**Cetacean genome size diversity**

Elias M. Oziolor1,a,\*, T. Ryan Gregory2,a, John W. Bickham3, Judy St. Leger4, Cole W. Matson1

1 Department of Environmental Science, Center for Reservoir and Aquatic Systems Research, and the Institute for Biomedical Studies, Baylor University, Waco, TX 76798, USA

2 Department of Integrative Biology, University of Guelph, Guelph, Ontario N1G 2W1, Canada

3 Department of Wildlife and Fisheries Sciences, Texas A&M University, College Station, TX 77845, USA

4 Animal Care – Research and Pathology, SeaWorld Parks and Entertainment, San Diego, CA 92109, USA

a These authors contributed equally to this manuscript.

\*Corresponding author: Elias M. Oziolor

Email: emoziolor@ucdavis.edu

**Keywords**: Body mass, brain size, Cetacea, C-value, encephalization quotient, evolution, genome size, whale

Sirenians and cetaceans are unique among mammals for exhibiting a purely aquatic lifestyle, resulting in dramatic morphological, physiological, and behavioral changes (Gatesy *et al.* 2013). Freed from the constraints of gravity by the buoyancy of water, cetaceans have evolved a range of body sizes, from the relatively diminutive Maui's dolphin (*Cephalorhynchus hectori maui*) at 1 m in length and 50 kg in mass to the largest animal ever known, the blue whale (*Balaenoptera musculus*) at ~30 m and almost 200,000 kg (Jefferson 1993, Montgomery *et al.* 2013). Modern dolphins, porpoises, and other whales (classically placed in the order “Cetacea”) are nested phylogenetically within the order Artiodactyla of even-toed ungulates (Gatesy *et al*. 2013). A revised order Cetartiodactyla reflects these new phylogenetic insights and emphasizes the close evolutionary affinity of whales to their terrestrial relatives (Gatesy *et al*. 2013).

Cetacea contains 88 extant species, divided into the 70 toothed whales (Odontoceti) and the 18 baleen whales (Mysticeti). Due to the protected status of many whales combined with inherent challenges in sampling, several aspects of their biology remain poorly described. Large-scale genomic analyses have become available for cetaceans only quite recently and include members of both toothed and baleen whales (McGowen *et al.* 2012, Keane *et al.* 2016). However, estimates of total genome size are still quite rare. At present, the Animal Genome Size Database contains data for only five species of cetaceans, most of which come from a single study published more than 10 yr ago (Du and Wang 2006).

Genome size has been shown to correlate with various cellular, morphological, developmental, and physiological parameters in mammals and other animals (Gregory 2005, Smith and Gregory 2009), but it has not been possible to examine potential relationships such as these in cetaceans. In other taxa, it has been observed that genome size, through the intermediate of cell size, is related to resting metabolic rate, diving ability, migration distance, cold tolerance, longevity, relative brain size, body size, and other such organism-level parameters (Gregory 2001, 2005; Andrews and Gregory 2009; Smith and Gregory 2009). We provide flow cytometric estimates for 18 species of cetaceans (primarily odontocetes) as well as the hippopotamus (*Hippopotamus amphibious*), the closest living relative of cetaceans, and another predominantly aquatic mammal, the bearded seal (*Erignathus barbatus*), to provide information for future comparisons that explore the evolutionary relationships of genome size in cetaceans.

Spleens were collected during postmortem evaluations from stranded or captive animals. Samples from stranded animals were collected under NOAA's responsibility to the MMPA 1972 under Section 109(h), and a Stranding Agreement as part of the Marine Mammal Health and Stranding Response Act. Samples from animals under managed care, or captivity, were collected during necropsy examinations from animals held under USDA/APHIS Registration No: 58-C-0077. All samples (Table S1, S2) were stored at -80˚C until analysis.

Spleen samples were used to establish genome size with a flow cytometry protocol originally developed by Vindelov and Christensen (1994) and modified by Oziolor *et al*. (2014). A minimum of 10,000 cells were used to establish a mean genome size of each sample. Information about the available cetaceans with previously established genome sizes by flow cytometry (7 individuals; 5 species) was added to this analysis from the Animal Genome Size Database. Chicken blood was used as a standard (accepted as 1.25 picograms (pg)) and was co-prepared with samples.

Haploid nuclear DNA contents in mammals range approximately 5-fold from 1.63 pg in Carriker's round-eared bat (*Lophostoma carrikeri*) to 8.40 pg in the polyploid red viscacha rat (*Tympanoctomys barrerae*), averaging 3.21 pg ± 0.85 SD (Gregory 2017). Some groups, most notably bats, have particularly small and constrained genomes, whereas others, such as rodents, have genomes that cover a much wider range (Gregory 2017) (Fig. 1). The expanded data set presented here suggests that cetaceans have genome sizes that are typical for mammals. As shown in Figure 1, the cetaceans studied to date have genome sizes that fall within the range observed among terrestrial artiodactyls, though they exhibit a larger average size than other members of the Cetartiodactyla (3.68 pg ± 0.40 SD for cetaceans *vs*. 3.39 pg ± 0.55 SD in terrestrial artiodactyls). The hippopotamus, which is the closest living relative of cetaceans, was found in the present study to have a genome size of 3.39 pg – *i.e*., average for an artiodactyl.

The smallest genomes so far reported for Cetartiodactyla (~2.7 pg) are found in species such as giraffes (*Giraffa camelopardalis*), camels (*Camelus bactrianus*), and llamas (*Lama glama*) (Gregory 2017). Interestingly, camels and llamas have erythrocytes that are not only enucleate like those of other adult mammals but are atypical in shape. Whereas most mammals (including cetaceans) have erythrocytes that are shaped as biconcave discs, those of some terrestrial artiodactyls may be tiny, spherical, oval, or crescent-shaped (Gulliver 1875). Previous relationships have shown the strong relationship between enucleate cell size and genome size in mammals (Gregory 2000, 2001). Thus, it is possible that achieving these unique erythrocyte morphologies requires smaller than average nuclei in progenitor cells, and thus smaller than average genomes. Unfortunately, data on cetacean genome sizes and cell sizes are insufficient for a detailed assessment of any correlation between these two parameters, but across mammals (Gregory 2000) and vertebrates more broadly (Gregory 2001), there is a strong positive relationship between genome size and overall cell size. However, it can be said that cetacean erythrocyte sizes are in the range that would be expected given their genome sizes (Hawkey 1975, Gregory 2000).

Aside from cetaceans, only a few other aquatic mammals have been studied for genome size thus far (Gregory 2017). However, even with this limited information, it appears that there is significant variability in genome size across species that share an aquatic lifestyle. For example, the bearded seal genome size was estimated in the present study as 2.96 pg, which is similar to published estimates for the spotted seal (*Phoca largha*; 2.94 pg) and California sea lion (*Zalophus californianus*; 3.15 pg) (Du and Wang 2006). All are smaller than the mammalian average. By contrast, the Florida manatee (*Trichechus manatus*) has a much larger estimated genome size of 4.67 pg (Redi *et al.* 2007) .

Cetaceans are remarkable for many reasons, though they remain a challenge to study. Very little information has been available on the genome sizes of cetaceans, in part because of the difficulty in obtaining material for analysis. The addition of 18 new estimates for cetaceans, as well as their closest living relative, the hippopotamus, represents an important step toward filling this gap in the animal genome size dataset. It is important to note that it is difficult to make any informed phylogenetically independent contrasts in cetaceans with the data we have available because of the necessity to obtain large enough sample sizes of each species represented. This is a direction that should be prioritized for future investigations.

**ACKNOWLEDGEMENTS**

We thank the Baylor University Molecular Biosciences Center for providing access to the flow cytometer used in this study. We thank Lexi Mena and Nancy Stedman of SeaWorld Parks and Entertainment, Megan Stolen and Wendy Durden of the Hubbs-SeaWorld Research Institute and Frances Gulland from The Marine Mammal Center for their contributions in sample collection and management of samples. EMO performed the genome size estimates. EMO and TRG analyzed the data and wrote the manuscript. JWB advised on analysis and edited manuscript. JSL organized and provided samples as well as edited manuscript. CWM organized study reviewed the manuscript and troubleshot data acquisition.

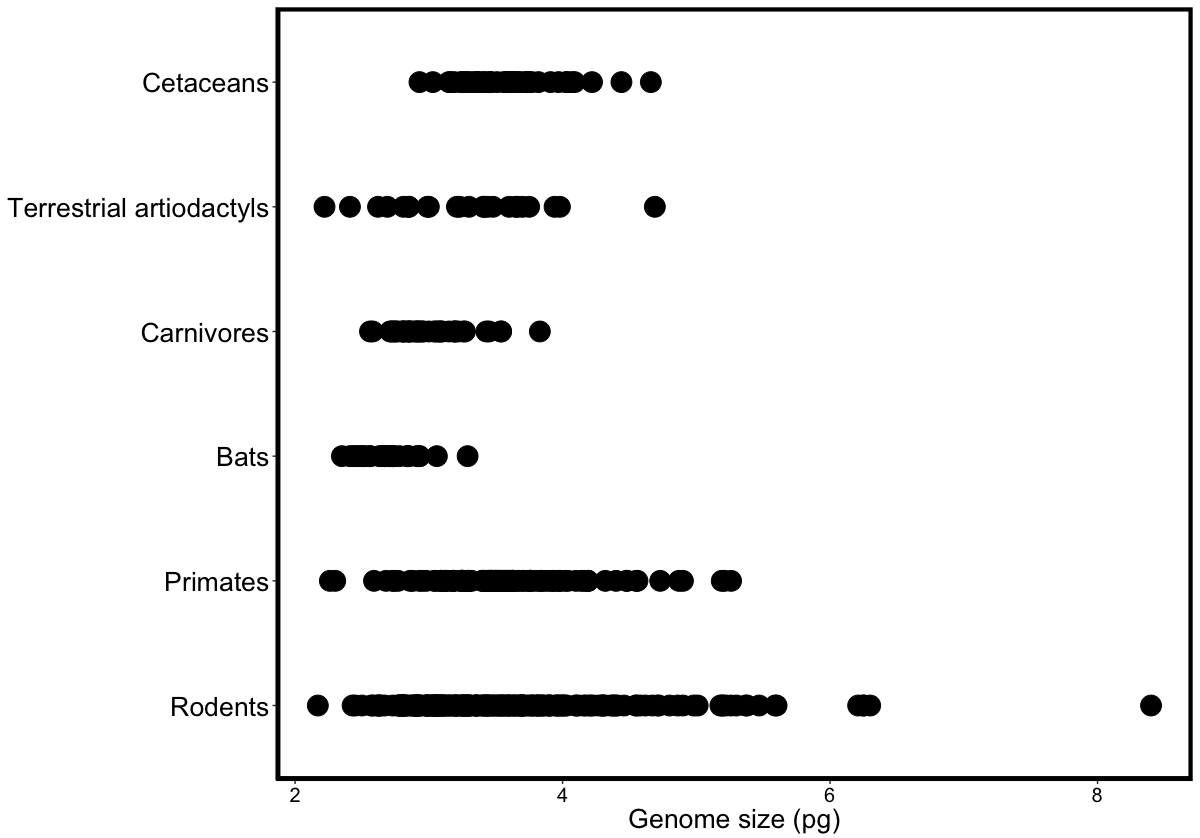
**DATA ACCESSIBILITY**

All our data are accessible in Supplementary Information.

**FUNDING**

Funding for this project was provided by start-up funds to CWM from Baylor University.

**FIGURES**



**Figure 1.** Diversity in genome size estimates for some major groups of mammals, based on data from the Animal Genome Size Database (Gregory 2017) and the present study (Table S1, S2). Each dot represents a single estimate. Numbers of species/individuals used are as follows: cetaceans (*n*=24/*n*=47); terrestrial artiodactyls (*n*=20/*n*=24); carnivores (*n*=28/*n*=43); bats (*n*=25/*n*=28); primates (*n*=66/*n*=110); rodents (*n*=90/*n*=218)

**REFERENCES**

Andrews, C. B. and T. R. Gregory. 2009. Genome size is inversely correlated with relative brain size in parrots and cockatoos. Genome 52:261-267.

Du, B. and D. Wang. 2006. C-values of seven marine mammal species determined by flow cytometry. Zoological Science 23:1017-1020.

Gatesy, J., J. H. Geisler, J. Chang*, et al.* 2013. A phylogenetic blueprint for a modern whale. Molecular Phylogenetics and Evolution 66:479-506.

Gregory, T. R. 2000. Nucleotypic effects without nuclei: Genome size and erythrocyte size in mammals. Genome 43:895-901.

Gregory, T. R. 2001. The bigger the C-value, the larger the cell: Genome size and red blood cell size in vertebrates. Blood Cells Molecules and Diseases 27:830-843.

Gregory, T. R. 2005. Genome size evolution in animals. Elsevier Science Bv, Sara Burgerhartstraat 25, Po Box 211, 1000 Ae Amsterdam, Netherlands.

Gregory, T. R. 2017. Animal genome size database. www.genomesize.com.

Gulliver, G. 1875. Observations on the sizes and shapes of the red corpuscles of the blood of vertebrates, with drawings of them to a uniform scale, and extended and revised tables of measurements. Proceedings of the Zoological Society of London, Part I: 474-495.

Hawkey, C. M. 1975. Comparative mammalian haematology. The Whitefriars Press Ltd., Great Britain.

Jefferson, T. A., Leatherwood S., Webber, M.A. 1993. FAO Species identification guide. Marine mammals of the world., Food and Agriculture Organization of the United Nations or of the United Nations Environment Programme. Rome.

Keane, M., J. Semeiks, A. E. Webb*, et al.* 2016. Insights into the evolution of longevity from the bowhead whale genome. Cell Reports 10:112-122.

McGowen, M. R., L. I. Grossman and D. E. Wildman. 2012. Dolphin genome provides evidence for adaptive evolution of nervous system genes and a molecular rate slowdown. Proceedings of the Royal Society B-Biological Sciences 279:3643-3651.

Montgomery, S. H., J. H. Geisler, M. R. Mcgowen, C. Fox, L. Marino and J. Gatesy. 2013. The evolutionary history of cetacean brain and body size. Evolution 67:3339-3353.

Oziolor, E. M., E. Bigorgne, L. Aguilar, S. Usenko and C. W. Matson. 2014. Evolved resistance to PCB- and PAH-induced cardiac teratogenesis, and reduced CYP1A activity in Gulf killifish (*Fundulus grandis*) populations from the Houston Ship Channel, Texas. Aquatic Toxicology 150:210-219.

Redi, C. A., S. Garagna, M. Zuccotti and E. Capanna. 2007. Genome size: A novel genomic signature in support of Afrotheria. Journal of Molecular Evolution 64:484-487.

Smith, J. D. L. and T. R. Gregory. 2009. The genome sizes of megabats (Chiroptera: *Pteropodidae*) are remarkably constrained. Biology Letters 5:347-351.

Vindelov, L. L. and I. J. Christensen. 1994. Detergent and proteolytic enzyme-based techniques for nuclear isolation and DNA content analysis. Methods in Cell Biology, Vol 41 41:219-229.