{ACTB}) did not show different degrees of methylation across the two populations (Fig. 2b).

Of the genes examined, the CpG sites in the promoter regions of two (\textit{BDNF} and \textit{NR3C1}) showed similar degrees of methylation across populations (Fig. 3a,b). However, site-specific differences in methylation patterns were seen in the promoter region of \textit{CRF} (Fig. 3c). The largest differences were in position −101 that is hypermethylated in Southern Resident individuals relative to the Northern Residents and in position −95 that is hypomethylated in Southern relative to North Residents (Fig. 3c).
Discussion

Methylation patterns at stress response genes differed between Northern and Southern Resident killer whales that could not be explained by age or sex and are likely due to differences in stress exposure. Many studies have demonstrated that methylation patterns at some CpG sites change with age in very predictable ways and can be used in age-predictive models. We chose a paired sample approach to help elucidate the possible effect of population, while also accounting for potential effects of age and sex. Our results suggest the CpG sites in the genes targeted in our study are not affected by age or sex, but rather methylation difference reflect differences in the environments experienced by Resident killer whales.

While genetically distinct, these two populations of killer whales are remarkably similar. They have the same dietary preferences, adjacent but overlapping distributions, and very similar physical appearances and acoustic behaviors. One of the primary differences in the environments experienced by these two populations is the degree of exposure to anthropogenic stress. Southern Residents spend much of the year in the interior waters near the busy port cities of Vancouver and Victoria, Canada; and Seattle, USA, and therefore, the whales in this area are exposed to high levels of pollution, heavy vessel traffic from recreational boaters, a large commercial whale watching fleet, and large cargo vessels causing both acoustic and physical disturbance. In contrast, Northern Residents inhabit quieter, more pristine waters to the north with less commercial and recreational vessel presence.

The patterns we identified in CRF (hypermethylation at position −101 and hypomethylation at position −95) have also been observed in some tissues of rats exposed to chronic variable mild stress and since the CRF promoter region in killer whales has over 95% similarity with other mammalian species, this suggests that they serve the same functions across species. Additionally, altered methylation patterns in response to stress have previously been identified at position −101 in other species, but the direction of these methylation changes has differed between tissue types and studies.

Only one of these two sites (position −101) falls within a previously identified transcription factor binding region: that for the aryl hydrocarbon receptor (Ahr). The Ahr binding region (position −101) is particularly interesting because, in addition to being involved in stress response, aryl hydrocarbon receptors are also involved in the metabolism of polycyclic aromatic hydrocarbons (PAHs), contaminants to which killer whales are frequently exposed but which do not tend to bioaccumulate and exposure is therefore difficult to assess.

Current practices for understanding the cumulative effects of multiple stressors on individuals rely on models that incorporate estimated effects of independent stressors and make assumptions about how these stressors interact and are experienced by the individual. The potential to use changes in the amount of methylation at...
stress-related genes within a population over time or between populations, may therefore represent a stronger approach for understanding the true biological impact and fitness consequences of multiple stressors experienced by individuals. This study represents the first step in demonstrating the application of this method for a mammalian population, which may prove to be an effective long term monitoring tool for wildlife populations.

For Southern Residents, the biological implications of these changes in methylation are yet to be determined. Further investigations into the biological or physiological implications of these methylation differences are needed, but our results suggest that long-term monitoring of levels of methylation in stress-related genes may provide a useful tool for quantifying the synergistic effects of multiple stressors, and help track the effectiveness of threat mitigation efforts on reducing stress on individual killer whales.

The cumulative and synergistic effects of anthropogenic-induced stress are likely influencing a wide range of wildlife populations, but quantifying such effects has not yet been possible. Assessing DNA methylation of genes involved in stress response may represent a promising option, as suggested by the application of this approach to critically endangered Southern Resident killer whales. Although we do not yet know the downstream biological implications of the differential methylation patterns observed, this represents a crucial first step in making such quantification possible.

Methods

Library preparation and sequencing. DNA samples for this study are part of a long-term tissue collection not collected explicitly for this study. Skin samples used were collected from free-swimming whales for previous work using a pneumatic dart system, except for one sample that was collected from a freshly deceased carcass. During biopsy sampling, biopsy darts collect a small piece of skin from the dorsal surface of the animal, near its dorsal fin. This biopsy method is consistent between populations. A paired sample approach was implemented when selecting which of the existing samples to use so that for each Southern Resident sample with sufficient high-quality DNA (n = 17), one or two Northern Residents of the same sex and similar age were included (n = 30) (Fig. 4).
The DNA was previously extracted using standard phenol:chloroform methods\(^4\) using the skin adjacent to the skin-blubber interface of each sample. \(400\) ng of DNA from each individual was subjected to sodium bisulfite conversion using Epitect Bisulfite Kits (Qiagen). This process converted unmethylated cytosines (C) in CpG sites to uracil (U), whereas methylated cytosines in CpG sites remained C.

We identified three genes in the HPA axis and its regulation that had demonstrated methylation changes in response to stressors in controlled studies of laboratory animals: brain derived neurotropic factor (BDNF), corticotropin releasing factor (CRF) and glucocorticoid receptor (NR3C1) as well as a common control gene, \(\beta\)-actin (ACTB) whose methylation patterns should not be influenced by age, sex, or exposure to external stressors\(^3\), rather interpopulation differences here could represent genetic differences in methylation/demethylation enzymes. We identified reference sequences for the promoter region of each of these genes from the published killer whale genome (GenBank accession GCA_000331955.1). We did in-silico sodium bisulfite conversion of our reference sequences and designed primers accounting for the DNA being single stranded and non-complementary post bisulfite treatment, avoiding CpG sites within the primer sequence if possible (see supplementary material). We performed Sanger sequencing to test that: (i) we had amplified the correct region, and (ii) the resulting sequences would be absent of cytosine except in potentially methylated sites.

Each bisulfite-treated sample was amplified using each primer pair. Our subsequent library preparation protocol was a modified version of that available for 16S sequencing\(^5\).

We cleaned our PCR products with Ampure XP beads using bead:sample ratios optimized for each primer set and final elution volumes adjusted for each sample as determined by concentration on a pre-cleanup agarose gel. We performed an indexing PCR to bind sample-specific indexes using the Nextera xt Index kit (Illumina) and cleaned the products. The final libraries were quantified, pooled and sequenced on an Illumina MiSeq 250 × 250 PE run at the BRC Sequencing Core at the University of British Columbia (Vancouver, Canada).

### Processing amplicon sequences.

Each index was unique to a single individual but included amplicons from multiple loci. Separating loci and trimming primer sequences was performed in CUTADAPT 2.6 with an allowable error rate of 0.2\(^4\). Leading and trailing bases with a Q < 20 were trimmed and reads were trimmed when the mean quality in a 5-base sliding window fell below \(Q = 30\) in TRIMMOMATIC 0.36\(^6\). Paired reads were merged using FLASH 1.2.11\(^1\). Aligned and indexed bam files were compiled using BWA 0.7.17-r1188 and SAMTOOLS 1.7\(^2\),\(^3\) and aligned against the reference genes used to design our primers.

Despite normalizing our library concentrations prior to sequencing, we had differential amplification and sequencing success of each locus resulting in an uneven distribution of reads across loci. As a result, we determined minimum read depth thresholds independently for each locus. Our goal was to identify percent methylation at each site and therefore wanted to optimize read depth to increase confidence in our measures of percent methylation while still maintaining relatively consistent certainty within each locus. We visualized read depths using histograms of read depth across samples and determined the minimum threshold for each locus as: ACTB—10 000, BDNF—20 000, CRF—100, NR3C1—500 (table S2 and figure S3). Individuals were excluded from the analysis if a minimum read depth threshold was not met. Due to low read depths of merged pair-end reads for NR3C1, forward unpaired reads were also trimmed and included in the aligned bam file to increase read depth.

### Analyses.

Using the reference sequence for each locus, we identified potentially methylated sites (CpG sites). Methylated CpG sites in our reads would have escaped bisulfite conversion and remained as C bases whereas
unmethylated CpG sites would be represented by Ts. For every individual at each gene, we calculated the proportion of C/T base calls as a metric of percent methylation at each CpG site.

We built a Bayesian regression model using the percent methylation of each site as the predicted variable where individual, age, sex, population and CpG site were the predictor variables, as well as the interaction between specific CpG sites and population. We also allowed for different standard deviations in each population. We built similar models investigating other pairwise interactions and none had an effect on our results, so we left them out of our final model. (See supplementary material for additional details of the model and its performance).

To better visualize and understand the differences between populations, we plotted the difference in the posterior probabilities of percent methylation at each site calculated as that from Northern Residents subtracted from that for the Southern Residents.

**Approval and ethics statement.** DNA samples for this study are part of a long-term tissue collection not collected explicitly for this study. Therefore, this study did not require any interaction with live animals. All biopsies were collected under relevant guidelines and regulations set out by permits issued by Fisheries and Oceans Canada and/or the US National Marine Fisheries Service. Sample collection was carried out by a number of individuals representing multiple institutions. All biopsies were collected with permission and approval from their respective ethical committee or institutional review boards. The collection of most samples was performed under approval of the University of British Columbia and their Animal Care and Use Program.

**Code availability**

Code and methylation data are available at [https://github.com/carlacrossman/StressMethylation](https://github.com/carlacrossman/StressMethylation).

Received: 2 February 2021; Accepted: 29 July 2021
Published online: 19 August 2021

**References**

1. Schipper, J. et al. The status of the world’s land and marine mammals: Diversity, threat, and knowledge. *Science* **322**, 225–230 (2008).
2. Cushman, S. A. Effects of habitat loss and fragmentation on amphibians: A review and prospectus. *Biol. Conserv.* **128**, 231–240 (2006).
3. Díaz, S. et al. Pervasive human-driven decline of life on Earth points to the need for transformative change. *Science* **366**, 6471 (2019).
4. Geary, W. L., Nimmo, D. G., Doherty, T. S., Ritchie, E. G. & Tulloch, A. I. Threat webs: Reframing the co-occurrence and interactions of threats to biodiversity. *J. Appl. Ecol.* **56**, 1992–1997 (2019).
5. Halpern, B. S. et al. A global map of human impact on marine ecosystems. *Science* **319**, 948–952 (2008).
6. Johnson, C. N., et al. Biodiversity losses and conservation responses in the Anthropocene. *Science* **356**, 270–275 (2017).
7. Di Prisco, G. et al. A mutualistic symbiosis between a parasitic mite and a pathogenic virus undermines honey bee immunity and health. *Proc. Natl. Acad. Sci. USA.* **113**, 3203–3208 (2016).
8. Grant, E. H. C. et al. Identifying management-relevant research priorities for responding to disease-associated amphibian declines. *Glob. Ecol. Conserv.* **6**, 00441 (2018).
9. Schindler, D. W. The cumulative effects of climate warming and other human stresses on Canadian freshwaters in the new millennium. *Can. J. Fish. Aquat. Sci.* **58**, 18–29 (2001).
10. Geary, W. L., Nimmo, D. G., Doherty, T. S., Ritchie, E. G. & Tulloch, A. I. Threat webs: Reframing the co-occurrence and interactions of threats to biodiversity. *J. Appl. Ecol.* **56**, 1992–1997 (2019).
11. Frid, A. & Dill, L. Human-caused disturbance stimuli as a form of predation risk. *Conserv. Ecol.* **6**, 11 (2002).
12. Rolland, R. M., Hunt, K. E., Kraus, S. D. & Wasser, S. K. Assessing reproductive status of right whales (*Eubalaena glacialis*) using fecal hormone metabolites. *Gen. Comp. Endocrinol.* **142**, 308–317 (2005).
13. Sheriff, M. J., Dantzer, B., Delehanty, B., Palme, R. & Boonstra, R. Measuring stress in wildlife: Techniques for quantifying glucocorticoids. *Oecologia* **166**, 869–887 (2011).
14. Crain, C. M., Kroeker, K. & Halpern, B. S. Interactive and cumulative effects of multiple human stressors in marine systems. *Ecol. Lett.* **11**, 1304–1315 (2008).
15. Beal, A., Rodriguez-Casariego, J., Rivera-Casas, C., Suarez-Ulloa, V. & Eirin-Lopez, J. M. Environmental epigenetics and its applications in marine organisms. In *Population Genomics: Marine Organisms* (eds. Oleksiak, M. F. & Rajora, O. P.) 325–359. [https://doi.org/10.1007/13836_2018_28] (Springer, 2018).
16. Eirin-Lopez, J. M. & Putnam, H. M. Marine environmental epigenetics. *Ann. Rev. Mar. Sci.* **11**, 335–368 (2019).
17. Laird, P. W. The power and the promise of DNA methylation markers. *Nat. Rev. Cancer* **3**, 253–266 (2003).
18. Bird, A. P. CpG-rich islands and the function of DNA methylation. *Nat. Rev. Genet.* **3**, 295–304 (2002).
19. Matosin, N., Cruceanu, C. & Binder, E. B. Preclinical and clinical evidence of DNA methylation changes in response to trauma and chronic stress. *Chronic Stress* **1**, 247054701771076 (2017).
20. Radtke, K. M. et al. Transregenerational impact of intimate partner violence on methylation in the promoter of the glucocorticoid receptor. *Transl. Psychiatry* **1**, e21–e26 (2011).
21. Muller, B. R. & Bale, T. L. Sex-specific programming of offspring emotionality after stress early in pregnancy. *J. Neurosci.* **28**, 9053–9065 (2008).
22. Elliott, E., Ezra-Nevo, G., Regev, L., Neufeld-Cohen, A. & Chen, A. Resilience to social stress coincides with functional DNA methylation of the Crf gene in adult mice. *Nat. Neurosci.* **13**, 1351–1353 (2010).
23. Boersma, G. J. et al. Prenatal stress decreases Bdnf expression and increases methylation of Bdnf exon IV in rats. *Epigenetics* **9**, 437–447 (2013).
24. Turecki, G. & Meaney, M. J. Effects of the social environment and stress on glucocorticoid receptor gene methylation: A systematic review. *Biol. Psychiatry* **79**, 87–96 (2016).
25. Sterrenburg, L. et al. Chronic stress induces sex-specific alterations in methylation and expression of corticotropin-releasing factor gene in the rat. *PLoS ONE* **6**, 1–14 (2011).
26. Reeder, D. A. M. & Kramer, K. M. Stress in free-ranging mammals: Integrating physiology, ecology, and natural history. *J. Mammal.* **86**, 225–235 (2005).
28. Jeanneveau, F. D. et al. BDNF and glucocorticoids regulate corticotrophin-releasing hormone (CRH) homeostasis in the hypothalamus. Proc. Natl. Acad. Sci. USA. 109, 1303–1310 (2012).
29. Smith, S. M. & Vale, W. W. The role of the hypothalamic-pituitary-adrenal axis in neuroendocrine responses to stress. Dialogues Clin. Neurosci. 8, 383–395 (2006).
30. Turner, J. D. & Muller, C. P. Structure of the glucocorticoid receptor (NR3C1) gene 5’ untranslated region: Identification, and tissue distribution of multiple new human exon 1. J. Mol. Endocrinol. 35, 283–292 (2005).
31. Bukusic, J., Schaufeli, W., Claes, S. & Godderis, L. Stress, burnout and depression: A systematic review on DNA methylation mechanisms. J. Psychosom. Res. 92, 34–44 (2017).
32. Center for Whale Research. Population. https://www.whaleresearch.com. Accessed 11 Jan 2021 (2020).
33. Fisheries and Oceans Canada. Recovery Strategy for the Northern and Southern Resident Killer Whales (Orcinus orca) in Canada (Proposed). Species at Risk Act Recovery Strategy Series, Fisheries & Oceans Canada, Ottawa, x + 84 pp (2018).
34. DFO. Population Status Update for the Northern Resident Killer Whale (Orcinus orca) in 2018. DFO Can. Sci. Advis. Sec. Sci. Resp. 2019/025, (2019).
35. Bigg, M. A., Olesiuk, P. F., Ellis, G. M., Ford, J. K. B. & Balcomb, K. C. Social organization and genealogy of resident killer whales (Orcinus orca) in the coastal waters of British Columbia and Washington State. Reports Int. Whal. Comm. 12, 383–405 (1990).
36. Ford, J. K. B. & Ellis, G. M. Selective foraging by fish-eating killer whales Orcinus orca in British Columbia. Mar. Ecol. Prog. Ser. 316, 185–199 (2006).
37. Chen, L. H. et al. Selection of reference genes for RT-qPCR studies in blood of beluga whales (Delphinapterus leucas). PeerJ 4, e1810 (2016).
38. Horvath, S. & Raj, K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat. Rev. Genet. 19, 371–384 (2018).
39. Hoelzel, A. R., Dahlheim, M. E. & Stern, S. J. Low genetic variation among killer whales (Orcinus orca) in the eastern north Pacific and genetic differentiation between foraging specialists. J. Hered. 89, 121–128 (1998).
40. Yao, M., Stenzel-Poore, M. & Denver, R. J. Structural and functional conservation of vertebrate corticotropin-releasing factor genes: Evidence for a critical role for a conserved cyclic AMP response element. Endocrinology 148, 2518–2531 (2007).
41. Aguinga, L. M., Yang, W., Yaggie, R. E., Schaeffer, A. J. & Klumpp, D. J. Acyloxyacyl hydrolase modulates depressive-like behaviors through aryl hydrocarbon receptor. Am. J. Physiol. Regul. Integr. Comp. Physiol. 317, R289–R300 (2019).
42. Hankinson, O. The aryl hydrocarbon receptor complex. Annu. Rev. Pharmacol. Toxicol. 35, 307–340 (1995).
43. Lundin, J. I. et al. Pre-oil spill baseline profiling for contaminants in Southern Resident killer whale fecal samples indicates possible exposure to vessel exhaust. Mar. Pollut. Bull. 136, 448–453 (2018).
44. MacDonald, L. H. Evaluating and managing cumulative effects: Process and constraints. Environ. Manag. 26, 299–315 (2000).
45. National Academies of Sciences Engineering and Medicine. Approaches to Understanding the Cumulative Effects of Stressors on Marine Mammals. https://doi.org/10.17226/23479 (National Academies Press, 2017).
46. Barrett-Lennard, L. G., Smith, T. G. & Ellis, G. M. A cetacean biopsy system using lightweight pneumatic darts, and its effect on the behavior of killer whales. Mar. Mammal Sci. 12, 14–27 (1996).
47. Sambrook, J., Fritsch, E. F. & Maniatis, H. Molecular Cloning: A Laboratory Manual (Cold Springs Harbor Laboratory Press, 1989).
48. Illumina. 16S Metagenomic Sequencing Library Preparation. Illumina.com 1–28 (2013).
49. Martin, M. Catadappt removes adapter sequences from high-throughput sequencing reads. EMNet. 17, 10–12 (2011).
50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114–2120 (2014).
51. Magoci, T. & Salzberg, S. L. FLASH: Fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963 (2011).
52. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760 (2009).
53. Li, H. et al. The sequence alignment/map format and SAMtools. Bioinformatics 25, 2078–2079 (2009).

Acknowledgements
We thank G. Ellis, C. Matkin and J. Towers for their assistance collecting biopsy samples.

Author contributions
C.A.C. and L.G.B.L. conceptualized the study and acquired funding. L.G.B.L. provided samples. T.R.F. and C.A.C. designed the study, conducted formal analyses, and wrote the manuscript. L.G.B.L. contributed to reviewing the manuscript. T.R.F. supervised the project.

Funding
This work was supported by a grant from the Fisheries and Oceans Canada Ocean and Freshwater Science Contribution Program, and stipend for CAC was provided by the Nova Scotia Graduate Scholarship Program.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-96255-1.

Correspondence and requests for materials should be addressed to C.A.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
