



Evaluation of fecal hormones for noninvasive research on reproduction and stress in humpback whales (*Megaptera novaeangliae*)

Kathleen E. Hunt^{a,c,*}, Jooke Robbins^b, C. Loren Buck^c, Martine Bérubé^{b,d}, Rosalind M. Rolland^a

^a Anderson Cabot Center for Ocean Life, New England Aquarium, Central Wharf, New England, Boston, MA 02110, USA

^b Center for Coastal Studies, 5 Hobway Avenue, Provincetown, MA 02657, USA

^c Department of Biological Sciences, Northern Arizona University, Flagstaff, AZ 86011, USA

^d Marine Evolution and Conservation, Groningen Institute of Evolutionary Life Sciences, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

ARTICLE INFO

Keywords:

Cetacea
Hormones
Endocrinology
Noninvasive
Fecal hormones
Stress
Reproduction
Thyroid

ABSTRACT

Fecal hormone analysis shows high potential for noninvasive assessment of population-level patterns in stress and reproduction of marine mammals. However, the marine environment presents unique challenges for fecal sample collection. Data are still lacking on collection methodology and assay validations for most species, particularly for those mysticete whales that have variable diets. In this study we tested collection techniques for fecal samples of free-swimming humpback whales (*Megaptera novaeangliae*), and validated immunoassays for five steroid and thyroid hormones. Resulting data were used for preliminary physiological validations, i.e., comparisons to independently confirmed sex and reproductive state. Pregnant females had significantly higher fecal progesterins and glucocorticoids than did other demographic categories of whales. Two possible cases of previously undetected pregnancies were noted. Males had significantly higher fecal testosterone metabolites than nonpregnant females. Fecal glucocorticoids were significantly elevated in pregnant females and mature males compared to nonpregnant females. Calf fecal samples had elevated concentrations of all fecal hormones. Fecal thyroid hormones showed a significant seasonal decline from spring to summer. Though sample sizes were small, and sampling was necessarily opportunistic, these patterns indicate that noninvasive fecal hormone analysis may facilitate studies of reproduction, stress and potentially energetics in humpback whales.

1. Introduction

Baleen whales are challenging to study, and even fundamental traits such as sex, age and reproductive state are often unknown. All species of baleen whales were heavily depleted by commercial whaling during the past several centuries, and today are exposed to various other human impacts (e.g., entanglement in fishing gear, vessel strikes, shipping noise, boat disturbance, etc.; Thomas and Reeves, 2015). These anthropogenic effects may cause direct mortality, but can also affect individuals and populations in more subtle sublethal ways, such as via long-term detrimental impacts on reproduction, health and physiological stress (Fair and Becker, 2000; Thomas and Reeves, 2015). Unfortunately, there are few validated, quantitative techniques for assessing reproduction, health and physiological stress levels of free-ranging large whales (reviewed in Hunt et al., 2013). Such techniques are essential to correctly interpret trends in reproduction and stress,

forecast potential effects on populations, and identify and minimize human impacts (Fair and Becker, 2000; Madliger et al., 2018). Endocrine tools are especially desirable (Atkinson et al., 2015; Madliger et al., 2018; NAS, 2016), but traditional endocrine sampling requires blood samples, which currently cannot be collected from free-swimming large whales (reviewed in Hunt et al., 2013).

Fecal hormone analysis is an alternative, noninvasive technique that is now widely employed for studies of stress and reproductive physiology in terrestrial wildlife (Brown, 2018; Kersey and Dehnhard, 2014). In most vertebrates, steroid and thyroid hormones are removed from circulation by the liver, with a portion excreted into bile and then into the gut lumen. Within the gut, a given steroid hormone is metabolized to a suite of various breakdown products (“fecal hormone metabolites”, FHM) which are eventually excreted in feces (Brown, 2018; Heistermann et al., 2006; Palme et al., 1996; Wasser et al., 2000). Mammalian feces thus typically contain high concentrations of

Abbreviations: EIA, enzyme immunoassay; fE, fecal estrogen metabolites; fGC, fecal glucocorticoid metabolites; FHM, fecal hormone metabolites; fP, fecal progesterin metabolites; fT, fecal testosterone metabolites; fThy, fecal thyroid hormone metabolites; NARW, North Atlantic right whale; RIA, radioimmunoassay

* Corresponding author at: Department of Biological Sciences, Northern Arizona University, Box 4185, Flagstaff, AZ 86011, USA.

E-mail address: Kathleen.Hunt@nau.edu (K.E. Hunt).

<https://doi.org/10.1016/j.ygcen.2019.04.004>

Received 16 December 2018; Received in revised form 22 March 2019; Accepted 1 April 2019

Available online 02 April 2019

0016-6480/ © 2019 Elsevier Inc. All rights reserved.

immunoreactive metabolites of all five classes of steroid hormones (proggestins [fP], testosterone and related androgens [fT], estrogens [fE], glucocorticoids [fGC], and mineralocorticoids) as well as some thyroid hormones (fThy; metabolites of thyroxine and tri-iodothyronine) (Brown, 2018; Burgess et al., 2017; Wasser et al., 2000, 2010). These FHM are often detectable by commercially available immunoassays designed for the parent hormones. Immunoreactive FHM content of feces is thought to represent not a single point in time, but rather an integrated measure of all hormone secreted over the previous hours or days, depending on species-specific gut passage time (Amaral, 2010; Hunt et al., 2004; Wasser et al., 2010). Thus, FHM data appear ideally suited for assessment of long-term endocrine trends, and for some research questions may even be superior to plasma, such as for identification of periods of chronic stress as opposed to acute stress (Dantzer et al., 2014; Dickens and Romero, 2013). In terrestrial wildlife, such measures have enabled non-invasive assessment of animal sex, reproductive state, age class, nutritional state and physiological stress level, without ever having to catch the animals (Brown, 2018; Wasser et al., 2017).

However, interpretation of fecal hormone data can be complex. For most species, the exact chemical identity of the FHMs is unknown, due to species-specific differences in steroid metabolism, gut chemistry, gut flora, and resulting assay antibody affinities (Heistermann et al., 2006; Hunt et al., 2006; Palme et al., 1996). Further, FHM concentrations can vary not just with the animal's physiological state but also with a multitude of other potentially confounding factors, including age, sex, individual variation, diet, and season, as well as methodological effects of collection, extraction, storage, and sample mass (Goymann, 2012; Hayward et al., 2010; Hunt and Wasser, 2003; Pappano et al., 2010). Thus, fecal hormone assays must be validated carefully. Initial validations should include both assay validations (can the assay detect and quantify the FHMs; Hunt et al., 2017) and also physiological validations (do FHM data reflect the animal's physiological state; Grotjan and Keel, 1996; Rolland et al., 2005; Wasser et al., 2000, 2010). In terrestrial taxa, injections of pituitary hormones and/or radiolabelled hormones can also be employed, but such experiments are not possible for large-bodied cetaceans (Hunt et al., 2013), and so physiological validations in cetaceans typically focus on comparison to known physiological state. An additional complication for fecal hormone studies of cetaceans, as compared to similar methodology in terrestrial mammals, is that sample collection rate for cetaceans tends to be low. Cetacean fecal samples disperse and sink rapidly, and collection opportunities are typically limited to the sea surface, which is a small portion of the whales' habitat. Sample collection is further restricted by weather, sea state, accessibility to sufficient numbers of whales, and associated budgetary and logistical constraints. Even a considerable field effort frequently yields sampling rates of < 30 samples/yr, some of these samples proving too small for analysis (e.g., Lemos et al., 2017; Rolland et al., 2005; Valenzuela-Molina et al., 2018). Thus, fecal hormone studies of cetaceans must contend with the statistical limitations characteristic of small-*n* studies. Nonetheless, even very-small-*n* studies can have value (Ploutz-Snyder et al., 2014), and, fortunately, effect sizes of steroid and thyroid hormones for the major questions of interest (pregnancy diagnosis, identification of acute or chronic stress, etc.) are typically great enough for acceptable statistical power even with a small-*n* study design (e.g., Hunt et al., 2006; Rolland et al., 2005; Valenzuela-Molina et al., 2018).

In mysticetes, fecal hormone assay validations have been published for only two species, North Atlantic right whale (*Eubalaena glacialis*, NARW) and blue whales (*Balaenoptera musculus*). In NARW, a series of comprehensive assay validations and physiological validations have been performed (Burgess et al., 2017; Corkeron et al., 2017; Hunt et al., 2006; Rolland et al., 2005, 2017). In blue whales, progesterone and glucocorticoid assays have been validated for females only (Valenzuela-Molina et al., 2018). However, both of these species have relatively low dietary variation during the seasons studied and also have fecal

consistency highly amenable to sample collection, with relatively large, semi-solid clumps of feces that float at the water surface (Rolland et al., 2005; Valenzuela-Molina et al., 2018). Other baleen whales have a more piscivorous and/or more variable diet, and this may cause differences in fecal consistency among sampling instances, individuals, populations and species. In several piscivorous cetaceans, feces consist of fine, unbound particles that rapidly disperse in the water column, forming a sinking “fecal plume” that poses considerable difficulties for collection. In such circumstances, recovering adequate sample mass can be a concern, since very small fecal samples can produce anomalously high apparent hormone concentrations, even when results are calculated on a per gram basis (the “small sample effect”; Hayward et al., 2010).

In this study we sought to determine whether the fecal hormone techniques previously developed for NARW are viable for mysticete whales with more piscivorous or variable diets and highly dispersed feces. To test this approach, we collected fecal samples from free-swimming North Atlantic humpback whales (*Megaptera novaeangliae*) of a well-studied population in the Gulf of Maine. This population is found in the same region as NARW and is exposed to many of the same anthropogenic stressors, such as chronic ocean noise, entanglement and shipstrike. However, humpback whales also differ from NARW in many biological respects, including that they produce more liquid, dispersed fecal plumes and feed at a higher trophic level, consuming primarily small schooling fish such as sand lance (*Ammodytes* spp.), Atlantic herring (*Clupea harengus*), mackerel (*Scomber scombrus*) and occasionally euphausiids (*Meganyctiphanes norvegica*; Ichii and Kato, 1991; Matthews, 1937; Overholtz and Nicolas, 1979).

The specific goals of this study were to: (1) identify a *practical method of collection* of fecal plume particles, one that consistently produces enough sample for multiple hormone assays, with little, if any, “small sample effect”; (2) perform *assay validations* for five hormone classes (fP, fT, fE, fGC, and fThy) using commercially available immunoassays; (3) where sample sizes allowed, perform *physiological validations* based on predicted differences across seasons, sexes, age classes, and reproductive states of known individual humpbacks. Physiological validations relied on the following predictions, based on previous literature (Hunt et al., 2006; Rolland et al., 2005, 2012, 2017; Wasser et al., 2017): (1) fP was predicted to be higher in pregnant females than in adult females not known to be pregnant; (2) fE was predicted to be higher in lactating mothers than other mature females; (3) fT and fT/fE ratio were predicted to be higher in adult males compared to other demographic classes (fT/fP ratio was also explored, but with no specific predictions as to potential demographic differences, due to the fact that progesterone is a precursor to testosterone); (4) fGC was predicted to be higher in pregnant females and adult males than in nonpregnant females, and was also predicted to be elevated in any whales experiencing elevated stress, such as from anthropogenic interactions; (5) while there was little basis for prediction of fThy patterns, we anticipated seasonal trends, such as could result from reduced thermal energetic demands on feeding grounds in mid-summer (based on Ayres et al., 2012; Behringer et al., 2018; Cristobal-Azkarate et al., 2016; Gesquiere et al., 2018; Schaebs et al., 2016; Thompson et al., 2017; Wasser et al., 2017).

2. Methods

2.1. Study population

Fecal plumes were sampled in the Gulf of Maine, a seasonal feeding ground off the northeastern United States. Humpback whales feed in these waters from late March through December and then migrate to the Caribbean to mate and calve. Individual humpback whales were identified by their pigmentation of ventral flukes and shape of dorsal fin (Katona and Whitehead, 1981). The identity of each individual was photographically documented at the time of fecal collection, and

individual sex, age class and life history were obtained from the Gulf of Maine Humpback Whale Catalog (Center for Coastal Studies, Massachusetts, USA). The sex of individuals was determined using ZFX/ZFY specific primers located on the sex chromosome (Bérubé and Palsbøll, 1996a,b; Palsbøll et al., 1992), observation of the genital slit (Glockner, 1983) or both. Exact age was known for individuals that were first catalogued in their first year of life as dependent calves, whereas independent whales were assumed to be at least one year old at the time of first sighting. The earliest age at sexual maturity in this species is four years (Clapham, 1992; Robbins, 2007). Females were considered sub-adults from age four until they produced their first calf (Robbins, 2007). There are no outward indicators of pregnancy in this species and so females were known to have been pregnant only if they returned with a calf the following summer. It was not possible to definitively identify nonpregnancy, because females that did not return with a calf the following year might have lost a fetus or a calf before it could be observed. Finally, this species mates in winter, followed by an 11–12 month gestation period (Chittleborough, 1958). Our sampling effort was temporally and geographically distinct from the mating and calving grounds of this migratory population. Thus, females were already pregnant or not when first encountered in the Gulf of Maine, and those that were pregnant likely ranged from early to late term when sampled.

2.2. Fecal collection

Fecal samples were collected during two periods: Aug 22 – Oct 18, 2012, and May 18 – Sept 29, 2013. During surveys, observers scanned for fecal plumes at all times, and sample collection was attempted whenever a plume was observed in close proximity to the vessel. Opportunistic fecal sampling also occurred when possible if fecal plumes were sighted while vessels were engaged in other research.

Early trials revealed that some humpback fecal samples passed through 300 μm mesh nets (the mesh size used successfully with NARW; Rolland et al., 2005). Mesh sizes of 150 μm , 100 μm and 50 μm were therefore also tested, using customized plankton mesh nets (Sea-Gear Corp., Melbourne, FL, USA). 50 μm mesh proved to have too much resistance (i.e., net was difficult to move quickly through the water). Sampling progressed with 100 μm and 150 μm nets for further testing, along with a modified bailer to collect an unfiltered sample (e.g., a volume of water including fecal particulates). The bailer or “scoop” method has been used successfully with some killer whale samples (Ayres et al., 2012). In three cases we were able to subsample a single fecal plume (i.e., from a single whale) with at least two, and in one case all three, of these collection devices (100 μm net, 150 μm net, and bailer). For these three cases, subsamples from the same plume were compared for freeze-dried mass and hormone content per g of dried fecal powder. Subsequent samples were collected using nets, with the mesh size (100 μm or 150 μm) chosen based on the specific circumstances, such as the apparent consistency of the sample.

Fecal samples were chilled on board the boat in a refrigerator or cooler and frozen at $-80\text{ }^{\circ}\text{C}$ upon return to shore. Samples were later shipped to the New England Aquarium (Boston, MA, USA) and stored at $-80\text{ }^{\circ}\text{C}$ until analysis within one year. We also collected three reference samples of seawater in 2013 to evaluate the possibility that low concentrations of immunoreactive hormone might be present routinely in seawater of the Gulf of Maine.

2.3. Extraction of hormones from feces

Most humpback fecal samples consisted of small fecal particles (often < 1 g) suspended in a large volume of seawater (often > 100 mL). In terrestrial fecal hormone studies, samples are commonly freeze-dried after collection to remove variation in fecal weight due to water content (Hunt and Wasser, 2003). In the humpback samples, the ratio of seawater to fecal particulates was so high that freeze-drying would have caused substantial inflation of apparent

sample mass due to the contribution of dried salts. Pilot trials confirmed that apparent dried fecal mass often doubled or tripled due to salt mass, causing corresponding large decreases in apparent ng/g hormone concentrations (data not shown). To minimize this effect, all samples were thawed and centrifuged for 30 min at 4000g, and any overlying seawater was decanted or pipetted to a separate container. The remaining fecal pellet was then refrozen and freeze-dried under vacuum at $-20\text{ }^{\circ}\text{C}$ for seven days or until dry. Once fully dry, samples were brought to room temperature, pulverized, mixed well, and a subsample of fecal powder weighed to the nearest 0.0001g for hormone extraction.

Based on reports that small samples can have inflated hormone content, potentially due to higher ratio of extraction solvent to sample mass (Hayward et al., 2010), we used a consistent 20:1 ratio (mL 90% methanol:g dried fecal powder) for extraction. Choice of solvent and the 20:1 ratio (modified to increase extraction efficiency from the 10:1 ratio of Wasser et al., 2000) were based on previous experience with fecal samples of North Atlantic right whale and numerous other species (Hunt et al., 2004, 2006; Hunt and Wasser, 2003; Rolland et al., 2005, 2012, 2017; Wasser et al., 2010). The 20:1 ratio reflects a balance between extraction efficiency (higher solvent:sample ratio desirable) and hormone detectability (lower solvent:sample ratio desirable) while also avoiding a dry-down step (increases variability; see Hunt et al., 2004). Samples with dried fecal mass < 0.0005 g were not extracted or analyzed, since the resulting extract volume would have been too small to assay. For all samples > 0.0005 g, the appropriate volume of 90% methanol was added to the weighed fecal powder in a 16 \times 100 mm borosilicate glass tube, which was capped and vortexed 1.0 h, centrifuged at 4000g for 15 min, and supernatant pipetted to a cryovial and stored at $-20\text{ }^{\circ}\text{C}$ overnight. The supernatant was decanted to second cryovial the next day to remove any remaining fecal particulates, producing a “1:1” or full-strength extract. Extracts were stored at $-20\text{ }^{\circ}\text{C}$ until assay.

2.4. Hormone assays and assay validations

The fP assay was an in-house ^3H radioimmunoassay (RIA) using progesterone antibody “Munro CL425” (University of California, Davis CA, USA). Two testosterone assays were tested, the first an in-house ^3H RIA using the “Niswender #250” antibody (Niswender lab, Colorado State University) and the second a colorimetric enzyme immunoassay (EIA; #K032-H5 from Arbor Assays, Ann Arbor, MI, USA). The fE and fGC assays were commercial double-antibody ^{125}I RIAs, and the fThy assay a coated-tube ^{125}I RIA (corticosterone assay #02-102103, “total estrogens” RIA #07-140202, and “total T3” assay #06B-254215 respectively, MP Biomedicals, Costa Mesa, CA, USA). Assay protocols, sensitivities and antibody cross-reactivities have been previously described in Rolland et al. (2005) for fP, fT, and fE RIAs; Hunt et al. (2006) for fGC and fT RIAs; Hunt et al. (2017) for the fT EIA; and Wasser et al. (2010) for the fThy RIA. For all assays, inter- and intra-assay variation were both < 10%.

Assay protocols employed standard QA/QC, including assay of non-specific binding tubes and zero-dose tubes in quadruplicate, and assay of all standards, controls and samples in duplicate, with final results averaged. Any samples with a coefficient of variation > 10% between duplicates, or that fell outside 10–90% bound on the standard curve (20–80% for the fE assay, based on assay-specific reliability), were re-diluted accordingly and re-assayed to confirm results. Any sample with results more than 2 standard deviations outside the mean for all samples in this study, or with anomalous results for the whale's category (age class, sex and reproductive state) was re-assayed to confirm results.

Assay validations employed standard parallelism and accuracy tests (Grotjan and Keel, 1996). The parallelism test involved assay of twelve serial dilutions of a pooled humpback whale fecal extract created by mixing equal volumes of all samples with samples with dried sample mass > 0.1 g (i.e., samples that produced ample extract), with results graphed as percent-bound vs. log[relative dose] and compared to the

binding curve of known-dose hormone standards run in the same assay. Parallelism of the linear portions of the two curves indicates that the fecal metabolites bind well to the assay antibody, signifying that the fecal extract likely contains either the parent hormone itself or a closely related metabolite(s).

Based on parallelism results, a dilution was selected for each assay that fell near 50% bound on the standard curve, the area of greatest assay precision. This dilution was tested in an accuracy test (aka “matrix effect” test; Grotjan and Keel, 1996), consisting of assay of a set of known-dose standards that were spiked with equal volumes of a diluted pool of fecal extract, run alongside a second standard curve spiked only with assay buffer. The best-fit regression line of apparent dose vs. known dose was then inspected for linearity and a slope close to 1.0 (see Statistical Analysis). Following validations, all samples were diluted in appropriate assay buffer for each individual assay.

2.5. Physiological validations: age, sex and reproductive state

Fecal samples were categorized based on the demographic traits and life history of the source animal, and results were compared across demographic categories. When multiple fecal samples were obtained from the same whale on the same day, only the sample with largest dried fecal mass was selected for statistical analysis. When samples were obtained from the same individual on multiple days, only the first sample was included in demographic comparisons.

The pregnancy status of each mature female was initially determined from life history data, based on whether she was sighted with or without a calf the following year. Adult females were thus classed retrospectively as “Resighted-With-Calf” ($n = 4$); “Resighted-Without-Calf” ($n = 8$); or “Not Resighted” ($n = 2$). It was expected that the majority of females in the Resighted-Without-Calf group were likely not pregnant, but that this group could include some pregnant females that lost the fetus or calf (Pallin et al., 2018). fP data were inspected for an fP threshold that could discriminate the majority of Resighted-With-Calf whales from all other adult whales of both sexes, via iterative testing of all possible fP values across the full range of sample concentrations. Once this fP threshold was identified, any female whales with a higher fP were tentatively categorized as “pregnant” in subsequent physiological validations of other hormones (fT, fE, fGC, fThy) (see also Pallin et al., 2018, for other potential causes of high progesterone, e.g. luteal activity).

fGC physiological validations focused on predictions of elevated fGC in two demographic categories known to entail high energetic burden and physiological stress in other whale species, namely adult males and pregnant females (Hunt et al., 2006), though note that adult males might not have had elevated fGC in this study due to sampling outside the breeding season. Additionally, fGC concentrations were opportunistically examined from three whales exposed to potential anthropogenic stressors. One sample was obtained from an individual while it was entangled in fishing gear. This whale was of unknown sex and age (though was visually judged to be “small”) and had no prior sightings. Although the duration of its entanglement and the extent of its injuries were not known, its body condition was poor. Three other samples were obtained from two mature males that were satellite-tagged as part of an independent study on the health effects of implantable tags on humpback whales (Robbins et al., 2013). One of these whales was fecal-sampled the day of tagging and each of the two following days. The other was fecal-sampled two years after tagging, but still had a swelling at the tag site and a possibility of embedded tag parts (Robbins et al., 2013). Results from entangled and satellite-tagged whales are provided primarily as anecdotal information for a growing body of data on human impacts, and were excluded from initial comparisons of fGC across demographic categories.

2.6. Statistical analysis

The effectiveness of fecal collection methods (150 μm net, 100 μm net, and bailer) was explored based on collected sample mass and ng/g hormone concentrations in the three plumes from which multiple samples were obtained. All remaining net-collected samples were also tested for a small-sample effect via linear regressions of all five hormones (ng/g) against the dried sample mass (g) used in the extraction. Assay validations were assessed with F tests for parallelism, i.e., testing for equality of slope on the linear portions of the two curves, while accuracy was assessed for slope within 0.7–1.3 (ideal slope = 1.0) and linearity with linear regression (ideal $r^2 > 0.95$). For physiological validations, most hormone data were not normally distributed and were log-transformed before analysis. If transformation was successful in normalizing data distribution, parametric tests were used; otherwise, nonparametric tests were used. All categories of whales with $n > 4$ (pregnant females, nonpregnant females, and adult males) were compared for fP, fT, fGC, fThy and for fT/fE and fT/fP ratios, using with t-tests or Mann Whitney tests for two-group comparisons, and ANOVAs or Kruskal-Wallis test for three-group comparisons as appropriate. fGC data for the (few) entangled and tagged whales are also reported as anecdotal descriptive data, in reference to appropriate group means (i.e., same age/sex class, where known). Finally, fThy data was assessed not only for demographic differences but also for potential seasonal trends via regression of fThy on Julian date. As fThy data had a high proportion of nondetectable hormone results, these relationships were explored twice, once with the nondetectable values coded as zero's and a second time with nondetectable results excluded. Since seasonal hormone cycles are often nonlinear, this analysis began with linear regression and proceeded to higher-order polynomials (2nd through 6th order examined in sequence) to assess whether nonlinear regression produced a markedly improved fit to the data. Statistical analyses were performed with Prism 6 or SPSS 11 for Mac OSX, with alpha (significance threshold) set at 0.05, with selection of tests following recommendations for small- n experimental designs and non-normal data as in Cohen (1988), Siegel and Castellan (1988), Ruxton and Neuhäuser (2010) and Murphy (2017). One-tailed tests (see Murphy, 2017; Ruxton and Neuhäuser, 2010, Siegel and Castellan, 1988) were used for comparisons of reproductive state to reproductive steroids, and of reproductive state to fGCs, since these analyses had a strong *a priori* predictions regarding direction of effect (based on consistent patterns reported for fecal hormones of cetaceans, other marine mammals, and mammals generally). All other comparisons employed two-tailed tests.

3. Results

3.1. Fecal sample collection and processing

Four fecal samples were collected with nets in 2012 and 57 were obtained with nets in 2013. Four samples that had been collected opportunistically and archived before the start of the study, in 2003, 2007, and 2009, were added to the dataset for a total sample size of 65 fecal samples. Seven samples produced dried pellets < 0.0005 g and were excluded from analysis. Thus, the final sample size was 58 fecal samples, of which 50 produced enough extract to assay for all five hormones. The remaining eight samples were assayed for varying numbers of hormones depending on available volume of extract, with the progesterin, glucocorticoid and thyroid assays given priority.

None of the five hormones (fP, fT, fE, fGC, fThy) showed a significant relationship to dried sample mass ($P > 0.40$ for all comparisons), indicating lack of the “small sample effect” noted in other studies (e.g., Hayward et al., 2010). 50% of samples had at least the maximum 0.1000 g of freeze-dried powder needed for a “full size” sample extraction, i.e., 2.000 mL of extraction solvent added to 0.1000 g of fecal powder, while the other 50% of samples ranged between 0.0035 g and 0.0815 g. Overall, mean sample mass extracted for

assays (i.e., not including samples < 0.0005 g that were not assayed) was $0.0596 \text{ g} \pm 0.0380 \text{ SD}$, median = 0.0540 g, range 0.0035–0.1300 g.

The nets, regardless of mesh size, collected a larger mass of fecal particulates than the bailer. Where sample mass was sufficient for assay, net-collected samples had higher hormone concentration than bailer samples collected from the same fecal plume, even when results were expressed on a ng/g basis. Subsequent analyses used only net-collected samples.

All three seawater samples tested had extremely low or undetectable amounts of all hormones. fGC, fE, and fThy were uniformly undetectable in seawater. Minute concentrations of immunoreactive fP and fT were noted (< 1 ng/mL fP, < 2 pg/mL fT), but absolute levels were near the limit of assay sensitivity (> 98% bound for both assays), and were two (fT) to three (fP) orders of magnitude lower than concentrations of all fecal samples. Subtracting these “background” (spurious) immunoreactive concentrations from fecal results did not alter any patterns in the data or the results of any statistical tests (data not shown). Due to the orders-of-magnitude difference between seawater immunoreactive hormone vs. fecal immunoreactive hormone, final analyses did not utilize a background correction.

3.2. Assay validations

Parallelism was good for all five assays (Table 1), but accuracy tests did not all immediately provide satisfactory results. The thyroid-hormone RIA demonstrated good accuracy on the first trial. However, results of the other four RIAs — fP, fE, fGC, and fT — initially demonstrated poor accuracy, with a slope outside the acceptable range of 0.7–1.3 (generally < 0.7). Visual and olfactory inspection of the samples used for these tests indicated that several samples included relatively large amounts of algae. New sample pools were created with “non-algae samples” only, samples that did not appear green and that did not smell like algae, following which the fP, fE and fGC assays demonstrated good accuracy. However, the fT RIA still had poor accuracy, consistently demonstrating a “shallow” accuracy slope of approximately 0.6–0.7. As a result a new fT assay was tested, the Arbor Assays fT EIA; parallelism and accuracy were good for this assay and all subsequent fT assays used the EIA. Final accuracy results are presented in Table 1.

3.3. Physiological validations

Twenty-five samples were excluded from physiological validations for the following reasons: sample was collected with a bailer (found to

Table 1

Statistical results of assay validations with humpback fecal extract in five immunoassays. Parallelism results should indicate no significant difference of slope between binding curves of diluted samples and hormone standards ($P > 0.05$); accuracy results should indicate slope of observed vs. expected dose near 1.0 (acceptable range 0.7–1.3) and straight line with $r^2 > 0.95$. fP = fecal progesterin metabolite assay; fE = fecal estrogen metabolite assay; fT = fecal testosterone metabolite assay; fGC = fecal glucocorticoid metabolite assay; fThy = fecal thyroid hormone metabolite assay.

Assay	Parallelism results	Accuracy results
fP	$F_{1,5} = 0.0247$ $P = 0.8812$	Slope = 0.828 $r^2 = 0.9898$
fE	$F_{1,10} = 0.2742$ $P = 0.6119$	Slope = 0.992 $r^2 = 0.9802$
fT	$F_{1,9} = 2.5649$ $P = 0.1437$	Slope = 1.102 $r^2 = 0.9989$
fGC	$F_{1,16} = 0.1617$ $P = 0.6929$	Slope = 1.185 $r^2 = 0.9963$
fThy	$F_{1,14} = 1.6272$ $P = 0.2228$	Slope = 1.085 $r^2 = 0.9994$

have lowered hormone content; $n = 4$); sample contained large amounts of algae ($n = 3$); sample from calf (possible transfer of maternal steroids via milk fat; $n = 3$), unknown individual ($n = 4$), individual of unknown sex ($n = 1$), or multiple samples from the same individual ($n = 10$). The final dataset for physiological validations totaled 33 humpback fecal samples, one each from 14 adult nonlactating females, one lactating female, one subadult female, and 17 mature males.

Resighted-With-Calf females had significantly higher fP than Resighted-Without-Calf females ($P = 0.0021$, $t_{10} = 3.684$; pregnant samples, $n = 4$, median = 1167 ng/g, mean \pm SEM = 1448 ± 410 ng/g; Resighted-Without-Calf samples, $n = 8$, median = 275 ng/g, mean \pm SEM = 387 ± 134 ng/g). An fP threshold value of 800 ng/g was superior at discriminating Resighted-With-Calf fecal samples from all other adult fecal samples (Fig. 1). Only two other samples in the entire humpback fecal sample dataset had fP > 800 ng/g. Both were from mature females; one was Resighted-Without-Calf and the other was Not-Resighted (Fig. 1), and both could have been pregnant. Based on these results, the putative pregnancy threshold was set at fP = 800 ng/g. For subsequent analyses of other hormones, the two additional females with fP > 800 ng/g were tentatively reclassified as “pregnant”, while all females with fP < 800 ng/g were classified as “nonpregnant”. Group means, medians and ranges for all whales after reclassification are shown in Table 2.

fE did not differ among males, nonpregnant females, and pregnant females, whether analyzed in a three-group comparison or a two-group comparison of all males to all females ($P > 0.05$ for all tests; Table 2). The single lactating mother had nondetectable fE (Table 2).

Adult males had significantly higher fT and fT/fE ratio than nonpregnant females (fT, $P = 0.0204$, $t_{25} = 2.158$, Fig. 2; fT/fE ratio, $P = 0.0325$, $U = 39$), but did not differ in fT/fP ratio from nonpregnant females ($P = 0.1115$, $U = 60$, Fig. 3). When pregnant females ($n = 6$) were added to produce a three-group analysis (i.e., adult males, nonpregnant females, and pregnant females; Fig. 3), there was still a significant difference in fT concentration ($P = 0.0330$, $KW = 5.437$; Fig. 2) and fT/fE ratio ($P = 0.0475$, $KW = 4.709$) and a significant difference appeared in fT/fP ratio as well ($P = 0.0008$, $KW = 12.95$). Post-hoc tests revealed that pregnant females differed from the other two groups for fT/fP ratio (Dunn's multiple comparisons test, $P < 0.05$); no other pairwise comparisons indicated significant differences.

Pregnant females had significantly higher fGC than nonpregnant females ($P = 0.0224$, $U = 10.000$; Fig. 4). Similarly, adult males had higher fGC than nonpregnant females ($P = 0.0061$, $U = 24$; Fig. 4). A Kruskal-Wallis test on all three groups also indicated significant differences ($P = 0.0367$, two-tailed, $KW = 6.610$), though in post-hoc comparisons only the nonpregnant females were different from adult males ($P < 0.05$). The lactating mother had higher fGC than most other nonpregnant females, 37 ng/g (Fig. 4, Table 2). Finally, the entangled whale had an fGC of 51 ng/g (Fig. 4). This was mildly elevated relative to the mean of other sampled whales, though none were in the same likely age class as the entangled whale. The two satellite-tagged whales, both of which were adult males, had fGC that fell near the median fGC for adult males of 21.2 ng/g (Table 2). One tagged whale had fecal glucocorticoids of 22 ng/g the hour after he received a satellite tag (this sample likely represents pre-tagging physiological state, due to gut transit time), 29 ng/g on the next day, and 26 ng/g on the third day. The other whale had an fGC value of 18 ng/g two years after tagging, while exhibiting a tag-site swelling and a possibility of embedded tag parts (Robbins et al., 2013).

For thyroid hormones, 45% of samples (15 of 33) had nondetectable fThy, in contrast to much lower nondetectability rates in the other assays (seven nondetectable samples for fE, two for fT, one for fGC, and zero for fP). There were no significant differences among adult males, nonpregnant females, or pregnant females in fThy concentration ($P > 0.05$ for all comparisons), regardless of whether samples with

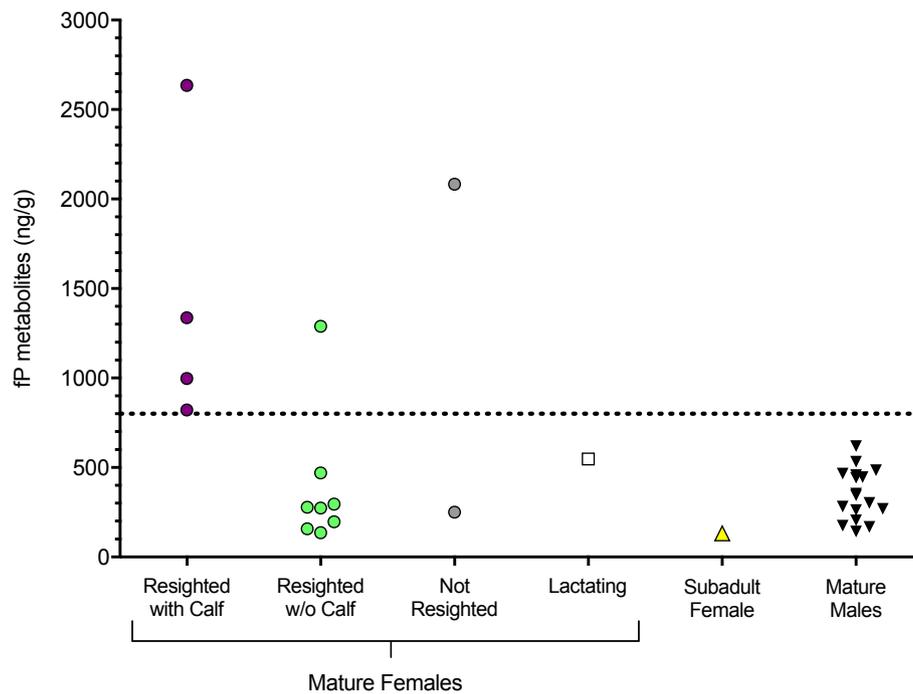


Fig. 1. Fecal progesterin metabolites of humpback whales categorized by sex, age class and reproductive state. Dashed line at 800 ng/g discriminates samples from known-pregnant females from all other adult samples (including duplicate samples not shown here) with maximum accuracy.

undetectable fThy were coded as zero's or were excluded from analysis. There were also no significant differences among these three reproductive categories in the proportion of samples with detectable vs. nondetectable hormone ($P > 0.05$ for all comparisons). However, fThy declined significantly as the year progressed, with higher fThy levels in spring and lower fThy levels in summer (Fig. 5; $P < 0.0001$, $F_{1,33} = 20.04$, $r^2 = 0.3379$). Goodness-of-fit improved markedly with a second-order (quadratic) fit ($r^2 = 0.6430$), reflecting a slight increase in fThy in early fall; additional polynomial fits did not markedly improve fit. The best-fit linear regression line and best-fit second-order polynomial nonlinear regression line are shown in Fig. 5. All polynomial fits revealed that fThy reached a nadir in approximately the middle of

August. These results did not change in direction or significance if nondetectable results were excluded.

3.4. Calf samples

The three fecal samples collected from calves had very high concentrations of all hormones (Table 2). Anecdotally, calf samples also had different consistency than samples from adults, with calf samples characterized by lighter color and larger “chunks” that floated rather than sinking. Mean fP for the three calves was in the range of pregnant females, the mean fGC was second only to the adult males, and the mean calf fT, fE and fThy were the highest for any category of whale

Table 2

Descriptive statistics of fecal metabolites of progestins (fP), testosterone metabolites & related androgens (fT), estrogen metabolites (fE), glucocorticoid metabolites (fGC) and thyroid hormone metabolites (fThy) for humpback whales of known age class, sex, and reproductive state. Data presented as mean \pm SEM (n) in first row, median in second row, range in third row. All data are in ng/g (nanograms of immunoreactive fecal hormone metabolites per g of dried mixed fecal powder). Sample sizes vary because not all samples were large enough to assay for all hormones. Nondetectable hormone values were assigned zeros.

	fP	fT	fE	fGC ⁴	fThy
Pregnant Females ¹	1527 \pm 283 (6) 1313 821–2635	32 \pm 10 (6) 28.0 11–69	9.2 \pm 5.2 (5) 3.3 0.0–28.7	46.9 \pm 22.4 (6) 26.0 10.7–157.0	18.2 \pm 12.1 (6) 7.0 0–76.5
Nonpregnant, Nonlactating Adult Females ²	257 \pm 37 (8) 262 136–470	15 \pm 3.4 13 5–32	18.0 \pm 3.1 (8) 15.1 7.6–35.7	17.4 \pm 6.9 (8) 11.6 0.9–63.4	16.2 \pm 5.3 (8) 19.3 0–40.1
Lactating Female	548 (1)	12 (1)	0.0 (1)	37.3 (1)	0.0 (1)
Subadult Female	132 (1)	6 (1)	12.3 (1)	4.7 (1)	18.9 (1)
Adult Males	350 \pm 34 (17) 344 142–620	31 \pm 10 (17) 20 8–175	9.7 \pm 2.1 15.1 591 \pm 508 (2)	98.9 \pm 51.0 (16) 21.2 0.0–694.9	17.3 \pm 5.5 (17) 15.4 0.0–95.6
Calves ³	892 \pm 241 (3) 1039 421–1215	198 \pm 91 (3) 192 43–358	591 \pm 508 (2) 591 83–1099	91.2 \pm 29.0 (3) 70.9 54.3–148.4	329.3 \pm 247.4 (2) 329.3 82.0–576.7

¹ Pregnant female group includes six adult females with fP above the putative pregnancy threshold (800 ng/g); four were resighted the next year with a calf, one was resighted the next year with no calf, and one was not resighted.

² Nonpregnant, nonlactating adult female group includes eight adult females with fP below the putative pregnancy threshold (800 ng/g); seven were resighted the next year with no calf, and one was not resighted.

³ Calves include two males and a third calf of unknown sex.

⁴ fGC data excludes samples from whales exposed to a known or possible anthropogenic stressor prior to sample collection.

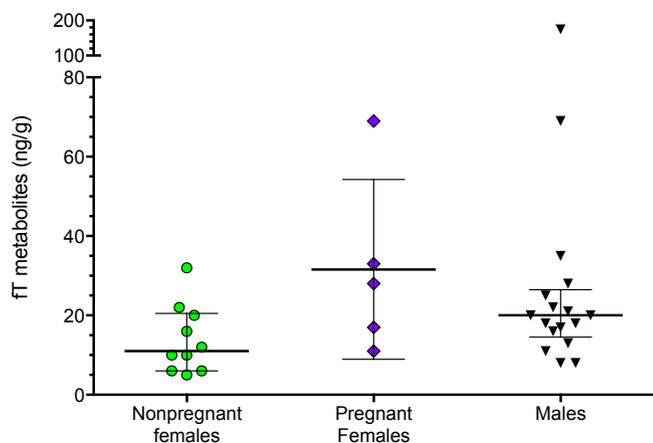


Fig. 2. Fecal testosterone metabolites of adult male humpback whales as compared to all nonpregnant females (including lactating and subadult females) and pregnant females. Pregnant or nonpregnant status was determined by fP threshold value. Bars represent medians \pm interquartile ranges.

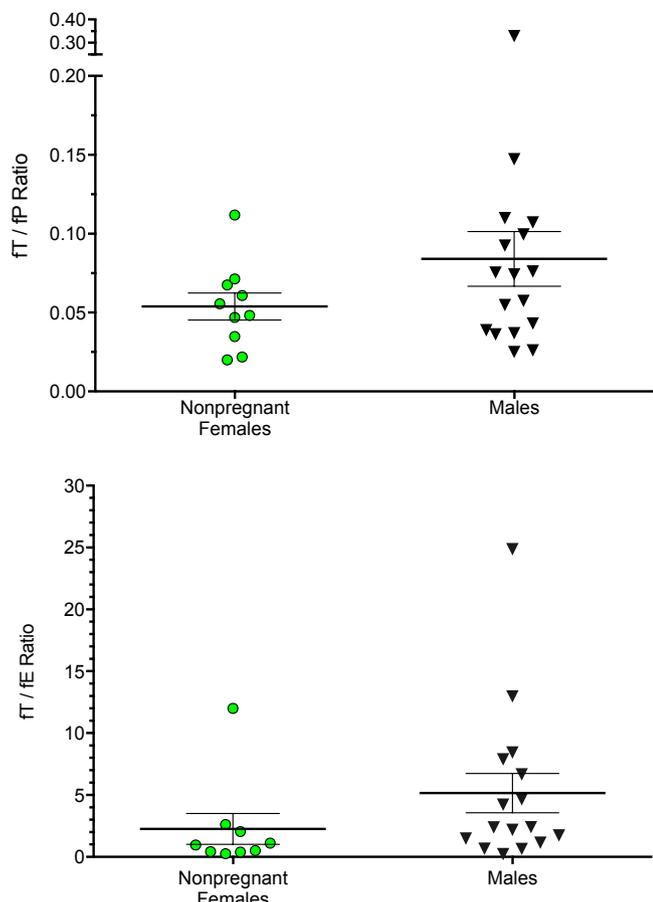


Fig. 3. Ratios of fecal testosterone metabolites to fecal progestin metabolites (fT/fP; top) and fecal testosterone metabolites to fecal estrogen metabolites (fT/fE; bottom) in all nonpregnant females as compared to adult males. Bars represent means \pm SEM.

(for example, calf fT concentrations greatly exceeded those of even the adult males). Two calves were later identified as male and the third was of unknown sex.

4. Discussion

Fecal metabolites of steroid and thyroid hormones are readily

detectable in humpback whale feces via immunoassay, provided that samples are collected and processed appropriately and that assays are validated. Both assay validations and physiological validations in this study were successful in that concentrations of fecal metabolites of reproductive steroids, adrenal glucocorticoids, and thyroid hormones correlated in expected ways with either known physiological state of identified individual whales or, for thyroid hormones, with season of collection. We caution, however, that these conclusions must be regarded as preliminary and tentative, given the small sample sizes and the fact that some demographic groups were insufficiently sampled despite being present in the sampled population. The latter is a fundamental limitation when opportunistically sampling shed materials. Possible interpretations of our results are outlined below, but we emphasize the need for continued data collection to achieve greater sample sizes.

4.1. Sample collection and processing

Although based on limited data, our tests indicate that nets may be superior to bailer/scoop methods for collection of highly particulate feces. Not only did nets collect a greater mass of fecal particulates from the same fecal plume, but net-collected samples also had higher hormone concentration even on a per-gram basis. These patterns may indicate that some hormone migrates from bailer-collected fecal samples to the surrounding water in the bailer. If so, it may be detrimental to store cetacean fecal samples in a large volume of seawater for extended periods of time; prompt centrifugation or decanting of overlying water might be advisable. We recommend further testing of this question. Finally, the 100 μ m and 150 μ m nets performed similarly, with both nets collecting sufficient fecal mass for analysis and with similar hormone concentrations in the resulting samples.

Once samples were collected, our fecal processing methods incorporated two steps not previously described for marine mammal fecal analysis: removal of overlying water via centrifugation, and use of a constant solvent:feces ratio (mL of extraction solvent to g of dried fecal powder) to minimize the potential effects of variable sample size on extraction efficiency. The lack of any relationship between sample mass and measured hormone concentration (no small sample effect) is encouraging, given reports that the small sample effect may occur widely in non-plasma sample types (Berk et al., 2016; Hayward et al., 2010), though we recognize the limitations of our current data set. Removal of overlying seawater may be necessary to prevent dried salts from inflating the apparent fecal mass, but this procedure may also remove some hormone that might have migrated from the fecal pellets into the seawater. Future studies could test this possibility via analysis of the seawater that is poured off of each sample.

4.2. Assay validations

Despite good parallelism, four of the five hormone assays initially demonstrated poor accuracy. In three of these cases, screening samples for algal contamination eliminated the issue. It may, therefore, be advisable to routinely inspect all cetacean fecal samples for the potential presence of non-fecal material that could interfere with assay accuracy. However, our fT RIA continued to have poor accuracy even after algal contamination was eliminated as a factor. This result was surprising given this assay's excellent parallelism and the fact that this fT assay has been used successfully for feces in another baleen whale (NARW; Rolland et al., 2005) as well as multiple other mammalian taxa (e.g. Brown, 1997; Hesterman et al., 2005). The same fT assay also performs poorly for humpback respiratory vapor (Hunt et al., unpublished data). Substituting another assay (the fT EIA) ultimately solved this problem. Our results exemplify the importance of testing assay accuracy, and not just parallelism, when studying an unusual sample type in a new species.

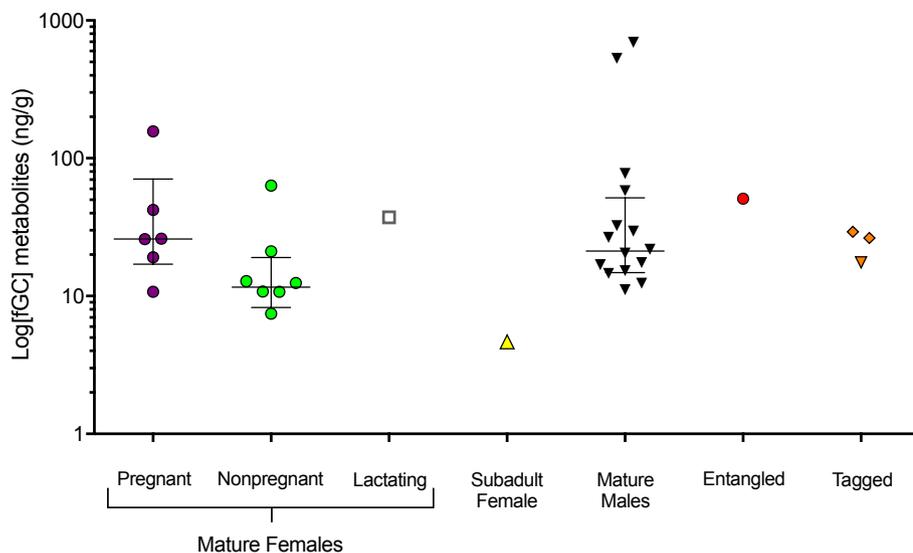


Fig. 4. Fecal glucocorticoid metabolites of humpback whales categorized by sex, age class and reproductive state. Pregnant or nonpregnant status was determined by fP threshold value. Bars represent medians \pm interquartile ranges. “Tagged” category includes two samples from a whale satellite-tagged 1 or 2 days prior to sampling (orange diamonds) and a single sample from a whale satellite-tagged 2 years prior to sampling, though still with a swelling at the site of the tag (orange triangle). Note logarithmic scale of y-axis.

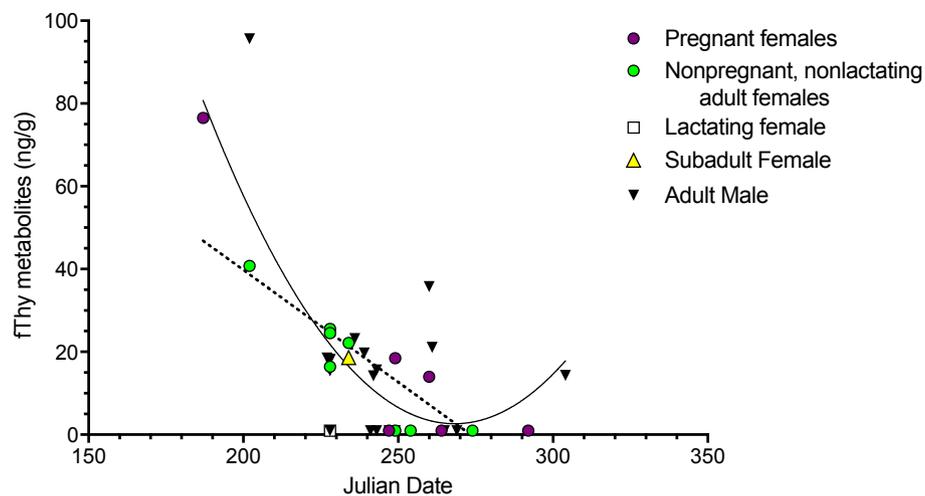


Fig. 5. Fecal thyroid hormone metabolites (fThy) of humpback whales relative to Julian date of sample collection. Best-fit linear regression line (dashed) and best-fit quadratic line (solid) are shown. Pregnant or nonpregnant status was determined by fP threshold value.

4.3. Physiological validations

Despite small sample sizes, most hormones showed the predicted differences among demographic categories and season. fP was higher in Resighted-With-Calf females than in Resighted-Without-Calf females, as expected. There were two instances of high fP in samples from females who were not known to have calved, but observational data could not exclude pregnancy in these cases. Rather, this result may indicate loss of pregnancy or death of the calf, though pseudopregnancy and extended luteal phases are also possibilities (Pallin et al., 2018). Routine endocrine evaluation via feces, blubber or respiratory vapor (as in NARW; Burgess et al., 2018) could more accurately estimate humpback pregnancy rate and calf mortality than has been possible from sightings data alone.

Fecal estrogens showed no differences between pregnant females, nonpregnant females, and adult males. Fecal estrogens have been reported in NARW to be higher in lactating mothers as compared to resting females, and also tend, in that species, to be higher in mature females than immature females. We were unable to test for these patterns in humpbacks because our sample set included no juvenile females and only one lactating mother. Greater sample sizes will be necessary to further explore the utility of fE analysis for identification of sex, age class or reproductive state.

Both fT and fT/fE ratio were both higher in adult males than in nonpregnant females, as expected, but there was pronounced overlap in fT and fE concentration across all categories of adult whales. Post-hoc analyses suggested that pregnant females may have elevated fecal androgens; one pregnant female, in fact, had higher fT than all adult males sampled. Significantly elevated fecal androgens in pregnant females have also been documented in NARW (Rolland et al., 2005). A two-step analysis of data, first excluding high-progesterone (likely pregnant) samples and then assigning likely sex to the remaining samples may prove fruitful for determining sex of unknown individuals (also see methods of Corkeron et al., 2017). Sample sizes precluded evaluation of such an approach for humpback whales, but future studies could explore this possibility. We were also unable in this study to assess fT as an index of male maturity, since no juvenile males were sampled, but this question too can be explored in future studies.

Glucocorticoids showed expected correlations with reproductive states. fGC was significantly higher in pregnant females and mature males as compared to nonpregnant females, a pattern seen in many other mammals (Dantzer et al., 2010; Fortune et al., 2013). fGC was also mildly elevated in a health-impaired entangled whale. Though the elevation was mild and this is only a single anecdotal case, similar patterns have been reported for entangled NARW (Hunt et al., 2006; Rolland et al., 2017) and wounded southern right whales, *Eubalaena*

australis (Fernández Ajó et al., 2018). These findings suggest that fGC measures may be broadly applicable as a measure of adrenal activity in mysticete whales. With this in mind, the relatively low fecal glucocorticoids measured in the two satellite-tagged whales might indicate that tagging was not a major physiological stressor at the time of sampling. However, caution is always advisable when interpreting low fGC, since long-term chronic stress can depress glucocorticoid concentrations (Dickens and Romero, 2013). Additional samples will be necessary to interpret these results and determine the normal range of fGC with greater certainty. Likewise, more samples from whales known to be stressed or in poor health will help determine whether fecal glucocorticoids do in fact reflect stress in humpback whales, and if so, what degree of elevation would indicate a biologically significant event.

Little information exists on thyroid hormones in the mysticete whales, with published data largely limited to reports of detectability of thyroxine and tri-iodothyronine in respiratory vapor (NARW; Hunt et al., 2014) and baleen (multiple species; Hunt et al., 2017; Lysiak et al., 2018). Analysis of thyroid hormone metabolites from mysticete feces has not been reported before. In odontocete plasma and feces, thyroid hormones can demonstrate seasonal patterns in nutritional state, thermal environment and related metabolic and energetic needs (beluga, St Aubin and Geraci, 1989; killer whale, Ayres et al., 2012), while in some species there may be effects of age, sex or reproductive state (bottlenose dolphin, St Aubin et al., 1996). In this humpback whale population, fThy showed a significant relationship with Julian date, i.e., a potential seasonal cycle with a nadir in late summer, with no differences attributable to age-class, sex or reproductive state. A nearly identical pattern has been reported in fThy analyses of southern resident killer whales (*Orcinus orca*) feeding in the North Pacific (Ayres et al., 2012), i.e. a U-shaped curve characterized by a nadir in mid-August. This result has been interpreted as evidence that southern resident killer whales may have access to a high-quality food source in spring, when fThy was highest. Seasonal peaks in thyroid hormones may also correlate with changes in energy expenditure (Behringer et al., 2018). For example, thyroid hormones in wild vertebrates frequent exhibit positive correlations with thermoregulatory requirements (e.g., higher in winter), seasonal gametogenesis, pregnancy, juvenile growth, and migration (reviewed in Behringer et al., 2018; see also Pérez et al., 2016). Any of these patterns might occur in mysticete whales, which have life cycles characterized by annual migrations, seasonal changes in thermal environment, seasonal breeding, long gestation, and a prolonged period of juvenile growth (Chivers, 2009). In this study, sample sizes were too small to discriminate among these possibilities. We recommend that fThy be further investigated in future studies of mysticetes, especially in relation to season, migratory movement, reproductive state, age class, prey availability, and individual body condition.

4.4. High hormones in samples from calves

The three calf fecal samples all had remarkably high fecal hormone content, often exceeding concentrations seen for adults. Fecal hormone concentrations from nursing mammals are rarely reported, but available evidence suggests that it is not uncommon for nursing young to have very high fecal hormone content (Faya et al., 2013; Hunt, unpublished data; Thompson and Lampl, 2013; Thompson et al., 2010a,b). These patterns may reflect milk transfer of maternal hormones. Steroid and thyroid hormones are lipophilic and may be transferred to the calf via milk fat, and then concentrated in the calf's feces. Alternatively, some of these hormones may be of calf origin. It may be relevant that at least two of the sampled calves were male. Neonatal male mammals exhibit a period of extremely high gonadal steroid production shortly after birth, the “neonatal testosterone surge,” whose function is still unclear. The neonatal testosterone surge occurs in a wide variety of taxa including humans (*Homo sapiens*), horses (*Equus caballus*), rats (*Rattus norvegicus*), cats (*Felis catus*), and at least

one artiodactyl, red deer (*Cervus elaphus*) (Corbier et al., 1991; Pavitt et al., 2014). Female neonates may show a corresponding, though somewhat smaller, neonatal surge of estrogens. In cats, for example, both sexes of kittens have higher fecal concentrations of both androgens and estrogens than adults during the first weeks of life (Faya et al., 2013). Data on other classes of hormones in neonatal mammals are lacking. As fecal samples from nursing young are commonly excluded from analysis in many fecal hormone studies, it is difficult to ascertain whether this pattern is widespread. We encourage other researchers to publish fecal hormone data from neonatal mammals and nursing young.

4.5. Conclusions

Comparative aspects of endocrinology of the mysticete whales have been understudied for many decades, due primarily to lack of validated field methods and analytic techniques. Yet the mysticetes have a unique combination of life history traits that may make them useful for study of basic comparative questions of mammalian endocrinology, should suitable methodology be developed. Among their unusual traits are long lifespan (approaching and in some cases exceeding 100 yrs); long gestation length and multi-year inter-calving cycles; annual migrations; strongly seasonal reproduction (i.e., seasonal calving, seasonal testosterone cycles in males; Hunt et al., 2018); and seasonal fasting in many species (Chivers, 2009). Finally, the physiological diving response as well as the requirements of osmotic balance in a marine environment lead to some unique aspects of catecholamine release, adrenal activity and some alterations in the renin-angiotensin-aldosterone system, as compared to terrestrial mammals (Atkinson et al., 2015). This unique combination of traits could make the mysticete whales especially informative for comparative endocrinology studies.

Our study shows that fecal hormone analysis may be broadly useful for such studies in mysticetes. Despite very small sample mass and concomitant difficulty of sample collection in humpback whales, we were able to retrieve enough sample for routine assay of five hormones via immunoassay, including fecal thyroid hormones (not before reported for any baleen whale). We caution, however, that sample collection and processing techniques warrant close attention. Fecal sampling is, by its very nature, an opportunistic technique and so individuals cannot be sampled at will. Methodological details such as collection device, algal contamination, removal of excess water, and solvent:sample ratio may all introduce important variation if not evaluated and controlled. Further, for each new species it is essential to perform not only assay validations (accuracy as well as parallelism), but also physiological validations from individuals of known state, before the technique can be reliably applied to questions of conservation and ecological interest. Overall, however, our results indicate that fecal hormone quantification may be broadly useful even for those aquatic animals that produce highly particulate and dispersed feces, potentially enabling completely noninvasive assessment of stress physiology, reproductive state, energetic physiology, and other aspects of comparative endocrinology.

Funding

This work was supported by (1) Pacific Life Foundation Marine Mammal Research Fund at The Ocean Foundation, (2) the Anderson Cabot Center for Ocean Life at the New England Aquarium, (3) the Center for Bioengineering Innovation at Northern Arizona University and (4) parallel research by the Center for Coastal Studies.

Acknowledgments

We are indebted to the research staff of the Center for Coastal Studies (CCS), especially David Mattila, Scott Landry and Jennifer Tackaberry. We also thank Gulf of Maine Humpback Whale Catalog contributors and collaborators, especially Boston Harbor Cruises, Blue

Ocean Society, Brier Island Whale and Seabird Cruises, the Dolphin Fleet, Whale and Dolphin Conservation and Whale Center of New England. The CCS Marine Animal Entanglement Response program provided the sample from the entangled whale. Samples from tagged whales were from parallel research by CCS, the Australian Antarctic Division, the Marine Mammal Laboratory/Alaska Fisheries Science Center and the Marine Mammal Center supported by National Oceanographic and Atmospheric Administration (NOAA), ExxonMobile Exploration Company, the National Fish and Wildlife Foundation, the National Oceanographic Partnership Program and the Massachusetts Environmental Trust. We particularly thank Alex Zerbini for his involvement. We are also grateful to Jodie Treloar and Danielle Dillon for technical assistance with hormone assays, and Per Palsbøll and Wenzhi Hao for molecular genetic sexing. Research was conducted under NOAA research permits # 16325, 14245, 633-1778, 932-1905 and approved by the Institutional Animal Care and Use Committee (IACUC) of the New England Aquarium (protocol #2012-05).

References

- Amaral, R., 2010. Use of alternative matrices to monitor steroid hormones in aquatic mammals: a review. *Aquatic Mamm.* 36, 162–171.
- Atkinson, S., Crocker, D., Houser, D., Mashburn, K., 2015. Stress physiology in marine mammals: how well do they fit the terrestrial model? *J. Comp. Physiol. B* 185, 463–486.
- Ayres, K., Booth, R., Hempelmann, J., Koski, K., Emmons, C., Baird, R., Balcomb-Bartok, K., Hanson, M., Ford, M., Wasser, S.K., 2012. Distinguishing the impacts of inadequate prey and vessel traffic on an endangered killer whale (*Orcinus orca*) population. *PLoS One*. <https://doi.org/10.1371/journal.pone.0036842>.
- Behringer, V., Deimel, C., Hohmann, G., Negrey, J., Schaebs, F.S., Deschner, T., 2018. Applications for non-invasive thyroid hormone measurements in mammalian ecology, growth, and maintenance. *Horm. Behav.* 105, 66–85.
- Berk, S., McGettrick, J., Hansen, W., Breuner, C., 2016. Methodological considerations for measuring glucocorticoid metabolites in feathers. *Conserv. Physiol.* 4, cow020.
- Bérubé, M., Palsbøll, P., 1996a. Identification of sex in cetaceans by multiplexing with three ZFX and ZFY specific primers. *Molec. Ecol.* 5, 283–287.
- Bérubé, M., Palsbøll, P., 1996b. Identification of sex in cetaceans by multiplexing with three ZFX and ZFY specific primers: erratum. *Molec. Ecol.* 5, 602.
- Brown, J., 1997. Fecal steroid profiles in black-footed ferrets exposed to natural photoperiod. *J. Wildl. Manage.* 61, 1428–1436.
- Brown, J., 2018. Comparative ovarian function and reproductive monitoring of endangered mammals. *Theriogenology* 109, 2–13.
- Burgess, E.A., Hunt, K.E., Rolland, R.M., Kraus, S.D., 2017. Adrenal responses of large whales: integrating fecal aldosterone as a complementary biomarker. *Gen. Comp. Endocrinol.* 252, 103–110.
- Burgess, E.A., Hunt, K.E., Kraus, S.D., Rolland, R.M., 2018. Quantifying hormones in exhaled breath for physiological assessment of large whales at sea. *Sci. Rep.* 8, 10031. <https://doi.org/10.1038/s41598-018-28200-8>.
- Chivers, S.J., 2009. Cetacean life history. In: Perrin, W.F., Würsig, B., Thewissen, J.G.M. (Eds.), *Encyclopedia of Marine Mammals*, 2nd ed. Academic Press, New York, pp. 215–220.
- Chittleborough, R.G., 1958. The breeding cycle of the female humpback whale, *Megaptera nodosa* (Bonnaterre). *Aust. J. Mar. Freshwater Res.* 9 (1), 1–18.
- Clapham, P., 1992. Age at attainment of sexual maturity in humpback whales, *Megaptera novaeangliae*. *Can. J. Zool.* 70, 1470–1472.
- Cohen, J., 1988. *Statistical Power Analysis for The Behavioral Sciences*, second ed. Lawrence Erlbaum Associates, New Jersey.
- Corbier, P., Edwards, D., Roffi, J., 1991. The neonatal testosterone surge: a comparative study. *Arch. Internat. Physiol. Biochim. Biophys.* 100, 127–131.
- Corkeron, P.J., Rolland, R.M., Hunt, K.E., Kraus, S.D., 2017. A right whale PooTree: fecal hormones and classification trees identify reproductive states in North Atlantic right whales (*Eubalaena glacialis*). *Cons. Physiol.* 5, cow006.
- Cristobal-Azkarate, J., Marechal, L., Semple, S., Majolo, B., MacLarnon, A., 2016. Metabolic strategies in wild male Barbary macaques: evidence from faecal measurement of thyroid hormone. *Biol. Lett.* <https://doi.org/10.1098/rsbl.2016.0168>.
- Dantzer, B., McAdam, A.G., Palme, R., Fletcher, Q.E., Boutin, S., Humphries, M.M., Boonstra, R., 2010. Fecal cortisol metabolite levels in free-ranging North American red squirrels: assay validation and the effects of reproductive condition. *Gen. Comp. Endocrinol.* 167, 279–286.
- Dantzer, B., Fletcher, Q., Boonstra, R., Sheriff, M., 2014. Measures of physiological stress: a transparent or opaque window into the status, management and conservation of species? *Conserv. Physiol.* 2, cow023.
- Dickens, M.J., Romero, L.M., 2013. A consensus endocrine profile for chronically stressed wild animals does not exist. *Gen. Comp. Endocrinol.* 191, 177–189.
- Fair, P., Becker, P., 2000. Review of stress in marine mammals. *J. Aquat. Eco. Stress Recov.* 7, 335–354.
- Faya, M., Carranza, A., Miotti, R., Ponchon, T., Furian, P., Cobello, C., 2013. Fecal estradiol-17 β and testosterone in prepubertal domestic cats. *Theriogenology* 80, 584–586.
- Fernández Ajó, A., Hunt, K.E., Uhart, M., Rowntree, V., Sironi, M., Marón, C.F., Di Martino, M., Buck, C.L., 2018. Lifetime glucocorticoid profiles in baleen of right whale calves: potential relationships to chronic stress of repeated wounding by Kelp Gulls. *Conserv. Physiol.* 6: coy045, [10.1093/conphys/coy045](https://doi.org/10.1093/conphys/coy045).
- Fortune, S.M.E., Trites, A.W., Mayo, C.A., Rosen, D.A.S., Hamilton, P.K., 2013. Energetic requirements of North Atlantic right whales and implications for species recovery. *Mar. Ecol. Progr. Ser.* 478, 253–272.
- Gesquiere, L., Pugh, M., Alberts, S., Markham, A., 2018. Estimation of energetic condition in wild baboons using fecal thyroid hormone determination. *Gen. Comp. Endocrinol.* 260, 9–17.
- Glockner, D.A., 1983. Determining the sex of humpback whales (*Megaptera novaeangliae*) in their natural environment. *Commun. Behav. Whales* 447–464.
- Goymann, W., 2012. On the use of non-invasive hormone research in uncontrolled, natural environments: the problem with sex, diet, metabolic rate and the individual. *Meth. Ecol. Evol.* 3, 757–765.
- Grotjan, H.E., Keel, B.A., 1996. Data interpretation and quality control. In: Diamandis, E.P., Christopoulos, T.K. (Eds.), *Immunoassay*. Academic Press, San Diego, CA, pp. 51–95.
- Hayward, L., Booth, R., Wasser, S., 2010. Eliminating the artificial effect of sample mass on avian fecal hormone metabolite concentration. *Gen. Comp. Endocrinol.* 169, 117–122.
- Heistermann, M., Palme, R., Ganswindt, A., 2006. Comparison of different enzyme immunoassays for assessment of adrenocortical activity in primates based on fecal analysis. *Am. J. Primatol.* 68, 257–273.
- Hesterman, H., Wasser, S., Cockrem, J., 2005. Longitudinal monitoring of fecal testosterone in male Malayan sun bears (*U. malayanus*). *Zoo Biol.* 24, 403–417.
- Hunt, K., Lysiak, N., Moore, M., Seton, R., Torres, L., Buck, C., 2017. Multiple steroid and thyroid hormones detected in baleen from eight whale species. *Conserv. Physiol.* cow061.
- Hunt, K.E., Lysiak, N.S., Matthews, C.J.D., Lowe, C., Fernández-Ajo, A., Dillon, D., Willing, C., Heide-Jørgensen, M.P., Ferguson, S.H., Moore, M.J., Buck, C.L., 2018. Multi-year patterns in testosterone, cortisol and corticosterone in baleen from adult males of three whale species. *Conserv. Physiol.* <https://doi.org/10.1093/conphys/coy049>. 6: coy049.
- Hunt, K., Trites, A., Wasser, S., 2004. Validation of a fecal glucocorticoid assay for Steller sea lions (*Eumetopias jubatus*). *Physiol. Behav.* 80, 595–601.
- Hunt, K., Wasser, S., 2003. Effect of long-term preservation methods on fecal glucocorticoid concentrations of grizzly bear and African elephant. *Physiol. Biochem. Zool.* 76, 918–928.
- Hunt, K.E., Moore, M.J., Rolland, R.M., Kellar, N.M., Hall, A.J., Kershaw, J., Raverty, S.A., Davis, C.E., Yeates, L.C., Fauquier, D.A., 2013. Overcoming the challenges of studying conservation physiology in large whales: a review of available methods. *Conserv. Physiol.* 1, cot006.
- Hunt, K.E., Rolland, R.M., Kraus, S.D., 2014. Detection of steroid and thyroid hormones via immunoassay of North Atlantic right whale (*Eubalaena glacialis*) respiratory vapor. *Mar. Mamm. Sci.* 30, 796–809.
- Hunt, K.E., Rolland, R.M., Kraus, S.D., Wasser, S.K., 2006. Analysis of fecal glucocorticoids in the North Atlantic Right Whale (*Eubalaena glacialis*). *Gen. Comp. Endocrinol.* 148, 260–272.
- Ichii, T., Kato, H., 1991. Food and daily food consumption of southern minke, 442 whales in the Antarctic. *Polar Biol.* 11, 479–487.
- Katona, S., Whitehead, H., 1981. Identifying humpback whales using their natural markings. *Polar Record* 20, 439–444.
- Kersey, D., Dehnhard, M., 2014. The use of noninvasive and minimally invasive methods in endocrinology for threatened mammal species conservation. *Gen. Comp. Endocrinol.* 203, 296–306.
- Lemos, L., Green, A., Smith, A., Burnett, J., Chandler, T., Torres, L.G., 2017. Combining traditional and novel techniques to link body condition and hormone variability in gray whales. *Proceedings of the 22nd Biennial Society for Marine Mammalogy Conference on the Biology of Marine Mammals*. October 2017, Halifax, Canada.
- Lysiak, N., Trumble, S., Knowlton, A., Moore, M., 2018. Characterizing the duration and severity of fishing gear entanglement on a North Atlantic right whale (*Eubalaena glacialis*) using stable isotopes, steroid and thyroid hormones in baleen. *Front. Mar. Sci.* <https://doi.org/10.3389/fmars.2018.00168>.
- Madliger, C., Love, O., Hultine, K., Cooke, S., 2018. The conservation physiology toolbox: status and opportunities. *Conserv. Physiol.* 6, coy029.
- Matthews, L., 1937. The humpback whale, *Megaptera nodosa*. *Discov. Rep.* 17, 7–92.
- Murphy, R., 2017. On the use of one-sided statistical tests in biomedical research. *Clin. Exp. Pharmacol. Physiol.* 45, 109–114.
- National Academies of Sciences, 2016. In: *Approaches to Understanding the Cumulative Effects of Stressors of Marine Mammals*. National Academies of Sciences Committee on the Assessment of the Cumulative Effects of Anthropogenic Stressors on Marine Mammals; Ocean Studies Board; Division on Earth and Life Studies, Washington, D.C., pp. 250.
- Overholtz, W., Nicolas, J., 1979. Apparent feeding by the fin whale, *Balaenoptera physalus*, and humpback whale, *Megaptera novaeangliae*, on the American sand lance, *Ammodytes americanus*, in the northwest Atlantic. *Fish. Bull.* 77, 285–287.
- Pallin, L., Robbins, J., Kellar, N.M., Berube, M., Friedlaender, A., 2018. Validation of a blubber-based endocrine pregnancy test for humpback whales. *Conserv. Physiol.* 6, coy031.
- Palme, R., Fischer, P., Schildorfer, H., Ismail, M., 1996. Excretion of infused ¹⁴C-steroid hormones via faeces and urine in domestic livestock. *Anim. Reprod. Sci.* 43, 43–63.
- Palsbøll, P., Veder, A., Bakke, I., Raafat El-Gewely, M., 1992. Determination of gender in cetaceans by the polymerase chain reaction. *Can. J. Zool.* 70, 2166–2170.
- Pappano, D., Roberts, E., Beehner, J., 2010. Testing extraction and storage parameters for a fecal hormone method. *Am. J. Primatol.* 72, 934–941.
- Pavitt, A., Wailing, C., McNeilly, A., Pemberton, J., Kruuk, L., 2014. Variation in early-life

- testosterone within a wild population of red deer. *Funct. Ecol.* 28, 1224–1234.
- Pérez, J.H., Furlow, J.D., Wingfield, J.C., Ramenofsky, M., 2016. Regulation of vernal migration in Gambel's white-crowned sparrows: role of thyroxine and tri-iodothyronine. *Horm. Behav.* 84, 50–56.
- Ploutz-Snyder, R.J., Fiedler, J., Feiveson, A.H., 2014. Justifying small-n research in scientifically amazing settings: challenging the notion that only “big-n” studies are worthwhile. *J. Appl. Physiol.* 116, 1251–1252.
- Robbins, J., 2007. Structure and Dynamics of the Gulf of Maine humpback Whale Population, Biology. St. Andrews University Ph.D. thesis.
- Robbins, J., Zerbini, A.N., Gales, N., Gulland, F.M.D., Double, M., Clapham, P.J., Andrews-Goff, V., Kennedy, A.S., Landry, S., Mattila, D.K., Tackaberry, J., 2013. Satellite tag effectiveness and impacts on large whales: preliminary results of a case study with Gulf of Maine humpback whales. Report to the Scientific Committee of the International Whaling Commission. SC-65a-SH05.
- Rolland, R., McLellan, W.A., Moore, M.J., Harms, C.A., Burgess, E.A., Hunt, K.E., 2017. Fecal glucocorticoids and anthropogenic injury and mortality in North Atlantic right whales (*Eubalaena glacialis*). *Endang. Sp. Res.* 34, 417–429.
- Rolland, R.M., Hunt, K.E., Kraus, S.D., Wasser, S.K., 2005. Assessing reproductive status of right whales (*Eubalaena glacialis*) using fecal hormone metabolites. *Gen. Comp. Endocrinol.* 142, 308–317.
- Rolland, R.M., Parks, S.E., Hunt, K.E., Castellote, M., Corkeron, P.J., Nowacek, D.P., Wasser, S.K., Kraus, S.D., 2012. Evidence that ship noise increases stress in right whales. *Proc. Roy. Soc. B: Biol. Sci.* 279, 2363–2368.
- Ruxton, G.D., Neuhäuser, M., 2010. When should we use one-tailed hypothesis testing? *Meth. Ecol. Evol.* 1, 114–117.
- Schaeb, F., Wolf, T., Behringer, V., Deschner, T., 2016. Fecal thyroid hormones allow for the non-invasive monitoring of energy intake in capuchin monkeys. *J. Endocrinol. JOE-16*.
- Siegel, S., Castellan Jr., N.J., 1988. *Nonparametric Statistics for the Behavioral Sciences*. McGraw-Hill, New York.
- St Aubin, D., Geraci, J., 1989. Seasonal variation in thyroid morphology and secretion in the white whale, *Delphinapterus leucas*. *Can. J. Zool.* 67, 263–267.
- St Aubin, D., Ridgway, S., Wells, R., Rhinehart, H., 1996. Dolphin thyroid and adrenal hormones: circulating levels in wild and semidomesticated *Tursiops truncatus*, and influence of sex, age and season. *Mar. Mamm. Sci.* 12, 1–13.
- Thomas, P., Reeves, R., 2015. Status of the world's baleen whales. *Mar. Mamm. Sci.* 32, 682–734.
- Thompson, A., Lampl, M., 2013. Prenatal and postnatal energetic conditions and sex steroids levels across the first year of life. *Am. J. Human Biol.* 25, 643–654.
- Thompson, A., Whitten, P., Johnson, M., Lampl, M., 2010a. Non-invasive methods for estradiol recovery from infant fecal samples. *Front. Physiol. Syst. Biol.* 1, 161.
- Thompson, A., Whitten, P., Lampl, M., 2010b. Measurement of testosterone in infant fecal samples. *Am. J. Human Biol.* 23, 820–822.
- Thompson, C., Powell, B., Williams, S., Hanya, G., Glander, K., Vinyard, C., 2017. Thyroid hormone fluctuations indicate a thermoregulatory function in both a tropical (*Alouatta palliata*) and seasonally cold-habitat (*Macaca fuscata*) primate. *Am. J. Primatol.* <https://doi.org/10.1002/ajp.22714>.
- Valenzuela-Molina, M., Atkinson, S., Mashburn, K., Gendron, D., Brownell, R., 2018. Fecal steroid hormones reveal reproductive state in female blue whales sampled in the Gulf of California, Mexico. *Gen. Comp. Endocrinol.* 261, 127–135.
- Wasser, S., Azkarate, J., Booth, R., Hayward, L., Hunt, K., Ayres, K., Vynne, C., Gobush, K., Canales-Espinosa, D., Rodriguez-Luna, E., 2010. Non-invasive measurement of thyroid hormone in feces of a diverse array of avian and mammalian species. *Gen. Comp. Endocrinol.* 168, 1–7.
- Wasser, S., Lundin, J., Ayres, K., Seely, E., Giles, D., Balcomb, K., Hempelmann, J., Parsons, K., Booth, R., 2017. Population growth is limited by nutritional impacts on pregnancy success in endangered Southern Resident killer whales (*Orcinus orca*). *PLoS One.* <https://doi.org/10.1371/journal.pone.0179824>.
- Wasser, S.K., Hunt, K.E., Brown, J.L., Cooper, K., Crockett, C.M., Bechert, U., Millspaugh, J.J., Larson, S., Monfort, S.L., 2000. A generalized fecal glucocorticoid assay for use in a diverse array of nondomestic mammalian and avian species. *Gen. Comp. Endocrinol.* 120, 260–275.