Steroid Hormone Variation and Stress Responses in Short-finned Pilot Whales

by

Jillian Wisse

University Program in Ecology
Duke University

Date:_____

Approved:

_____
Douglas P. Nowacek, Supervisor

_____
Andrew J. Read

_____
Christine M. Drea

_____
Ashley S.P. Boggs

_____
Nicholas M. Kellar

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the University Program in Ecology in the Graduate School
of Duke University

2023

ABSTRACT

Steroid Hormone Variation and Stress Responses in Short-finned Pilot Whales

by

Jillian Wisse

University Program in Ecology
Duke University

Date:_____
Approved:

_____
Douglas P. Nowacek, Supervisor

_____
Andrew J. Read

_____
Christine M. Drea

_____
Ashley S.P. Boggs

_____
Nicholas M. Kellar

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the University Program in Ecology in the Graduate School
of Duke University

2023

Abstract

As humans continue to introduce stressors into the marine environment, we are obligated to understand how our behaviors impact wildlife. For many cetacean populations, anthropogenic noise poses a significant conservation threat, but the ability to monitor these animals is constrained by their often-remote habitats and limited time at the surface. Researchers have developed innovative solutions to overcome these challenges, including the development of techniques that enable physiological sampling with minimal disturbance. As frontiers, these methods require careful development and validation before they can be used reliably in experimental studies.

In this dissertation, I employ one of these innovative techniques, remote blubber biopsy, for steroid hormone measurement in short-finned pilot whales (*Globicephala macrorhynchus*). As mediators of reproduction and stress, steroid hormones provide information that is advantageous for wildlife monitoring. Because a validated method for measuring these compounds in blubber from short-finned pilot whales did not yet exist, I adapted a liquid chromatography-tandem mass spectrometry (LC-MS/MS) method to simultaneously quantify 11 steroid hormones of interest. Before proceeding with its application, I tested its analytical and biological validity with blubber from three stranded short-finned pilot whales. Next, I applied the validated LC-MS/MS method to an archive of blubber biopsies, collected from free-ranging short-finned pilot whales in the western North Atlantic. Leveraging the comprehensive nature of LC-MS/MS profiling, I investigated relationships between hormones and characterized steroid hormone profiles across demographic groups and seasons. In the fourth chapter, I conducted an acoustic

response study, using the previously established methods to collect and quantify steroid hormones after exposure to simulated mid-frequency active sonar (MFAS). I modeled the responses of cortisol and cortisone over time to gain insight into steroid perfusion rates in cetacean blubber and asked whether demography contributed to these responses.

Together, my results substantiate pilot whale blubber as a reliable matrix for measuring steroid hormones and reflecting relevant physiological states, like stress and pregnancy. Although the described LC-MS/MS method struggled to detect some steroids of interest in this matrix, it accomplished concurrent multi-group steroid measurements. This dissertation shows the relevance of multi-steroid profiling and offers reference points for baseline steroid concentrations in analytes relevant to behavior, reproduction, and stress. Lastly, post-exposure glucocorticoid measurements support the hypothesized relationship between noise and physiological response in short-finned pilot whales and demonstrate the applicability of blubber sampling in this context.

Contents

List of Tables

List of Figures

Acknowledgements

The research would not have been possible without the support of my mentors, collaborators, friends, and family. During various stages of my career, I've found community and growth in several lab groups, including those of Christine Drea, Janet Mann, Doug Nowacek, and Andy Read. These faculty and their groups provided inspiration, encouragement, and opportunities to work with more incredible animals than I could have ever dreamed possible.

I am forever grateful to my PhD advisor, Doug Nowacek, who never lost faith in my research questions or the value of the data. I feel incredibly lucky to have had his generous support in building collaborations, funding data collection and analysis, and navigating the hurdles of research. His sense of humor, empathy, and genuine enthusiasm for inter-disciplinary work are qualities I aspire to carry throughout my career.

Andy Read provided mentorship well-beyond what can reasonably be asked of a committee member. The feedback he provided on my work was invaluable and the opportunity to muse with him about the lives of whales was a privilege. I could not have completed my work without the archive of pilot whale biopsies, which he built with his team over two decades, and I look forward to working together to tease apart the questions these samples continue to generate. I'm indebted to the incredible field team at Duke, specifically Danielle Waples, Zach Swaim, and Heather Foley, who were fundamental in data collection and interpretation. They brought me on countless field adventures and taught me the careful art of biopsy collection. Among many who contributed to biopsy collection over the years, I am thankful for Will Cioffi, Matt

Dawson, Julian Dale, Nicola Quick, Kim Urian, Matt Bowers, Greg Merrill, and charter boat captains Reed Meredith (The KAHUNA) and Smitty (The Spray).

The data analysis conducted at NIST was only possible through Ashley Boggs, who generously taught me to conduct the LC-MS/MS analysis method. Her intentional mentorship formed me into a bench scientist, encouraged me to harness my personal strengths, and demonstrated what it looks like to 'walk the walk' in creating inclusive lab spaces. The folks at NIST, including John Kucklick, Kevin Huncik, Jared Ragland, Erin Legacki, Thomas Galligan, and Greta Dalle Luche provided support with logistics, instrumentation, and data analysis.

The blubber research conducted in this dissertation was inspired and enabled by the foundations Nick Kellar has laid through his scholarship. This dissertation would look much different without his willingness to discuss blubber, hormones, and behavior for hours on the fly. His encouragement and belief in the applications of these methods was a North Star throughout my Ph.D. work.

My interest in endocrinology was sparked while working in Christine Drea's lab as an undergraduate and I was fortunate to have her on my Ph.D. committee. The detailed feedback, insight, and support she and Lydia Greene provided during my Ph.D. helped ground my work within larger and evolutionary framework.

I am also grateful to Janet Mann's lab at Georgetown, which brought me into the world of marine mammalogy. During my time there, Ewa Krzyszczyk, Madison Miketa, Vivienne Foroughirad, Anni Jacoby, Meg Wallen, Caitlin Karniski, and Eric Patterson

and introduced me to the dolphins of Shark Bay and showed me what is possible with long-term data sets.

Choosing the Duke Marine Lab for my Ph.D. was one of the best choices I have ever made. As an Ecology student at the Marine Lab, I had two cohorts (UPE and MSC), both of which provided me with colleagues that challenged my ideas and introduced me to lifelong friends. The Superpod group helped me form research ideas and explore questions of ecology, evolution, and conservation. Will Cioffi, Vivienne Foroughirad, Ellie Heywood, Jeanne Shearer, Austin Allen, Joe Fader, Anni Jacoby, Heather Foley, Danielle Waples, Nicola Quick, Ashley Blawas, and Anne Harshbarger were particularly influential and offered constructive feedback. During my Ph.D., I participated in the Bass Connections: Learning from Whales team. Nicola Quick, Jason Somarelli, and Tom Schultz led the team, which continues to provide opportunities for professional growth and fuels my excitement for science. The Duke Marine Lab's interdisciplinary community created a space to question my own ideas. Many of the folks I interacted with during my time here left an indelible mark on me, including Alejandro Garcia Lozano, Hillary Smith, and Heather Vermuelen. Rachel Lo Piccolo, Danielle Wiggins, and many other administrators kept everything afloat and resolved many stressors along the way.

Lastly, my family and friends sustained me in ways I can never repay. My mom, Trish, whose love for animals sparked my own. My father, Jay, and sister, Carly, for being there every step of the way. My community, who encouraged me when the Ph.D. mountain felt insurmountable, who offered joy and respite when they were needed most. Thank you. I love you all.

1. Introduction

1.1 Motivation

The challenges faced by aquatic animals are difficult for humans to conceptualize. The physical properties of water flip the visual and auditory sensory experiences most of us are accustomed to on land. Water absorbs light, limiting its reach below the surface and constraining visibility. On the other hand, water is an excellent sound conductor; under the right conditions, water can carry sound waves across an entire ocean basin (Munk et al., 1994). It is in this environment that cetaceans evolved. From foraging to finding mates, they use sound for some of their most essential behaviors. Unfortunately, this relationship with sound also makes cetaceans especially susceptible to noise disturbance. In recent decades, human activities and the noise they generate have escalated in marine environments (Hildebrand, 2009), creating an increased potential for the disturbance of cetaceans.

The effective management of cetacean populations depends on our ability to measure responses to disturbance and model their downstream effects on individuals and populations. However, many of these species can be difficult to access and spend little time at the surface (Shearer et al., 2019), which has limited our ability to monitor health and model responses to change (Hunt et al., 2013). Recent advancements in tissue sampling and analysis have enabled baseline hormone monitoring in these taxa, setting the stage for more comprehensive physiological studies of stress responses in cetaceans.

Off the North Carolina coast, researchers have been studying short-finned pilot whales (*Globicephala macrorhynchus*) since the early 2000s, documenting sightings and conducting behavioral studies (Bowers et al., 2018; Quick et al., 2017). Around 2006,

they began collecting remote tissue biopsies, analyzing the skin for genetic sex and storing the blubber for future projects. In 2016, the Atlantic Behavioral Response Study began off the North Carolina coast with the goal of testing the behavioral responses of short-finned pilot whales and goose-beaked whales (*Ziphius cavirostris*) to simulated and real mid-frequency active naval sonar (MFAS).

Behavioral response studies (BRSs), like the Atlantic BRS, have predominated the past few decades of research assessing cetacean responses to acoustic stimuli (Harris 2018) and have established causal links between noise exposure and behavioral responses (Southall 2016). Although these behavioral responses result from physiological activation, the logistical constraints of physiological sampling in cetacean systems have historically constrained the implementation of physiological response studies (NAS 2005, Pirotta et al., 2018). It is within the context of recent advances in cetacean endocrine studies that I collaborated with the Atlantic BRS project to document endocrine responses to MFAS.

Because so little is known about baseline endocrinology in short-finned pilot whales, I began by developing a comprehensive method for multi-hormone profiling of their blubber (Chapter 2) and characterizing the baseline endocrinology of short-finned pilot whales in this population (Chapter 3). In Chapter 4, I document short-finned pilot whale responses to an acoustic stressor, including the temporal nature of the stress response in blubber.

1.2 Why hormones?

Hormones regulate physiology and behavior; without them, animals would be unable to reproduce or survive long enough to do so. They act as chemical messengers,

telling our cells which genes to activate, organizing the development of our bodies, timing our feeding habits, and coordinating social behavior (Adkins-Regan, 2005; Norris, 2013). Hormones can exert effects over long time scales, mediating ontological or seasonal shifts, and they can elicit rapid physiological and behavioral responses to environmental challenges (Adkins-Regan, 2005).

Steroid hormones are neuroendocrine compounds. The brain mediates their production via hypothalamic and pituitary pathways that signal endocrine tissue to produce and secrete steroid hormones (Norris, 2013). These hormones are synthesized via steroidogenesis, which begins with cholesterol and transforms hormones with enzymes along interconnected pathways, generally grouped as progestogens, androgens, estrogens, and corticoids (Figure 1) (Norris, 2013). This process primarily occurs in the gonads and adrenals (Norris, 2013), where tissues are specialized to produce specific hormones. Hormones then travel through the bloodstream, reaching cellular receptors in target tissues and eventually signaling back to the hypothalamus. Cellular receptors, like the glucocorticoid receptor and androgen receptor, can bind to multiple hormones within

these groups, and variation in binding affinities moderate the potency of specific hormones and the molecular cascades they initiate (Norris, 2013).

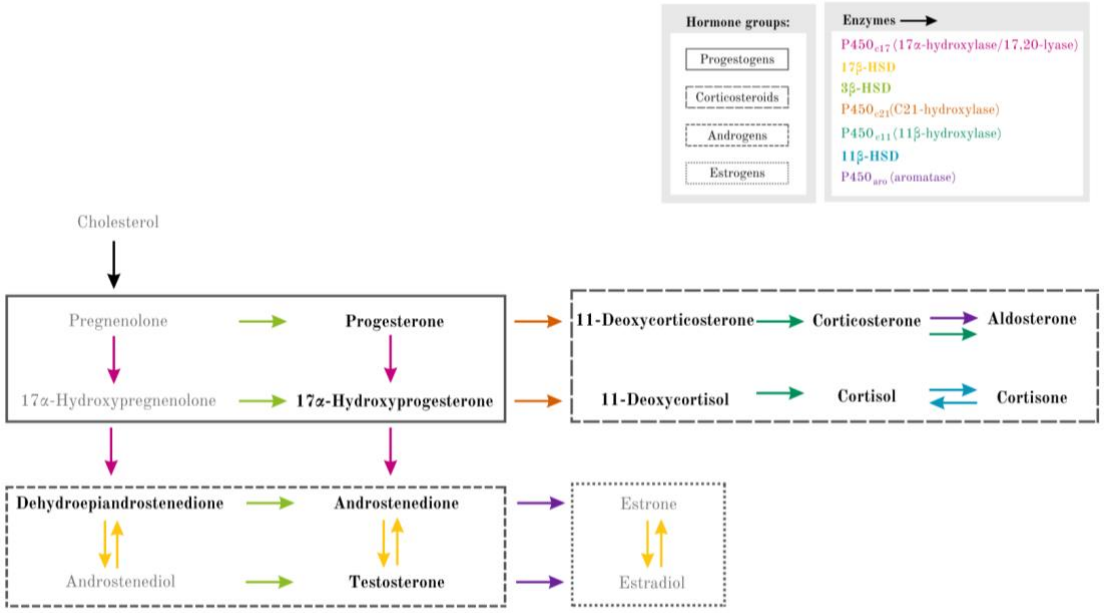

Figure 1. Synthesis pathways of steroid hormones, as mediated by enzymes. The steroid analytes monitored in this dissertation are shown in bold, with boxes designating steroid groups, and major enzymes represented as color-coded arrows. This figure does not include all steroids or enzymes and is based on figures from Miller and Auchus (2011) and Galligan et al. (2018).

Sex steroid (androgens, estrogens, and progestogens) are primarily synthesized in the testes and ovaries, although their synthesis occurs in many tissues (e.g., placentas, adrenal, and adipose). These steroids are under immense evolutionary selection (Norris, 2013), and their roles can be highly variable from one another. Despite popular narratives that have labeled androgens as "male hormones" and estrogens as "female hormones", both groups are present and active across sexes. Hormone concentrations and their effects can vary substantially between males and females (Adkins-Regan, 2005), so throughout this dissertation, I categorize animals by sex as a crude means of understanding the

factors that underlie hormone variation. Because the interplay between hormones, receptor pathways, and environmental factors is complex, sex as a category only scratches the surface of the drivers at play.

In males, androgens are primarily produced by the testes, rising as males approach sexual maturity, aiding in the development of secondary sexual characteristics (Norris, 2013). They can also be synthesized in the adrenals and ovaries (Abraham, 1974) and upregulated during pregnancy (Drea, 2009). Behaviorally, androgens are associated with competition, aiding resource acquisition and defense (van Anders, 2013).

In females, estrogens prepare ovaries for reproduction, then facilitate sexual and maternal behaviors. During the ovulatory cycle, estrogen increases leading up to ovulation, then declines as progesterone increases in preparation for implantation and pregnancy. Progesterone, produced by the corpora lutea and placenta, has long been recognized for supporting mammalian pregnancy (Short, 1960; Short, 1956). In humans, progesterone steadily increases throughout gestation, but that isn't the case for all mammals (Norris, 2013). The progestogens involved in pregnancy maintenance vary between species. In several species of cetaceans, progesterone appears to decrease during the second half of pregnancy, while other progestogens and androgens increase (Legacki et al., 2020).

The endocrine stress response is primarily mediated by three groups of hormones: catecholamines (e.g., epinephrine), glucocorticoids (e.g., cortisol), and mineralocorticoids (e.g., aldosterone), all produced by the adrenals. When an animal first perceives a stressor, catecholamines are the first adrenal hormones released. They contribute to the sympathetic fight-or-flight response by increasing heart rate and respiration and making

glucose quickly available during times of energetic demand (Tort & Teles, 2011). Glucocorticoids are released around the same time, inhibiting glucose use in peripheral tissue and mobilizing fat stores through lipolysis (Norris, 2013; Sapolsky et al., 2000). Ultimately, glucocorticoids make energy available to muscle and nervous tissue (Tort & Teles, 2011), allowing the body to remain alert and responsive. Activation of the stress response inhibits functions that do not immediately affect survival, including immune function and the hypothalamic-pituitary-gonadal axis (Sapolsky et al., 2000).

While the functional roles of hormone groups are evolutionarily conserved across vertebrates, the importance of specific hormones may be variable even between closely related species. Cortisol is the predominant glucocorticoid secreted in most mammals (Tort & Teles, 2011); however, in some animals, such as mice, corticosterone is the principal glucocorticoid (Adkins-Regan, 2005). Mineralocorticoids, like aldosterone, are not typically assumed to contribute substantially to stress responses. However, aldosterone's regulation of salt balance could be especially relevant in cetaceans, which live in saline environments and do not have access to fresh water (Atkinson et al., 2015). Aldosterone increases in response to acute stress in bottlenose dolphins and elephant seals (Champagne et al., 2018; McCormley et al., 2018). Additionally, androgens, dehydroepiandrosterone (DHEA) and dehydroepiandrosterone-sulfate (DHEAS) are produced in the adrenals and have been connected to aging, disease, and long-term stress (Gundlach et al., 2018; Maninger et al., 2010). Understanding acute and chronic stress in marine mammals necessitates comprehensive multi-hormone profiling.

Although stress can elicit negative health consequences, the stress response is not inherently bad. As a physiological mechanism, it accompanies everyday life events like

courtship, parturition, and hunting (Tort & Teles, 2011), and the ability to elicit an acute stress response is critical to an organism's survival. Under normal conditions, the stress response enacts short-term physiological and behavioral changes to resolve the challenge at hand and return to homeostasis. Endocrine secretion is prolonged when the stressor fails to resolve or frequently repeats. This chronic stress can impair growth, reproduction, and immune function (Tort & Teles, 2011). Ultimately, long-term exposure to stressors can have detrimental effects on the fitness of individuals and populations (Blas et al., 2007; Bortolotti et al., 2009).

1.3 Hormone Sampling and Measurement

Hormone measurements can help us figure out when animals reproduce and what portion of a population is reproductive. They can alert us to the environmental challenges faced by populations, and how sex, age, and sociality affect their responses to stress. Over the past several decades, hormones have become relatively widespread biomarkers for studying the behavior and health of wildlife (Amaral, 2010; de Mello & de Oliveira, 2016; Sheriff et al., 2011). Despite the maturation of hormone studies in terrestrial mammals, the application of hormone studies with cetaceans has been constrained by:

(1) The ability to collect samples

(2) Reliability of hormone measurements in samples

(3) Contextualization of hormone measurements

1.3.1 The ability to collect samples

Historically, endocrine studies of live animals have predominantly relied on blood sampling, which has limited focal subjects to those species that could be housed in captivity or temporarily restrained and handled in the wild. The necessity of handling has

limited the number of serum-based endocrine studies of cetaceans. However, blood sampling is not the only sampling method for endocrine studies. Once steroid hormones enter the bloodstream, they are readily transported to other tissues, metabolized, and excreted. Over the past four decades, researchers have explored tissue and excretion alternatives to blood (i.e., alternative matrices), showing hormones to be present and measurable in excrement like feces (Wasser et al., 1988) and urine (Carlstead et al., 1992), and in cornified epithelial tissues like feathers (Bortolotti et al., 2009), nails (Warnock et al., 2010), and hair (Bechshoft et al., 2012).

Cetacean researchers have adapted and expanded upon these methods, using dogs to locate fecal samples (Hunt et al., 2006), drones to capture blow (i.e., exhaled breath condensate (Acevedo-Whitehouse et al., 2010), and modified crossbows and rifles for remote blubber sampling (Lambertsen, 1987). Hormones have been measured in baleen (Hunt et al., 2014) and earwax (Trumble et al., 2013), which incorporate hormones across temporally associated layers; but these tissues can only be collected post-mortem and are specific to certain taxa (e.g., baleen from mysticetes). Over the past 30 years, remote blubber biopsy has become the most common technique for tissue sampling from free-ranging cetaceans (Hunt et al., 2013), often used for dietary, contaminant, and genetic studies (Noren & Mocklin, 2012). This method collects small skin and blubber plugs from free-ranging individuals without the need for capture, resulting in safer conditions for researchers and focal animals. Scientists recently began to analyze these blubber plugs for hormones and the past decade has seen a rapid increase in hormone studies using live cetacean blubber (Atkinson et al., 2019; Boggs et al., 2019; Champagne et al., 2017; Champagne et al., 2018; Dalle Luche et al., 2020; Dalle Luche et al., 2021;

Galligan et al., 2019; Galligan et al., 2020; Goertz et al., 2019; Graham et al., 2021; Kellar et al., 2017; Lemos et al., 2021; Melica et al., 2022; Mingramm et al., 2020; Pérez et al., 2011; Teerlink et al., 2018).

1.3.2 Reliable measurement of hormones in samples

Once samples have been obtained and extracted, researchers typically use one of two methods for hormone analysis: immunoassay or mass spectrometry. Enzyme immunoassay (EIA), a commonly used immunoassay technique, is less expensive, easier to learn, and more sensitive than mass spectrometry. However, each EIA is run for just one analyte and EIAs are susceptible to antibody cross-reactivity, making it challenging to identify which specific hormone metabolites are being quantified (Ghazal et al., 2022; Karashima & Osaka, 2022). Alternatively, mass spectrometry methods enable highly specific measurements of target analytes. Methods like liquid chromatography-tandem mass spectrometry (LC-MS/MS) separate and identify multiple target analytes, simultaneously measuring multiple hormones of interest in each run.

Regardless of the method chosen, analytical validations are necessary to test method reliability each time they are applied in a new context. Compounds found in each matrix have the potential to distort measurements and may vary by species, necessitating matrix- and species-specific testing. Analytical validation varies by assay method but typically includes assessments of accuracy or recovery, specificity, linearity, parallelism, and repeatability (Koren et al., 2018). These analytical validations have been conducted for the measurement of steroid hormones in baleen (Hunt et al., 2014), earplugs (Crain et al., 2020), blow (Burgess et al., 2018), and blubber (Boggs et al., 2017; Carbajal et al., 2021; Clark et al., 2016; Dalle Luche et al., 2019; Graham et al., 2021; Hayden et al.,

2017; Kellar et al., 2015; Trana et al., 2015; Trego et al., 2013) from many cetacean species, though not for short-finned pilot whales.

1.3.3 Contextualization of hormone measurements

Without context, hormone measurements lend little understanding to the physical state of the animals sampled. A critical step in method validation includes checking that hormone measurements reflect relevant endocrine states using physiological or biological validation. Physiological validation typically induces suppression or stimulation of biomarkers of interest. For steroid hormones, this is often conducted through administration of upstream compounds to trigger hormone release (e.g., adrenocorticotropic hormone (ACTH)), then assessing whether the expected changes were detected in analysis (Koren et al., 2018). Biological validation looks for correlation between relevant variables (e.g., sex, season, reproductive status) and hormone measurements (Koren et al., 2018). In many species, documenting higher progesterone in known-pregnant individuals compared to non-pregnant individuals lends biological validity to progesterone measurements.

Researchers must also consider how each matrix distorts physiological signals. While hormone concentrations in blood serum represent a relatively short snapshot of an animal's physiological condition, hormones take longer to reach and move through other downstream matrices. For example, in an animal that defecates every other day, a fecal sample represents the last 48 hours more closely than the previous ten minutes. In keratinous structures, like hair and baleen, hormones are integrated as the structure grows, so hormone concentrations vary along its length, representing different periods of time. Defecation and tissue growth rates vary between and within species, so researchers

must consider how behavioral states, like fasting, alter hormone deposition. Additionally, travel in the bloodstream, receptor binding, and metabolism vary between hormones and tissues (Norris, 2013), resulting in concentration differences across matrices (Champagne et al., 2018). High metabolite presence can lead to discrepancies between EIAs and LCMS, which vary in analyte specificity, and should be compared with caution.

Lag times between the onset of a physiological state and detection in matrices can be determined by directly observing or inducing a known change in physiological condition (e.g., ovulation, fasting, or stress) and measuring the subsequent hormonal shift. In cetaceans, there is particular interest in the temporal dynamic between stressors and hormone measurement in blubber. Similar to physiological validation, this can be experimentally tested with pharmaceutical administration, which has been conducted with bottlenose dolphins in captivity (Champagne et al., 2017). Experimental manipulations like this require handling and are not practical for most cetacean species; however, sampling after a known stressor can yield the same insight. This has been conducted in captive groups of bottlenose dolphins with blubber sampling occurring at a few discrete intervals post-stressor (Champagne et al., 2018), but offer limited resolution to the timing of peak and decline.

1.4 Short-finned Pilot Whales

Short-finned pilot whales are large members of the Delphinidae family, found in tropical and sub-tropical waters globally. They can dive more than 1,000 m and are typically seen in deep water along continental shelf breaks, where they primarily feed on meso- and bathypelagic cephalopods (Aguilar Soto et al., 2008; Mintzer et al., 2008; Shearer et al., 2022). Recent genetic data suggests short-finned pilot whales should be

divided into two subspecies: the Shiho subspecies, primarily found in the eastern Pacific Ocean, and the Naisa subspecies, which consists of populations in the central/western Pacific, Indian, and Atlantic Oceans (Van Cise et al., 2019). In the western North Atlantic population, individuals have been seen as far south as the Caribbean and north as Canada (Thorne et al., 2017), transiting vast distances within this range (Foley, 2018).

Short-finned pilot whales are highly social, found in small stable groups, occasionally joining large aggregations reaching up to 1,000s of individuals (Mahaffy et al., 2015). They have a matrilineal social structure with female natal philopatry (Mahaffy et al., 2015), and there is debate as to whether males disperse from their natal groups (Mahaffy et al., 2015). Adults exhibit high sexual size dimorphism, with males averaging 5.6 m and females 3.95 m (Pomeroy, 2011). Adult males often have a thicker, larger, and distinctively shaped dorsal fin, which aids in distinguishing them from females and subadults in the field. Relative testis size suggests males rely more heavily on sperm competition for reproductive advantage than would be expected based on their body size and degree of sexual dimorphism (Dines et al., 2015).

In the western Pacific, reproduction is diffusely seasonal, peaking in summer months but occurring throughout the year (Kasuya & Marsh, 1984). Demographic variation in testosterone and sperm production indicates that some adult males, independent of age, may not be reproductive in a given year (Kita et al., 1999). Similarly, not all females reproduce each year. Gestation lasts nearly 15 months, after which calves nurse for at least two years (Kasuya & Marsh, 1984). Males reach reproductive maturity around 15 and have been documented to live up to 46 years (Kasuya & Marsh, 1984). Females reach maturity around nine years of age and reproduce into their late-30s, when

they begin a post-reproductive phase, with females reportedly living as long as 63 years (Kasuya & Marsh, 1984).

1.5 Hormones, behavior, and noise

The endocrine system mediates behavior; however, behavior also influences endocrinology. As hormone measurements can give insight to pilot whale behavior, understanding the foundations of pilot whale behavior can also contextualize hormone measurements. For example, reproduction is largely mediated by androgens and estrogens, but the social flux and competition that often accompany reproduction can also result in corticosteroid elevation (Adkins-Regan, 2005). Depending on individuals' reproductive availability, they may experience different challenges during these times or respond to them differently. Understanding the contributors to physiological and behavioral variation allows us to establish more effective baselines for comparison when animals are exposed to environmental threats.

Of the environmental threats faced by cetaceans, noise pollution can be particularly insidious. Elevated background noise, which is the result of vessel traffic, seismic surveys, sonar, and other anthropogenic sources (Hildebrand, 2009), can lead to masking, or the degraded ability to detect or recognize a sound of interest due to the presence of another sound (Erbe et al., 2016). Noise can also act as acoustic stimuli and may be perceived as a threat, eliciting anti-predator responses (Frid & Dill, 2002). In cetaceans, noise can interfere with communication (Dahlheim & Castellote, 2016), disrupt foraging (Aguilar Soto et al., 2006), and lead to displacement (Lusseau, 2005), which has the potential to affect fitness (Council, 2005). However, these effects are not uniform. Species differences in movement ecology, predator risk, reproductive strategies,

and life histories can influence the likelihood and severity of disturbance from noise (Keen et al., 2021). Within species, variables like age, sex, and reproductive status interact with population characteristics, making some groups more vulnerable to disturbance. Our ability to establish relevant baselines and effectively monitor responses to noise relies on accounting for context.

As methods mature for measuring hormones in free-ranging cetaceans, new possibilities arise for monitoring responses to ocean noise. Hormones influence behavior, and may function as direct indicators of stimulus-response. In this dissertation, I use remote blubber biopsy to characterize the baseline variation of steroid hormones and how they respond to an acoustic stimulus. The application of cetacean hormone measurement in an open-water noise response study represents a frontier in cetacean stimulus-response studies. This dissertation walks through the careful methodological development, baseline characterization, and contextualization that underlie our use of blubber biopsy methods.

2. Multi-class steroid profiling in short-finned pilot whale blubber using liquid chromatography-tandem mass spectrometry (LC-MS/MS)

2.1 Introduction

Steroid hormones mediate physiology and behavior, enabling organismal survival and reproduction. The lipophilic nature of these small molecules aids in their transport into many tissues, enabling scientists to conduct wildlife studies through the collection of feces, hair, blubber, and other methods less invasive than capture. Whale and dolphin (i.e., cetacean) blubber, can be collected at sea by remote blubber biopsy and contains adrenal and gonadal steroids (Kellar et al., 2015; Pérez et al., 2011).

Scientists often use enzyme immunoassay (EIA) to quantify steroid hormones in wildlife samples. EIAs are sensitive, affordable, and simple to use; however, they are subject to antibody cross-reactivity (Ghazal et al., 2022; Karashima & Osaka, 2022) and are limited to one hormone per analysis. In contrast, mass spectrometry offers high specificity and the ability to concurrently measure multiple compounds (Stanczyk & Clarke, 2010). Enabled by efficient blubber extraction methods, e.g., Boggs et al. (2017), liquid chromatography-tandem mass spectrometry (LC-MS/MS) methods have been developed for quantifying steroid hormones in cetacean blubber (Hayden et al., 2017).

LC-MS/MS methods have been employed to reliably quantify corticosteroids, androgens, progestogens, and estrogens in blubber from dolphins (Boggs et al., 2017; Galligan, Schwacke, Houser, et al., 2018) and baleen whales (Dalle Luche et al., 2019; Hayden et al., 2017; Wittmaack et al., 2022). Good chromatographic separation has been achieved by running corticosteroids on a C18 column and separately running androgens and progestogens on a Biphenyl column (Boggs et al., 2017). Because many steroids

have identical precursor and product ions, using different chromatographic sorbents can reliably separate multiple classes of steroids from one sample extraction.

Though studies have successfully employed these methods for assessments of stress and reproductive hormones in blubber from free-ranging populations, the inter-specific variability of blubber tissue (Koopman, 2007) makes it necessary to test for matrix effects before applying existing methods to new species. In this study, we tested and adapted methods established in Boggs et al. (2017) for use with blubber from short-finned pilot whales (*Globicephala macrorhynchus*). We conducted a spike recovery experiment with blubber samples from stranded individuals to assess this method's accuracy and precision for application in this species' blubber tissue.

This study sought to expand existing methods by adding three steroid hormones connected to stress in cetaceans: aldosterone, dehydroepiandrosterone (DHEA), and dehydroepiandrosterone-sulfate (DHEAS). Aldosterone, a mineralocorticoid, regulates sodium balance and increases in response to stressors in marine mammal species (Champagne et al., 2018; McCormley et al., 2018). Aldosterone has been measured in cetacean serum and feces using immunoassay (Champagne et al., 2017) and recently, in the blubber of two stranded gray whales (Wittmaack et al., 2022). DHEA and its sulfonated version, DHEAS are prohormones primarily synthesized in the adrenal cortex and secreted in response to adrenocorticotropic hormone (ACTH). However, they are also synthesized in the gonads and are categorized as androgens. Changes in DHEAS and DHEA concentrations have been connected to chronic stress (Maninger et al., 2010) and disease (Gundlach et al., 2018), but neither of these hormones have previously been measured in cetacean blubber.

We added these steroid hormones through the optimization of an LC-MS/MS method using a C18 column. Isotopically labeled internal standards (IS) were added to samples to enable accurate analyte quantification by controlling for sample loss throughout processing. Ideally, methods include a matched IS for each analyte which helps verify peak identity, but the expense of these compounds impedes universal use. To guide decisions around which ISs should be used in future methods, we assessed and compared the performance of 10 isotopically labeled ISs for 11 analytes.

Therefore, the aims of this chapter are to (1) report an optimized method for simultaneous quantification of eight adrenal hormones, (2) assess the accuracy and precision of two LC-MS/MS methods applied to pilot whale blubber, and (3) compare the performances of isotopically labeled ISs for each analyte.

2.2 Materials and Methods

2.2.1 LC-MS/MS Method Optimization

We divided target analytes into two groups for liquid chromatography, referred to as the 'gonadal steroids' (17α-hydroxyprogesterone ($17OHP_4$), androstenedione (AE), testosterone (T), and progesterone(P_4)) and 'adrenal steroids' (aldosterone (ALD), cortisol (F), cortisone (E), corticosterone (B), 11-deoxycortisol (S), 11-decoxycorticosterone (11DOC), DHEA, and DHEAS) (Table 1). We sourced standards and internal standards from multiple suppliers (Supplementary Table A1). We assessed the four gonadal steroids using the LC-MS/MS method developed previously (Boggs et al., 2017) detailed in the methods section. Previous studies have detailed the separation and quantitation of five corticosteroids of interest (F, E, B, S, and 11DOC) in blubber (Boggs et al., 2019;

Dalle Luche et al., 2019). We adapted these methods to incorporate ALD, DHEA, and DHEAS, for the concurrent measurement of all eight adrenal steroids of interest.

Table 1: C18 Compound Optimization Parameters. Columns show retention time (RT), abbreviations (Abbr.), peak type (Quantitative or Qualitative), Precursor Ion (Pre. Ion), Product Ion (Pro. Ion), Declustering potential (DP), Entrance Potential (EP), Collision Energy (CE), and Cell Exit Potential (CXP).

Steroid Common Name	RT (min)	Abbr.	Peak type	Pre. Ion (m/z)	Pro. Ion (m/z)	DP (V)	EP (V)	CE (V)	CXP (V)
Adrenal Steroids (C18 Method)									
Aldosterone	16.5	ALD	Quant.	361.2	315.2	80	10	27	15
			Qual.	361.2	109.2	80	10	40	5
		ALD-d_4	IS	365.2	319.2	80	10	27	15
Cortisone	18.0	E	Quant.	361.4	163.3	100	5	35	12
			Qual.	361.4	121.3	50	15	60	10
		E-$^{13}C_3$	IS	364.4	166.3	100	5	35	12
Cortisol	19.6	F	Quant.	363.1	121.1	50	10	40	5
			Qual.	363.1	267.3	25	5	30	30
		F-$^{13}C_3$	IS	366.1	124.1	50	10	40	5
Corticosterone	23.7	B	Qual.	347.2	121.3	40	5	20	15
			Quant.	347.2	135	25	5	35	5
		B-d_4	IS	351.2	139	25	5	35	5
11-Deoxycortisol	24.4	S	Quant.	347.3	97.1	75	15	30	10
			Qual.	347.3	109.2	25	5	35	10
		S-$^{13}C_3$	IS	350.3	112.2	25	5	35	10
11-Deoxycorticosterone	28.2	11DOC	Quant.	331.7	109.3	75	10	35	5
			Qual.	331.7	97.1	25	10	25	5
		11DOC-$^{13}C_3$	IS	334.7	100.1	25	10	25	5
Dehydroepiandrosterone	30.0	DHEA	Quant.	271.2	253.3	80	10	20	20
			Qual.	271.2	213.2	80	10	27	15
Gonadal Steroids (Biphenyl Method)									
Androstenedione	10.8	AE	Quant.	287.2	97.2	100	15	30	12
			Qual.	287.2	109.2	100	15	30	20
		AE-$^{13}C_3$	IS	290.6	100.2	75	10	40	5
Testosterone	8.7	T	Quant.	289.1	109.2	100	10	30	5
			Qual.	289.1	97.2	100	15	30	12
		T-$^{13}C_3$	IS	292.6	112.2	75	10	40	5
Progesterone	12.8	P_4	Quant.	315.1	109.2	100	15	30	20
			Qual.	315.1	97.2	75	15	25	10
		P_4-$^{13}C_3$	IS	318.4	112.2	50	10	35	5
17α-hydroxyprogesterone	8.5	17OHP$_4$	Quant.	331.2	109.2	100	15	35	20
			Qual.	331.2	97.2	75	10	25	10

		17OHP$_4$-$^{13}C_3$	IS	334.6	100.1	40	5	40	5

We conducted optimization using direct injection of individual analytes (Supplementary Table A1) onto an AB Sciex (Framingham, MA) API 4000 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer. The Boggs et al. method (2017) used positive mode electrospray ionization (ESI) to identify mass transitions for corticosteroids (F, E, B, S, and 11DOC). Therefore, we used matched ionization methods to optimize ALD, DHEA, and DHEAS. We evaluated fragmentation patterns at varying collision energies until we identified two to seven candidate product ions for each analyte. We used tuning mode to optimize source (curtain gas = 276 kPa temperature = 600 ºC, ion source gas 1 = 209 kPa, ion source gas 2 = 414 kPa, interface heater = Off, collision gas- High, ion spray voltage = 4000 V) and compound parameters (Table 1). We optimized all eight adrenal steroids in positive and negative ionization mode; however, positive mode performed better for all analytes and therefore negative mode data will not be discussed. DHEAS did not ionize well in positive mode and was later removed from the final positive ionization method.

Using a mixture of the eight adrenal analytes, we created a chromatography method on a Zorbax Eclipse Plus C18 column (150 mm x 2.1 mm, 5 μm particle size) from Agilent with the 1200 Series HPLC system with a binary pump and autosampler from Agilent (Santa Clara, CA). We determined compound retention times through multiple-reaction monitoring (MRM), by monitoring two mass transitions per compound. We selected the peak with the largest area as the quantitative ion and the second largest peak as the qualitative ion. The resulting scheduled multiple reaction monitoring (sMRM)

method used a 240-second detection window for each mass transition, including six internal matched standards (Table 1, Figure 2).

2.2.2 Spike-retrieval experiment

2.2.2.1 Sample collection

We collected blubber samples post-mortem from two short-finned pilot whales, live-stranded in North Carolina, USA. Both carcasses were in fresh-dead condition (Code 1, i.e., meaning stranded alive then died or euthanized) when samples were collected. One sample came from an adult male (ID: RJM009), 480 cm in length, that stranded in 2009, hereafter referred to as the 'male sample'. The 'female sample' was obtained from a pregnant adult female (ID: RT48), 352 cm in length, that stranded in 2005.

2.2.2.2 Sample extraction

For each of the whales, we processed ten replicates, approximately 0.4 g of blubber each: five endogenous (un-spiked) and five spiked with 400 μL of a calibrant mixture of the 11 target analytes (masses in Supplementary Table A2). In addition to three blanks, we ran three un-spiked replicates from a standard reference material (NIST SRM 1945), taken from another pregnant female short-finned pilot whale. For quantification, we diluted neat standards in methanol to create ten calibration standards covering a physiologically relevant range of concentrations (Supplementary Table A2). We created an IS mix with ten isotopically labeled matched ISs (Supplementary Table A2), which we added to all samples (blanks, SRMs, calibrants, and blubber) before extraction. We prepared all stocks and samples gravimetrically and extracted all samples using a bead homogenization and QuEChERS extraction protocol previously detailed (Boggs et al., 2017).

2.2.2.3 LC-MS/MS

To analyze gonadal steroids, we reconstituted samples in HPLC grade methanol and used the biphenyl LC-MS/MS method (Boggs et al., 2017) (Table 1). We injected 10 μL of each sample onto the Restek (Bellefonte, PA) Ultra Biphenyl column (250 mm x 4.6 mm, 5 μm particle size) column and conducted separation using a gradient of acetonitrile and methanol (both with 0.1 % volume formic acid) over 36 minutes at a flow rate of 500 μL/min.

We analyzed adrenal steroids using the C18 LC-MS/MS method developed in this study and described above. Before injecting 10 μL of each sample onto a C18 column for adrenal steroid analysis, we conducted a solvent exchange, bringing samples to a 50:50 volume fraction of Milli-Q water and methanol (described in Galligan, Schwacke, Houser, et al. (2018)). Once injected, we separated steroids using a gradient of methanol and Milli-Q water (both with 0.1 % volume acetic acid) throughout the 52-minute method at a 250 μL/min flow rate (Table 2).

Table 2: C18 method mobile phase gradient

Step	Time(min)	Water	Methanol
0	0.0	80 %	20 %
1	2.0	80 %	20 %
2	7.0	52 %	48 %
3	10.0	50 %	50 %
4	15.0	47 %	53 %
5	20.0	42 %	58 %
6	24.0	35 %	65 %
7	32.0	25 %	75 %
8	35.0	10 %	90 %
9	35.1	0 %	100 %
10	40.0	0 %	100 %
11	40.1	80 %	20 %
12	52.0	80 %	20 %

2.2.2.4 Quantification

After mass spectrometry, we used Sciex Analyst software (Version: 1.6; Framingham, MA) to integrate peaks manually. To determine analyte concentrations, we first calculated peak area ratios between the target compound and the matched IS. We then used the peak area ratios of calibrants to interpolate linear regressions, establishing calibration curves (Supplementary Tables A3 and A4). For the biphenyl data, where there were high analyte concentrations in spiked and pregnant samples, we created separate high and low calibration curves to target relevant concentrations.

We defined each analyte's observed reporting limit (RL_{obs}) as the lowest calibrant used in the calibration curve. We calculated the limit of detection (LOD) by adding the mean of the three blanks with three times the standard deviation of the blanks. We compared the RL_{obs} and LOD and used the higher of the two as the reporting limit (RL). For calculations using analyte concentrations, we substituted values below RL with RL/2. We determined accuracy with the percent analyte recovered in the spike-recovery experiment and used relative standard deviation (RSD) to assess precision. We calculated percent recovery by comparing recovered analyte mass in spiked samples with expected analyte mass, using the following equation:

$$Percent\ Recovery = \frac{ab}{cb + de} \times 100\%$$

where ‘a’ is the analyte mass fraction (ng/g) measured in a spiked sample, ‘b’ is the mass (g) of that spiked sample, ‘c’ is the mean analyte mass fraction (ng/g) measured in endogenous (i.e., un-spiked) samples, ‘d’ is the analyte mass fraction (ng/g) in the calibrant used to spike the sample, ‘e’ is the mass of the spike (g). We considered percent

recovery between 70 % and 120 % and relative standard deviation (RSD) values below 15 % acceptable.

2.2.3 Internal Standard Comparison

Initially, we used matched ISs to determine peak area ratios and calculate analyte concentrations. We then compared the suitability of IS substitutions by calculating analyte concentrations from peak area ratios created with other ISs in the method. For example, cortisol (F) concentrations were calculated using F-$^{13}C_3$ for ratios, then compared with concentrations calculated using each IS monitored in the C18 method (Ald-d_4, E-$^{13}C_3$, B-d_4, S-$^{13}C_3$, 11DOC-$^{13}C_3$). We repeated this for each analyte, resulting in six comparisons for each adrenal steroid and four for the gonadal steroids. To assess the performance of these substitutions, we used the equation above to calculate percent recovery and RSD for each substitution. We compared endogenous concentrations between IS substitutions to determine whether substitutions substantially altered analyte concentrations calculated.

2.3 Results

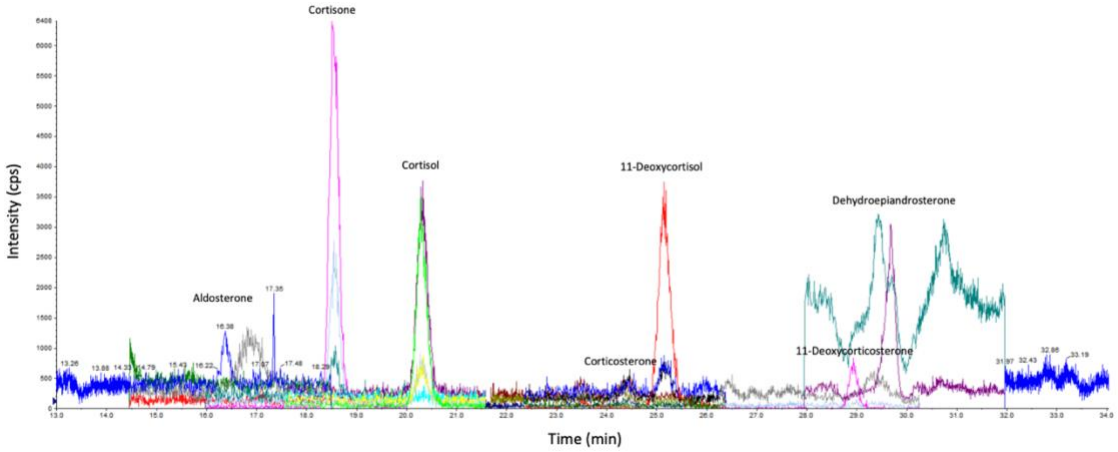

Figure 2. C18 SMRM peaks in blubber from a female short-finned pilot whale (ID: RT48)

2.3.1 Spike-retrieval experiment

When using matched internal standards, mean percent recoveries ranged from 83 % to 110 % (Figure 3). RSDs were between 1.4 % to 9.4 % for all compounds, except B (18.8 %). Injection of one SRM sample onto the C18 column failed; therefore, we determined the mean SRM concentrations for adrenal steroids from two samples. One un-spiked replicate from the male sample and two spiked replicates from the female sample were damaged during processing and excluded from analyses.

For un-spiked samples, several mean analyte concentrations fell below the RL. Mean endogenous ALD concentrations fell below the RL in all three samples and endogenous B and 11DOC could only be quantified in the SRM, which had higher concentrations of all adrenal steroids. All biphenyl analytes were quantified above the RL in un-spiked samples, except for T, which was not quantifiable in the female sample or the SRM (also female).

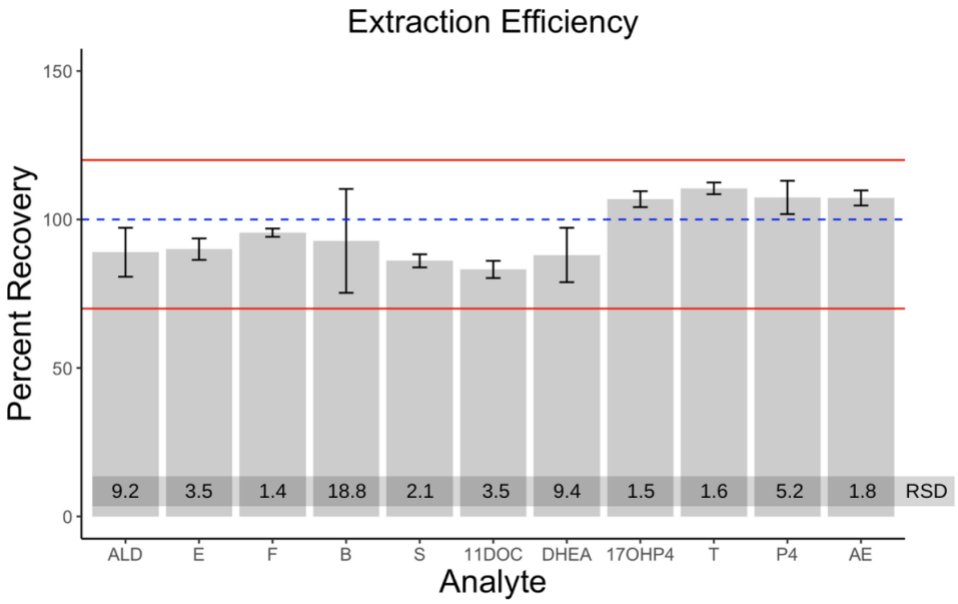

Figure 3: Mean extraction efficiencies for target adrenal and gonadal steroids in pilot whale blubber, determined by a spike-retrieval experiment with blubber samples from two individuals (RJM009 and RT48). Analytes were quantified using isotopically matched internal standards. DHEA, which did not have a matched internal standard, was quantified using S-C3. Error bars represent standard deviation and red lines delineate the acceptable percent recovery range (70% - 120%). RSDs are shown in the shaded bar along the bottom.

2.3.2 Internal Standard Comparison

In the biphenyl separation method, all analytes performed best with their matched ISs, though RSD and percent recovery remained in the acceptable range for several substitutions (Table 3). For 17OHP$_4$, the T-$^{13}C_3$ and AE-$^{13}C_3$ ISs yielded good percent recoveries (98.0 % and 108.9 % respectively) and RSDs (4.3 % and 1.4 % respectively), though measurements of endogenous concentrations decreased, especially at lower concentrations. For calculating T, the nearby 17OHP$_4$-$^{13}C_3$ IS peak performed well as an IS substitute (percent recovery = 110 % RSD = 3.7 %). For AE, the T-$^{13}C_3$ and 17OHP$_4$-$^{13}C_3$ substitutions showed good recovery (99.4 % and 107 % respectively) and consistent

endogenous AE measurements (Figure 4), but RSDs were not considered acceptable (18.9 % and 21.0 % respectively). For quantifying P_4, AE-$^{13}C_3$ was a good IS substitute, with acceptable percent recovery (109 %), RSD (5.3 %), and P_4 measurements consistent with concentrations determined with the matched IS. When quantified with T-$^{13}C_3$ and 17OHP$_4$-$^{13}C_3$, P_4 concentrations were 28.1 to 67.0 % lower than the concentration determined with P_4-$^{13}C_3$.

In the C18 method, the four $^{13}C_3$ ISs outperformed the two d_4 standards for quantification of every adrenal steroid, except ALD (Figure 5). The B-d_4 peaks had small peak areas compared to the other analytes, which likely caused higher variability in the IS ratios used to calculate concentrations. This is likely reflected by high RSD values (18.5 % - 23.0 %) when using B-d_4 as an IS for other adrenal steroids (Table 4). As a target compound, B was best quantified with $^{13}C_3$ ISs, which were characterized by larger, more distinct peaks than B-d_4. In comparison to an RSD of 18.8 % when calculated with B-d_4, RSDs for B fell below 4.5 % when calculated with S-$^{13}C_3$, F-$^{13}C_3$, E-$^{13}C_3$, and 11DOC-$^{13}C_3$. Though ALD performed best with ALD-d_4, this IS also had a relatively small and broad peak, leading to more variability when used as an IS substitute, with analyte RSDs ranging from 14.8 % to 16.5 %. Except for quantifying 11DOC with S-$^{13}C_3$, the accuracies and RSDs calculated for E, F, B, S, 11DOC, and DHEA using $^{13}C_3$ ISs all fell within acceptable ranges.

For the most part, individual differences were the primary drivers in analyte concentration. However, even when percent recovery and RSD fell within acceptable thresholds, IS substitutions altered observed analyte concentrations (Figure 5). In the female sample, concentrations of B were close to the RL, 0.869 ng/g. When calculated

with B-d_4, concentrations fell below RL but exceeded RL when calculated with S-$^{13}C_3$, 11DOC-$^{13}C_3$, and E-$^{13}C_3$. In both methods, substitutions that failed to meet acceptable RSD and/or percent recovery thresholds altered concentrations substantially. In the C18 method, concentrations varied from those calculated with matched ISs by as much as 63.3 %. On average, good substitutions varied much less from matched IS analyte concentrations (mean = 15.5 % median = 11.3 %) than poor substitutions (mean = 31.7 % median = 30.1 %).

Table 3: Comparison of IS reference performances for each biphenyl analyte, arranged by proximity to analyte peak. RSDs and accuracies shown are averages for all male and female spiked replicates and the concentration for each sample represents the mean concentration of endogenous replicates. The best performing IS for each analyte is shown in bold. Values exceeding thresholds for RSD (>15) and accuracy (<70 or >120%) are in gray. The slope of the line used for each calibration curve is shown, along with the correlation (R^2) of calibrants. Separate low and high calibration curves help account for the high concentrations of gonadal steroids seen in some samples.

Analyte	IS	T dist (min)	Accur.	RSD	RL (ng/g)	Male (ng/g)	Female. (ng/g)	SRM (ng/g)	Curve	R^2	Slope
17OHP$_4$	**17OHP$_4$-$^{13}C_3$**	**---**	**106.83**	**1.507**	**0.2088**	**0.3508**	**0.4586**	**3.195**	**low**	**0.9999**	**0.0646**
									high	**0.9990**	**0.0599**
	T-$^{13}C_3$	0.19	97.96	4.315	0.1112	0.2560	0.3208	3.628	low	0.9999	0.4976
									high	0.9997	0.5613
	AE-$^{13}C_3$	2.37	108.87	1.404	0.1056	0.1600	0.2868	3.075	low	0.9999	0.7898
									high	1.0000	0.0639
	P$_4$-$^{13}C_3$	3.71	222.00	59.909	0.1056	0.4575	0.8876	10.945	low	0.9999	0.1907
									high	1.0000	0.1976
T	**T-$^{13}C_3$**	**---**	**110.49**	**1.633**	**0.1505**	**4.878**	**<LOQ**	**<LOQ**	**low**	**1.0000**	**1.1497**
									high	**1.0000**	**1.1493**
	17OHP$_4$-$^{13}C_3$	0.19	110.08	3.730	0.1505	5.721	<LOQ	<LOQ	low	0.9999	0.1492

									high	0.9999	0.1493
	AE-$^{13}C_3$	2.19	120.59	26.390	0.1505	4.955	<LOQ	<LOQ	low	0.9998	2.1209
									high	0.9996	1.8151
	P_4-$^{13}C_3$	3.53	231.69	55.585	0.1505	7.339	<LOQ	<LOQ	low	1.0000	0.4616
									high	0.9999	0.4398
	AE-$^{13}C_3$	**---**	**107.23**	**1.818**	**1.260**	**10.514**	**12.97**	**11.120**	**low**	**0.9999**	**0.2268**
									high	**1.0000**	**0.2218**
	P_4-C3	1.35	186.23	29.194	1.260	14.287	22.35	22.217	low	1.0000	0.0530
									high	1.0000	0.0523
AE	T-$^{13}C_3$	2.18	99.41	18.897	1.260	9.693	10.86	9.180	low	0.9998	0.1381
									high	0.9999	0.1483
	17OHP$_4$-$^{13}C_3$	2.37	106.94	21.040	1.260	11.169	13.15	8.566	low	1.0000	0.0169
									high	0.9993	0.0178
	P_4-$^{13}C_3$	**---**	**107.42**	**5.236**	**0.07101**	**0.3653**	**112.20**	**348.7**	**low**	**0.9999**	**0.5318**
									high	**0.9995**	**0.4732**
	AE-$^{13}C_3$	1.34	108.53	5.272	0.07101	0.2320	114.40	351.6	low	1.0000	2.2952
									high	1.0000	0.4725
P_4	T-$^{13}C_3$	3.52	71.91	27.888	0.07101	0.2544	46.37	115.1	low	1.0000	1.3282
									high	1.0000	1.4811
	17OHP$_4$-$^{13}C_3$	3.71	77.08	34.630	0.07101	0.2625	64.16	139.2	low	1.0000	0.1740
									high	0.9998	0.1590

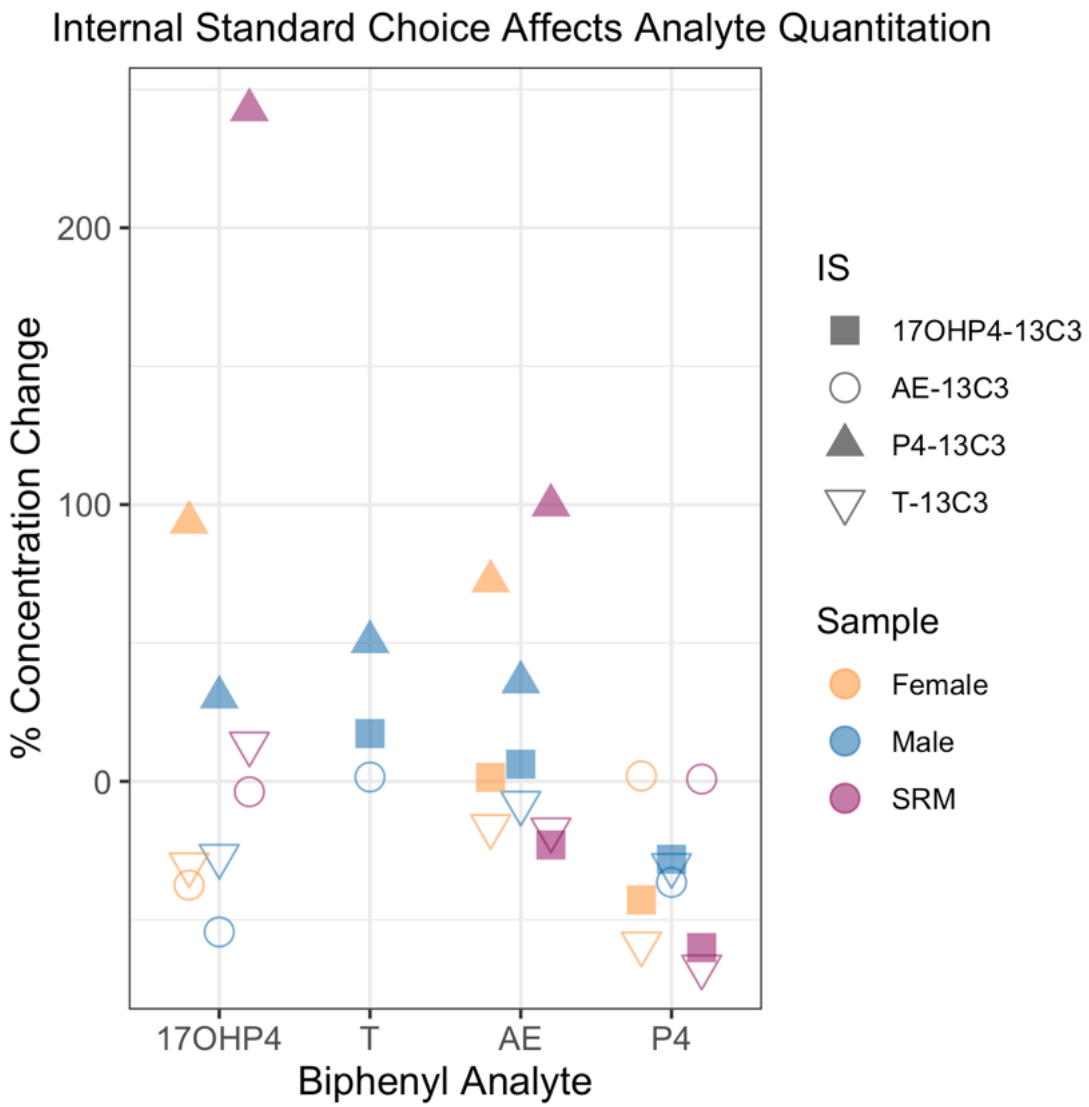

Figure 4: Change in mean endogenous concentrations of biphenyl (gonadal) analytes when calculated with varying IS references. Endogenous T was not quantifiable in the female sample or the SRM and is not shown on the plot.

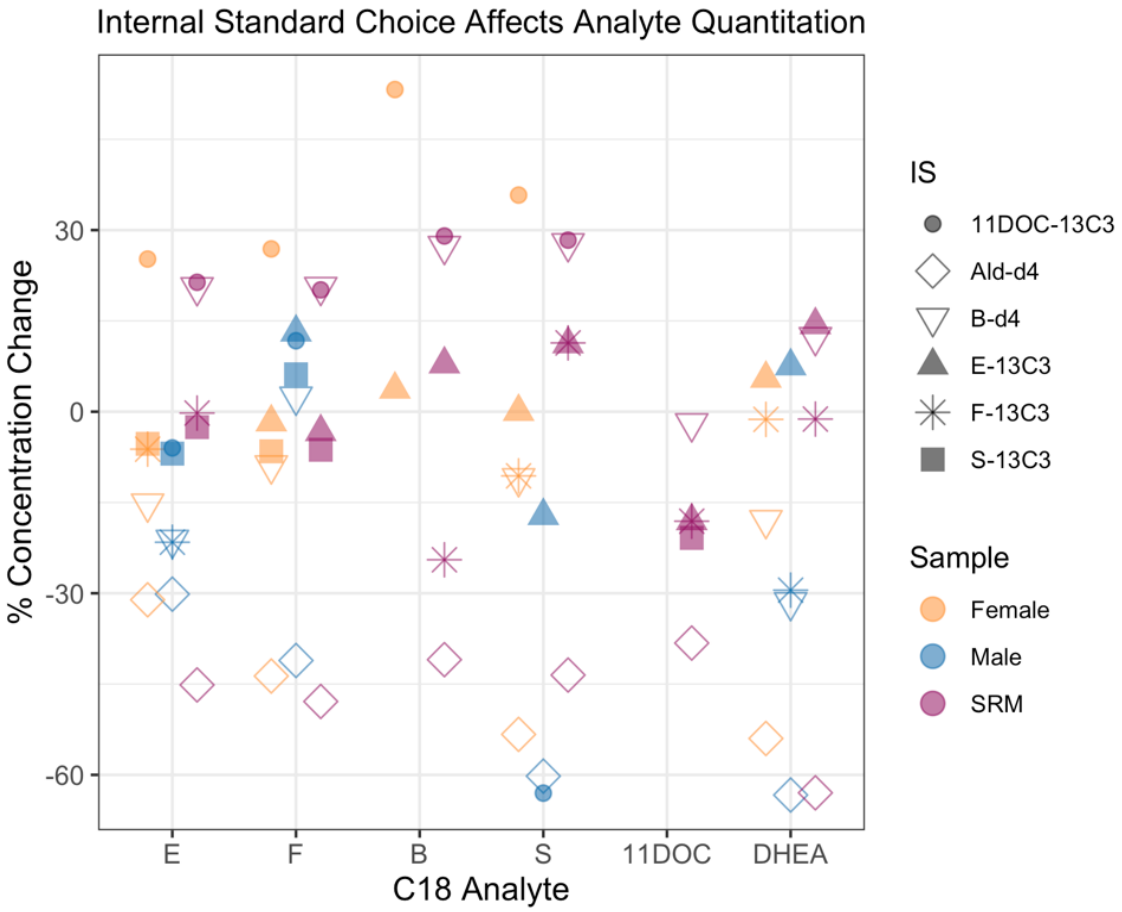

Figure 5: Changes in C18 (adrenal) analyte concentrations when calculated with varying IS references. Endogenous aldosterone concentrations were below the reporting limit for all three samples, regardless of IS used, and are not included below.

Table 4: Comparison of IS performances for each C18 analyte, arranged by proximity to analyte peak. RSDs and accuracies shown are averages for all male and female spiked replicates and the concentration for each sample represents the mean concentration of endogenous replicates. The best performing IS for each analyte is shown in bold. Values exceeding thresholds for accuracy (<70 or >120%) and RSD (>15) are in gray.

Analyte	IS reference	Peak distance (min)	Accuracy	RSD	RL (ng/g)	Male conc. (ng/g)	Female conc. (ng/g)	SRM conc. (ng/g)	R^2	Slope
ALD	**Ald-d_4**	---	**88.96**	**9.159**	**3.539**	**<RL**	**<RL**	**<RL**	**0.9994**	**1.1725**
	E-$^{13}C_3$	1.59	69.29	12.43	3.539	<RL	<RL	<RL	0.9337	0.1715
	F-$^{13}C_3$	3.31	87.58	11.43	3.539	<RL	<RL	<RL	0.9997	0.2846

	B-d_4	7.42	64.37	21.43	3.539	<RL	<RL	<RL	0.9954	7.4969
	S-$^{13}C_3$	8.17	61.67	9.553	3.539	<RL	<RL	<RL	0.9950	0.1848
	11DOC-$^{13}C_3$	12.01	114.2	10.33	3.539	<RL	<RL	<RL	0.9996	0.9914
E	**E-$^{13}C_3$**	**---**	**90.00**	**3.534**	**0.1035**	**0.9520**	**1.378**	**6.440**	**0.9998**	**0.9696**
	F-$^{13}C_3$	1.69	94.72	4.128	0.1035	0.7467	1.292	6.425	1.0000	1.6618
	Ald-d_4	1.74	88.09	15.60	0.1047	0.6650	0.9494	3.533	0.9998	7.1236
	B-d_4	5.8	91.33	19.45	0.1035	0.7505	1.168	7.764	1.0000	29.4480
	S-$^{13}C_3$	6.55	89.77	3.692	0.1035	0.8860	1.304	6.276	1.0000	0.6868
	11DOC-$^{13}C_3$	10.39	104.3	6.807	0.1035	0.8950	1.725	7.816	1.0000	6.1497
F	**F-$^{13}C_3$**	**---**	**95.56**	**1.443**	**0.05613**	**1.494**	**3.825**	**9.066**	**0.9999**	**0.9108**
	E-$^{13}C_3$	1.74	90.06	4.115	0.05613	1.690	3.758	8.763	1.0000	0.5455
	Ald-d_4	3.45	93.38	15.61	0.05613	0.8798	2.154	4.725	1.0000	3.7853
	B-d_4	4.09	92.00	18.97	0.05613	1.531	3.486	10.93	0.9999	15.9360
	S-$^{13}C_3$	4.84	89.09	3.344	0.05613	1.586	3.574	8.496	1.0000	0.3895
	11DOC-$^{13}C_3$	8.68	104.6	6.532	0.05613	1.669	4.853	10.89	1.0000	3.3828
B	B-d_4	---	92.79	18.85	0.8693	<RL	<RL	5.328	1.0000	3.7159
	S-$^{13}C_3$	**0.66**	**89.69**	**2.369**	**0.8693**	**<RL**	**0.8997**	**4.183**	**0.9997**	**0.0905**
	F-$^{13}C_3$	4.2	101.4	2.717	0.8693	<RL	<RL	3.159	0.9999	0.2058
	11DOC-$^{13}C_3$	4.5	106.0	4.482	0.8693	<RL	1.379	5.397	0.9999	0.7883
	E-$^{13}C_3$	5.92	96.74	2.540	0.8693	<RL	0.9329	4.511	1.0000	0.1185
	Ald-d_4	7.63	93.85	15.73	0.8693	<RL	<RL	2.469	0.9998	0.8768
S	**S-$^{13}C_3$**	**---**	**86.07**	**2.114**	**0.02738**	**0.1255**	**0.3834**	**5.969**	**1.0000**	**0.7621**
	B-d_4	0.76	89.43	18.47	0.02738	<RL	0.3413	7.630	1.0000	31.2410
	11DOC-$^{13}C_3$	3.83	103.4	5.938	0.02738	0.04642	0.5206	7.660	1.0000	6.6189
	F-$^{13}C_3$	4.87	98.57	5.013	0.02738	<RL	0.3427	6.647	1.0000	1.6906
	E-$^{13}C_3$	6.59	94.71	4.956	0.02738	0.1040	0.3834	6.636	0.9999	0.9812
	Ald-d_4	8.3	90.99	16.49	0.02738	0.04995	0.1790	3.372	1.0000	7.2227
11DOC	**11DOC-$^{13}C_3$**	**---**	**83.18**	**3.491**	**0.3664**	**<RL**	**<RL**	**5.611**	**1.0000**	**1.6794**
	S-$^{13}C_3$	3.84	69.37	3.926	0.3664	<RL	<RL	4.441	0.9998	0.1899

	B-d_4	4.59	71.76	20.12	0.3664	<RL	<RL	5.504	0.9999	7.9127
	F-$^{13}C_3$	8.7	74.12	5.565	0.3664	<RL	<RL	4.596	1.0000	0.4575
	E-$^{13}C_3$	10.42	70.07	4.042	0.3664	<RL	<RL	4.598	0.9999	0.2668
	Ald-d_4	12.13	104.4	14.81	0.3664	<RL	<RL	3.466	1.0000	1.2503
DHEA	11DOC-$^{13}C_3$	1.79	117.0	8.148	0.1082	<RL	<RL	<RL	1.0000	2.7035
	S-$^{13}C_3$	**5.63**	**88.05**	**9.368**	**0.1423**	**1.618**	**1.194**	**0.4861**	**1.0000**	**0.3218**
	B-d_4	6.38	71.36	22.96	0.06617	1.109	0.9804	0.5460	0.9997	17.3940
	F-$^{13}C_3$	10.49	93.49	10.29	0.1351	1.141	1.179	0.4802	1.0000	0.7788
	E-$^{13}C_3$	12.21	88.57	7.898	0.2530	1.739	1.260	0.5557	1.0000	0.4518
	Ald-d_4	13.92	83.47	15.52	0.1670	0.5932	0.5494	0.1800	1.0000	3.4153

2.4 Discussion

The Boggs et al. (2017) methods for quantifying steroid hormones in whale blubber can be applied to blubber samples from short-finned pilot whales. As the spike-recovery experiment demonstrated, this tissue matrix does not interfere with the reliable measurement of hormones when using this method. When using matched ISs, accuracy and RSD were acceptable for all 11 hormones measured, except B. This is rectified by substituting B-d_4 with an appropriate IS, like S-$^{13}C_3$. Accuracy was similar to Boggs et al. (2017) (84 % -112 %) and showed increased precision for some analytes when compared with Boggs et al. (2017) and Dalle Luche et al. (2019). This improvement was likely due to the inclusion of additional matched ISs for adrenal analytes. Observing expected differences in T and P_4 concentrations between adult male and known-pregnant female samples demonstrates the biological validity of this method, which is promising for the future application of this method to blubber biopsy samples from free-swimming individuals.

The C18 method optimized in this study is the first to validate the measurement of aldosterone and DHEA in whale blubber. Though endogenous aldosterone was not present in measurable quantities in the pilot whale blubber assessed, this is the first LC-MS/MS method to quantify DHEA in cetacean blubber. We expect adrenal steroids to be significantly higher in blubber from live-stranded whales than free-swimming individuals. Even so, analyte concentrations were below RL in some samples. What this method gains from specificity, it loses in sensitivity. This is especially true for ALD, which had a relatively high RL of 3.54 ng/g, which is similar to the LOD published in Wittmaack et al. (2022), 3.72 ng/g. Future studies should explore this further to determine whether resolution can be improved and whether ALD metabolites are abundant and measurable in this tissue.

The IS assessment conducted in this study highlights the necessity of careful IS selection when developing LC-MS/MS methods. Not all analytes from this study required matched ISs. In most cases, nearby ^{13}C-labeled IS peaks could be used as substitutes without sacrificing significant accuracy and precision. However, chromatographic proximity is not the sole indicator of a suitable IS substitution. Even using IS substitutions of nearby ^{13}C-labeled peaks affected performance and concentrations measured (see S/11DOC-$^{13}C_3$ in Table 4). Matrix interferences at the retention time of the analyte could be responsible for this effect. ALD-d_4 and B-d_4 standards performed poorly compared to ^{13}C-labeled standards. This could stem from hydrogen/deuterium exchange occurring during extraction or measurement, or may reflect the relatively low sensitivity of the method to these analytes. If using ALD-d_4 or B-d_4 as ISs, researchers should

consider using concentrations well above endogenous levels or reliable IS substitutions. Further studies should assess the use of ^{13}C-labeled ISs for ALD and B.

This is the first study to validate the measurement of multi-class steroid hormone profiling in short-finned pilot whales. Where prior studies used immunoassay to measure progesterone in only females (Yoshioka et al., 1989) or testosterone in only males (Kita et al., 1999), we present a comprehensive method to concurrently measure hormones regardless of sex. Comprehensive studies like this can maximize the information gleaned from small and difficult-to-obtain samples, such as blubber biopsies. Aside from testosterone, this is the first study to quantify steroid hormones in male short-finned pilot whales. This method did not prove effective in the measurement of aldosterone but was able to reliably detect DHEA, which has not been previously measured in short-finned pilot whales. While corticosterone is the dominant glucocorticoid in some smaller animals, like mice and birds, cortisol was the most abundant glucocorticoid observed in these samples. The predominance of cortisol echoes the patterns seen in cetaceans and other large mammals. The use of LC-MS/MS for highly-specific multiclass steroid profiling can inform which analytes researchers select for studies with single-target methods, like immunoassay.

3. Variation of Steroid Hormones in Short-Finned Pilot Whales

3.1 Introduction

In mammals, understanding hormone variation is key to understanding life history. Life history parameters (e.g., age of maturation, rates of reproduction, average lifespan) are foundations of effective wildlife management; however, these data are lacking for many cetacean populations, which can be challenging to observe and sample. The use of alternative matrices (e.g., blubber, blow, feces) has expanded access to hormone sampling in free-ranging cetaceans and promises novel insight into the growth, reproduction, and survival of populations. In this chapter, we use blubber biopsies to establish steroid hormone references for short-finned pilot whales (*Globicephala macrorhynchus*) and characterize variation within the context of life history.

3.1.1 Sources of Hormone Variation

Hormones reflect the physiological states of individuals, which can vary with sex, life history stage, and environmental context. When reporting hormone concentrations, including these variables creates relevant reference points. Sexes are easily assigned with genetic analyses and environmental context may be readily observed, but it is challenging to determine sexual maturity and pregnancy status in the field, especially without long-term observation or hands-on physiological assessment.

In mammalian wildlife studies, pregnancy is often assigned from elevated progesterone concentrations (Spencer & Bazer, 2002). Longitudinal studies of captive odontocetes (bottlenose dolphins, killer whales, beluga whales, and false killer whales), show progesterone concentrations peak shortly after conception and gradually decrease

throughout pregnancy (Funasaka, 2018; Katsumata et al., 2006; Legacki et al., 2020; O'Brien & Robeck, 2012). In some cases, early pregnancy progesterone concentrations overlap with the luteal progesterone peak, making it difficult to distinguish between pregnant and ovulating individuals (Dalle Luche et al., 2020; Funasaka, 2018; Legacki et al., 2020; Robeck et al., 2017; Robeck et al., 2021).

Fortunately, progesterone is not the only steroid hormone upregulated during pregnancy. Androgens like dehydroepiandrosterone (DHEA), androstenedione, and testosterone, increase above baseline during the second half of pregnancy in several odontocete species, including humpback whales, killer whales, and bottlenose dolphins. Thus, androgens can serve as secondary biomarkers for pregnancy confirmation (Dalle Luche et al., 2020; Legacki et al., 2020; Robeck et al., 2017; Steinman et al., 2021; Zhang et al., 2021). As moderators of energy allocation, glucocorticoid concentrations might be an avenue for identifying pregnant or lactating individuals. Glucocorticoids are elevated during pregnancy in some species (Hunt et al., 2006, Dalle Luche et al., 2020); however, in pregnant odontocetes (i.e., toothed whales) cortisol may only increase in the days leading up to parturition, and can be just as high, if not higher during ovulation (Robeck et al., 2017; Steinman et al., 2021; Steinman et al., 2016).

In addition to sex and life history stage, the biotic and abiotic environment contribute to hormone variation. For animals with seasonal reproductive cycles, physiological and behavioral demands shift throughout the year as do their hormonal mediators (Adkins-Regan, 2005). Annual rhythms are frequently described for sex steroids, but can also be reflected in basal and responsive glucocorticoid concentrations

(Romero, 2002). For populations with unknown seasonal cycles, fluctuations in steroid hormone concentrations can help identify the timing of peak mating and birthing periods.

3.1.2 Endocrine sampling

Endocrine sampling in cetaceans has historically been limited to blood serum collection (Hunt et al., 2013). In live cetaceans, serum is solely accessible through capture or captivity, which poses logistical and experimental constraints. Logistically, capture can be impractical and dangerous, especially in deep-water habitats. Experimentally, both capture and captivity can profoundly alter the production and regulation of hormones (Fair et al., 2014; St Aubin et al., 1996; Thomson & Geraci, 1986), which constrains the ability to discern basal physiological processes.

For the study of free-ranging cetaceans, recent developments with alternative matrices, such as feces, blow, and blubber, offer less invasive approaches than serum collection. The application of hormone matrices relies on the ability to collect samples and validate their reliability. Fecal collection is non-invasive and reflects biological states (Lemos et al., 2022; Rolland et al., 2012). While successful for many species, opportunities for fecal collection are rare for species like short-finned pilot whales, whose feces disperses quickly. Blow, or exhaled breath condensate, can be collected from free-ranging animals with the use of cantilever poles or drones (Hunt 2013); however, small sample volumes and the current inability to account for water contamination and sample concentration hinder interpretation (Mingramm et al., 2019).

Remote blubber biopsy has been used with cetaceans for a few decades and is well-established (Lambertsen, 1987; Noren & Mocklin, 2012). Several methods exist for quantifying hormones in blubber (Kellar et al., 2006; Boggs et al., 2017) and have been

used to reliably indicate biological states, from stress to reproductive maturity and pregnancy (Champagne et al., 2018; Kellar et al., 2014). Compared to serum samples collected with capture and restraint, blubber biopsy offers a clearer window into baseline hormone concentrations. Biopsy does require boat approaches, but the limited data available suggests sampling practices do not alter blubber cortisol (Mingramm et al., 2020.

3.1.3 Short-finned Pilot Whales

The ongoing development of hormone sampling and interpretation expand the potential for non-lethal methods to generate data that previously relied on lethal sampling. For short-finned pilot whales, the foundational life history data were generated through lethal collection, which we draw on to develop this study and contextualize our results. These foundational studies, conducted on specimens from the Japanese drive fishery, established parameters for sexual maturity, gestation, senescence, and seasonal patterns of reproduction (Kasuya & Marsh, 1984, Marsh & Kasuya, Kita 1999). For short-finned pilot whales in the western Pacific, reproduction is diffusely seasonal, indicated by increased testosterone, ovulation, and conception in the spring and summer months (Kasuya & Marsh, 1984). However, ovulation and sperm production are not constrained to these seasons, and reproduction can occur throughout the year (Kasuya & Marsh, 1984). Conception is followed by a long gestation period, estimated to last about 450 days, just under 15 months, with peak parturition occurring in early August (Kasuya & Marsh, 1984).

Hormone concentrations for short-finned pilot whales have only been published in two studies, both of which used lethal sampling methods. In the first, Yoshioka et al.

(1989) reported progesterone concentrations in serum from females and, in the second, Kita et al. (1999) reported testosterone concentrations in the serum and testes of males. Although limited in sample size, the data in Yoshioka et al. (1989) contradicts assumptions that progesterone will be higher in pregnant females. In the Yoshioka study, serum progesterone was elevated but relatively low in two mid-pregnancy animals (3.1 and 8.7 ng/mL) compared to the two females in the luteal phase of ovulatory cycling, which exhibited the highest progesterone concentrations (10.9 and 30.0 ng/mL). Yoshioka et al. (1989) did not measure additional steroid hormones, like androstenedione or testosterone, and the stress associated with capture and death may affect hormone concentrations; however, it serves as a warning for careful interpretation of progesterone data.

3.1.4 Research Approach

In this chapter, we investigate hormone variation in blubber using a multi-steroid approach. The LC-MS/MS method we employ was validated in Chapter 2, demonstrating reliable detection of ten endogenous steroid hormones in blubber from stranded short-finned pilot whales (Table 5). In applying this method, we report concentrations for many of these hormones for the first time in this species. The concentrations in this chapter represent ‘baselines’ for this population, so our first question asks (1) which samples should be considered baseline. These should not be confused with ‘basal’ samples, which implies animals are unstressed and cannot be verified in free-ranging individuals (Romero 2002). Rather, in this case, we determine which samples can reasonably represent the population.

In addition to establishing reference concentrations for hormones in pilot whale blubber, we describe steroid hormone relationships with sex, reproductive status, and seasonality. We view these on both an individual hormone basis and through multi-steroid approaches. For inter-hormone relationships, we expect to see positive correlations within steroid groups like androgens and glucocorticoids, which has been observed in similar cetacean studies (Dalle Luche et al., 2021; Galligan, Schwacke, Houser, et al., 2018).

We then ask (2) whether sex contributes to hormone variation, expecting sex differences will be most apparent in sex steroids, which are primarily produced in the gonads and placenta. These differences should be driven by adult males and females, but we are not able to distinguish between immature animals, maturing adults, and adults. Although we expect average androgen concentrations to be much lower in females than males, there may be limited overlap due to pregnant females.

Although we cannot definitively assign reproductive status, we ask (3) whether signatures of pregnancy contribute to variation. We expect to see elevated progesterone in some females, which are likely ovulating or pregnant. Because elevated androgens are indicative of later-stage gestation, we only expect to observe androgen elevation in a subset of possibly pregnant females.

Lastly, we ask (4) whether seasonality contributes to hormone variation. Seasonal patterns of reproductive behavior and physiology have not been described for these animals in the western North Atlantic, where our study is conducted. We expect seasonal patterns will mirror those described in the western Pacific population, which occurs at a similar latitude and in similar environments to the western North Atlantic population.

Given that these populations inhabit separate ocean basins, however, variation in environmental variables, such as prey availability could drive slight differences in reproductive timing (Ketterson et al., 2015). These populations also differ genetically, potentially enough to be considered a separate subspecies (Van Cise et al., 2019).

These analyses set the stage for future studies, establish sex-specific baselines for hormones of interest, and allow us to describe seasonal trends in sex steroids and glucocorticoids. This approach expands the existing conservation toolbox, improving our ability to detect pregnancy and identify peak mating and calving periods.

3.2 Materials and Methods

3.2.1 Sample Collection

We collected blubber samples (n = 114) from free-ranging short-finned pilot whales in the western North Atlantic. Nearly all samples (93 %) were collected off the coast of Cape Hatteras, North Carolina, near the continental shelf break, about 70 km offshore (Figure 6). The remaining 7 % of samples (n = 8) were collected off the coast of Jacksonville, Florida. While there is considerable separation between these locations, individuals tagged in Jacksonville have traveled to Cape Hatteras and further north (Foley, 2018), suggesting they are members of the same population.

Regular cetacean surveys have been conducted at the Cape Hatteras field site since 2006, and a photo-identification catalogue has been developed to track individuals, using distinctive patterns of nicks and notches on their dorsal fins. The photo-identification catalogue includes 1,300 individuals, approximately 35% of which have been sighted more than once. Although animals are seen year-round, we primarily

collected biopsy samples between May and August (Figure 7), when weather conditions are more favorable for fieldwork.

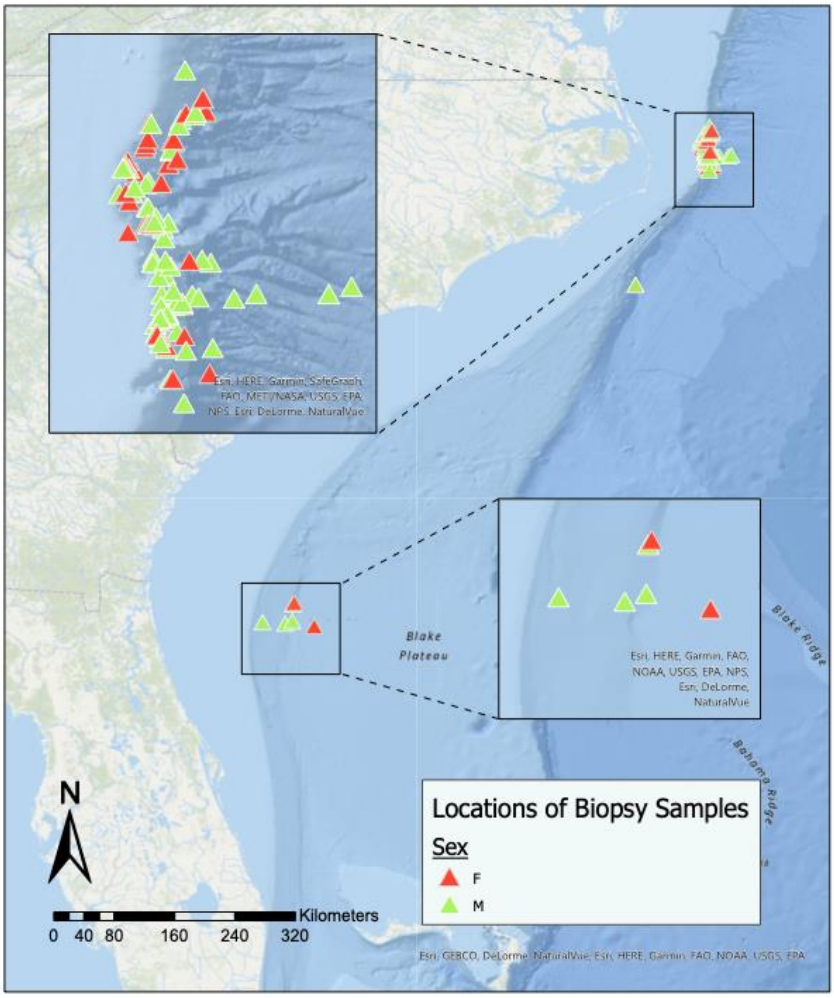

Figure 6: Pilot whale study site and biopsy collection locales off the North Carolina coast.

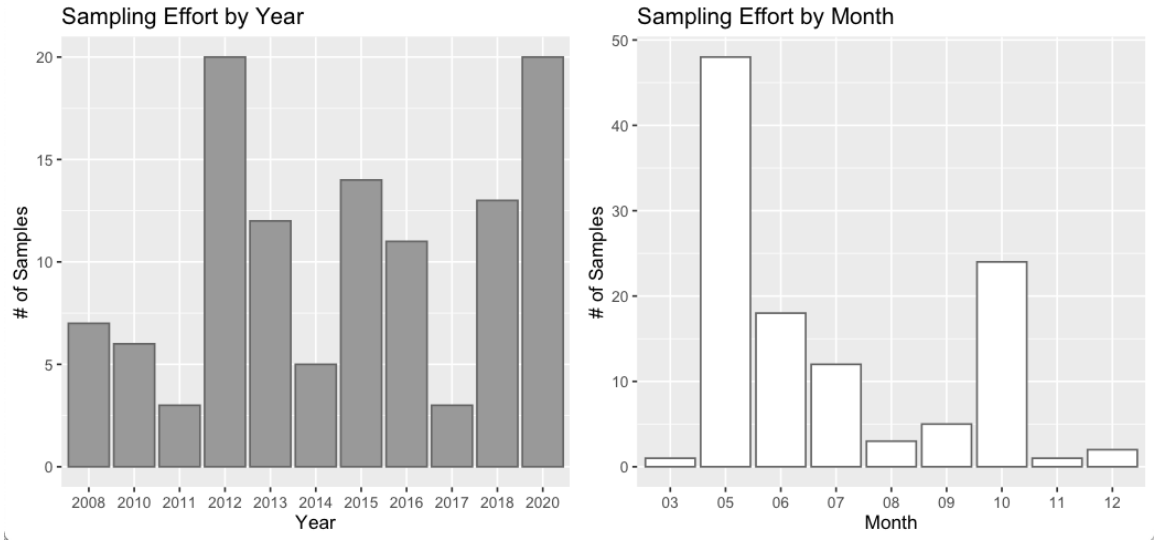

Figure 7: Temporal distribution of biopsy collection of short-finned pilot whales.

We used crossbows with modified bolts fitted with sterilized 25mm stainless steel sampling tips to collect remote blubber biopsies, which we stored on the research vessel on ice or flash-frozen in liquid nitrogen, before relocation to long-term storage in a -80°C freezer. Following permit requirements, we avoided repeated sampling of individuals and did not biopsy dependent calves or adult females with neonates. Despite precautions, photo-ID showed three individuals had been re-sampled; in these cases, only the first biopsy was included in analyses. There is a significant sex bias in our sampling (64% male versus 35% female), which reflects our preferential sampling of large, well-marked individuals and our avoidance of females with dependent calves.

3.2.2 Defining Baseline

These biopsy samples were collected over a twelve-year period and as part of numerous studies with a variety of objectives. During this period, researchers tagged subjects and conducted acoustic playbacks, potentially inducing stress responses prior to biopsy. Out of caution, we removed 74 biopsies collected after tagging, repeat biopsy, or sound exposure experiments. The remaining 40 biopsies (35%) in our potential baseline

subset included 27 males, 12 females, and one animal of undetermined sex. This subset included animals exposed to indirect, within-group tagging (n = 13) and biopsy (n = 13). To test whether these animals should be excluded from the baseline subset, we compared cortisol concentrations between each group and remaining baseline samples. Additionally, all animals were unavoidably exposed to research vessel presence while biopsies were collected. To determine whether individuals exposed to longer vessel presence should be removed, we tested the relationship between cortisol and survey sighting time.

3.2.3 Assigning sex for subjects

We assigned sex by extracting DNA from skin cells and using a polymerase chain reaction to amplify segments of DNA from the X and Y chromosomes based on the p2-3ez/p1-3ez (Aasen & Medrano, 1990; Fain, 1995) and Y53-3C/Y53-3D (Fain, 1995) primer systems. The presence of 447bp and 224bp target segments were visualized with gel electrophoresis, using one band to identify females (X) and two to identify males (XY). Early analyses were conducted by research partners at the National Oceanic and Atmospheric Administration; more recent analyses were conducted at the Duke University Marine Lab in Beaufort, North Carolina. One sample did not have clear results from DNA analysis and did not have enough material available for re-analysis. This sample was categorized as 'unknown sex' and was excluded from sex-specific analyses.

3.2.5 Assigning pregnancy status for subjects

We used progesterone to separate females into pregnant and non-pregnant categories. To establish a threshold for this distinction, we calculated upper tolerance limits for progesterone in males. These upper tolerance limits represent a threshold under

which we expect 95% of the represented population to lie (with 95 % confidence). We also created an initial group of non-pregnant females, using the highest male progesterone concentration as a cutoff for inclusion. We calculated the upper tolerance limit for this group of females, which we compared to the upper limit established for males. We then used the higher of the two as our threshold for categorizing females as possibly pregnant versus non-pregnant.

Because there are no existing data on pilot whale androgens during pregnancy, we could not use DHEA, androstenedione, or testosterone to definitively categorize pregnancy in this study. However, we used multi-variate analyses to assess the potential for elevated androgen concentrations to indicate pregnancy.

3.2.6 Hormone Analysis

Biopsies were kept frozen at -80 °C prior to processing at the National Institute of Standards and Technology, in Charleston, South Carolina. Prior to processing, we removed skin from the biopsy plug and weighed the remaining blubber samples, which weighed between 0.0603 g and 0.5379 g. We processed the samples in four batches between 2018 and 2020 (Table 6), extracting and analyzing them for 11 steroid hormones (Table 5).

To control for variable sample loss during extraction, we added isotope-labeled hormones to each sample and calibrant prior to extraction. We used matched internal standards (isotope-labeled hormones) for each target analyte, except for DHEA. A list of isotope-labeled internal standards and their paired hormones can be seen in Chapter 2. Isotope labeling does not alter elution time in the liquid chromatography step, but these

analytes are easily distinguished during mass spectrometry, due to differences in compound mass.

We conducted extractions with the previously established Quick, Easy, Cheap, Effective, Rugged, and Safe (i.e. QuEChERS) method (Boggs et al., 2017). After the first batch, we altered the final step of the QuEChERS protocol to reconstitute small samples (mass less than 0.2 g) in 60 % of the solvent volume used for larger samples. Dilution changes from altering this step were accounted for with the described use of labeled internal standards.

Table 5: List of steroid hormone analytes quantified using the described LC-MS/MS method. Analytes are listed under the common names used in this chapter, their common abbreviation and steroid class.

Steroid Class	Abbreviation	Analyte	Method
Glucocorticoids	F	Cortisol	C18
	E	Cortisone	C18
	B	Corticosterone	C18
	S	11-Deoxycortisol	C18
	11DOC	11-Deoxycorticosterone	C18
Mineralocorticoids	ALD	Aldosterone	C18
Androgens	DHEA	Dehydroepiandrosterone	C18
	T	Testosterone	Biphenyl
	AE	Androstenedione	Biphenyl
Progestogens	P_4	Progesterone	Biphenyl
	$17OHP_4$	17α-Hydroxyprogesterone	Biphenyl

3.2.7 LC-MS/MS Data Analysis

LC-MS/MS peaks were manually integrated in SciEx Analyst (Version: 1.6; Framingham, MA). Concentrations were calculated by finding the ratio of each target analyte to labeled internal standard in each sample and compared to a calibration curve using ratios from known-concentration calibration samples. One to three replicates of the standard reference material, SRM1945, were run with each batch to monitor consistency.

As described in Chapter 2, the LC-MS/MS method is not sensitive enough to detect all target compounds in all samples. Reporting limits (RLs) were determined for each analyte by batch, as detectability varied over the course of the study (Table 6). RLs were defined as the lowest calibrant used in the calibration curve and adjusted for the mass of each sample. The limit of detection (LOD) was calculated for each analyte by taking the mean concentration of three blank samples and adding three times the standard deviation of these measurements. If the LOD exceeded the concentration of the lowest calibrant, the LOD was used as the RL.

Two LC-MS/MS methods, described in Chapter 2, were used to analyze extracted samples. Sex steroids were separated using liquid chromatography on a biphenyl column and tandem mass spectrometry using an API 4000 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer AB Sciex (Framingham, MA) to implement the targeted monitoring method described in Boggs et al. (2017). Corticoids and DHEA were analyzed using a C18 chromatography column and the LC-MS/MS method developed in Chapter 2.

3.2.8 Statistics

All statistical analyses were performed using R version 4.1.1 (R Core Team 2021). To account for samples with analyte concentrations below the RL, we used NADA (version 1.6.1.1) and NADA2 (version 1.0.2) packages (Julian & Helsel, 2021; Lopaka, 2020). Summary statistics were calculated depending on the proportion of samples falling below the RL for each analyte (Helsel, 2012). For groups with detection frequencies above 50 %, we used NADA2::cfit to employ the Kaplan-Meier method. For

groups with detection frequencies between 20 % and 50 %, we used NADA::censtats, to estimate central tendency and spread based on robust regression order statistics.

Correlations between analytes were calculated using NADA::cenken, which computes the Kendall's tau correlation coefficient for censored data. Correlative analyses were only conducted when three or more samples were detected above the RL for a given analyte and group. These relationships between analytes were visualized using the corrplot package (version 0.92) (Taiyun Wei and Viliam Simko, 2021) 'corrplot' function. To explore multivariate relationships, we used u-scores, which order observations within each analyte, to conduct a principal components analysis (PCA) and cluster analysis, visualized as a PCA biplot and dendrogram. Uscores were calculated using NADA2::uscores, PCA was conducted with stats::princomp, and cluster analysis with stats::hclust.

To compare between groups, we used a Wilcoxon rank sum test. To ensure all non-detections were tied and ranked below detections, we first re-censored data at the highest RL for each analyte. We only performed comparisons for analytes with more than one sample detected above the RL. To investigate the effects of sighting time and seasonal changes, we used the cenGAM package (Fang, 2017), which builds on the mgcv package (Wood et al., 2016) to run Generalized Additive Models with censored data. For models with more than four quantifiable samples from each sex, we created a date-only model and models that incorporated sex as an effect. We used an analysis of variance (ANOVA) with chi-squared tests and compared Akaike information criterion (AIC) scores to select the best-performing model.

3.3 Results

3.3.1 Defining Baseline

We did not remove any of the 40 samples from the baseline subset. Cortisol concentrations of individuals exposed to within-group biopsy did not differ from individuals sampled first in their group ($p = 0.3676$), neither did individuals in groups where one or more individuals were tagged ($p = 0.1296$). For the 40 baseline samples, cortisol concentrations did not change significantly as animals were exposed to longer periods of vessel presence ($p= 0.102$, Figure 8).

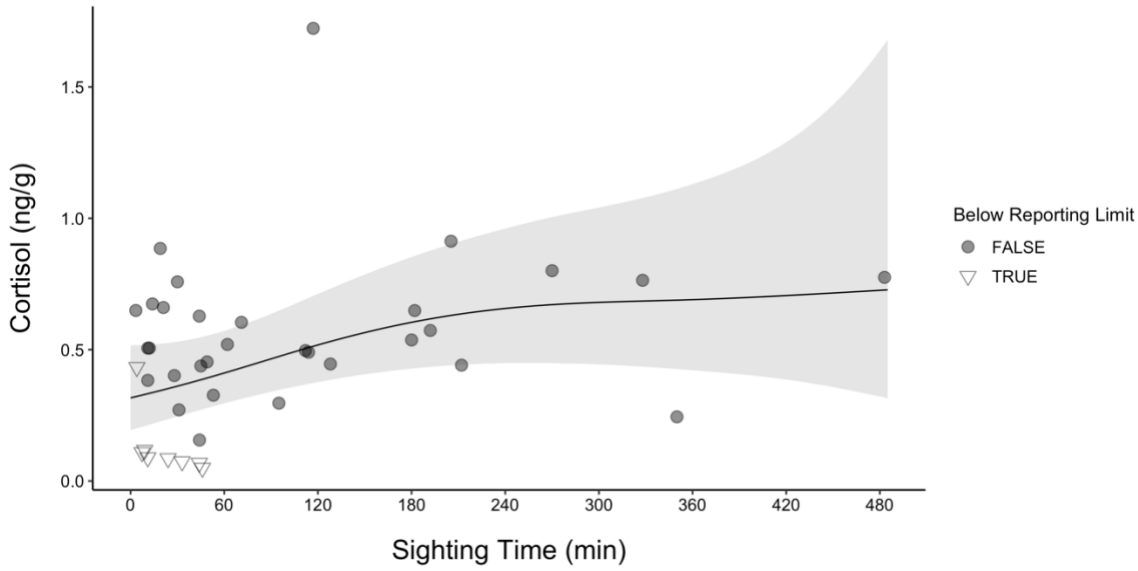

Figure 8. Cortisol concentrations of the 40 baseline samples. The line represents a non-significant relationship between vessel presence (or sighting time) and cortisol concentrations. Open triangles represent samples for which we could not quantify cortisol concentrations and are located at the reporting limit. True values for those concentrations lie between 0 and the reporting limit.

3.3.2 LC-MS/MS Detection Rates

Detection rates were highly variable for the 11 target analytes, ranging from 0 to 79 %. Relatively high reporting limits likely contributed to low detection rates for some of these analytes. RLs were particularly high for aldosterone and DHEA (mean RLs = 5.51 ng/g and 1.85 ng/g, respectively), and varied throughout the study (Table 6). We removed four analytes with low detectability from further analysis. This included aldosterone and 11-deoxycortisol, which were unquantifiable in all samples, and corticosterone and 11-deoxycorticosterone, which were quantified in only one sample each (Table 7). Of the 11 original target analytes, seven remained for characterization: two glucocorticoids, two progestogens, and three androgens.

In some cases, detectability varied not only between analytes, but also between sexes. Detection rates were the highest in the two remaining glucocorticoids, cortisol and cortisone, which were quantifiable in most samples (79 % and 64 % respectively), regardless of sex. For the five remaining sex steroids, detection varied strongly between males and females, with four of these analytes detected primarily in males. In the 12 biopsies from females, DHEA and 17α-hydroxyprogesterone were not quantifiable in any samples, while testosterone was quantifiable in only one, and androstenedione was quantifiable in three (Table 7). In the 27 biopsies from males, progesterone was only quantified in four samples (15%).

Table 6: Reporting limits for each analyte in ng/g. These vary by batch with calibration standards and the blank samples used to establish the limit of detection.

Analyte	Mean	Batch			
		Mar-18	Oct-18	Dec-18	Dec-20
ALD	**5.5079**	3.5394	3.5390	3.5394	11.4137

E	**0.0795**	0.1035	0.1040	0.0465	0.0639
F	**0.1758**	0.2187	0.2190	0.2187	0.0470
B	**0.9599**	1.8120	0.8690	0.1916	0.9670
S	**0.0981**	0.1067	0.1070	0.1067	0.0719
11DOC	**0.4176**	0.3664	0.3660	0.4227	0.5151
DHEA	**1.8455**	1.6903	3.6100	0.2362	N/A
17OHP$_4$	**0.4533**	0.4113	0.5178	0.4140	0.4699
T	**0.1369**	0.1505	0.1500	0.1500	0.0970
AE	**0.3334**	0.4309	0.4310	0.4310	0.0407
P4	**0.0684**	0.0710	0.0710	0.0710	0.0606

Table 7: Summary statistics for each analyte partitioned by sex. Columns show the number of samples analyzed (n), the percent of samples exceeding RLs (% > RL), and the mean inter-batch RL (Av. RL). The statistical methods used are specified as KM= Kaplan-Meier and ROS=robust regression order statistics, which were chosen depending on the proportion of samples censored below the RL. For groups with more than 80% of samples censored, mean, median, and standard deviation could not be reliably calculated and are not reported. Where calculated, significance levels are shown (p), with NS representing a lack of significant difference between sexes (alpha = 0.05).

Analyte	Sex	n	% > RL	Av. RL	Method	Median	Mean	Stdev	p
Aldosterone	all	40	0	5.508					
11-Deoxycortisol	all	40	0	0.0981					
Cortisol	F	12	100	0.1758	KM	0.4196	0.4314	0.1717	NS
	M	27	70		KM	0.5056	0.4730	0.3679	
Cortisone	F	12	83	0.0795	KM	0.1164	0.1371	0.0843	NS
	M	27	56		KM	0.1329	0.1661	0.0804	
Corticosterone	F	12	0	0.9599					NS
	M	27	4						
11-Deoxycort-icosterone	F	12	0	0.4176					NS
	M	27	4						
Progesterone	F	12	92	0.0684	KM	0.3097	7.725	15.45	p < 0.001
	M	27	15						
17-Hydroxy-progesterone	F	12	0	0.4533					NS
	M	27	11						
Testosterone	F	12	8	0.1369					p < 0.001
	M	27	81		KM	1.200	1.665	1.613	
Androstenedione	F	12	25	0.3334	ROS	0.0013	1.663	4.683	p < 0.01
	M	27	74		KM	2.178	4.087	5.059	
Dehydroepi-androsterone	F	6	0	1.845					NS
	M	22	68		ROS	0.8660	2.654	3.814	

3.3.3 Sex as a Contributor to Variation

3.3.3.1 Baseline Concentrations

Baseline sex steroid hormone concentrations differed significantly between males and females. Testosterone and androstenedione were significantly higher in males than females ($p = 0.00037$ and $p = 0.0099$, respectively), while progesterone was significantly higher in females than males ($p = 0.00042$) (Table 7). Because DHEA and 17α-hydroxyprogesterone were not quantifiable in any female samples, we could not test sex differences. On the other hand, glucocorticoid concentrations did not vary significantly between males and females (Figure 9).

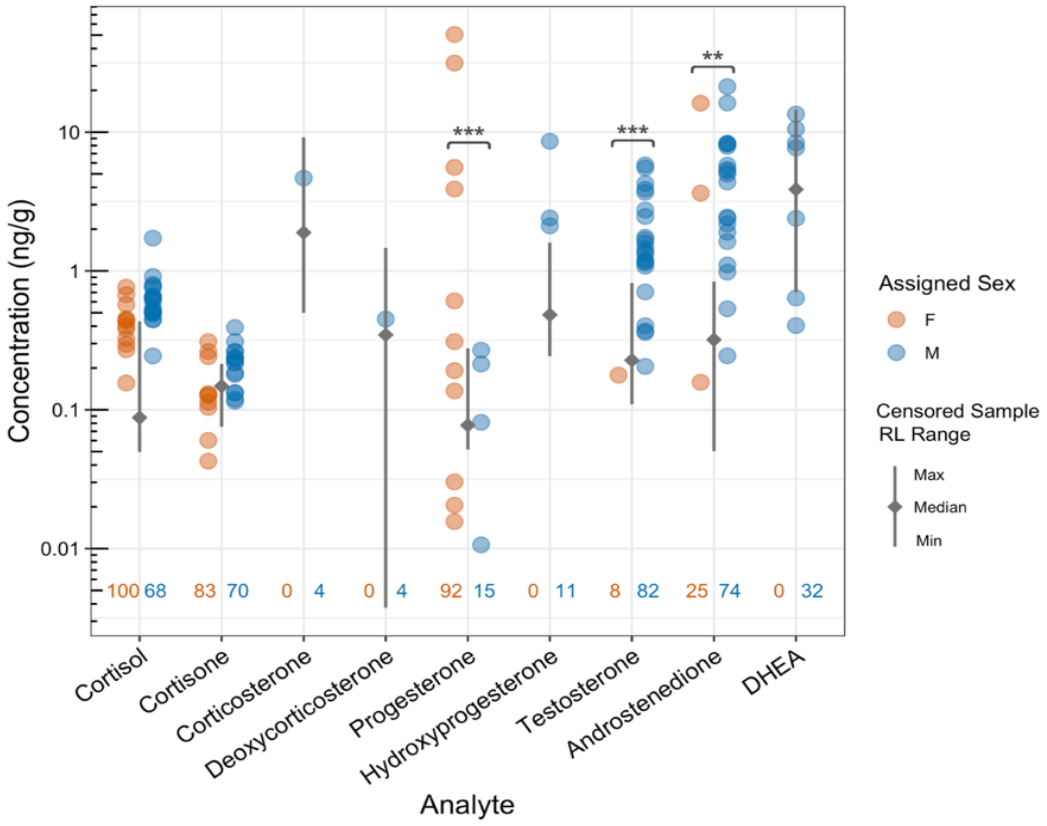

Figure 9: Baseline blubber hormone concentrations from short-finned pilot whales sampled off the North Carolina and Florida coasts. Analytes are split by sex, with females in orange and males in blue. Closed circles represent concentrations of samples above the RL. Vertical grey bars represent the range of sample-specific RLs

for censored samples. RLs vary based on sample mass and batch sensitivity, so low analyte concentrations are easier to detect in large samples. Detection frequencies (percentage of samples exceeding the RL) are shown above the bottom of the x-axis for each analyte by sex. Significance of sex differences are shown as ** $p < 0.01$ and *** $p < 0.001$.

3.3.3.2 Correlational Analysis

Correlation analyses confirmed significant relationships both within and between steroid groups (Figure 10). The two glucocorticoids, cortisol and cortisone, were positively correlated with one another, regardless of sex (females $p < 0.05$ males $p < 0.01$), while their relationships with other steroid groups depended on sex. Cortisone was significantly correlated with testosterone ($p < 0.05$) in males and with progesterone in females ($p < 0.05$). In males, two of the androgens, testosterone and androstenedione, had a strong positive correlation ($p < 0.001$). Our ability to test many of these relationships was limited by the lack of samples with quantifiable analytes, which especially affected our ability to test sex steroid relationships in females.

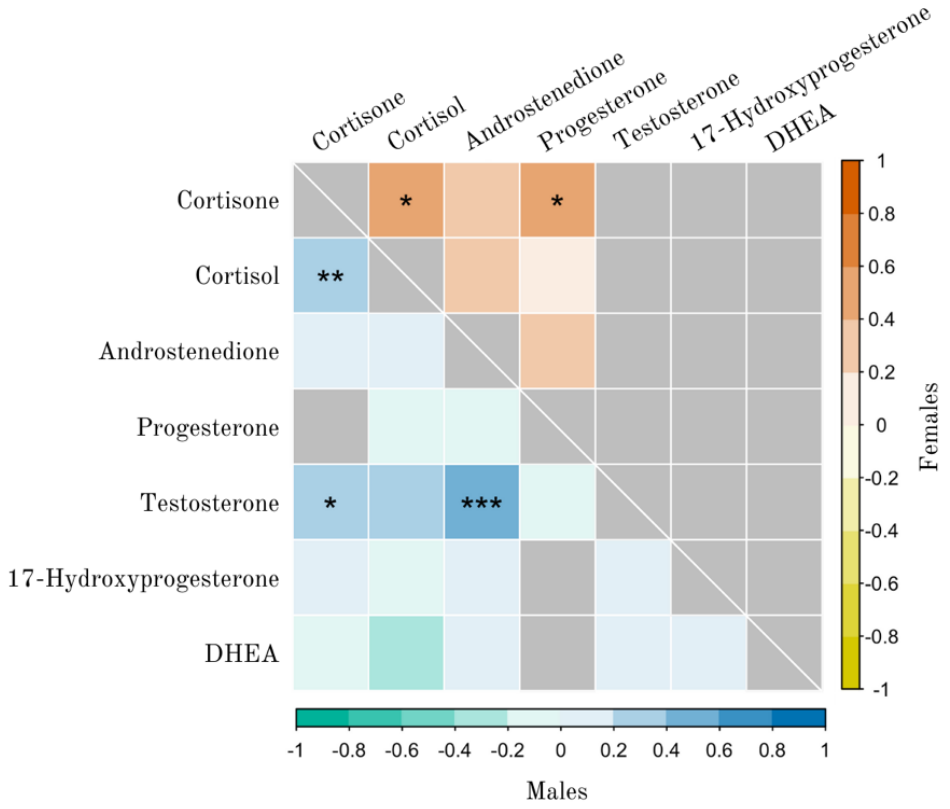

Figure 10: Correlogram for baseline samples. Squares on the bottom left portion represent males and the squares on the top right represent females. Colors indicate Kendall's tau correlation coefficient, with females in yellow-orange and males in green-blue. Asterisks designate significance thresholds (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Boxes in grey represent analyte pairs that were not included, as too few samples exceeded RLs.

3.3.4 Pregnancy Status as a Contributor to Variation

3.3.4.1 Pregnancy determination

The highest progesterone concentration measured in baseline males (n = 27) was 0.2688 ng/g; however, the highest progesterone concentration measured in all the male samples collected (n=70) was 0.4934 ng/g. We initially separated baseline females with progesterone concentrations below 0.4934 ng/g to represent non-pregnant females. The upper cutoff for this subset of females was 1.613 ng/g (90th percentile, 95% upper

confidence interval). We compared this to the upper limit established for males, 0.08280 ng/g and decided to use the non-pregnant female cutoff (1.613 ng/g) as a threshold for categorizing females as non-pregnant (i.e., likely immature or non-pregnant adults) or possibly pregnant (i.e., likely pregnant or in the post-ovulation luteal phase).

Of the 12 baseline females, eight were categorized as non-pregnant, all with progesterone concentrations below 0.6081 ng/g. The four animals categorized as possibly pregnant had progesterone concentrations ranging from 3.888 ng/g to 50.44 ng/g. Androstenedione was detectable above the RL in three of the 12 female samples. Two of these samples came from individuals categorized as possibly pregnant based on progesterone concentrations (ZTS-13-020 and HJF-15-001) and were elevated above the median male androstenedione concentration (2.178 ng/g).

3.3.4.2 Pregnancy status adds resolution to multi-hormone variation

Multivariate analyses showed variation in steroid hormone concentrations across different groups (Figure 11). There was some overlap between males and females, but clusters were significantly different between groups ($p < 0.002$). When females were divided into non-pregnant and possibly pregnant groups, the clusters gained resolution, with possibly pregnant females diverging further from males. In the PCA (Figure 11), there was no overlap between the ellipses representing male samples and possibly pregnant female samples.

The cluster analysis separated samples into three primary branches (Figure 11). The first branch only included the two females with elevated progesterone and androstenedione concentrations, ZTS-13-020 and HJF-15-001. The next branch was comprised of 18 individuals, all males. The last branch included 19 samples: nine males,

all eight non-pregnant females, and two possibly pregnant females. The males in this last branch fell into clusters characterized by low androstenedione, low testosterone, or in the case of AJR-026, low androgens and detectable progesterone (0.2688 ng/g).

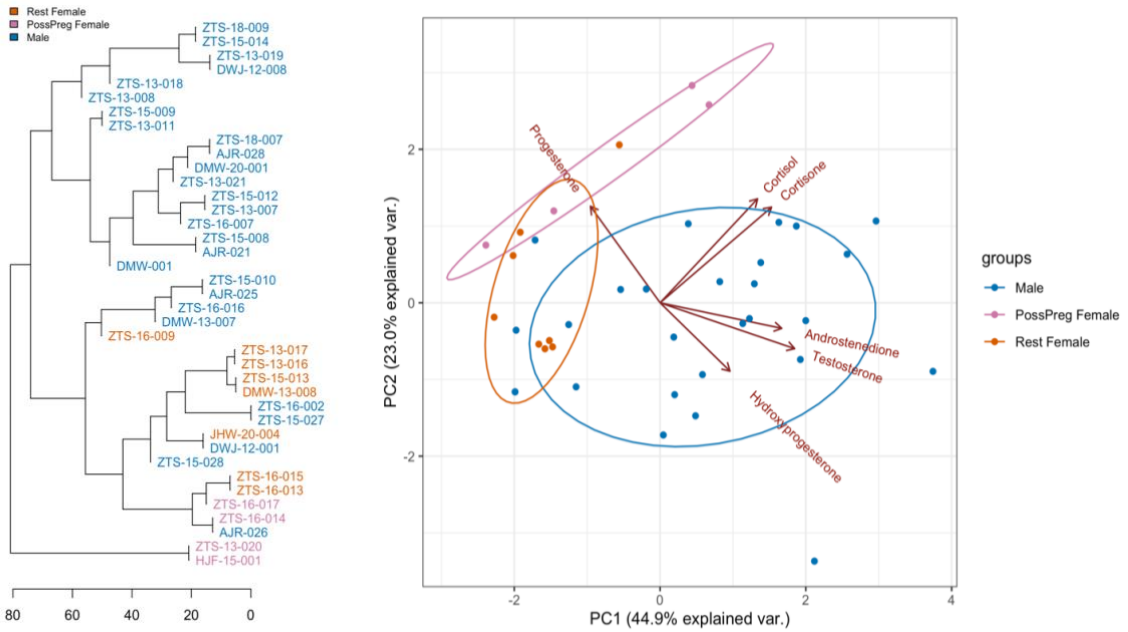

Figure 11: Multivariate relationships between samples. (a) Cluster analysis of u-scores and (b) PCA conducted on u-scores, showing the distribution of samples along the first two principal components. Ellipses capture each demographic group with 68% confidence. Males are shown in blue, resting females in orange, and possibly pregnant females in pink.

3.3.5 Season as a Contributor to Variation

For both glucocorticoids, cortisol and cortisone, seasonal variation was best modeled without including sex as an effect. Model fits peaked around May and June (Figure 12a and 12c) for both models, although neither model was significant (Table 8), showing ordinal date was not an effective predictor for these hormones. As sex steroid concentrations differed significantly between males and females (Figure 9), we modelled sex-specific seasonal variation for each analyte. However, low detection frequencies only

enabled the modeling of one sex for each analyte, so we were unable to include sex as an effect and could not compare seasonal analyte variation between sexes. In females, progesterone was not predicted by ordinal date (Table 8). High progesterone was only observed between May and August, but only two samples were collected outside of this time frame (Figure 12e), limiting our ability to test this relationship. Androstenedione, testosterone, and DHEA models only included samples from males, which spanned the months of March through December, with most samples collected in the early summer. Ordinal date was a significant driver of androstenedione concentrations ($p < 0.05$), which peaked between May and June (Figure 12b). We observed a similar pattern in testosterone, which was strongly affected by ordinal date ($p < 0.001$). Testosterone also peaked in May and June and tapered off into late fall (Figure 12d). DHEA measurements, which were characterized by high RLs for most of our study, were not predicted by ordinal date (Figure 12f). We did not quantify 17α-hydroxyprogesterone in enough samples to model its annual variation but note that all four samples detected above RL for this analyte were also collected between May and June.

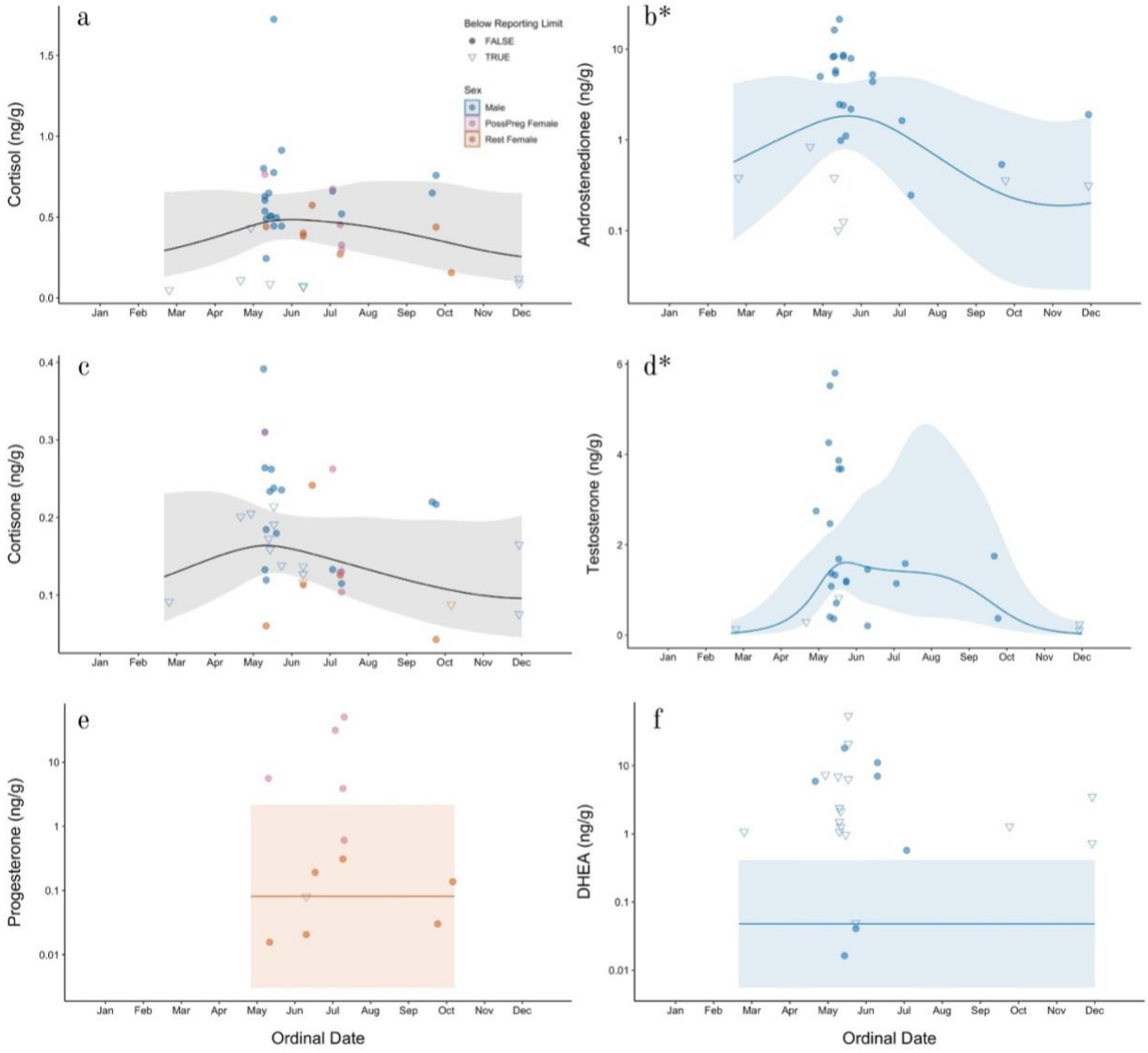

Figure 12: Annual variation in concentrations of (a) Cortisol, (b) Androstenedione, (c) Cortisone, (d) Testosterone, (e) Progesterone, and (f) DHEA . Asterisks indicate which hormones were significantly predicted by day of year. Points represent concentrations observed in males, possibly pregnant females, and resting females, which are shown in blue, pink, and orange, respectively. Samples with concentrations below RL are represented by upside-down triangles correspond with sample-specific RLs. The true concentrations of these samples fall between zero and the RL. Lines show model predictions based on day of the year and analyte concentrations.

Table 8: Selected generalized additive models (GAMs) for each hormone, showing which subset of data they were modeled for (Sex), the effective degrees of freedom

used in each model (edf), and the significance (p-value) of ordinal date as a predictor variable. *Significance ($p < 0.05$)

Analyte	Sex	Predictor	edf	p
Cortisol	all	s(OrdDate)	1.235	0.207
Cortisone	all	s(OrdDate)	1.177	0.141
Progesterone	females	s(OrdDate)	1.37E-05	0.627
Androstenedione	males	s(OrdDate)	1.361	0.0398*
Testosterone	males	s(OrdDate)	3.063	0.000798*
DHEA	males	s(OrdDate)	4.36E-05	0.449

3.4 Discussion

3.4.1 Baseline measurements

This is the first study to quantify blubber hormones in live, free-swimming short-finned pilot whales, and our measurements establish a reference point for hormone concentrations for this species. In our effort to choose representative "baseline" samples, we removed biopsies from individuals that were tagged or exposed to experimental playbacks. Determining whether tagging and playbacks elicit stress responses are beyond the scope of this chapter, but we removed these samples as a precaution. We included animals exposed to the vessel presence and within group tagging or biopsy based on the absence of evidence that these samples were different. It is possible that the tagging and biopsy of group members elicits a stress response in some scenarios and should be tested more rigorously than the analysis presented here.

In the 40 samples we categorized as baseline, we quantified cortisol, cortisone, androstenedione, testosterone, progesterone, 17α-hydroxyprogesterone, and DHEA; our analytical approach was not sensitive enough to reliably quantify endogenous aldosterone, 11-deoxycortisol, corticosterone, or 11-deoxycorticosterone. The decreased quantitation in biopsies compared to stranded specimens (Chapter 2) is not surprising and

likely reflects the lower corticoid concentrations measured in biopsy samples. For analytes with low detectability, future studies should consider using a more sensitive steroid quantitation method, like enzyme immunoassay or more sensitive mass spectrometers.

3.4.2 Sex

Sex was a critical factor for characterizing hormone variability, especially for sex steroids. Sex hormone concentrations varied between sexes, with testosterone and androstenedione significantly higher in males and progesterone significantly higher in females. These differences were reflected in detection rates, which were more than 30% lower for females in several sex steroids.

Glucocorticoid concentrations did not differ between males and females, and cortisol and cortisone had a significant positive correlation regardless of sex. However, many of the significant correlative relationships were contingent on sex. In males, there was a tight relationship between testosterone and androstenedione. Cortisone had significant relationships with hormones from other steroid groups, showing a positive correlation with progesterone in females and with testosterone in males. Cortisone, testosterone, and progesterone are primarily produced in different tissues, so we suspect these inter-group relationships may be the result of external factors, such as reproductive behavior. These analyses were limited by low detection rates but suggest relationships worth further investigation. Analysis of DHEA, which may be an indicator of chronic stress (Gundlach et al., 2018), was limited to a subset of males. The negative relationship between cortisol and DHEA was not significant but should be investigated as endocrine indications of acute and chronic stress.

Variation between and within sexes were apparent in multivariate analyses. Two-thirds of the males clustered separately from females on the dendrogram. The other third of males clustered with females and were characterized by low androstenedione and/or low testosterone. These low androgen concentrations likely reflect animals that have not yet reached sexual maturity (Kita et al., 1999).

3.4.3 Pregnancy Status

Unfortunately, we were unable to confidently confirm pregnancy in live animals. We confirmed pregnancy status in the two stranded females from Chapter 2, but our categorization of pregnancy status in live females via biopsy is not without potential error. If pregnancy concentrations of blubber progesterone can fall below the 1.613 ng/g cutoff, or if blubber progesterone exceeds 1.613 ng/g during the luteal phase, some of our female samples may have been miscategorized.

Longitudinal sampling of females in captivity has provided useful references for hormone variation throughout gestation, pointing to progestogens and androgens as markers of pregnancy (Dalle Luche et al., 2020; Legacki et al., 2020; Robeck et al., 2017; Steinman et al., 2021; Zhang et al., 2021). Indicative of mid-to-late gestation (Dalle Luche et al., 2020; Legacki et al., 2020), we observed high androstenedione concentrations in two of the biopsies from possibly pregnant females and in the two known-pregnant stranded individuals from Chapter 2. Pregnancy is a complex endocrine phenomenon that varies among species, and short-finned pilot whales do not appear to exhibit the elevated testosterone or DHEA observed in pregnant cetaceans from other species (Dalle Luche et al., 2020; Legacki et al., 2020; Robeck et al., 2017; Steinman et

al., 2021; Zhang et al., 2021). If these signals occur in pilot whales, we did not detect them in blubber.

There may also be a spatial component to the variation observed in this study. There were not enough samples from the Florida coast to compare with North Carolina samples, but it is worth noting that the four females sampled near Florida grouped together in the cluster analysis, along with one male (AJR-026), which were all characterized by elevated progesterone and low androstenedione. Two of these females were categorized as possibly pregnant, while the other two had the highest progesterone concentrations observed in the non-pregnant category. This split of this clustering indicates some luteal or early-pregnancy females may be falling below the established progesterone threshold. Short-finned pilot whales have been shown to travel long distances (Foley, 2018; Thorne et al., 2017), and it is possible that they use different areas for mating and calving. If such spatial and temporal patterns in reproduction can be resolved, endocrine research has the potential to inform research priorities and management strategies.

3.4.4 Seasonality

Seasonal shifts in behavior and reproduction likely contributed to the variation observed in male hormone concentrations. Of the nine males that clustered with females in the dendrogram (Figure 11), four were sampled between the months of October and March, when androgen concentrations appear to decrease in this population (Figure 12). Hormone concentrations in these western North Atlantic males appear to follow a similar pattern to their western Pacific counterparts, which exhibit diffuse seasonal reproduction peaking in the early summer. The males in our study exhibited elevated testosterone and

its precursor, androstenedione, in May and June. Progesterone concentrations in females did not show predictable seasonality, but this is unsurprising. Even if progesterone were only elevated during pregnancy, the 15-month gestation period in this species would lead to overlap between peaks of early and late pregnancies. This is complicated by progesterone elevation in luteal females and our inability to distinguish between pregnant and non-pregnant females. Elucidating this pattern would require improved pregnancy detection and increased sampling effort throughout the year.

Beyond its utility for identifying reproductive timing, understanding seasonal variability in steroid hormones is critical for studying stress responses. Annual rhythms in baseline glucocorticoid concentrations are thought to reflect seasonal demands for physiological and behavioral responses to adversity (Romero, 2002). In this study, sex and ordinal date were not significant predictors of baseline glucocorticoid concentrations, but we cannot rule them out as contributors. Both may interact with other factors, like body condition or reproductive status, and recent multi-parameter models show sampling month is an important contributor to glucocorticoid measurements in whales (Lemos et al., 2022). The magnitude of glucocorticoid response to stressors can also vary seasonally (Romero, 2002), so it is critical that researchers account for seasonality in experimental design.

Overall, our study establishes baseline reference concentrations for biologically important steroid hormones in short-finned pilot whales. For most of the hormones measured, short-finned pilot whales show similar concentration ranges to other odontocetes; but they also show some unexpected patterns, such as the relatively low progesterone concentrations observed in females. Sex, reproductive state, and time of

year contribute to the observed hormone variation in short-finned pilot whales. These factors should be considered in future research design and highlight the need for increased understanding of the behavior and physiology of cetaceans.

4. Hormone Responses of Pilot Whales to Simulated Midfrequency Active Sonar

4.1 Introduction

Over the last two decades, there has been a concerted effort to understand the impacts of noise on cetacean populations (Council, 2005; National Academies of Sciences, 2017; Southall et al., 2016). Cetaceans' reliance on sound makes them particularly vulnerable to noise pollution, which has rapidly escalated over the past several decades through the expansion of shipping, offshore development, energy exploration, and military sonar (Hatch & Wright, 2007; Hildebrand, 2009). While high-intensity noise events from these sources can have lethal consequences (Cox et al., 2006; Evans & England, 2001), most exposures result in sub-lethal effects (Council, 2005).

Even at these lower levels, noise exposure can interfere with cetacean communication (Scheifele et al., 2005), foraging (Goldbogen et al., 2013), and habitat use (Bryant et al., 1984; Morton & Symonds, 2002). These changes may be subtle, but can diminish the health of individuals and populations (New et al., 2014). The impact of sub-lethal disturbance on population-level health is challenging to investigate empirically, especially in long-lived species, which hinders our ability to predict and mitigate the effects of disturbance.

The Population Consequences of Disturbance (PCoD) model (Figure 13) provides a framework for understanding the relationship between anthropogenic disturbance and marine mammal population health (National Academies 2017). When disturbance becomes more frequent or severe, short-term behavioral and physiological responses can

accumulate to diminish the survival and reproduction of individuals. As more individuals are exposed to a disturbance, the potential for population-level impacts grows.

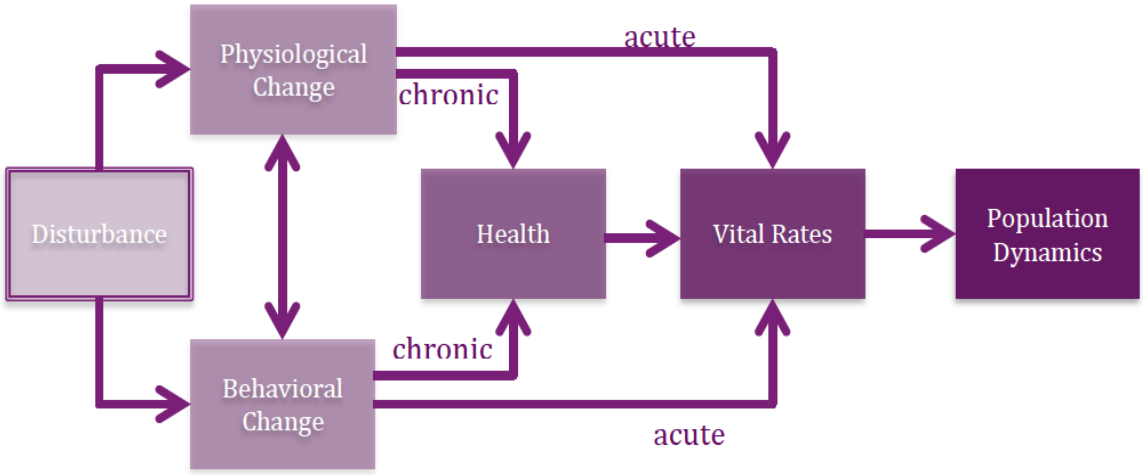

Figure 13: The PCoD model (figure adapted from National Academies 2017)

Although the PCOD model was initially proposed in 2005 (originally as the PCAD model (Council, 2005)), many of its links remain poorly understood. Without the ability to monitor direct connections between disturbance and individual fitness, current research tends to focus on intermediary links, particularly the initial relationship between disturbance and changes in behavior and physiology (e.g., injury, allostasis, movement, vocalization, and foraging) (National Academies of Sciences, 2017). Physiological responses have typically been challenging to study (Council, 2005; Pirotta et al., 2018), so the bulk of what we know about cetacean responses to noise disturbance is based on behavioral response studies (Dunlop et al., 2015; Goldbogen et al., 2013; Houser et al., 2013; Miller, 2012; Nowacek et al., 2004; Southall et al., 2016). These studies can be conducted opportunistically or as controlled exposure experiments (CEEs), the latter

enabling experimental manipulation of exposure parameters, such as received levels, which can contribute to dose-response functions (Harris et al., 2018).

Widespread behavioral response studies, enabled by the rapid development of multi-sensor tags (Johnson & Tyack, 2003) and analysis methods, have demonstrated varying responses of cetaceans to acoustic disturbance and the relevance of context (Southall et al., 2016). Acoustic characteristics of the stimulus, like received level, distance to the sound source, rise time in the signal, and dosage can impact whether a response occurs and the severity of response (DeRuiter et al., 2013). The acoustic context interacts with species-specific factors, like auditory ranges and predation risk (Cure et al., 2012; Visser et al., 2016), and individual factors, like an animal's behavioral state when exposed and previous exposure experience (Ellison et al., 2012; Harris et al., 2015).

In some cases, behavioral methods are not well-positioned to detect responses to noise. For example, when animals rely strongly on specific areas for activities like foraging and calving, they are more likely to remain in those areas. Rather than displaying avoidance, which is often relied upon to indicate disturbance, they are likely to endure the presence of a stressor (i.e., tolerance). Thus, the lack of an obvious behavioral change, in this case avoidance, does not confirm the lack of a response or decreasing responsiveness to an aversive yet non-injurious stimulus (i.e., habituation) (Gotz & Janik, 2011). In the case of distinguishing between tolerance and habituation, physiological markers, like glucocorticoids, may be useful indicators of a response.

When acting as a stressor, noise exposure can stimulate the hypothalamic-pituitary-adrenal (HPA) axis, which inhibits non-urgent functions (e.g., immunity,

growth, and reproduction) and mobilizes energy for physiological and behavioral responses (Sapolsky et al., 2000). As mediators of this response, glucocorticoids can be measured to detect physiological responses to stressors. Although glucocorticoids, like cortisol and cortisone, fluctuate naturally to accommodate energetic demands, chronic elevation can diminish survival and fecundity (Boonstra et al., 1998; Romero & Wikelski, 2001). Despite the relevance of glucocorticoids as biomarkers, their use in cetacean studies has been limited by the logistical challenges of sampling and analysis. Previously limited to serum sampling, experimental observations of hormone responses to noise were made in captive facilities. These studies yielded inconsistent results (Romano et al., 2004; Thomas et al., 1990), constrained by small sample sizes and the use of handling to obtain samples, which itself can elicit a stress response (Fair et al., 2014; St Aubin et al., 1996; Thomson & Geraci, 1986).

In recent years, alternative methods for hormone sampling in cetaceans have rapidly expanded, yielding opportunities to better identify and understand responses to noise. To date, two studies have connected glucocorticoid measurements in baleen whales with vessel traffic and elevated background noise, both opportunistically sampling feces (Lemos et al., 2022; Rolland et al., 2012). Feces cannot be reliably collected in many cetacean species, but blubber tissue is another promising matrix for these studies. Over the past few decades, remote blubber biopsy has become a reliable method for tissue collection without the need for capture (Lambertsen, 1987; Noren & Mocklin, 2012), and hormones in blubber tissue can be reliable indicators of biological states, like pregnancy and stress (Champagne et al., 2018; Kellar et al., 2014). Recent

bottlenose dolphin studies show blubber glucocorticoids increase within two hours of HPA activation (Champagne et al., 2018; Galligan et al., 2020), making this a feasible tissue for measuring potential physiological responses to CEEs in free-ranging cetaceans.

In this study, we ask if a presumed acute stress response to simulated naval sonar can be detected in blubber from short-finned pilot whales. While overt behavioral responses have not been observed in short-finned pilot whales following naval sonar exposure, short-finned pilot whales have stranded in association with sonar events (Hohn et al., 2006), and behavioral responses to sonar have been observed in long-finned pilot whales (Antunes et al., 2014; Rendell & Gordon, 1999), suggesting sonar is a likely stressor for short-finned pilot whales. We sought foundational insight into the link between acoustic disturbance and physiological response by conducting a simulated sonar exposure and measuring hormones in post-exposure blubber tissue samples.

4.2 Materials and Methods

4.2.1 Controlled Exposure Experiment (CEE)

We conducted this study off the US east coast, about 70 km east of Cape Hatteras, North Carolina. We conducted three simulated midfrequency active sonar (MFAS) exposures during two field days in October 2020. We used two research vessels: a charter fishing vessel that served as the platform for the acoustic source and a 9 m rigid hull Safe Boat® for collecting remote blubber biopsies. The acoustic source, described by Southall et al., (2012), is comprised of 15 elements in a vertical line array. Before exposure, the charter vessel was positioned about 500 to 1500 m from a focal group, targeting received levels between 120 to 150 dB re 1 μPa RMS (hereafter dB). Each exposure lasted 30

minutes, beginning with a source level of 150 dB at 1 m and ramping up to 200 dB at 1 m during the first 7 minutes. The transmission consisted of tones between 3.5 and 4.05 kHz and transmitted 1.6-second simulated MFAS signals played every 25 seconds. The signal was designed to emulate some of the tactical midfrequency sonar systems used by the U.S Navy.

Following each sound exposure, the rigid hull vessel team biopsied up to three individuals in the focal group they had been tracking during exposure, then left the focal group to search the area for nearby groups to biopsy (i.e., non-focal groups), continuing this process for two hours after the exposure ended. After the end of the sound exposure period, the charter vessel team also located non-focal groups and collected biopsies. No individuals were re-sampled and no more than three individuals were biopsied in each group. Within each group (focal and non-focal), all biopsies were collected within 30 minutes of biopsy initiation to minimize the potential for disturbance from sampling. Groups are hereafter referred to by names combining the CEE (i.e., A, B, or C) and the order of group encounter for each CEE (i.e., 1, 2, 3, etc.). For example, group C2 is the second group we encountered after the third CEE.

On the first day, we conducted two CEEs. The second CEE was conducted 15 km SSE from the first exposure location. Using the propagation model detailed below (section 4.2.4), we estimate received levels at the second CEE location were 95.33 dB during the first CEE. Due to weather conditions, we could only collect one biopsy after the second exposure, which came from a non-focal group (B1) 106 minutes after the playback began.

In total, we conducted three exposures over two days and collected 18 biopsies from 10 groups. Following permit requirements, we did not expose groups containing neonates and did not biopsy dependent calves. Biopsies were collected using crossbows and modified bolts with 25mm stainless steel sampling tips. Most of these biopsies were taken from the dorsal-lateral region below the dorsal fin, which we target for consistency, but a few of these biopsies were taken from the dorsal fin.

4.2.2 Hormone Analysis

Biopsies were immediately stored on ice and transferred to temporary -20°C storage, then long-term storage at -80°C. Samples were extracted and analyzed at the National Institute of Standards and Technology in Charleston, South Carolina. After removing skin from the biopsy plug, we weighed and extracted the entire blubber portion using the Quick, Easy, Cheap, Effective, Rugged, and Safe (i.e., QuEChERS) method (Boggs et al., 2017), as described in Chapter 2. To account for sample loss during processing, we added isotope-labeled matched internal standards before extraction.

We used LC-MS/MS to quantify ten steroid hormones of interest (Chapter 2) but focused our analyses on two glucocorticoids that could be reliably quantified: cortisol and cortisone. These analytes were separated with liquid chromatography using a C18 column, before electrospray ionization and scheduled multiple reaction monitoring with an AB Sciex (Framingham, MA) API 4000 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometer. Analyte peaks were confirmed with matched internal standards and manually integrated in Sciex Analyst (Version: 1.6; Framingham, MA). Analyte concentrations were determined with calibration curves, established by ratios of known

concentration calibrants and matched internal standards. The lowest concentration calibrant determined the reporting limit for each analyte unless it was exceeded by the limit of detection, calculated as the mean of three blank samples plus three times the standard deviation of these measurements. To account for varying masses between samples, reporting limits were mass-adjusted by sample. We refer to samples with analyte concentrations below their reporting limit as being censored and account for the uncertainty of their concentrations with censored statistics (Helsel, 2012).

4.2.3 Demography of sampled animals

We assigned sex to samples using DNA extraction and polymerase chain reactions to amplify DNA segments from the X and Y chromosomes, which we visualized with gel electrophoresis (Aasen & Medrano, 1990; Fain, 1995). Of 18 samples, seven had Y chromosome segments and were assigned male; the other 11 were assigned as female. Three of the 11 females had progesterone concentrations indicative of pregnancy or ovulation, and one female was just below the 1.6 ng/g cutoff established in Chapter 2. One of the high-progesterone females also had elevated androstenedione, which may indicate late-stage gestation (Dalle Luche et al., 2020; Legacki et al., 2020).

4.2.4 Determination of exposure

To determine whether animals were exposed to the simulated sonar at levels above ambient noise, we estimated RLs for each group, which we compared with background noise for that day. To calculate RLs, we used the RLtool (Margolina pers. comm.) in MATLAB (version 2016b), which uses bathymetry and oceanic conditions to run propagation models for a given animal location and depth. We calculated background

noise with the wind-to-noise function (Margolina pers. comm.) in MATLAB, which estimates background noise in the frequency of choice based on wind speeds. Wind speeds for each location were determined using E.U. Copernicus Marine Service Information (https://doi.org/10.48670/moi-00305). We did not verify received levels (RLs) with boat-based or tag-based hydrophones, nor do we know the locations of non-focal animals during exposure.

We used two methods to estimate animal location during exposure. For focal animals, which we followed with the rigid hull vessel throughout the exposure, we used the vessel track to represent their location throughout the CEE. We calculated RLs for the closest and farthest points during the exposure. For non-focal animals, which we encountered after the CEE, we estimated their farthest likely position from the sound source during the playback, which we used to determine a minimum likely RL. To choose these positions, we created a radius around the location where we first sighted each group, calculated by assuming directional transit at 2 m/s (Bowers, 2016; Tyson Moore et al., 2020; Williams & Maresh, 2015) between the end of the CEE and their initial sighting. Along that radius, we chose the point farthest from the sound source at the end of the playback (Figure 14).

Because RLs can vary throughout the water column, we ran the propagation models for each position at two depths, representing RLs experienced at the surface and during diving. We used a 10 m depth for surface calculations to avoid issues with interference at the surface. To calculate RLs encountered at depth, we used 500 m, the mean daytime foraging dive depth for this population (Shearer pers. comm.). In one case

(Group B1), the seafloor was less than 500 m deep, so we used a position 10 m above the seafloor (465 m).

Figure 14: Group C2's farthest potential location from exposure. The boat tracks are represented by white (source vessel) and dark red (biopsy vessel) lines, with the location of the sound source at the start and end of the third CEE marked with red and white airhorns. We first encountered Group C2 at the site labeled with a blue "2" about 49 minutes after the playback ended. Directional travel at two m/s during this time would be 5.87 km, designated by a red circle around the initial sighting location of that group. The red triangle marks the farthest point from the end of the CEE and the site used to determine minimum RLs and likelihood of exposure for this group.

4.2.5 Statistical analysis

Based on observations of glucocorticoid perfusion in bottlenose dolphin blubber (Champagne et al., 2018; Galligan et al., 2020), we expected an increase in blubber glucocorticoids to be detectable within the two-and-a-half-hour post-exposure period, but

we were unsure whether this peak would occur early enough to include a subsequent decline in blubber glucocorticoids. To account for potential non-linearity in cortisol and cortisone responses, we used the 'gam' function of the mgcv package (Wood et al., 2016), which models terms without the assumption of linearity. While the gam function is typically used for Generalized Additive Models (GAMs) with multiple variables, we used it to estimate smooths for a single variable, time. We did not specify the number of knots allowed and used smoothing parameters selected with Restricted Maximum Likelihood (REML), which is more resistant to over-fitting (Wood et al., 2016). Cortisol concentrations exceeded the reporting limit in all samples. To model cortisol, we ran the gam function with the inputs described above and the gamma distribution family, which works well with non-negative data.

Our LC-MS/MS method was not sensitive enough to quantify cortisone above the reporting limit in all 18 samples. Eight of these samples were censored, so we used censored statistical methods to assess cortisone responses to playback experiments. These methods are preferable to those which throw out censored samples or make substitutions (Helsel, 2012). We used the cenGAM package (Fang, 2017), which builds on the NADA, NADA2 (Julian & Helsel, 2021; Lopaka, 2020), and mgcv (Wood et al., 2016) packages to run the gam function with censored data. Instead of the gamma distribution, we used tobit1, which incorporates censorship thresholds into the model.

While other variables like demography, behavioral state, and degree of exposure could influence glucocorticoid release after exposure, we limited models to one factor per 10 samples (Bolker et al., 2009). Although we could not include sex as a factor, we did

run female-specific analyses for cortisol and cortisone with the 11 female samples. To assess models, we conducted an analysis of variance (ANOVA) with chi-squared tests and compared Akaike information criterion (AIC) scores between the smoothed, linear, and null models.

4.3 Results

RL analyses confirmed that all 18 individuals were exposed to the playback at received levels above background noise (Table 9). For focal groups, distance to the sound source ranged from 0.22 to 2.80 km during playbacks. Focal group C1 was initially 0.85 km from the source but passed within 0.22 km during the exposure, resulting in an estimated 174.7 dB exposure at the surface. RLs varied throughout the water column and were generally higher at depth, but not at this location, where RL estimates at depth were 0. A 500 m dive at this location may have occurred at a sharp enough angle below the source, which would avoid the modeled sound beam path (Jensen et al., 2011). Aside from the extremes during C1's close pass, focal animals were estimated to receive between 108.4 and 136.5 dB during playbacks.

All non-focal groups were sighted within 6 km of the playback source, but these initial sightings occurred 30-90 minutes after playbacks ended. We estimated hypothetical maximum travel radii and distances from the playback source, which ranged between 5.70 to 15.76 km. According to transmission loss models, RLs estimated for these locations were between 126.6 and 88.3 dB. Because the actual locations during

CEEs are unknown for non-focal groups, it is possible that they were exposed at RLs higher than these estimates.

Table 9: Comparison of background noise levels with RL estimates (all in dB) for each group sampled. The number of individuals biopsied in each group is noted in the 'n' column. Distances for focal animals, which were followed during the playback, represent their closest and farthest locations to the sound source during the playback. For non-focal groups, which were encountered after playbacks ended, distances represent the farthest hypothetical position from the sound source during exposure. The transmission loss model used these distances to calculate estimated RLs for each group at the surface (10 m depth) and dive depth (500 m).

CEE	Group	Focal	n	Distance (km)	RL at Surface	RL at Depth	Background noise
1	A1	focal	3	1.47	127.90	135.67	80.47
				2.80	108.39	129.68	80.47
1	A2	non-focal	1	5.70	126.55	120.04	80.49
1	A3	non-focal	1	13.96	100.89	115.19	80.26
1	A4	non-focal	1	15.22	93.45	111.59	80.25
2	B1	non-focal	1	12.41	109.26	98.87	79.28
3	C1	focal	3	0.22	174.69	0	70.18
				1.21	128.81	136.51	70.34
3	C2	non-focal	3	9.92	95.78	120.35	72.45
3	C3	non-focal	1	11.45	101.05	114.91	73.17
3	C4	non-focal	2	14.76	88.28	110.95	72.37
3	C5	non-focal	2	15.76	91.67	111.55	73.44

Glucocorticoid analyses show cortisol and cortisone responses were best modeled with non-linear smoothing functions of time, which characterize an increase and subsequent decrease in glucocorticoids during the 140 minutes following playback onset (Figure 15). Between the null, linear, and smoothed models, AIC scores were lowest in the smoothed models for cortisol and cortisone (Table 10). We used ANOVAs to rank models based on residual degrees of freedom and residual deviance, which were compared with chi-square tests. For cortisol, the smoothed model ranked higher than the

linear model (p = 0.0811), which ranked higher than the null model, but not significantly (p=0.4762). The selected smooth model (Figure 15a) used 2.139 effective degrees of freedom and explained 29.9% of the deviance but was not considered significant (p = 0.1550). Results were similar for cortisone, with the smoothed model ranking higher than the linear model (p = 0.0573) and the null model ranking last but not significantly (p = 0.7006). We selected the smooth model (Figure 15b), which used 2.285 effective degrees of freedom and was not significant (p = 0.149). In the models chosen, cortisol peaks at 83 minutes post-exposure onset and cortisone peaks slightly after, at 90 minutes.

Table 10: Statistics for models describing glucocorticoid variation in response to time after simulated MFAS exposure. Columns show the number of samples used to create each model, the degrees of freedom, adjusted r-squared, the significance (p) of time's contribution to each model, AIC score, and rankings generated with ANOVA comparisons and significance compared to the next-best model (Rank (p)).

Hormone	n	Model	df	r-sq. (adj)	p	AIC	Rank (p)
Cortisol	18	smooth	4.65	0.136	0.1550	-31.278	1 (0.0811)
		linear	3.00	-0.0494	0.4645	-28.807	2 (0.4762)
		null	2.00	0	---	-30.412	3
Cortisone	18	smooth	4.85	0.119	0.1490	-29.885	1 (0.0573)
		linear	3.00	-0.0496	0.5889	-27.099	2 (0.7006)
		null	2.00	0	---	-28.951	3

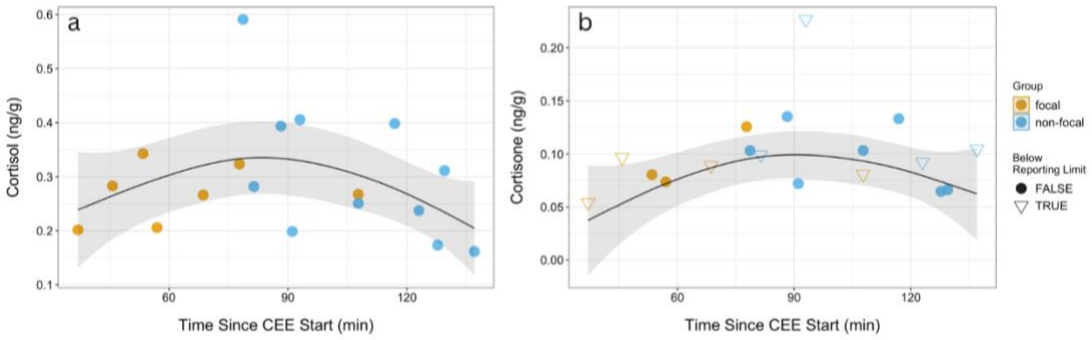

Figure 15: Glucocorticoid concentrations in pilot whale blubber following the onset of simulated MFAS. The colors of each point indicate whether they were part of focal (yellow) or non-focal (blue) groups. For cortisone (b), samples with concentrations below the reporting limit are plotted at mass-adjusted reporting limit as open triangles. True concentrations for these samples are unknown but fall between zero and the reporting limit. This uncertainty is accounted for in the analysis methods. Black lines predict responses based on the highest-performing model for each analyte, with 95% confidence intervals shown in gray shading.

Because of the small number of male samples, we could not include sex as a factor in the full models. However, we tested glucocorticoid responses in the female group, for which we had 11 biopsies. Using the same model selection procedure, we compared null, linear, and smooth models (Table 11). For cortisol, the selected smooth model only had one effective degree of freedom, resulting in a linear model. Because the smooth model was equivalent to the linear model, we excluded it from ANOVA rankings and comparisons. Between the null and linear model for cortisol in females, the linear model had a lower AIC score and performed significantly better in ANOVA (0.0039). The linear model explained 49% of deviance and was a significant predictor of cortisol (0.0104). Samples from females were collected 68 to 137 minutes after the playback started, during which time, cortisol decreased (Figure 16a).

Analysis of cortisone response in females was limited by sample censorship. Five of the 11 samples were censored (i.e., below mass-adjusted reporting limits). Using censored statistics, we compared null, linear, and smooth models. For cortisone, the AIC score was lowest in the null model. ANOVA comparisons were not significant but ranked the smooth model highest ($p = 0.2405$) and the linear model second ($p = 0.2512$). The smooth model used 1.867 effective degrees of freedom and was not significant ($p = 0.4040$). The later peak in the full cortisone model also carries into the female-specific model, with the highest cortisone predicted at 90 minutes (Figure 16b).

Table 11: Statistics for models describing female-specific glucocorticoid response to time since onset of simulated MFAS exposure. Columns show the number of samples used to create each model, the degrees of freedom, adjusted r-squared, the significance (p) of time's contribution to the model, AIC score, and rankings generated with ANOVA comparisons and significance compared to the next-best model (Rank (p)).

Hormone	n	Model	df	r-sq. (adj)	p	AIC	Rank (p)
Cortisol	11	smooth	3.00	0.269	0.0104	-18.627	---
		linear	3.00	0.269	0.0104	-18.627	1 (0.0039)
		null	2.00	0	---	-13.247	2
Cortisone	11	smooth	4.35	0.0633	0.4040	-13.644	1 (0.2405)
		linear	3.00	-0.039	0.3263	-13.729	2 (0.3512)
		null	2.00	0	---	-14.860	3

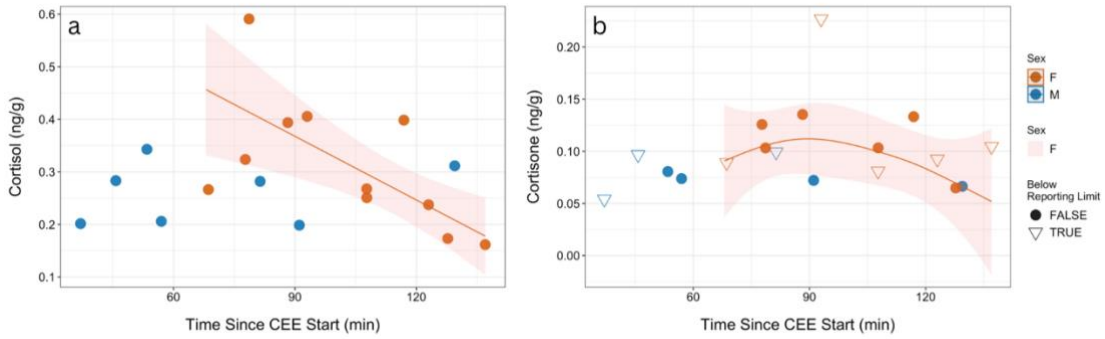

Figure 16: Glucocorticoid concentrations in pilot whale blubber following the onset of simulated MFAS. Models were created for females only and are represented by orange lines with 95% confidence intervals shaded in light orange. The colors of each point represent the sex of individuals sampled, with females in orange and males in blue. For cortisone (b), samples with concentrations below the reporting limit are plotted at mass-adjusted reporting limits as open triangles. True concentrations for these samples are unknown but fall between zero and the reporting limit, which is accounted for in the analysis methods.

4.4 Discussion

This experiment aimed to test whether an endocrine response could be detected in pilot whales exposed to simulated MFAS. As blubber biopsy has not been previously employed in CEEs for hormone monitoring, this study provides a test of this application. Post-exposure glucocorticoid variation is complicated by the time lag between stressor onset and hormone integration in the blubber layer. The timing of this phenomenon is poorly understood, as it likely varies with conditions that affect perfusion (e.g., water temperature, vascularization, and diving behavior)(Champagne et al., 2018). Similarly, we do not understand the persistence of steroid hormones in blubber tissue. By documenting temporal changes in post-exposure blubber glucocorticoids, a secondary

aim for this study was to further elucidate the timing of glucocorticoid fluctuations in blubber.

Ultimately, we were able to show evidence consistent with elevated glucocorticoid concentrations in individuals that were exposed to this stimulus. This response was most apparent in cortisol measurements from females, which decreased during the sampling period, 69-137 minutes post-exposure onset. This trend was echoed in the full dataset of males and females, which was best explained by a smooth model, although it was not significant. This model, which spans 37-137 minutes after exposures began, suggests an increase in cortisol until 83 minutes, followed by a subsequent decrease.

Models for cortisone were not significant but were best fitted with smooth models in the full and female datasets. Both of these models predicted peak cortisone at 90 minutes post-exposure. Agreement in the timing of these responses across cortisol and cortisone analyses is indicative of a consistent pattern of glucocorticoid activity. The short offset seen between these models may disappear with increased sampling; however, it could also be a reflection of metabolic relationships between these analytes. Cortisol, which peaks slightly earlier in our models, is a precursor of cortisone and may metabolize into cortisone within the blubber layer (Galligan, Schwacke, McFee, et al., 2018).

The initial increase and subsequent decrease of glucocorticoids are consistent with the hypothesized response of hormone integration into the blubber layer and subsequent return to baseline. However, the observed decline of glucocorticoids may not be solely determined by time. During post-exposure sampling, we moved farther away from the

sound source with time, increasingly sampling animals that likely received lower doses of exposure. If animals sampled longer post-exposure were also exposed to lower RLs and mounted smaller HPA responses, a post-peak glucocorticoid decline could appear earlier or stronger than time alone would predict.

The female-specific model does not capture the initial cortisol increase described by the larger model. This is likely due to the later sampling period seen in females, which begins very close to the peak seen in the larger model. With additional samples and earlier sampling relative to the peak, a smooth model with more degrees of freedom might provide a better fit than the linear model selected in this study. We see this in the cortisone models, where the 90-minute peak falls deeper into the female-specific sample set, and is reflected in the female-only analysis, which was best fit with a smooth model. However, these cortisone models are sample-limited, not significant at the $p < 0.05$ level, and thus should be interpreted cautiously.

There may also be a sex-dependent response, which has been observed in other systems (Keen et al., 2021). Effects like the disruption of acoustic communication through noise masking could be especially stressful for mothers with dependent calves. Pilot whales with dependent calves would not likely be pregnant (Marsh & Kasuya, 1984), and the three highest cortisol concentrations are observed in females that do not have elevated progesterone or androstenedione. Following exposure, males were generally sampled earlier than females and showed no apparent trend in cortisol or cortisone during the post-exposure period. Unfortunately, the sample size in this study is too small to disentangle the effects of sex, reproductive status, and time.

Seasonality can impact baseline glucocorticoid concentrations and responsiveness to stressors (Romero, 2002). On average, cortisol concentrations in this experiment (mean = 0.31 ng/g) were lower than baseline samples that were primarily collected during summer months (mean = 0.45 ng/g). Baseline seasonal analysis suggested weak seasonal variation in glucocorticoids, but was limited in winter coverage and not significant (Chapter 3). Future efforts to expand year-round sampling would help establish seasonal baselines for comparisons in experiments like this and for long-term monitoring of population health.

Physiological responses to stressors are notoriously challenging to monitor in free-ranging cetaceans, and this is the first study to report experimental evidence supporting an adrenal response to noise in cetaceans. Although glucocorticoids are widely studied biomarkers of stress, their connection to health and reproduction is limited in cetaceans. Future efforts to connect hormone responses with known links in the PCoD model could broaden our ability to monitor disturbance and predict population impacts.

This study is one of few to track changes in cetacean blubber cortisol following exposure to a stressor, and the first to measure the endocrine response in short-finned pilot whales. While the models created are limited in certainty and resolution, they offer insight into the occurrence and timing of this biological response. Researchers designing stress response studies with blubber collection may want to target sampling in the 30-90 minute range, with the caveat that characteristics like smaller body size, higher vascularization, and warmer temperatures may shift this period earlier.

In our study, the temporal variable is correlated with received levels. Limited sample size and uncertainty around non-focal group locations constrained our ability to look at time and exposure separately. Future studies may be able to increase replicates and leverage tagging or aerial tracking for more precise RLs. These efforts would help to establish dose-response curves, which are critical for managing marine mammal populations in areas with sonar activity.

5. Conclusions

Remote blubber biopsy is not a new sampling technique, but advancements in our ability to measure and understand hormones in blubber have propelled its application in cetacean endocrine studies. As recent efforts have narrowed down the timing of steroid hormone perfusion into blubber (Champagne et al., 2018; Galligan et al., 2020), they have set the stage for monitoring acute endocrine shifts in free-ranging cetaceans. The ability to monitor acute hormone responses to stressors, like noise, represents a leap forward in understanding the physiological processes that drive cetacean behavior and the health consequences of sub-lethal disturbance.

In this dissertation, I set out to characterize steroid hormone profiles of short-finned pilot whales, both under normal conditions and in response to an experimental stressor. This required that I could (1) obtain samples, (2) reliably measure hormones in these samples, and (3) contextualize these measurements. We decided to use remote blubber biopsy for tissue sampling as it provided the most effective tradeoff between reliability and invasiveness. Over the past two decades, our field team has collected over 100 biopsies from short-finned pilot whales, demonstrating the feasibility of employing this method in our study system. However, the methods used to measure hormones in these samples required development.

In the absence of an established method for pilot whales, we developed and validated a new method for multi-class steroid measurements in this tissue. Adapted from the Boggs et al. (2017) LC-MS/MS method for measuring stress and sex steroid hormones in bottlenose dolphin (*Tursiops truncatus*) blubber, we created a new method for 11 analytes of interest, including aldosterone and DHEA, which had not been

previously measured in cetacean blubber. To assess the reliability of hormone measurement in blubber from this species, we used replicate samples from stranded animals to conduct a spike-recovery experiment. The precision and accuracy measurements from this experiment demonstrated analytical validity for steroid detection in this tissue and guided our choices of matched internal standards for each analyte.

The analyses conducted in Chapter 2 also helped demonstrate the biological validity of measurements in this tissue. Using this method, we obtained patterns expected for endogenous testosterone and progesterone measurements, observing differences between males and pregnant females. We also observed elevated glucocorticoids in stranded individuals compared to subsequent analyses of blubber from live, presumed healthy free-ranging individuals. While this may be indicative of stress associated with stranding death (Kellar et al., 2015), we did not compare hormone concentrations between deep and superficial blubber tissue and could not rule out sampling depth as a contributor to the differences we observed (Loseto et al., 2017; Trana et al., 2015).

While the LC-MS/MS method delivered highly specific quantitation of a suite of steroid hormones, it lacked the sensitivity to quantify endogenous steroids that occur at low concentrations in some analytes. For example, none of the samples we analyzed, from stranded or live individuals, exceeded the high reporting limit for aldosterone (mean RL = 5.5 ng/g). This issue was more common in blubber biopsies from free-ranging animals, which generally had lower steroid hormone concentrations than blubber from stranded individuals. Researchers looking to quantify steroids in blubber with this LC-MS/MS method should consider further method optimization for analytes like aldosterone, corticosterone, and DHEA. Depending on the study being conducted,

methods that offer higher sensitivity (e.g., enzyme immunoassay) may be preferable and the highly specific LC-MS/MS findings can help inform assay choice and interpretation of results.

Our goals in Chapter 3 were to (1) establish reference concentrations for steroid hormones in short-finned pilot whales, as deposited in blubber, and (2) characterize steroid hormone variation within the context of sex, reproductive status, and seasonality. We leveraged an existing archive of biopsy samples to measure hormone concentrations in live, free-ranging short-finned pilot whales. Using the LC-MS/MS method developed in Chapter 2, we observed high detection rates for cortisol and cortisone in both males and females. However, detection rates were sex-specific for progesterone, 17α-hydroxyprogesterone, testosterone, androstenedione, and DHEA. Except for progesterone and androstenedione, these analytes were almost solely detected in samples from males. Observed hormone concentrations were similar to LC-MS/MS measurements from bottlenose dolphin biopsies (Boggs et al., 2019) but tended to be lower in short-finned pilot whales. Median cortisol concentrations were similar in these species, but cortisol only exceeded 1.0 ng/g in one pilot whale sample, while more than 25% of bottlenose dolphin samples exceeded that threshold.

Because the LC-MS/MS method quantifies multiple analytes in each sample, we were able to pursue multi-steroid investigations. We characterized hormone profiles across sex and suspected pregnancy groups using cluster analyses, which divided samples into three prominent groups. Primarily driven by progesterone and androgen (androstenedione and testosterone) concentrations, these groups likely represented (a) reproductive males, (b) ovulating or pregnant females, and (c) non-reproductive animals.

The third group, which included males and females, may have included immature animals and adults that were not actively upregulating progesterone or androgens for reproduction (e.g., anestrus, post-reproductive, or sampled outside peak reproductive season). More sensitive methods, the inclusion of additional hormones (e.g., progestins, estrogens, protein hormones), and additional sampling across seasons would add resolution to these classifications.

Our capacity to characterize pregnancy in females was constrained by our inability to ground-truth pregnancy in animals sampled. In general, progesterone concentrations are the highest in pregnant cetaceans (Trego et al., 2013); however, in some cases, progesterone delineation is not so clear between pregnant and non-pregnant individuals (Robeck et al., 2017; Robeck et al., 2021). During pregnancy, progesterone concentrations can vary substantially between species (Trego et al., 2013), and limited data have shown low progesterone in pregnant short-finned pilot whales (Yoshioka et al., 1989). While distinctively higher than progesterone measurements in immature and resting females, serum progesterone in pregnant females was lower than progesterone in ovulating females, likely driven by luteal versus placental progesterone production (Yoshioka et al., 1989).

In our Chapter 2 analysis, progesterone concentrations in the two known-pregnant females (112 and 349 ng/g) were similar to or higher than those reported for other odontocetes (Kellar et al., 2006; Pérez et al., 2011; Trego et al., 2013). In comparison, progesterone concentrations in free-ranging individuals were substantially lower (median = 0.31, max = 50.4 ng/g). For assigning pregnancy based on blubber progesterone, other studies have used thresholds ranging from 4 ng/g to 100 ng/g, depending on the species

(Boggs et al., 2019; Dalle Luche et al., 2020; Kellar et al., 2013; Trego et al., 2013). A 100 ng/g threshold would categorize all 12 of the females as non-pregnant. It is possible that we did not biopsy any pregnant individuals; however, that would indicate that we sampled 2 to 4 luteal individuals, which seems unlikely.

Other studies have shown that androgens, which rise during the second half of gestation in many cetaceans (Dalle Luche et al., 2020; Legacki et al., 2020; Robeck et al., 2017; Zhang et al., 2021), may be useful biomarkers for pregnancy detection. The results of our studies support that phenomenon, but only with androstenedione. Both known-pregnant stranded females had androstenedione concentrations similar to males. Testosterone was not quantifiable in either of the known-pregnant stranded females and was only quantifiable in one female biopsy, which was categorized as possibly pregnant based on progesterone. We were able to quantify DHEA in both stranded females, but not in biopsies from females, which may be a result of high reporting limits. Of the 12 biopsies from females, we categorized four as possibly pregnant (progesterone > 1.6 ng/g), two of which had high androstenedione concentrations and clustered separately from all other males and females. Elevated androstenedione lends support to our pregnancy classification but does not provide a clean cut-off for pregnancy delineation.

Clearly, progesterone activity can vary substantially between species, and we caution researchers from using pregnancy thresholds established for other species. These thresholds are also subject to variation based on the quantitation methods used to establish them. Thresholds established for measurements made with immunoassay may be elevated by cross-reactivity of progesterone metabolites, and LC-MS/MS thresholds may be lower than applicable for immunoassay measurements. Regardless of the chosen

analytical method, future studies should prioritize the measurement of progestogens and androstenedione in blubber biopsies from pilot whales of known reproductive status. Hormone concentrations may vary substantially between live and dead animals, but samples from strandings offer an opportunity to ground-truth hormone measurements with corpora lutea, fetal lengths, and milk presence in mammary glands. In free-ranging animals, pregnancy detection could be possible in populations with high resight rates and by using drones for photogrammetry (Cheney et al., 2022).

In Chapter 3, we also looked at seasonality as a driver of hormone variation. Previous studies have shown diffuse seasonal reproduction in short-finned pilot whales (Kasuya & Marsh, 1984). Although these studies were conducted in a genetically and geographically distinct population from our system (Van Cise et al., 2019), our results showed similar patterns in seasonality. In males, androgens were elevated in summer months, which suggests higher reproductive activity may occur between May and July. If this corresponds with peak conception, as seen in the Western Pacific (Marsh & Kasuya, 1984), we would expect calving to peak 15 months later in September. Limited sampling between November and April in our study constrained the power of these models and our understanding of pilot whale physiology and behavior during those months. Future efforts should be made to conduct surveys and collect biopsies during this time to gain better resolution to these seasonal patterns. Peak mating and calving may represent times of peak vulnerability in this population (Frid & Dill, 2002), and population managers should consider seasonality in strategies for mitigating disturbance.

In Chapter 4, we asked if exposure to simulated MFAS would elicit a stress response in short-finned pilot whales that could be detected in blubber. Blubber cortisol

increases have been observed in captive bottlenose dolphins following hydrocortisone dosage (Champagne et al., 2017), bottlenose dolphins after capture and handling (Champagne et al., 2018; Galligan et al., 2020), and in stranded dolphins (Agusti et al., 2022; Kellar et al., 2015). These studies counter previous ideas that steroids may take days to integrate into the blubber layer, showing increases in cortisol within the first two hours following stressor onset in bottlenose dolphins (Champagne et al., 2018; Galligan et al., 2020). While these studies do not identify a specific, predictable lag time between stressor onset and glucocorticoid integration in cetacean blubber, they do offer a coarse understanding of the time frame of this response. Our results identified a similar time window, suggesting glucocorticoids may peak in pilot whale blubber as early as 80 minutes after stressor onset.

Unfortunately, our study did not have the power to disentangle the effect of time from the received level. Additional experiments should focus on animals with known locations and received levels to offer more resolution to the temporal response in blubber and enable the creation of dose-response curves, which are used for species management (Finneran & Jenkins, 2012). Collecting more samples in the 30-minutes following stressor onset and extending sampling to continue for several hours post-exposure would further elucidate the onset and persistence of glucocorticoids in blubber tissue.

The glucocorticoid changes we observed after simulated MFAS exposure are consistent with the hypothesis that military sonar can elicit a physiological stress response in short-finned pilot whales. To date, no studies have published glucocorticoid measurements from free-ranging cetaceans after sonar exposure. However, one such study is underway off the California coast, testing this response in common dolphins

(*Delphinus delphis*) (Kellar pers. comm.), and is likely to provide an insightful comparison with our results. Previously, the connection between noise and glucocorticoids in cetaceans has only been substantiated in North Atlantic right whales (*Eubalaena glacialis*), when fecal glucocorticoids decreased concurrently with shipping traffic and background noise following the events of September 11th, 2001 (Rolland et al., 2012). The ability to conduct disturbance experiments and detect endocrine responses in free-ranging cetaceans represents a shift in our capacity to understand and predict the impacts of disturbance on individuals, their physiology, and the contributions of these to population health.

Because the relationship between endocrinology and behavior is reciprocal, factors like season, sex, and reproductive status can alter glucocorticoid measurements. In Chapter 3, we observed correlations between cortisone and testosterone in males and progesterone in females. Additionally, seasonal increases in male androgen concentrations indicate diffusely seasonal reproduction, which is likely accompanied by behavioral shifts. Seasonal variation was not a significant driver of baseline glucocorticoid concentrations on its own, but models with better seasonal coverage may provide better resolution to this relationship and its interaction with other factors, like sex and reproductive status. In Chapter 4, the observed stress response was most apparent in cortisol measurements from females and concentrations were highest in resting females, highlighting the nuance that reproductive information can add to disturbance response studies.

As a whole, this dissertation demonstrates that steroid hormone measurements in blubber are relevant indicators of physiological states in short-finned pilot whales.

Throughout these studies, sex, reproductive status, and seasonality influence hormone profiles and add context to our findings. Enabled by a multi-steroid approach, we characterized hormone variation as a reference for comparative studies or future monitoring of this population. As humans continue to use the ocean for shipping, miliary sonar, and offshore energy, there is a critical need to understand the consequences of sound disturbance to cetaceans. Endocrine monitoring of cetaceans is a rapidly maturing field and, as demonstrated in this dissertation, can be effectively applied in stress response studies of free-ranging cetaceans.

Appendix A

Supplementary Table A1. List of neat standards and manufacturers/sources.

Reference Standard	Abbreviation	Company
Androstenedione	AE	Steraloids
Androstenedione-$^{13}C_3$	AE-$^{13}C_3$	Cerilliant
Testosterone	T	Sigma-Aldrich
Testosterone-$^{13}C_3$	T-$^{13}C_3$	Cerilliant
Progesterone	P_4	Sigma-Aldrich
Progesterone-$^{13}C_3$	P_4-$^{13}C_3$	Cambridge Isotopes
17α-hydroxyprogesterone	$17OHP_4$	Sigma-Aldrich
17α-hydroxyprogesterone -$^{13}C_3$	$17OHP_4$-$^{13}C_3$	Cerilliant
Dehydroepiandrosterone	DHEA	Sigma
Dehydroepiandrosterone sulfate	DHEAS	Sigma
Aldosterone	ALD	Acros Organics
Aldosterone-2H4	ALD-d_4	IsoSciences
Cortisone	E	Sigma-Aldrich
Cortisone-$^{13}C_3$	E-$^{13}C_3$	Sigma-Aldrich
Cortisol	F	Sigma-Aldrich
Cortisol-$^{13}C_3$	F-$^{13}C_3$	IsoSciences
Corticosterone	B	Sigma-Aldrich
Corticosterone-$^{2}H_4$	B- d_4	IsoSciences
11-Deoxycortisol	S	Steraloids
11-Deoxycortisol-$^{13}C_3$	S-$^{13}C_3$	IsoSciences
11-Deoxycorticosterone	11DOC	Steraloids
11-Deoxycorticosterone-$^{13}C_3$	11DOC-$^{13}C_3$	IsoSciences

Supplementary Table A2. Analyte masses used in the spike-recovery experiment.

Group	Analyte	Mass in High Calibrant (ng)	Mass in Low Calibrant (ng)	Av. Mass in Spike (ng)	Matched IS	Average IS Mass per Sample (ng)
Adrenal (C18)	ALD	127.9	0.01114	11.70	Ald-d_4	1.977
	E	124.6	0.01085	11.40	E-$^{13}C_3$	1.230
	F	263.1	0.02290	24.07	F-$^{13}C_3$	1.206
	B	122.0	0.01062	11.16	B-d_4	1.621
	S	128.3	0.01117	11.74	S-$^{13}C_3$	0.5971
	11DOC	150.7	0.01312	13.79	11DOC-$^{13}C_3$	1.148
	DHEA	130.5	0.01136	11.94	S-$^{13}C_3$	0.5971
Gonadal (Biphenyl)	P_4	332.8	0.02897	30.45	P_4-$^{13}C_3$	9.661
	$17OHP_4$	494.9	0.04309	45.28	$17OHP_4$-$^{13}C_3$	0.3177
	T	705.4	0.06141	64.53	T-$^{13}C_3$	2.943
	AE	518.4	0.04514	47.43	AE-$^{13}C_3$	2.247

Supplementary Table A3. Calibration curves used to determine Biphenyl analyte concentrations with each IS reference.

Biphenyl Analyte	Curve Level	IS	R^2	Slope	Intercept	# Cals	High Cal (ng/g)	Low Cal (ng/g)	LOD (ng/g)	RL (ng/g)
17OHP$_4$	low	17OHP$_4$-$^{13}C_3$	0.9999	0.0646	0.0195	6	53.49	0.1056	0.2088	0.2088
		T-$^{13}C_3$	0.9999	0.4976	0.0179	6	53.49	0.1056	0.1112	0.1112
		AE-$^{13}C_3$	0.9999	0.7898	0.0672	4	53.49	0.1056	0.05006	0.1056
		P$_4$-$^{13}C_3$	0.9999	0.1907	0.0027	5	53.49	0.1056	0.05096	0.1056
	high	17OHP$_4$-$^{13}C_3$	0.9990	0.0599	0.7871	3	608.1	53.49		
		T-$^{13}C_3$	0.9997	0.5613	-0.1729	4	309.5	3.527		
		AE-$^{13}C_3$	1.0000	0.0639	0.0313	4	309.5	3.527		
		P$_4$-$^{13}C_3$	1.0000	0.1976	-0.0016	4	309.5	3.527		
T	low	17OHP$_4$-$^{13}C_3$	0.9999	0.1492	-0.0034	6	76.24	0.1505	0.1092	0.1505
		T-$^{13}C_3$	1.0000	1.1497	0.0035	6	76.24	0.1505	0.06447	0.1505
		AE-$^{13}C_3$	0.9998	2.1209	0.0146	4	5.027	0.1505	0.05388	0.1505
		P$_4$-$^{13}C_3$	1.0000	0.4616	0.0010	5	19.52	0.1505	0.03936	0.1505
	high	17OHP$_4$-$^{13}C_3$	0.9999	0.1493	-0.0098	4	76.24	1.715		
		T-$^{13}C_3$	1.0000	1.1493	0.0069	4	76.24	1.715		
		AE-$^{13}C_3$	0.9996	1.8151	0.3279	4	76.24	1.715		
		P$_4$-$^{13}C_3$	0.9999	0.4398	0.0085	4	76.24	1.715		
AE	low	17OHP$_4$-$^{13}C_3$	1.0000	0.0169	0.0011	5	324.2	0.4309	0.1284	1.260
		T-$^{13}C_3$	0.9998	0.1381	-0.0027	5	56.04	0.4309	0.145	1.260
		AE-$^{13}C_3$	0.9999	0.2268	0.0074	3	14.34	1.260	0.02202	1.260
		P$_4$-$^{13}C_3$	1.0000	0.0530	0.0000	4	56.03	1.260	0.07775	1.260
	high	17OHP$_4$-$^{13}C_3$	0.9993	0.0178	-0.0549	5	637.0	3.695		
		T-$^{13}C_3$	0.9999	0.1483	-0.0283	5	324.2	1.260		
		AE-$^{13}C_3$	1.0000	0.2218	0.0049	4	324.2	3.695		
		P$_4$-$^{13}C_3$	1.0000	0.0523	0.0006	4	324.2	3.695		
P$_4$	low	17OHP$_4$-$^{13}C_3$	1.0000	0.1740	0.0039	4	35.97	0.07101	0.05674	0.07101
		T-$^{13}C_3$	1.0000	1.3282	-0.0027	4	9.208	0.07101	0.05385	0.07101
		AE-$^{13}C_3$	1.0000	2.2952	0.0160	4	9.208	0.07101	0.03381	0.07101
		P$_4$-$^{13}C_3$	0.9999	0.5318	-0.0001	5	9.208	0.07101	0.04205	0.07101
	high	17OHP$_4$-$^{13}C_3$	0.9998	0.1590	0.3676	3	208.1	9.208		
		T-$^{13}C_3$	1.0000	1.4811	-0.3063	4	408.9	2.372		
		AE-$^{13}C_3$	1.0000	0.4725	0.0293	4	408.9	2.372		
		P$_4$-$^{13}C_3$	0.9995	0.4732	0.0739	4	408.9	9.208		

Supplementary Table A4. Calibration curves used to determine C18 analyte concentrations with each IS reference.

C18 Analyte	IS	R^2	Slope	Intercept	# Cals	High Cal (ng/g)	Low Cal (ng/g)	LOD (ng/g)	RL (ng/g)
ALD	ALD-d_4	0.9994	1.1725	-0.0782	3	40.43	3.539	0.7282	3.539
	E-$^{13}C_3$	0.9337	0.1715	0.2976	4	80.00	3.539	-1.215	3.539
	F-$^{13}C_3$	0.9997	0.2846	0.2251	3	80.00	3.539	-0.1839	3.539
	B-d_4	0.9954	7.4969	-1.9959	3	40.43	3.539	0.8031	3.539
	S-$^{13}C_3$	0.9950	0.1848	-0.1933	3	40.43	3.539	1.044	3.539
	11DOC-$^{13}C_3$	0.9996	0.9914	0.6806	3	80.00	3.539	-0.2787	3.539
E	ALD-d_4	0.9998	0.9696	0.0056	7	77.91	0.1035	0.06630	0.1035
	E-$^{13}C_3$	1.0000	1.6618	0.0664	9	153.1	0.02658	0.01889	0.1035
	F-$^{13}C_3$	0.9998	7.1236	-0.2867	5	77.91	0.1035	0.1047	0.1047
	B-d_4	1.0000	29.4480	0.1203	6	153.1	0.1035	0.02904	0.1035
	S-$^{13}C_3$	1.0000	0.6868	0.0172	6	77.91	0.1035	0.07529	0.1035
	11DOC-$^{13}C_3$	1.0000	6.1497	0.0943	4	39.38	0.1035	0.07229	0.1035
F	ALD-d_4	0.9999	0.9108	0.0255	7	83.16	0.05613	-0.01301	0.05613
	E-$^{13}C_3$	1.0000	0.5455	0.0052	5	83.16	0.05613	0.01019	0.05613
	F-$^{13}C_3$	1.0000	3.7853	0.0588	5	164.5	0.05613	-0.01316	0.05613
	B-d_4	0.9999	15.9360	0.2493	7	83.16	0.05613	-0.01018	0.05613
	S-$^{13}C_3$	1.0000	0.3895	0.0085	5	83.16	0.05613	0.004823	0.05613
	11DOC-$^{13}C_3$	1.0000	3.3828	0.067	4	83.16	0.05613	-0.01112	0.05613
B	ALD-d_4	1.0000	3.7159	0.3096	3	38.55	0.8693	-0.08554	0.8693
	E-$^{13}C_3$	0.9997	0.0905	0.0148	5	38.55	0.2965	-0.05182	0.8693
	F-$^{13}C_3$	0.9999	0.2058	0.0531	6	149.9	0.8693	-0.7170	0.8693
	B-d_4	0.9999	0.7883	0.0734	4	38.55	0.2965	-0.05844	0.8693
	S-$^{13}C_3$	1.0000	0.1185	0.0143	4	76.28	0.8693	-0.1024	0.8693
	11DOC-$^{13}C_3$	0.9998	0.8768	0.0053	3	76.28	0.8693	0.02783	0.8693
S	ALD-d_4	1.0000	0.7621	0.0109	4	40.56	0.02738	-0.003203	0.02738
	E-$^{13}C_3$	1.0000	31.2410	0.4487	5	40.56	0.02738	-0.01729	0.02738
	F-$^{13}C_3$	1.0000	6.6189	0.0655	4	40.56	0.02738	-0.006203	0.02738
	B-d_4	1.0000	1.6906	0.0732	5	80.26	0.02738	-0.04220	0.02738
	S-$^{13}C_3$	0.9999	0.9812	0.0276	5	80.26	0.02738	-0.02478	0.02738
	11DOC-$^{13}C_3$	1.0000	7.2227	0.1232	4	80.26	0.02738	-0.02775	0.02738
11DOC	ALD-d_4	1.0000	1.6794	0.0226	3	47.65	0.3664	0.04527	0.3664
	E-$^{13}C_3$	0.9998	0.1899	0.0053	5	94.27	0.3664	0.04600	0.3664
	F-$^{13}C_3$	0.9999	7.9127	0.1956	5	47.65	0.3664	0.01584	0.3664
	B-d_4	1.0000	0.4575	0.0124	6	370.5	0.3664	0.01632	0.3664
	S-$^{13}C_3$	0.9999	0.2668	0.0023	5	94.27	0.3664	0.04347	0.3664
	11DOC-$^{13}C_3$	1.0000	1.2503	0.0628	3	47.65	0.3664	-0.04338	0.3664
DHEA	ALD-d_4	1.0000	2.7035	0.5443	6	81.60	0.02784	0.1082	0.1082

	E-$^{13}C_3$	1.0000	0.3218	0.1334	6	81.60	0.02784	0.1423	0.1423
	F-$^{13}C_3$	0.9997	17.3940	1.5849	6	41.24	0.02784	0.06617	0.06617
	B-d_4	1.0000	0.7788	0.1865	6	81.60	0.02784	0.1351	0.1351
	S-$^{13}C_3$	1.0000	0.4518	0.0940	4	81.60	0.02784	0.2530	0.2530
	11DOC-$^{13}C_3$	1.0000	3.4153	0.3328	4	160.3	0.02784	0.16700	0.16700

References

Aasen, E., & Medrano, J. F. (1990). Amplification of the ZFY and ZFX genes for sex identification in humans, cattle, sheep and goats. *Biotechnology (N Y)*, *8*(12), 1279-1281. https://doi.org/10.1038/nbt1290-1279

Abraham, G. E. (1974). Ovarian and Adrenal Contribution to Peripheral Androgens During the Menstrual Cycle. *The Journal of Clinical Endocrinology & Metabolism*, *39*(2), 340-346. https://doi.org/10.1210/jcem-39-2-340

Acevedo-Whitehouse, K., Rocha-Gosselin, A., & Gendron, D. (2010). A novel non-invasive tool for disease surveillance of free-ranging whales and its relevance to conservation programs. *Animal Conservation*, *13*(2), 217-225. https://doi.org/10.1111/j.1469-1795.2009.00326.x

Adkins-Regan, E. (2005). *Hormones and animal social behavior*. Princeton University Press. https://search.library.duke.edu/search?id=DUKE003482664

Aguilar Soto, N., Johnson, M., Madsen, P. T., Tyack, P. L., Bocconcelli, A., & Fabrizio Borsani, J. (2006). Does Intense Ship Noise Disrupt Foraging in Deep-Diving Cuvier's Beaked Whales (Ziphius Cavirostris)? *Marine Mammal Science*, *22*(3), 690-699. https://doi.org/10.1111/j.1748-7692.2006.00044.x

Aguilar Soto, N., Johnson, M. P., Madsen, P. T., Diaz, F., Dominguez, I., Brito, A., & Tyack, P. (2008). Cheetahs of the deep sea: deep foraging sprints in short-finned pilot whales off Tenerife (Canary Islands). *J Anim Ecol*, *77*(5), 936-947. https://doi.org/10.1111/j.1365-2656.2008.01393.x

Agusti, C., Carbajal, A., Olvera-Maneu, S., Domingo, M., & Lopez-Bejar, M. (2022). Blubber and serum cortisol concentrations as indicators of the stress response and overall health status in striped dolphins. *Comp Biochem Physiol A Mol Integr Physiol*, *272*, 111268. https://doi.org/10.1016/j.cbpa.2022.111268

Amaral, R. S. (2010). Use of Alternative Matrices to Monitor Steroid Hormones in Aquatic Mammals: A Review. *Aquatic Mammals*, *36*(2), 162-171. https://doi.org/10.1578/am.36.2.2010.162

Antunes, R., Kvadsheim, P. H., Lam, F. P., Tyack, P. L., Thomas, L., Wensveen, P. J., & Miller, P. J. (2014). High thresholds for avoidance of sonar by free-ranging long-finned pilot whales (Globicephala melas). *Mar Pollut Bull*, *83*(1), 165-180. https://doi.org/10.1016/j.marpolbul.2014.03.056

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*(5), 463-486. https://doi.org/10.1007/s00360-015-0901-0

Atkinson, S., Gendron, D., Branch, T. A., Mashburn, K. L., Melica, V., Enriquez-Paredes, L. E., & Brownell, R. L. (2019). Pregnancy rate and biomarker validations from the blubber of eastern North Pacific blue whales. *Marine Mammal Science*, *36*(1), 6-28. https://doi.org/10.1111/mms.12616

Bechshoft, T. O., Riget, F. F., Sonne, C., Letcher, R. J., Muir, D. C., Novak, M. A., Henchey, E., Meyer, J. S., Eulaers, I., Jaspers, V. L., Eens, M., Covaci, A., & Dietz, R. (2012). Measuring environmental stress in East Greenland polar bears, 1892-1927 and 1988-2009: what does hair cortisol tell us? *Environ Int*, *45*, 15-21. https://doi.org/10.1016/j.envint.2012.04.005

Blas, J., Bortolotti, G. R., Tella, J. L., Baos, R., & Marchant, T. A. (2007). Stress response during development predicts fitness in a wild, long lived vertebrate. *Proc Natl Acad Sci U S A*, *104*(21), 8880-8884. https://doi.org/10.1073/pnas.0700232104

Boggs, A. S. P., Ragland, J. M., Zolman, E. S., Schock, T. B., Morey, J. S., Galligan, T. M., Dalle Luche, G., Balmer, B. C., Wells, R. S., Kucklick, J. R., & Schwacke, L. H. (2019). Remote blubber sampling paired with liquid chromatography tandem mass spectrometry for steroidal endocrinology in free-ranging bottlenose dolphins (Tursiops truncatus). *Gen Comp Endocrinol*, *281*, 164-172. https://doi.org/10.1016/j.ygcen.2019.06.006

Boggs, A. S. P., Schock, T. B., Schwacke, L. H., Galligan, T. M., Morey, J. S., McFee, W. E., & Kucklick, J. R. (2017). Rapid and reliable steroid hormone profiling in Tursiops truncatus blubber using liquid chromatography tandem mass spectrometry (LC-MS/MS). *Anal Bioanal Chem*, *409*(21), 5019-5029. https://doi.org/10.1007/s00216-017-0446-z

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H., & White, J. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends Ecol Evol*, *24*(3), 127-135. https://doi.org/10.1016/j.tree.2008.10.008

Boonstra, R., Hik, D., Singleton, G. R., & Tinnikov, A. (1998). The Impact of Predator-Induced Stress on the Snowshoe Hare Cycle. *Ecological Monographs*, *68*(3), 371-394. https://doi.org/10.2307/2657244

Bortolotti, G. R., Marchant, T., Blas, J., & Cabezas, S. (2009). Tracking stress: localisation, deposition and stability of corticosterone in feathers. *J Exp Biol*, *212*(Pt 10), 1477-1482. https://doi.org/10.1242/jeb.022152

Bowers, M. T. (2016). <Bowers_2016_dissertation.pdf>.

Bowers, M. T., Friedlaender, A. S., Janik, V. M., Nowacek, D. P., Quick, N. J., Southall, B. L., & Read, A. J. (2018). Selective reactions to different killer whale call categories in two delphinid species. *The Journal of Experimental Biology*, *221*(11). https://doi.org/10.1242/jeb.162479

Bryant, P. J., Lafferty, C. M., & Lafferty, S. K. (1984). Reoccupation of Laguna Guerrero Negro, Baja California, Mexico, by gray whales. *The gray whale, Eschrichtius robustus*, 375-387.

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*(1), 10031. https://doi.org/10.1038/s41598-018-28200-8

Carbajal, A., Agustí, C., Domingo, M., Olvera-Maneu, S., & Lopez-Bejar, M. (2021). Influence of sample location on blubber cortisol concentration in striped dolphins (Stenella coeruleoalba): The importance of the reference denominator. *Marine Mammal Science*, *38*(2), 756-764. https://doi.org/10.1111/mms.12862

Carlstead, K., Brown, J. L., Monfort, S. L., Killens, R., & Wildt, D. E. (1992). Urinary monitoring of adrenal responses to psychological stressors in domestic and nondomestic felids. *Zoo Biology*, *11*(3), 165-176.

Champagne, C. D., Kellar, N. M., Crocker, D. E., Wasser, S. K., Booth, R. K., Trego, M. L., & Houser, D. S. (2017). Blubber cortisol qualitatively reflects circulating cortisol concentrations in bottlenose dolphins. *Marine Mammal Science*, *33*(1), 134-153. https://doi.org/10.1111/mms.12352

Champagne, C. D., Kellar, N. M., Trego, M. L., Brendan, D., Rudy, B., Wasser, S. K., Booth, R. K., Crocker, D. E., & Houser, D. S. (2018). Comprehensive Endocrine Response to Acute Stress in the Bottlenose Dolphin from Serum, Blubber, and Feces. *Gen Comp Endocrinol*. https://doi.org/10.1016/j.ygcen.2018.05.015

Cheney, B. J., Dale, J., Thompson, P. M., Quick, N. J., Scales, K., & Bouchet, P. (2022). Spy in the sky: a method to identify pregnant small cetaceans. *Remote Sensing in Ecology and Conservation*, *8*(4), 492-505. https://doi.org/10.1002/rse2.258

Clark, C. T., Fleming, A. H., Calambokidis, J., Kellar, N. M., Allen, C. D., Catelani, K. N., Robbins, M., Beaulieu, N. E., Steel, D., & Harvey, J. T. (2016). Heavy with child? Pregnancy status and stable isotope ratios as determined from biopsies of humpback whales. *Conserv Physiol*, *4*(1), 1-13. https://doi.org/10.1093/conphys/cow050

Council, N. R. (2005). *Marine mammal populations and ocean noise: determining when noise causes biologically significant effects*. National Academies Press.

Cox, T., Ragen, T., Read, A., Vos, E., & others, a. (2006). Understanding the impacts of anthropogenic sound on beaked whales. *J Cetacean Res Manag*, *7*, 177−187.

Crain, D. D., Thomas, A., Mansouri, F., Potter, C. W., Usenko, S., & Trumble, S. J. (2020). Hormone comparison between right and left baleen whale earplugs. *Conserv Physiol*, *8*(1), coaa055. https://doi.org/10.1093/conphys/coaa055

Cure, C., Antunes, R., Samarra, F., Alves, A. C., Visser, F., Kvadsheim, P. H., & Miller, P. J. (2012). Pilot whales attracted to killer whale sounds: acoustically-mediated interspecific interactions in cetaceans. *PLoS One*, *7*(12), e52201. https://doi.org/10.1371/journal.pone.0052201

Dahlheim, M., & Castellote, M. (2016). Changes in the acoustic behavior of gray whales Eschrichtius robustus in response to noise. *Endangered Species Research*, *31*, 227-242. https://www.int-res.com/abstracts/esr/v31/p227-242/

Dalle Luche, G., Bengtson Nash, S., Kucklick, J. R., Mingramm, F. M. J., & Boggs, A. S. P. (2019). Liquid chromatography tandem mass spectrometry for the quantification of steroid hormone profiles in blubber from stranded humpback whales (Megaptera

novaeangliae). *Conserv Physiol*, *7*(1), coz030. https://doi.org/10.1093/conphys/coz030

Dalle Luche, G., Boggs, A. S. P., Kucklick, J. R., Gross, J., Hawker, D. W., & Bengtson Nash, S. (2020). Androstenedione and testosterone but not progesterone are potential biomarkers of pregnancy in Humpback Whales (Megaptera novaeangliae) approaching parturition. *Sci Rep*, *10*(1), 2954. https://doi.org/10.1038/s41598-020-58933-4

Dalle Luche, G., Boggs, A. S. P., Kucklick, J. R., Hawker, D. W., Wisse, J. H., & Bengtson Nash, S. (2021). Steroid hormone profiles and body conditions of migrating male humpback whales (Megaptera novaeangliae). *Gen Comp Endocrinol*, *313*, 113888. https://doi.org/10.1016/j.ygcen.2021.113888

de Mello, D. M. D., & de Oliveira, C. A. (2016). Biological matrices for sampling free-ranging cetaceans and the implications of their use for reproductive endocrine monitoring. *Mammal Review*, *46*(2), 77-91. https://doi.org/10.1111/mam.12055

DeRuiter, S. L., Southall, B. L., Calambokidis, J., Zimmer, W. M., Sadykova, D., Falcone, E. A., Friedlaender, A. S., Joseph, J. E., Moretti, D., Schorr, G. S., Thomas, L., & Tyack, P. L. (2013). First direct measurements of behavioural responses by Cuvier's beaked whales to mid-frequency active sonar. *Biol Lett*, *9*(4), 20130223. https://doi.org/10.1098/rsbl.2013.0223

Dines, J. P., Mesnick, S. L., Ralls, K., May-Collado, L., Agnarsson, I., & Dean, M. D. (2015). A trade-off between precopulatory and postcopulatory trait investment in male cetaceans. *Evolution*, *69*(6), 1560-1572. https://doi.org/10.1111/evo.12676

Drea, C. M. (2009). Endocrine Mediators of Masculinization in Female Mammals. *Current Directions in Psychological Science*, *18*(4), 221-226. https://doi.org/10.1111/j.1467-8721.2009.01640.x

Dunlop, R. A., Noad, M. J., McCauley, R. D., Kniest, E., Paton, D., & Cato, D. H. (2015). The Behavioural Response of Humpback Whales (Megaptera novaeangliae) to a 20 Cubic Inch Air Gun. *Aquatic Mammals*, *41*(4), 412-433. https://doi.org/10.1578/am.41.4.2015.412

Ellison, W. T., Southall, B. L., Clark, C. W., & Frankel, A. S. (2012). A new context-based approach to assess marine mammal behavioral responses to anthropogenic sounds. *Conserv Biol*, *26*(1), 21-28. https://doi.org/10.1111/j.1523-1739.2011.01803.x

Erbe, C., Reichmuth, C., Cunningham, K., Lucke, K., & Dooling, R. (2016). Communication masking in marine mammals: A review and research strategy. *Marine Pollution Bulletin*, *103*(1), 15-38. https://doi.org/https://doi.org/10.1016/j.marpolbul.2015.12.007

Evans, D., & England, G. (2001). Joint interim report : Bahamas marine mammal stranding event of [15]-16 March 2000 [Miscellaneous]. https://repository.library.noaa.gov/view/noaa/16198

Fain, S. a. J. L. (1995). Gender identification of humans and mammalian wild. *Proc Am Acad Forensic Sci*, 1:34.

Fair, P. A., Schaefer, A. M., Romano, T. A., Bossart, G. D., Lamb, S. V., & Reif, J. S. (2014). Stress response of wild bottlenose dolphins (Tursiops truncatus) during capture-release health assessment studies. *Gen Comp Endocrinol*, *206*, 203-212. https://doi.org/10.1016/j.ygcen.2014.07.002

Fang, Z. (2017). cenGAM: Censored Regression with Smooth Terms. %7Bhttps://CRAN.R-project.org/package=cenGAM%7D,

Finneran, J., & Jenkins, A. (2012). *Criteria and thresholds for US Navy acoustic and explosive effects analysis*.

Foley, H. J. (2018). *Spatial Ecology and Movement Patterns of Deep-Diving Odontocetes in the Western North Atlantic Ocean* North Carolina State University]. ProQuest Dissertations Publishing.

Frid, A., & Dill, L. M. (2002). Human-caused Disturbance Stimuli as a Form of Predation Risk. *Conservation Ecology*, *6*(1), 11.

Funasaka, N. (2018). Long-term monitoring of circulating progesterone and its relationship to peripheral white blood cells in female false killer whales.

Galligan, T. M., Balmer, B. C., Schwacke, L. H., Bolton, J. L., Quigley, B. M., Rosel, P. E., Ylitalo, G. M., & Boggs, A. S. P. (2019). Examining the relationships between blubber steroid hormones and persistent organic pollutants in common bottlenose dolphins. *Environ Pollut*, *249*, 982-991. https://doi.org/10.1016/j.envpol.2019.03.083

Galligan, T. M., Boggs, A. S. P., Balmer, B. C., Rowles, T., Smith, C. R., Townsend, F., Wells, R. S., Kellar, N. M., Zolman, E. S., & Schwacke, L. H. (2020). Blubber steroid hormone profiles as indicators of physiological state in free-ranging common bottlenose dolphins (Tursiops truncatus). *Comp Biochem Physiol A Mol Integr Physiol*, *239*, 110583. https://doi.org/10.1016/j.cbpa.2019.110583

Galligan, T. M., Schwacke, L. H., Houser, D. S., Wells, R. S., Rowles, T., & Boggs, A. S. P. (2018). Characterization of Circulating Steroid Hormone Profiles in the Bottlenose Dolphin (Tursiops truncatus) by Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS). *Gen Comp Endocrinol*. https://doi.org/10.1016/j.ygcen.2018.04.003

Galligan, T. M., Schwacke, L. H., McFee, W. E., & Boggs, A. S. P. (2018). Evidence for cortisol-cortisone metabolism by marine mammal blubber. *Mar Biol*, *165*. https://doi.org/10.1007/s00227-018-3373-4

Ghazal, K., Brabant, S., Prie, D., & Piketty, M. L. (2022). Hormone Immunoassay Interference: A 2021 Update. *Ann Lab Med*, *42*(1), 3-23. https://doi.org/10.3343/alm.2022.42.1.3

Goertz, C. E. C., Burek-Huntington, K., Royer, K., Quakenbush, L., Clauss, T., Hobbs, R., & Kellar, N. M. (2019). Comparing progesterone in blubber and serum to assess

pregnancy in wild beluga whales (Delphinapterus leucas). *Conserv Physiol*, *7*(1), coz071. https://doi.org/10.1093/conphys/coz071

Goldbogen, J. A., Southall, B. L., DeRuiter, S. L., Calambokidis, J., Friedlaender, A. S., Hazen, E. L., Falcone, E. A., Schorr, G. S., Douglas, A., Moretti, D. J., Kyburg, C., McKenna, M. F., & Tyack, P. L. (2013). Blue whales respond to simulated mid-frequency military sonar. *Proc Biol Sci*, *280*(1765), 20130657. https://doi.org/10.1098/rspb.2013.0657

Gotz, T., & Janik, V. M. (2011). Repeated elicitation of the acoustic startle reflex leads to sensitisation in subsequent avoidance behaviour and induces fear conditioning. *BMC Neurosci*, *12*, 30. https://doi.org/10.1186/1471-2202-12-30

Graham, K. M., Burgess, E. A., & Rolland, R. M. (2021). Stress and reproductive events detected in North Atlantic right whale blubber using a simplified hormone extraction protocol. *Conserv Physiol*, *9*(1), coaa133. https://doi.org/10.1093/conphys/coaa133

Gundlach, N. H., Schmicke, M., Ludes-Wehrmeister, E., Ulrich, S. A., Araujo, M. G., & Siebert, U. (2018). New Approach to Stress Research in Phocids-Potential of Dehydroepiandrosterone and Cortisol/Dehydroepiandrosterone Ratio as Markers for Stress in Harbor Seals (Phoca Vitulina) and Gray Seals (Halichoerus Grypus). *J Zoo Wildl Med*, *49*(3), 556-563. https://doi.org/10.1638/2017-0191.1

Harris, C. M., Sadykova, D., DeRuiter, S. L., Tyack, P. L., Miller, P. J. O., Kvadsheim, P. H., Lam, F. P. A., & Thomas, L. (2015). Dose response severity functions for acoustic disturbance in cetaceans using recurrent event survival analysis. *Ecosphere*, *6*(11). https://doi.org/10.1890/es15-00242.1

Harris, C. M., Thomas, L., Falcone, E. A., Hildebrand, J., Houser, D., Kvadsheim, P. H., Lam, F. P. A., Miller, P. J. O., Moretti, D. J., Read, A. J., Slabbekoorn, H., Southall, B. L., Tyack, P. L., Wartzok, D., Janik, V. M., & Blanchard, J. (2018). Marine mammals and sonar: Dose-response studies, the risk-disturbance hypothesis and the role of exposure context. *Journal of Applied Ecology*, *55*(1), 396-404. https://doi.org/10.1111/1365-2664.12955

Hatch, L., & Wright, A. J. (2007). A Brief Review of Anthropogenic Sound in the Oceans. *International Journal of Comparative Psychology*, *20*(2). https://doi.org/10.46867/ijcp.2007.20.02.12

Hayden, M., Bhawal, R., Escobedo, J., Harmon, C., O'Hara, T. M., Klein, D., San-Francisco, S., Zabet-Moghaddam, M., & Godard-Codding, C. A. J. (2017). Nanospray liquid chromatography/tandem mass spectrometry analysis of steroids from gray whale blubber [Article]. *Rapid Communications in Mass Spectrometry*, *31*(13), 1088-1094. https://doi.org/10.1002/rcm.7884

Helsel, D. R. (2012). *Statistics for censored environmental data using Minitab and R* (2nd ed.). Wiley.

Hildebrand, J. A. (2009). Anthropogenic and natural sources of ambient noise in the ocean. *Marine Ecology Progress Series*, *395*, 5-20.

Hohn, A. A., Rotstein, D. S., Harms, C. A., & Southall, B. L. (2006). Report on marine mammal unusual mortality event UMESE0501Sp: Multispecies mass stranding of pilot whales (Globicephala macrorhynchus), minke whale (Balaenoptera acutorostrata), and dwarf sperm whales (Kogia sima) in North Carolina on 15-16 January 2005.

Houser, D. S., Martin, S. W., & Finneran, J. J. (2013). Exposure amplitude and repetition affect bottlenose dolphin behavioral responses to simulated mid-frequency sonar signals. *Journal of Experimental Marine Biology and Ecology*, *443*, 123-133. https://doi.org/10.1016/j.jembe.2013.02.043

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., Rowles, T. K., & Kraus, S. D. (2013). Overcoming the challenges of studying conservation physiology in large whales: a review of available methods. *Conservation Physiology*, *1*(1), cot006-cot006. https://doi.org/10.1093/conphys/cot006

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). *General and Comparative Endocrinology*, *148*(2), 260-272. https://doi.org/10.1016/j.ygcen.2006.03.012

Hunt, K. E., Stimmelmayr, R., George, C., Hanns, C., Suydam, R., Brower, H., Jr., & Rolland, R. M. (2014). Baleen hormones: a novel tool for retrospective assessment of stress and reproduction in bowhead whales (Balaena mysticetus). *Conserv Physiol*, *2*(1), cou030. https://doi.org/10.1093/conphys/cou030

Jensen, F. B., Kuperman, W. A., Porter, M. B., Schmidt, H., & Tolstoy, A. (2011). *Computational ocean acoustics* (Vol. 794). Springer.

Johnson, M. P., & Tyack, P. L. (2003). A digital acoustic recording tag for measuring the response of wild marine mammals to sound. *IEEE Journal of Oceanic Engineering*, *28*(1), 3-12. https://doi.org/10.1109/joe.2002.808212

Julian, P., & Helsel, D. (2021). *NADA2: Data Analysis for Censored Environmental Data*. In https://github.com/SwampThingPaul/NADA2

Karashima, S., & Osaka, I. (2022). Rapidity and Precision of Steroid Hormone Measurement. *J Clin Med*, *11*(4). https://doi.org/10.3390/jcm11040956

Kasuya, T., & Marsh, H. (1984). Life history and reproductive biology of the short-finned pilot whale (Globicephala macrorhynchus) off the pacific coast of Japan. *Report of the International Whaling Commission, Special*, *6*, 259-310.

Katsumata, E., Jaroenporn, S., Katsumata, H., Konno, S., Maeda, Y., Watanabe, G., & Taya, K. (2006). Body Temperature and Circulating Progesterone Levels before and after Parturition in Killer Whales (Orcinus orca). *Journal of Reproduction and Development*, *52*(1), 65-71. https://doi.org/10.1262/jrd.17063

Keen, K. A., Beltran, R. S., Pirotta, E., & Costa, D. P. (2021). Emerging themes in Population Consequences of Disturbance models. *Proc Biol Sci*, *288*(1957), 20210325. https://doi.org/10.1098/rspb.2021.0325

Kellar, N. M., Catelani, K. N., Robbins, M. N., Trego, M. L., Allen, C. D., Danil, K., & Chivers, S. J. (2015). Blubber cortisol: a potential tool for assessing stress response in free-ranging dolphins without effects due to sampling. *PLoS One*, *10*(2), e0115257. https://doi.org/10.1371/journal.pone.0115257

Kellar, N. M., Keliher, J., Trego, M. L., Catelani, K. N., Hanns, C., George, J. C. C., & Rosa, C. (2013). Variation of bowhead whale progesterone concentrations across demographic groups and sample matrices. *Endangered Species Research*, *22*(1), 61-72. https://doi.org/10.3354/esr00537

Kellar, N. M., Speakman, T. R., Smith, C. R., Lane, S. M., Balmer, B. C., Trego, M. L., Catelani, K. N., Robbins, M. N., Allen, C. D., Wells, R. S., Zolman, E. S., Rowles, T. K., & Schwacke, L. H. (2017). Low reproductive success rates of common bottlenose dolphins Tursiops truncatus in the northern Gulf of Mexico following the Deepwater Horizon disaster (2010-2015). *Endangered Species Research*, *33*, 143-158. https://doi.org/10.3354/esr00775

Kellar, N. M., Trego, M. L., Chivers, S. J., Archer, F. I., & Perryman, W. L. (2014). From Progesterone in Biopsies to Estimates of Pregnancy Rates: Large Scale Reproductive Patterns of Two Sympatric Species of Common Dolphin,Delphinusspp. off California, USA and Baja, Mexico. *Bulletin, Southern California Academy of Sciences*, *113*(2), 58-80. https://doi.org/10.3160/0038-3872-113.2.58

Kellar, N. M., Trego, M. L., Marks, C. I., & Dizon, A. E. (2006). Determining pregnancy from blubber in three species of delphinids. *Marine Mammal Science*, *22*(1), 1-16.

Ketterson, E. D., Fudickar, A. M., Atwell, J. W., & Greives, T. J. (2015). Seasonal timing and population divergence: when to breed, when to migrate. *Current Opinion in Behavioral Sciences*, *6*, 50-58. https://doi.org/10.1016/j.cobeha.2015.09.001

Kita, S., Yoshioka, M., & Kashiwagi, M. (1999). Relationship between sexual maturity and serum and testis testosterone concentrations in short-finned pilot whales globicephala macrorhynchus. *Fisheries Science*, *65*(6), 878-883.

Koopman, H. N. (2007). Phylogenetic, ecological, and ontogenetic factors influencing the biochemical structure of the blubber of odontocetes. *Marine Biology*, *151*(1), 277-291.

Koren, L., Bryan, H., Matas, D., Tinman, S., Fahlman, Å., Whiteside, D., Smits, J., Wynne-Edwards, K., & Bieber, C. (2018). Towards the validation of endogenous steroid testing in wildlife hair. *Journal of Applied Ecology*, *56*(3), 547-561. https://doi.org/10.1111/1365-2664.13306

Lambertsen, R. H. (1987). A Biopsy System for Large Whales and Its Use for Cytogenetics. *Journal of Mammalogy*, *68*(2), 443-445. https://doi.org/10.2307/1381495

Legacki, E. L., Robeck, T. R., Steinman, K. J., & Conley, A. J. (2020). Comparative analysis of steroids in cyclic and pregnant killer whales, beluga whales and bottlenose dolphins by liquid chromatography tandem mass spectrometry. *Gen Comp Endocrinol*, *285*, 113273. https://doi.org/10.1016/j.ygcen.2019.113273

Lemos, L. S., Haxel, J. H., Olsen, A., Burnett, J. D., Smith, A., Chandler, T. E., Nieukirk, S. L., Larson, S. E., Hunt, K. E., & Torres, L. G. (2022). Effects of vessel traffic and ocean noise on gray whale stress hormones. *Sci Rep*, *12*(1), 18580. https://doi.org/10.1038/s41598-022-14510-5

Lemos, L. S., Olsen, A., Smith, A., Burnett, J. D., Chandler, T. E., Larson, S., Hunt, K. E., & Torres, L. G. (2021). Stressed and slim or relaxed and chubby? A simultaneous assessment of gray whale body condition and hormone variability. *Marine Mammal Science*, *38*(2), 801-811. https://doi.org/10.1111/mms.12877

Lopaka, L. (2020). *NADA: Nondetects and Data Analysis for Environmental Data*. In https://CRAN.R-project.org/package=NADA

Loseto, L. L., Pleskach, K., Hoover, C., Tomy, G. T., Desforges, J.-P., Halldorson, T., & Ross, P. S. (2017). Cortisol levels in beluga whales (Delphinapterus leucas): Setting a benchmark for Marine Protected Area monitoring&#xD. *Arctic Science*. https://doi.org/10.1139/as-2017-0020

Lusseau, D. (2005). Residency pattern of bottlenose dolphins Tursiops spp. in Milford Sound, New Zealand, is related to boat traffic. *Marine Ecology Progress Series*, *295*, 265-272. https://www.int-res.com/abstracts/meps/v295/p265-272/

Mahaffy, S. D., Baird, R. W., McSweeney, D. J., Webster, D. L., & Schorr, G. S. (2015). High site fidelity, strong associations, and long-term bonds: Short-finned pilot whales off the island of Hawai‘i. *Marine Mammal Science*, *31*(4), 1427-1451. https://doi.org/10.1111/mms.12234

Maninger, N., Capitanio, J. P., Mason, W. A., Ruys, J. D., & Mendoza, S. P. (2010). Acute and chronic stress increase DHEAS concentrations in rhesus monkeys. *Psychoneuroendocrinology*, *35*(7), 1055-1062. https://doi.org/10.1016/j.psyneuen.2010.01.006

Marsh, H., & Kasuya, T. (1984). Changes in the ovaries of the short-finned pilot whale, Globicephala macrorhynchus, with age and reproductive activity. *Report of the International Whaling Commission, Special*, *6*, 311-335.

McCormley, M. C., Champagne, C. D., Deyarmin, J. S., Stephan, A. P., Crocker, D. E., Houser, D. S., & Khudyakov, J. I. (2018). Repeated adrenocorticotropic hormone administration alters adrenal and thyroid hormones in free-ranging elephant seals. *Conserv Physiol*, *6*(1), coy040. https://doi.org/10.1093/conphys/coy040

Melica, V., Atkinson, S., Calambokidis, J., Gendron, D., Lang, A., & Scordino, J. (2022). Naturally stressed? Glucocorticoid profiles in blubber of blue and gray whales in response to life history parameters. *Marine Mammal Science*. https://doi.org/10.1111/mms.12954

Miller, P. (2012). The Severity of Behavioral Changes Observed During Experimental Exposures of Killer (Orcinus orca), Long-Finned Pilot (Globicephala melas), and Sperm (Physeter macrocephalus) Whales to Naval Sonar. *Aquatic Mammals*, *38*(4), 362-401. https://doi.org/10.1578/am.38.4.2012.362

Miller, W. L., & Auchus, R. J. (2011). The molecular biology, biochemistry, and physiology of human steroidogenesis and its disorders. *Endocr Rev*, *32*(1), 81-151. https://doi.org/10.1210/er.2010-0013

Mingramm, F. M. J., Dunlop, R. A., Blyde, D., Whitworth, D. J., & Keeley, T. (2019). Evaluation of respiratory vapour and blubber samples for use in endocrine assessments of bottlenose dolphins (Tursiops spp.). *Gen Comp Endocrinol*, *274*, 37-49. https://doi.org/10.1016/j.ygcen.2018.12.015

Mingramm, F. M. J., Keeley, T., Whitworth, D. J., & Dunlop, R. A. (2020). Blubber cortisol levels in humpback whales (Megaptera novaeangliae): A measure of physiological stress without effects from sampling. *Gen Comp Endocrinol*, *291*, 113436. https://doi.org/10.1016/j.ygcen.2020.113436

Mintzer, V. J., Gannon, D. P., Barros, N. B., & Read, A. J. (2008). Stomach contents of mass-stranded short-finned pilot whales (Globicephala macrorhynchus) from North Carolina. *Marine Mammal Science*, *24*(2), 290-302.

Morton, A. B., & Symonds, H. K. (2002). Displacement of Orcinus orca (L.) by high amplitude sound in British Columbia, Canada. *ICES Journal of Marine Science*, *59*(1), 71-80.

Munk, W. H., Spindel, R. C., Baggeroer, A., & Birdsall, T. G. (1994). The heard island feasibility test. *The Journal of the Acoustical Society of America*, *96*(4), 2330-2342.

National Academies of Sciences, E., and Medicine. (2017). *Approaches to Understanding the Cumulative Effects of Stressors on Marine Mammals*. The National Academies Press. https://doi.org/10.17226/23479

New, L. F., Clark, J. S., Costa, D. P., Fleishman, E., Hindell, M. A., Klanjšček, T., Lusseau, D., Kraus, S., McMahon, C. R., Robinson, P. W., Schick, R. S., Schwarz, L. K., Simmons, S. E., Thomas, L., Tyack, P., & Harwood, J. (2014). Using short-term measures of behaviour to estimate long-term fitness of southern elephant seals. *Marine Ecology Progress Series*, *496*, 99-108. https://doi.org/10.3354/meps10547

Noren, D. P., & Mocklin, J. A. (2012). Review of cetacean biopsy techniques: Factors contributing to successful sample collection and physiological and behavioral impacts. *Marine Mammal Science*, *28*(1), 154-199.

Norris, D. O. (2013). *Vertebrate endocrinology*. Elsevier/Academic Press. https://search.library.duke.edu/search?id=DUKE005912860

Nowacek, D. P., Johnson, M. P., & Tyack, P. L. (2004). North Atlantic right whales (Eubalaena glacialis) ignore ships but respond to alerting stimuli. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *271*(1536), 227-231. https://doi.org/10.1098/rspb.2003.2570

O'Brien, J. K., & Robeck, T. R. (2012). The relationship of maternal characteristics and circulating progesterone concentrations with reproductive outcome in the bottlenose dolphin (Tursiops truncatus) after artificial insemination, with and without ovulation induction, and natural breeding. *Theriogenology*, *78*(3), 469-482. https://doi.org/10.1016/j.theriogenology.2012.02.011

Pérez, S., García-López, Á., De Stephanis, R., Giménez, J., García-Tiscar, S., Verborgh, P., Mancera, J. M., & Martínez-Rodriguez, G. (2011). Use of blubber levels of progesterone to determine pregnancy in free-ranging live cetaceans. *Marine Biology*, *158*(7), 1677-1680. https://doi.org/10.1007/s00227-011-1676-9

Pirotta, E., Booth, C. G., Costa, D. P., Fleishman, E., Kraus, S. D., Lusseau, D., Moretti, D., New, L. F., Schick, R. S., Schwarz, L. K., Simmons, S. E., Thomas, L., Tyack, P. L., Weise, M. J., Wells, R. S., & Harwood, J. (2018). Understanding the population consequences of disturbance. *Ecol Evol*, *8*(19), 9934-9946. https://doi.org/10.1002/ece3.4458

Pomeroy, P. (2011). Reproductive cycles of marine mammals. *Anim Reprod Sci*, *124*(3-4), 184-193. https://doi.org/10.1016/j.anireprosci.2010.08.021

Quick, N. J., Isojunno, S., Sadykova, D., Bowers, M., Nowacek, D. P., & Read, A. J. (2017). Hidden Markov models reveal complexity in the diving behaviour of short-finned pilot whales. *Sci Rep*, *7*, 45765. https://doi.org/10.1038/srep45765

Rendell, L., & Gordon, J. (1999). VOCAL RESPONSE OF LONG-FINNED PILOT WHALES (GLOBICEPHALA MELAS) TO MILITARY SONAR IN THE LIGURIAN SEA. *Marine Mammal Science*, *15*(1), 198-204.

Robeck, T. R., Steinman, K. J., & O'Brien, J. K. (2017). Characterization and longitudinal monitoring of serum androgens and glucocorticoids during normal pregnancy in the killer whale (Orcinus orca). *Gen Comp Endocrinol*, *247*, 116-129. https://doi.org/10.1016/j.ygcen.2017.01.023

Robeck, T. R., Steinman, K. J., Parry, C. B., Gomez, F. M., & Jensen, E. D. (2021). Comparisons of Serum Progesterone and Progestagen Concentrations in Normal and Abnormal Bottlenose Dolphin (Tursiops truncatus) Pregnancies. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.630563

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 Biol Sci*, *279*(1737), 2363-2368. https://doi.org/10.1098/rspb.2011.2429

Romano, T. A., Keogh, M. J., Kelly, C., Feng, P., Berk, L., Schlundt, C. E., Carder, D. A., & Finneran, J. J. (2004). Anthropogenic sound and marine mammal health: measures of the nervous and immune systems before and after intense sound exposure. *Canadian Journal of Fisheries and Aquatic Sciences*, *61*(7), 1124-1134. https://doi.org/10.1139/f04-055

Romero, L. M. (2002). Seasonal changes in plasma glucocorticoid concentrations in free-living vertebrates. *Gen Comp Endocrinol*, *128*, 1-24.

Romero, L. M., & Wikelski, M. (2001). Corticosterone levels predict survival probabilities of Galapagos marine iguanas during El Nino events. *Proc Natl Acad Sci U S A*, *98*(13), 7366-7370. https://doi.org/10.1073/pnas.131091498

Sapolsky, R. M., Romero, L. M., & Munck, A. U. (2000). How do glucocorticoids influence stress responses? Integrating permissive, suppressive, stimulatory, and

preparative actions. *Endocrine Reviews*, *21*(1), 55-89. https://doi.org/10.1210/er.21.1.55

Scheifele, P. M., Andrew, S., Cooper, R., Darre, M., Musiek, F., & Max, L. (2005). Indication of a Lombard vocal response in the St. Lawrence River beluga. *The Journal of the Acoustical Society of America*, *117*(3), 1486-1492.

Shearer, J. M., Jensen, F. H., Quick, N. J., Friedlaender, A., Southall, B., Nowacek, D. P., Bowers, M., Foley, H. J., Swaim, Z. T., Waples, D. M., & Read, A. J. (2022). Short-finned pilot whales exhibit behavioral plasticity in foraging strategies mediated by their environment. *Marine Ecology Progress Series*, *695*, 1-14. https://doi.org/10.3354/meps14132

Shearer, J. M., Quick, N. J., Cioffi, W. R., Baird, R. W., Webster, D. L., Foley, H. J., Swaim, Z. T., Waples, D. M., Bell, J. T., & Read, A. J. (2019). Diving behaviour of Cuvier's beaked whales (Ziphius cavirostris) off Cape Hatteras, North Carolina. *R Soc Open Sci*, *6*(2), 181728. https://doi.org/10.1098/rsos.181728

Sheriff, M. J., Dantzer, B., Delehanty, B., Palme, R., & Boonstra, R. (2011). Measuring stress in wildlife: techniques for quantifying glucocorticoids. *Oecologia*, *166*(4), 869-887. https://doi.org/10.1007/s00442-011-1943-y

Short, R. (1960). Blood progesterone levels in relation to parturition. *Reproduction*, *1*(1), 61-70.

Short, R. V. (1956). Progesterone in the Placentæ of Domestic Animals. *Nature*, *178*(4536), 743-744. https://doi.org/10.1038/178743b0

Southall, B. L., Moretti, D., Abraham, B., Calambokidis, J., DeRuiter, S. L., & Tyack, P. L. (2012). Marine mammal behavioral response studies in Southern California: Advances in technology and experimental methods. *Marine Technology Society Journal*, *46*(4), 48-59. https://doi.org/10.4031/MTSJ.46.4.1

Southall, B. L., Nowacek, D. P., Miller, P. J. O., & Tyack, P. L. (2016). Experimental field studies to measure behavioral responses of cetaceans to sonar. *Endangered Species Research*, *31*, 293-315. https://doi.org/10.3354/esr00764

Spencer, T. E., & Bazer, F. W. (2002). Biology of progesterone action during pregnancy recognition and maintenance of pregnancy. *Front Biosci*, *7*, d1879-1898. https://doi.org/10.2741/spencer

St Aubin, D. J., Ridgway, S. H., Wells, R. S., & Rhinehart, H. (1996). Dolphin thyroid and adrenal hormones: Circulating levels in wild and semidomesticated Tursiops truncatus, and influence of sex, age, and season. *Marine Mammal Science*, *12*(1), 1-13. https://doi.org/10.1111/j.1748-7692.1996.tb00301.x

Stanczyk, F. Z., & Clarke, N. J. (2010). Advantages and challenges of mass spectrometry assays for steroid hormones. *J Steroid Biochem Mol Biol*, *121*(3-5), 491-495. https://doi.org/10.1016/j.jsbmb.2010.05.001

Steinman, K. J., Montano, G. A., & Robeck, T. R. (2021). Characterization of Circulating Androgens, Cortisol and Estrogens During Normal, Abnormal and False Pregnancy in

Bottlenose Dolphins (Tursiops truncatus) Under Managed Care. *Frontiers in Marine Science, 8*. https://doi.org/10.3389/fmars.2021.737926

Steinman, K. J., Robeck, T. R., & O'Brien, J. K. (2016). Characterization of estrogens, testosterone, and cortisol in normal bottlenose dolphin (Tursiops truncatus) pregnancy. *Gen Comp Endocrinol, 226*, 102-112. https://doi.org/10.1016/j.ygcen.2015.12.019

Taiyun Wei and Viliam Simko. (2021). *R package 'corrplot': Visualization of a Correlation Matrix*. In https://github.com/taiyun/corrplot

Teerlink, S., Horstmann, L., & Witteveen, B. (2018). Humpback Whale (Megaptera novaeangliae) Blubber Steroid Hormone Concentration to Evaluate Chronic Stress Response from Whale-Watching Vessels. *Aquatic Mammals*, *44*(4), 411-425. https://doi.org/10.1578/am.44.4.2018.411

Thomas, J. A., Kastelein, R. A., & Awbrey, F. T. (1990). Behavior and blood catecholamines of captive belugas during playbacks of noise from an oil drilling platform. *Zoo Biology, 9*(5), 393-402.

Thomson, C., & Geraci, J. (1986). Cortisol, aldosterone, and leucocytes in the stress response of bottlenose dolphins, Tursiops truncatus. *Canadian Journal of Fisheries and Aquatic Sciences*, *43*(5), 1010-1016.

Thorne, L. H., Foley, H. J., Baird, R. W., Webster, D. L., Swaim, Z. T., & Read, A. J. (2017). Movement and foraging behavior of short-finned pilot whales in the Mid-Atlantic Bight: importance of bathymetric features and implications for management. *Marine Ecology Progress Series*, *584*, 245-257.

Tort, L., & Teles, M. (2011). The endocrine response to stress-a comparative view. In *Basic and Clinical Endocrinology Up-to-Date*. InTech.

Trana, M. R., Roth, J. D., Tomy, G. T., Anderson, W. G., & Ferguson, S. H. (2015). Influence of sample degradation and tissue depth on blubber cortisol in beluga whales. *Journal of Experimental Marine Biology and Ecology*, *462*, 8-13.

Trego, M. L., Kellar, N. M., & Danil, K. (2013). Validation of blubber progesterone concentrations for pregnancy determination in three dolphin species and a porpoise. *PLoS One, 8*(7), e69709. https://doi.org/10.1371/journal.pone.0069709

Trumble, S. J., Robinson, E. M., Berman-Kowalewski, M., Potter, C. W., & Usenko, S. (2013). Blue whale earplug reveals lifetime contaminant exposure and hormone profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(42), 16922-16926. http://www.jstor.org/stable/23750687

Tyson Moore, R. B., Douglas, D. C., Nollens, H. H., Croft, L., & Wells, R. S. (2020). Post-Release Monitoring of a Stranded and Rehabilitated Short-Finned Pilot Whale (Globicephala macrorhynchus) Reveals Current-Assisted Travel. *Aquatic Mammals*, *46*(2), 200-214. https://doi.org/10.1578/am.46.2.2020.200

van Anders, S. M. (2013). Beyond masculinity: testosterone, gender/sex, and human social behavior in a comparative context. *Front Neuroendocrinol*, *34*(3), 198-210. https://doi.org/10.1016/j.yfrne.2013.07.001

Van Cise, A. M., Baird, R. W., Baker, C. S., Cerchio, S., Claridge, D., Fielding, R., Hancock-Hanser, B., Marrero, J., Martien, K. K., Mignucci-Giannoni, A. A., Oleson, E. M., Oremus, M., Poole, M. M., Rosel, P. E., Taylor, B. L., & Morin, P. A. (2019). Oceanographic barriers, divergence, and admixture: Phylogeography and taxonomy of two putative subspecies of short-finned pilot whale. *Mol Ecol*, *28*(11), 2886-2902. https://doi.org/10.1111/mec.15107

Visser, F., Cure, C., Kvadsheim, P. H., Lam, F. P., Tyack, P. L., & Miller, P. J. (2016). Disturbance-specific social responses in long-finned pilot whales, Globicephala melas. *Sci Rep*, *6*, 28641. https://doi.org/10.1038/srep28641

Warnock, F., McElwee, K., Seo, R. J., McIsaac, S., Seim, D., Ramirez-Aponte, T., Macritchie, K. A., & Young, A. H. (2010). Measuring cortisol and DHEA in fingernails: a pilot study. *Neuropsychiatric Disease and Treatment*, *6*, 1.

Wasser, S. K., Risler, L., & Steiner, R. A. (1988). Excreted Steroids in Primate Feces Over the Menstrual Cycle and Pregnancy1. *Biology of Reproduction*, *39*(4), 862-872. https://doi.org/10.1095/biolreprod39.4.862

Williams, T. M., & Maresh, J. L. (2015). Exercise energetics. *Marine mammal physiology: Requisites for ocean living*, 47-68.

Wittmaack, C., Urbán Ramírez, J., Bernot-Simon, D., Martínez-Aguilar, S., Subbiah, S., Surles, J. G., Looney, M., Kumar, N., Halaska, B., Duignan, P. J., Knauss, M., Burns, K., & Godard-Codding, C. A. J. (2022). Small Blubber Samples (50 mg) Sufficient for Analyses of 10 Stress and Reproductive Steroid Hormones in Gray and Fin Whales via Liquid Chromatography Mass Spectrometry. *Frontiers in Marine Science*, *8*. https://doi.org/10.3389/fmars.2021.808764

Wood, S. N., Pya, N., & Säfken, B. (2016). Smoothing Parameter and Model Selection for General Smooth Models (with discussion). *Journal of the American Statistical Association*, *111*, 1548-1575. https://doi.org/10.1080/01621459.2016.1180986

Yoshioka, M., Aida, K., & Hanyu, I. (1989). Correlation of serum progesterone levels with reproductive status in female striped dolphins and short-finned pilot whales. *日本水産学会誌*, *55*(3), 475-478.

Zhang, P., Wei, Z., Hui, S. W., Abel, G., Martelli, P., Hao, Y., & Li, S. (2021). Sexual maturity, seasonal estrus, and gestation in female Indo-Pacific bottlenose dolphins Tursiops aduncus inferred from serum reproductive hormones. *Integr Zool*, *16*(4), 575-585. https://doi.org/10.1111/1749-4877.12491