A method for determining sex and chromosome copy number: sex-by-sequencing reveals the first two species of marine mammals with XXY chromosome condition

ANTHONY L. Einfeldt,* DARA N. ORBACH, AND LAURA J. FEYRER

Department of Biology, Dalhousie University, 1355 Oxford Street, Halifax, Nova Scotia B3H 4J1, Canada (ALE, DNO, LJF)

* Correspondent: tony.einfeldt@gmail.com

Molecular assays of sex determination typically rely on qualitative evaluation of sex-linked markers, which can lead to uncertainty when results contradict morphological identifiers of sex. To investigate whether disagreement between phenotypic and genotypic assays of sex could be underpinned by variation in sex chromosome copy number, we developed a quantitative genetic method to determine sex that is broadly applicable to mammals with XY sex determination. We sequenced a region of the zinc-finger gene \( \text{ZF} \), which has fixed genetic differences between the X and Y chromosomes, and screened 173 cetacean specimens for \( \text{ZFX} - \text{ZFY} \) haplotype identity and read depth. Using a subset of 90 male specimens, we demonstrate that haplotype read depth is an accurate estimator of chromosome copy number. We identified three specimens representing two different cetacean species that had external female morphological traits, Y chromosome haplotypes, and ratios of \( \text{ZFX} : \text{ZFY} \) haplotypes that were above the 1:1 value expected for genetic males. These results provide the first evidence of XXY aneuploidy in cetaceans. Investigation of the reproductive tract of one specimen, a True’s beaked whale (\( \text{Mesoplodon mirus} \)), revealed an intersex phenotype; despite having external characteristics typically diagnostic for the female sex, a penis and testes were present. Our results suggest that intersex phenotypes may be associated with XXY aneuploidy, and that this phenomenon may be underestimated due to it not being detectable by qualitative assays for determining sex.

Key words: chromosomal anomalies, disorder of sex development, genetic sex determination, Hyperoodon ampullatus, marine mammal, Mesoplodon mirus, northern bottlenose whale, True’s beaked whale, XXY aneuploidy

Accurately determining the sex of individuals has important consequences for understanding population parameters, behavior, ecology, and evolution. Although sexual dimorphism and observations of genitalia can be used to identify sexes in some species, there are inherent challenges associated with sex determination in the field. Molecular techniques for determining sex provide a reliable alternative method to determine an individual’s sex from tissue samples (e.g., Palsbøll et al. 1992; Bérubé and Palsbøll 1996; Shaw et al. 2003; Morin et al. 2005). Due to their molecular basis and perceived lack of ambiguity, molecular assays of sex are often considered more reliable and robust than observational methods.

Mismatches between molecular genotype and sexual phenotype are, however, not uncommon. For example, PCR amplification of the \( \text{ZFX} - \text{ZFY} \) homologous regions of sex chromosomes has had mismatches between phenotypic and genotypic sex at rates as high as 7.0% in beluga whales (\( \text{Delphinapterus leucas} \)—Shaw et al. 2003; Petersen et al. 2012) and 10.4% in beavers (\( \text{Castor canadensis} \)—Williams et al. 2004). If phenotypic information is unavailable (e.g., with skin biopsies or non-invasive sloughed skin sampling—Palsbøll et al. 1997), molecular assays of sex do not provide information about errors in sex determination and could skew demographic estimates of population parameters (Robertson and Gemmel 2006). When phenotypic information is available, discrepancies between molecular and morphological results for sex determination are typically attributed to human errors in field observations, sample labeling, or technical errors in molecular methods. This overlooks the possibility that apparently failed assays and mismatches between genotype and phenotype may be of biological importance.

Genetic sex determination assays assume that individuals are either XY males or XX females, which can be false at the level...
of phenotype or genotype. At the phenotypic level, intersex individuals with gonads, sex hormones, or genitalia that do not conform to the definitions of male or female sex characteristics have been documented across a growing number of taxa, including fish, amphibians, reptiles, and mammals (Abdel-Moneim et al. 2015; Raudsepp and Chowdhary 2016). In cetaceans, individuals with both male and female gonads or genital phenotypes have been identified in fin whales (Balaenoptera physalus—Bannister 1963), beluga whales (Delphinapterus leucas—De Guise et al. 1994), striped dolphins (Stenella coeruleoalba—Nishiwaki 1953), bowhead whales (Balaena mysticetus—Tarpley et al. 1995), and short-beaked common dolphins (Delphinus delphis—Murphy et al. 2011). There is little known about the developmental factors influencing these conditions, but phenotypic variations have previously been attributed to hormonal disturbance during early pregnancy and androgen expression failure (De Guise et al. 1995; Tarpley et al. 1995).

At the genotypic level, individuals that are not strictly XX or XY have been genetically confirmed in many wild and domestic terrestrial mammals (e.g., horses, cats, dogs, tigers—Hauffe et al. 2010; Raudsepp and Chowdhary 2016). Due in part to the difficulty of identifying and characterizing variations in sex chromosomes, the range of the diversity and effects on sexual phenotype are not well known. In humans, variation in copy number of sex-chromosome-linked genes can result from incomplete segregation (non-disjunction) during meiosis I, meiosis II, or post-zygotic mitosis (Hall et al. 2006), from incomplete segregation (non-disjunction) during meiosis I, meiosis II, or post-zygotic mitosis (Hall et al. 2006), and the most common sex chromosome aneuploidy present in humans is the XXY genotype (Jacobs and Strong 1959; Skuse et al. 2018). Individuals with this aneuploidy often exhibit male secondary sexual characteristics with some feminization, due to the presence of the testes-determining gene SRY (Visootsak and Graham 2006). Mosaicism of multiple cell lines with different karyotypes occurs in ~15–20% of human XXY males, but is likely underestimated (Samplaski et al. 2014; Mohd Nor and Jalaludin 2016). In domesticated mammals, the majority of Y chromosome abnormalities are in mosaic form, with different cell types carrying different combinations of Y and X aneuploidies (Raudsepp and Chowdhary 2016).

Establishment of non-invasive molecular methods to detect sex chromosome copy numbers would provide more sensitive, efficient, and cost-effective means to understand the extent of sex chromosome anomalies in live animals and wild populations, reduce uncertainties regarding technical or human errors, and support investigations on the causes of atypical sexual development.

To identify variation in copy number in sex chromosomes, using an assay applicable to a wide range of marine mammals, we modified a molecular protocol for sex determination (Konrad et al. 2017) to provide information on the sequence identity and relative quantity of sex chromosome haplotypes present in tissue samples. Konrad et al.’s (2017) method uses PCR amplification and digestion of a 94-bp region of the ZFX–ZFY gene, with a sex-specific TaqI cut-site polymorphism that cuts a region of ZFX (TCAG) but not the homologous region in ZFY (TTAG). The short amplified fragment length makes the method suitable for highly degraded DNA (Morin et al. 2005; Konrad et al. 2017). The success of this amplification-digestion technique relies on a stable cut-site with low levels of polymorphism surrounding it. We modified this technique by directly sequencing amplified DNA from the same primers with the Illumina platform.

Materials and Methods

Sampling.—We amplified and sequenced the ZFX–ZFY marker from samples of two True’s beaked whales (Mesoplodon mirus), six Sowerby’s beaked whales (Mesoplodon bidens), one Cuvier’s beaked whale (Ziphius cavirostris), one beluga whale that had both male and female genitalia, and 163 northern bottlenose whales (Hyperoodon ampullatus; Table 1). An additional 19 whales were assayed but did not pass downstream sequence-quality filters. Samples of True’s, Sowerby’s, and Cuvier’s beaked whales and beluga whales were taken from stranded, naturally deceased animals in Atlantic Canada and Québec. Tissues from northern bottlenose whales were obtained from animals whaled in Labrador in 1971, and biopsies collected from live animals from Nova Scotia, Newfoundland, and the Davis Strait.

Laboratory procedures.—Genomic DNA was extracted using a phenol–chloroform protocol (Sambrook and Russell 2006) on archived gum tissue scrapings and a glass binding filtration protocol (modified from Elphinstone et al. 2003) on alcohol or dimethyl sulfoxide (DMSO)-preserved tissues of recent collection. To determine identity of sex chromosomes and copy number, we amplified a 94-bp region of the homologous sex-linked zinc-finger protein ZFX–ZFY using PCR conditions and primers described by Konrad et al. (2017) and ligated forward and reverse Illumina index adapters in unique combinations to the PCR products of each sample. As an additional test for the presence or absence of the Y chromosome in whales with external female phenotypes from which we

Table 1.—Species sequenced for a 59-bp portion of the sex-linked zinc-finger protein locus ZFX–ZFY. Total number of individuals, number of individuals with each external phenotype, number of individuals with mismatched external sex phenotype and genetic sex assay based on the presence of Y-linked haplotype, and number of individuals with ZFX/ZFY ratios above the expected 1:1 value are indicated for each species.

<table>
<thead>
<tr>
<th>Species</th>
<th>n</th>
<th>Male</th>
<th>Female</th>
<th>Unknown</th>
<th>Intersex</th>
<th>Phenotype–genotype mismatches</th>
<th>High ZFX/ZFY ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperoodon ampullatus</td>
<td>163</td>
<td>86</td>
<td>71</td>
<td>6</td>
<td>6</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Delphinapterus leucas</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesoplodon bidens</td>
<td>6</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mesoplodon mirus</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ziphius cavirostris</td>
<td>1</td>
<td>2</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>173</td>
<td>90</td>
<td>76</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
</tbody>
</table>
recovered ZFY haplotypes, we assayed for the SRY gene using PCR conditions and the forward primer described by Richard et al. (1994—5′-CATGTCGTGTCGTCGATGTC-3′) and a reverse primer (5′-AGTCTCTGTGCCCTCCTCGAGAAAT-3′) designed from SRY sequence from a killer whale (Orcinus orca; GenBank AB108526.2), with index adapters in unique combinations for each sample ligated to PCR products. We used DNA from seven whales with external female phenotypes and in which no ZFY haplotype was detected as controls (one True’s beaked whale, four Sowerby’s beaked whales, and two northern bottlenose whales). For both ZFX–ZFY and SRY assays, we sequenced single-end reads on an Illumina MiSeq at the Marine Gene Probe Laboratory (Halifax, Canada).

Analyses.—We demultiplexed sequences using bcl2fastq2 (Illumina) and trimmed all Illumina indexes, adapter content, and PCR primers to produce sequences of 59 bp for the ZFX–ZFY marker and 107 bp for the SRY marker. Samples were interpreted as having an SRY gene if SRY haplotype sequences were detected. For the ZFX–ZFY assay, we identified whether sequences were derived from an X chromosome (ZFX) or Y chromosome (ZFY) in silico based on a single base pair difference at the 21st bp position, which corresponds to a change in the TaqI restriction enzyme cut-site (ZFX: TCGA; ZFY: TTGA). We examined the relationship between the number of ZFX–ZFY haplotype reads and non-target (i.e., non-ZFX–ZFY sequences) or sequence-error reads (i.e., reads that aligned to ZFX–ZFY sequences but occurred at <10% of the frequency of the most common haplotype in each sample). We found a significant negative linear relationship between error and the number of ZFX–ZFY reads (linear model r² = 0.155, P < 0.001), and filtered out individuals with below read depth corresponding to 20% read error rate (<188 ZFX–ZFY reads). As each cell (except blood cells) in males was expected to have one copy each of the X and Y chromosomes, the corresponding ZFX and ZFY sequences in males were expected to have a 1:1 ratio of ZFX:ZFY haplotypes. To determine whether read depth could be used as an indicator of chromosome copy number, we tested whether the ratio of ZFX to ZFY haplotypes in males were normally distributed around the expected 1:1 ratio using the Shapiro–Wilk’s test of normality.

Morphological assessment.—For one anomalous specimen described below (True’s beaked whale, Field Code: TBW-2017-268), we necropsied the reproductive tract. Photographs were collected of the external genital opening. The presence of a single long urogenital opening with parallel mammary glands or two disjunct urogenital openings and no mammary glands were used to characterize the specimen as female or male, respectively. The frozen-thawed excised reproductive tract was then carefully examined and all blubber and extraneous muscle tissues were removed to isolate any gonads, accessory glands, ducts, bladder, and genitalia. The paired gonads were opened in cross-section to visually confirm if they were ovaries or testes. We searched for the presence of any vaginal, cervical, or uterine structures to indicate the specimen was a female, or for the presence of a penis tip or shaft to indicate the specimen was a male. The presence of reproductive tract structures was noted relative to standard positioning in female and male cetaceans.

Results

Sex identification by genotype assays.—We detected differing levels of sequence divergence between the ZFX and ZFY haplotypes in different species, with ZFX differing from ZFY by five variable sites (H. amphiatus, M. bidens, Z. cafferrostris) or four variable sites (M. mirus). Because ZFX is differentiated from ZFY at multiple positions, haplotype sequences have improved confidence in identifying the presence or absence of sex chromosomes when compared to assays that rely on a single base pair polymorphism in the same region. Within species, ZFX and ZFY sequences were conserved.

We successfully sequenced the ZFX–ZFY gene for 173 individuals, and 167 of these individuals had external phenotype information available (Table 1). One beluga whale that had previously been described as having both male and female genitalia (Field Code: DL-02-89—De Guise et al. 1994) tested positive for both the ZFY and SRY genes. Of the 166 individuals, whose sex was determined successfully by sequencing and with either external male or female phenotypes, 163 sex assignments were in agreement with previous phenotypic descriptions. The exceptions were two northern bottlenose whales (adult, NBW-1971-19; and juvenile, NBW-2018-16) and one True’s beaked whale (calf, TBW-2017-268) that had female external phenotypes but also had ZFY haplotypes. We recovered SRY sequences for all three anomalous individuals, confirming the presence of a Y chromosome. These results suggest that the mismatch between external phenotypes and genotypes in the three anomalous individuals may be due to phenotyping error or biological processes not accounted for by binary sexing assays.

ZFX:ZFY ratios.—To determine variation in sex chromosome copy number in the interssex beluga whale and three anomalous beaked whales with mismatches between external phenotype and genotype, we compared their ratios of ZFX:ZFY haplotype read counts to the distribution found in individuals with male external phenotypes (n = 90). The ratio of ZFX:ZFY haplotypes in all externally phenotypic male samples fit a normal distribution with a mean of 0.9799 ± 0.0213 (95% confidence interval; Shapiro–Wilk’s test P = 0.23; Fig. 1). These results indicate that the ratio of ZFX:ZFY haplotype reads is an accurate indicator of the expected number of sex chromosomes. The beluga whale had a ZFX:ZFY ratio of 0.97 (481/465 reads), suggesting it has a single copy of both the X and Y chromosomes. All three anomalous beaked whales with external female phenotypes that tested positive for a Y chromosome had ZFX:ZFY ratios above the 1:1 ratio expected for XY individuals (Fig. 1). The two northern bottlenose whales had ZFX:ZFY ratios of 1.5 (119/79 reads) and 2.0 (240/120 reads), respectively. The True’s beaked whale had additional reproductive tract tissue available, and was further investigated with two sequencing replicates for each of three DNA samples isolated from separate tissue biopsies. We observed ZFX:ZFY...
ratios of 2.17 (581/274 reads) and 2.25 (750/333 reads) in the first tissue sample, 1.66 (164/99 reads) and 0.86 (88/102 reads) in the second tissue sample, and 3.16 (275/87 reads) and 3.20 (467/146 reads) in the third tissue sample.

Morphological assessment.—Our examination of the reproductive tract from the True’s beaked whale calf indicated this individual had both male and female phenotypic characteristics. Externally, the calf appeared to be a female because of two pronounced mammary slits and one long and continuous urogenital slit, which are traits that are typically sex-diagnostic for female cetaceans (Fig. 2A). There was no clitoris present at the opening of the urogenital canal (Fig. 2B). Internally, a small penis measuring less than 10 cm from the shaft to the tip and testes were present (Fig. 2C). There was no evidence of spermatogenesis from the testes. The pelvic bones were found in the typical position for males at the crurae of the penis. No internal female genital structures were found. The positioning of reproductive tract structures conformed to typical male cetaceans (D. N. Orbach, pers. obs.). Although diagnosing genitalia as developmentally “normal” is difficult given that the reproductive tracts of True’s beaked whales have not been previously published, no anomalies were evident in the penis gross morphology.

**Discussion**

Sequencing an amplified region of the \(ZFX-ZFY\) gene granted several major improvements over presence–absence and amplification–restriction sex assays, including: 1) generating direct sequences of the \(ZFX\) and \(ZFY\) haplotypes, removing reliance on enzymatic digestion and gel electrophoresis, and reducing ambiguity in genotypeing due to the presence of several polymorphic sites differentiating \(ZFX\) and \(ZFY\), 2) allowing simultaneous genotyping of the \(ZFX-ZFY\) region with other genetic markers by multiplexing, and 3) generating sequence depth per haplotype information that serves as a proxy of the relative amount of each sequence present in extracted DNA (e.g., Trost et al. 2018). Direct sequencing correctly identified the external sex characteristics of all but three of the 173 whales screened in this study. If these samples were analyzed using presence–absence and amplification–restriction sex assays, the apparent conflict between each of these three samples having an external female phenotype and a Y chromosome would likely have been attributed to observer error, and the individuals would have been designated as male based on the molecular results. In contrast, the unambiguous read-depth information produced by direct sequencing revealed that the apparent mismatch between external phenotype and genotype in these three samples are not errors, but result from sex chromosome aneuploidies. These results demonstrate that the sex-by-sequencing method provides reduced uncertainty compared other molecular sexing assays, and produces additional information about chromosome copy number that is of that is fundamental importance to understanding species’ biology.

Two of the three anomalous individuals we detected were northern bottlenose whales. One individual was harvested during commercial whaling in 1971 in the Davis Strait, and had previously been identified as female based on melon form and genital phenotype characteristics by whalers. This individual had also been identified as female based on the absence of banding in a gel electrophoresis-based assay for the \(SRY\) gene (Richard et al. 1994) as part of a study by Dalebout et al. (2006). However, electrophoresis-based sex assays are prone to Type II error, especially in old and highly degraded DNA (Morin et al. 2000).
which may have led to incorrect sex genotyping when this individual was initially assayed. The second anomalous northern bottlenose whale was encountered during surveys in the Davis Strait in 2018. This individual was identified in the field as a juvenile female based on melon morphology (Gowans et al. 2000). The presence of a Y chromosome in both of these individuals is at odds with their externally female genotypes, and with most presence–absence-based sex assays would be attributed to observer or laboratory error. However, the high $ZFX$:$ZFY$ ratios that we detected in both these individuals support an alternative hypothesis that these whales have XXY aneuploidy. This could explain the apparent mismatch between genotype and phenotype, as sex chromosome aneuploidies in other mammals affect sexual development. Further testing of this hypothesis was not possible, as neither whale was available for morphological inspection, and direct observation of sex chromosomes requires fresh tissue undergoing meiosis.

The hypothesis that XXY aneuploidy is associated with an external female phenotype in beaked whales was supported by the third anomalous whale. This True’s beaked whale had external female characteristics, $ZFX$–$ZFY$ and $SRY$ genes, and a high $ZFX$:$ZFY$ ratio. It stranded in the Magdalen Islands in 2017 and its reproductive tract was excised and preserved frozen. Despite having external traits that are typically sex-diagnostic for female cetaceans (two pronounced mammary slits and one long continuous urogenital slit), this individual had an intersex
phenotype with internal traits typical of males (a penis, testes, and pelvic bones positioned at the crurae of the penis). While the phenomenon of mammary slits in male cetaceans has not to our knowledge been published in the scientific literature, we are aware of anecdotes of its occurrence in other cases. We suggest that it occurs infrequently, has not been well documented, and is worth considering whether other such individuals may be intersex. Genetically, the True’s beaked whale had variable ZFX:ZFY ratios from different tissue samples within its reproductive tract, which were mostly above the 1:1 ratio expected for genetic males. These results suggest that in addition to XXY aneuploidy, the True’s beaked whale had genetic mosaicism of multiple cell lines with different karyotypes. Little is known about the phenotypic effects of sex chromosome aneuploidy in cetaceans, but in humans XXY aneuploidy usually results in a predominantly male phenotype, although female characteristics can occur despite the presence of the male-determining SRY gene (Thangaraj et al. 1998; Saavedra-Castillo et al. 2005). The external female and internal male sex characteristics of the True’s beaked whale suggests there could be underlying differences between humans and cetaceans in sexual development, despite sharing a conserved XX–XY sex determination system.

The variations in sex chromosome copy number in three whales and the intersex phenotype observed in the True’s beaked whale calf are consistent with XXY aneuploidy and mosaicism of multiple cell lines with different karyotypes, but other explanations cannot be entirely ruled out. Abnormal ratios of homologous regions of sex chromosomes have also been observed due to translocations of part of the Y chromosome in hematopoietic cells can also lead to sex chromosomes will reduce error and provide additional read-depth information provided by sex-by-sequencing identified these individuals as having variations in sex chromosome copy number, resulting in a 0% error rate. Given the non-zero error rates of assigning phenotypes from genotypes in other marine mammals, we anticipate that a combination of adopting a third phenotypic category for intersex individuals and accounting for copy number variations in sex chromosomes will reduce error and provide further insight into the factors affecting sexual determination.

**Acknowledgments**

Biopsy samples from *H. ampullatus* were collected under permit from the Department of Fisheries and Oceans (DFO), Species at Risk division. Biopsies collected in the Davis Strait were contributed by S. Fergusson, DFO Arctic. Samples and data collected from harvested *H. ampullatus* were provided by N. Øien at the Institute for Marine Research in Norway. *M. bidens* and *Z. cavirostris* were collected by the Marine Animal Response network. *M. mirus* tissue samples were collected by the Québec Marine Mammal Emergency Response Network. Tissue samples from stranded animals were stored by D. McAlpine as part of the collection at the Museum of New Brunswick. T. Frasier at St. Mary’s University provided tissue samples from the intersex *D. leucas* specimen collected by the Québec Marine Mammal Emergency Response Network previously described in De Guise et al. (1994).

**Literature Cited**


De Guise et al. 1994). Muscle tissue from this individual tested positive for SRY and had a ZFX:ZFY ratio of 0.97, suggesting it had a single copy of both the X and Y chromosomes. The results from our method support De Guise et al.’s (1995) conclusions that environmental factors have had adverse effects on sexual development in beluga whales. This further demonstrates the utility of sex-by-sequencing in ruling out chromosomal variations as a potential cause of abnormalities in sexual phenotypes, which is an important test in systems where alterations in sexual and functional development are suspected to be chemically induced (Colborn and Clement 1992).

The sex-by-sequencing method that we present here has demonstrated utility in resolving sexual phenotype via low-cost molecular methods, with improved explanatory power over presence–absence based assays of sex chromosomes. Although sex determination of three out of the 173 whales screened in this study would typically be considered erroneous, either in field observations of phenotype or in genotype from presence–absence based methods, the additional read-depth information provided by sex-by-sequencing identified these individuals as having variations in sex chromosome copy number, resulting in a 0% error rate. Given the non-zero error rates of assigning phenotypes from genotypes in other marine mammals, we anticipate that a combination of adopting a third phenotypic category for intersex individuals and accounting for copy number variations in sex chromosomes will reduce error and provide further insight into the factors affecting sexual determination.


Submitted 1 April 2019. Accepted 06 August 2019.

Associate Editor was Aleta Hohn.