Cantaurus, Vol. 20, 14-16, May 2012 © McPherson College Division of Science and Technology

Genetic Variability and Single Nucleotide Polymorphisms in Humpback Whale DNA from Wilhelmina Bay, Antarctic Peninsula

Shelly Brandt

ABSTRACT

Genetic variation in humpback whale DNA is a powerful tool for determining the uniqueness of a given population. In order to develop a greater understanding of the species success, standard DNA sequencing techniques are used. This method produces a series of nucleotide base sequences that encode for the various traits in an organism. To determine the uniqueness of an individual, single nucleotide polymorphisms (snips) in the DNA are observed. A single nucleotide polymorphism is a variation in a DNA sequence, that occurs when a single nucleotide base (A, T,C,G) in the genome differs between members of a given biological species or that of paired chromosomes. If the DNA contains large quantities of single nucleotide polymorphisms, it is indicative that the individual is genetically diverse. Such variation is a promising sign that there is not excessive inbreeding occurring in a specific population. This may also lead to other implications of the population's ecological relationships within the environment as far as competition for food or mating. Lack of genetic variation may be a red flag that a population is under stress from the environment. By studying specific trends in the DNA, may reinforce the probability of observing these warning signs prior to the populations decline. Therefore such research is beneficial for the preservation of humpback whale populations in the Antarctic Peninsula.

Keywords: *humpback whale, megaptera novaengliae, single nucleotide polymorphisms, DNA sequencing*

INTRODUCTION

Humpback Whales (Megaptera novaengliae) are derived from lineages of ancient origin. They are considered to have an extensive life-span overlapping generations despite the ongoing prohibition against whale hunting. (Baker et al. 1993)Humpback Whales were once estimated to be approximately 125,000 individuals, distributed throughout three predominant oceanic populations; North Pacific, North Atlantic and the Southern oceans. Since then whale populations have been drastically reduced to <5000 individuals.(Nowacek et al. 2011) Excessive hunting over the last 200 years has reduced some regional sub-populations to near extinction, resulting in less than 300 individuals. (Baker et al. 1993.)

This drastic change in population also described as a population bottleneck, has greatly impacted genetic variability in Hump Back Whales. The decrease in genetic diversity is a prominent concern to conservation geneticists. The loss of such diversity may be influenced by evolutionary factors related to inbreeding, loss of reproductive fitness, and randomized mating. Further interpretation may also suggest a correlation between commercial hunting and its effects on genetic variation. (Perry et al. 1993)These factors may lead to elevated extinction risks, and increased genetic mutations. (Frankham 2012) Genetic diversity is required for populations to evolve and adapt. Favoring natural selection and ensuring the species survival. (Palumbi et al. 1994)

Phylogenetic reconstruction and gene flow analysis is used to determine whether genetic variation is effected by post-exploitation migration, or simply a remnant of past population genetic variation. The use of PCR (Polymerase Chain Reaction) amplifies DNA fragments, resulting in the detection and interpretation of genetic variation in a natural population. Expected nucleotide divergence is anticipated to be within >0.196% in each geographic subpopulation. This suggests that the DNA from each subpopulation will be identical to one another. The DNA will then be sequenced to determine the specific nucleotide sequence of each whale. These nucleotide sequences are then analyzed. The result of which helps to determine if there is a significant correlation between the DNA of each sample taken, and that of the overall species of Megaptera novangliae. Single nucleotide polymorphisms are a single base change in the nucleotide sequence of the genome. This is the single nucleotide variation between individuals that determines their uniqueness. If these factors produce a significant correlation, and the data is closely related, implications can be made that suggest the whale population is undergoing severe environmental effects. These factors are influenced by: excessive inbreeding or lack of reproductive diversity, commercial whaling, or a significant decrease in food source population.

MATERIALS AND METHODS

The samples for genetic analysis were derived from 28 individual humpback whales from Wilhelmina Bay

Antarctic Peninsula. Research was conducted with the assistance of the National Science Foundation (NSF) in collaboration with Duke University. DNA was collected from biopsy samples, acquired using a cross bow and a small cutter tip. This method removes a small sample of skin and blubber roughly the size of a pencil eraser. DNA is contained at -81 degrees C. Each micro centrifuge tube contains approximately >10ul.

DNA was sequenced using standard DNA sequencing procedures. The forward primer sequence: Dlp1.5L (5' TCA CCC AAA GCT GRA RTT CTA 3') and a reverse primer sequence: Dlp5-H (5' CCA TCG WGA TGT CTT ATT TAA GRG GAA 3') required for the initial reaction. The forward and reverse primers required synthesis. Sequencing followed standard protocol. DNA and primers did not require premixing prior to sequencing. NCBI database and Finch TV were used to analyze the data.

RESULTS

DNA sequencing revealed errors in the DNA.50% of the DNA sequenced using ABI sequence analyzer, and sequencher caused the DNA to appear as though all of the reactions had essentially failed and therefore would yield no quality sequencing reads.

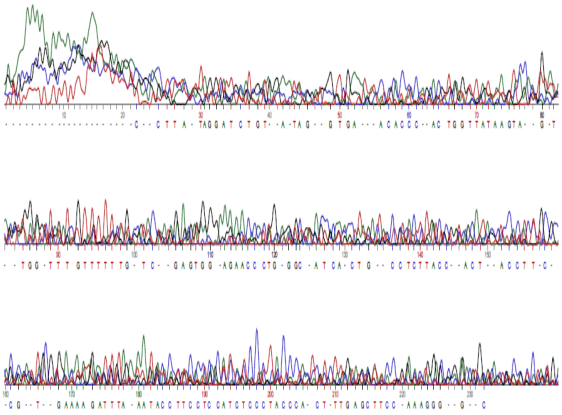

Figure 1: DNA sequence sample B156c, displaying spacing and nucleotide base sequencing.

The remaining functional DNA was analyzed through NCBI data base. Initial blasting sequences at highly similar sequences(megablast) resulted in no correlations to any nucleotide sequences in the database. Secondary blasting at somewhat similar sequence levels(Blastn) resulted in low/moderate percentages of query coverage. The samples were closet in relation to that of homo sapien(human), and mus musculus(house mouse.)

Table 1: Results of DNA blast

Sample	Megablast	Blastn	NCBI Database	Genome
B133a	0	18%	0	5%
B138b	0	0	33%	0
B144a	0	26%	19%	17%
B156c	0	12%	55%	16%
WAM532	0	21%	53%	0

Table 2: Sequence matches

Sample	Match(99%)	% Fit
B133a	Homo Sapien	18%
B138b	0	0
B144a	Homo Sapien	26%
B156c	Mus Musculus	12%
WAM532	Mus Musculus	21%

The data was then blasted against a randomly selected FASTA sequence for Megaptera novangliae. The resulting data determined that there was a <60% match between the sequences. The initial nucleotide sequences were blasted against the Megaptera novangliae genome.

The resulting data concludes that there was a <20% correlation between the DNA and that of the Humpback Whale genome. And in some cases the data determined that there was <10% correlation between the DNA and the genome. Invalidating the hypothesis, and resulting in the implication that the DNA may be contaminated. This leads to possibility that the DNA samples may be favoring that of another species. This may also imply that the DNA is unnaturally diverse, and that there is a great deal of genetic variation. Statistics show that there is less than a 20% of the DNA being that of humpback whale, and an 80% chance that the DNA is not that of a humpback whale. The difficulty being that you are looking for a 1 in 10,000 ratio of finding a correlation between the DNA sample and that of the entire species genome.

The ability to determine Single Nucleotide Polymorphisms was compromised due to error in the DNA sequencing process that resulted in incomplete nucleotide sequences.

DISCUSSION

The lack of correlation between the DNA samples and that of NCBI data may be influenced by potential contamination of the DNA samples prior to shipping. Heat may have also played a role in the formation of secondary structures in the DNA that would have complicated the sequencing process.

There is the possibility that the primers were

inadequate for sequencing, resulting in the inability to read the DNA properly.

ACKNOWLEDGEMENTS

Special thanks to Duke University. Tom Schultz and Douglas P. Nowacek and to the individuals at MWG operon Eurofins.

LITERATURE CITED

Baker C,Perry A,Bannister J, Weinrich T et. Al (1993) Abundant mitochondrial DNA variation and world – wide population structure in humpback whales. Proc.Natl.Acad.Sci (1993) vol 90: 8239-8243

Baker CS,Palumbi SR, Lambersten M, Weinrich J et al. (1990) Influence of seasonal migration on geographic distribution of mitochondrial DNA halotypes in humpback whales.Nature 344: 238-239.

Frankham R(2012) How closely does genetic diversity in finite populations conform to predictions of neutral theory? Large deficits in regions of low recombination. Heredity 108: 167-178

Mangin B, Siberchicot A, Nicolas S, Doligez A (2012) Novel measures of linkage disequilibrium that correct the bias due to population structure and relatedness. Heredity 108: 285-291

Nowacek DP, Friedlaender AS, Halpin PN, Hazen EL, Johnston DW, et al. (2011) Super-Aggregations of Krill and Humpback Whales in Wilhelmina Bay, Antarctic Peninsula. PLoS ONE 6(4): e19173. doi:10.1371/journal.pone.0019173

Palumbi AR, Cipriano F(1998) Species identification using genetic tools: the value of nuclear and mitochondrial gene sequences in whale conservation. J Hered (1998) 89(5): 459-464 doi:10.1093/jhered/89.5.459

Palumbi SR, Baker CS(1994) Contrasting population structure from nuclear intron sequences and mtDNA of humpback whales. Mol Biol Evol (1994) 11(3): 426-435