Identification of SNP loci in gray whales using NGS sequencing approaches and gray whale samples from across the North Pacific

Aimee R. Lang, Barb Taylor

Rationale:

Considerable information has been derived from genetic studies of the population structure of gray whales. The picture that has emerged suggests that gray whales exhibit maternally-directed site fidelity to at least some feeding grounds (LeDuc et al. 2002, Frasier et al. 2011, Lang et al. 2011, 2014). Significant nuclear (nucDNA) differences have not been detected between whales feeding in the Bering and Chukchi Seas, suggesting that interbreeding between these groups may occur while on migration or in wintering areas (Lang et al. 2014). In contrast, significant nuclear differences have been detected in comparison of whales feeding off Sakhalin with whales feeding in the Bering and Chukchi Seas or while on migratory routes in the eastern North Pacific (Lang et al. 2011). This finding was initially interpreted as being consistent with interbreeding of Sakhalin whales while migrating and overwintering in the western North Pacific. However, recent information obtained from genetics, telemetry, and photo-identification studies have revealed that some Sakhalin whales migrate to and overwinter in the eastern North Pacific (Lang 2010, Mate et al. 2011, Weller et al. 2012, Urban et al. 2013). These findings have led to the need for a reappraisal of the population structure and movements of North Pacific gray whales by the International Whaling Commission's Scientific Committee, and this process was begun at an intersessional workshop this spring (SC/65b/Rep08).

As interest in gray whale population structure has grown, the number of laboratories involved in collecting and analyzing gray whale samples has also increased. Currently, gray whale sample collection and analysis has been and/or is being conducted in labs located in at least five countries (US, Canada, Mexico, Russia, and Japan). While numerous samples from gray whales exist (summarized in SC/65b/Rep08), sampling coverage and intensity in some areas (e.g., the northern feeding ground) is low, and obtaining reasonable sampling coverage is likely a long-term project. In contrast, there are other regions where a moderate to high proportion of whales using the area have been sampled (e.g., PCFG, Sakhalin). To maximize the use of existing samples in the future, however, data should be generated in such a way that it facilitates collaboration amongst labs and that it can be utilized in future analyses as additional samples become available. While this process is quite straightforward for mtDNA sequence data, to date all nuclear analyses of population structure in gray whales have relied on microsatellite data (Lang et al. 2011, 2014; D'Intino et al. In press), which is widely recognized as being problematic in comparisons across labs and over time.

In light of this issue, one of the recommendations made at the Rangewide Workshop (SC/65b/Rep08) was to develop a panel of Single Nucleotide Polymorphism (SNP) markers for use with gray whales. SNP data is sequenced-based and is

straightforward to compare across labs and over time. In addition, SNPs have been shown to be successful in genotyping low-quality samples (Morin and McCarthy 2007), such as bone or baleen, which would facilitate the analysis of any historical samples that are identified. Recent advances in technology have provided efficient and cost-effective methods to identify and genotype SNPs. Given the increased interest in gray whale population structure, the available technology, and the value of collaborations (both now and in the future) to study wide-ranging species, it is timely to begin work on the identification and genotyping of SNP loci in gray whales. An additional benefit of this work is that it will allow comparison of results generated by analysis of more traditional markers (mtDNA control region and microsatellites) with those generated using the methods outlined below. As such, the proposed work also has value in increasing our current understanding of the population structure of gray whales in the North Pacific, as well as in increasing our understanding of the utility of this approach to assess population structure in cetaceans.

Objectives:

1) Utilize next generation sequencing approach to identify a panel of SNPs for use with gray whales.
2) Conduct genotyping by sequencing for 200 samples from three feeding areas to evaluate the consistency of results with previous analyses.

Methods:

Approximately ~200 samples will be chosen from the ~350 existing samples that were collected on the Sakhalin feeding ground, the PCFG feeding ground, and the Northern feeding ground. These samples were previously analyzed using mtDNA control region sequences and 12 microsatellite loci. DNA will be extracted using the recommended protocol (Qiagen Dneasy Extraction Kit). DNA will be quantified using an intercalating dye (PicoGreen). Approximately 30 uL of DNA at 50-100 ng/uL will be needed per sample. DNA will be submitted for genotyping by sequencing (GBS) to the Institute of Genomic Diversity at Cornell University. IGD uses a highly-multiplexed system for constructing reduced representation libraries for the Illumina next-generation sequencing platform (Ellshire et al. 2011). This technique utilizes restriction enzymes to reduce genome complexity and avoid the repetitive fraction of the genome, and results in hundreds to thousands of genotyped SNP markers. Following sequencing, a custom SNP pipeline analysis will be used to identify SNPs for further analysis.

Expected outcomes:

- Identification of a panel of SNPs that can be used by other labs as well as with any historical samples that are available
- Assessment of population structure of gray whales utilizing SNPs; results can be compared to those generated using microsatellites to evaluate whether

both approaches provide consistent results.

Collaborators: Dave Weller, Jon Scordino

Timeline:
Sample selection and preparation: Oct - Dec 2014
Sequencing (Institute of Genomic Diversity): Jan – Feb 2015
Analysis and summary of results: Mar - August 2015

- A progress report will be submitted at SC/66a

Budget:
Total: $18K
Supplies: $10.3K

- Includes NGS sequencing conducted by the Institute of Genomic Diversity at Cornell University
- Includes one-time optimization fee ($700) for sequencing new species

Labor: $7.7K

- Includes labor by technician to prepare samples for processing by IGD: DNA extraction, archiving, quantification, and preparation of dilutions

References:

D'Intino A.M., Darling J.D., Urban R. J., Frasier T.R. 2013. Lack of nuclear differentiation suggests reproductive connectivity between the "southern feeding group" and larger population of eastern North Pacific gray whales, despite previous detection of mitochondrial differences. *Journal of Cetacean Research & Management* 13:97-104.

Elshire R.J., Glaubitz J.C., Sun Q., Poland J.A., Kawamoto K., Buckler E.S., Mitchell S.E. 2011. A robust, simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLOS One* 6:e19379.

Frasier T.R., Koroscil S.M., White B.N., Darling J.D. 2011. Assessment of population substructure in relation to summer feeding ground use in the eastern North Pacific gray whale. *Endangered Species Research* 14:39-48.

International Whaling Commission. 2014. Report of the Workshop on the Rangewide Review of the Population Structure and Status of North Pacific Gray Whales. SC/65b/Rep08.

Lang A.R. 2010. The population genetics of gray whales (*Eschrichtius robustus*) in the North Pacific: University of California San Diego.

Lang A.R., Calambokidis J., Scordino J., Pease V.L., Klimek A., Burkanov V.N., Gearin P., Litovka D.I., Robertson K.M., Mate B.R., Jacobsen J.K., Taylor B.L. 2014. Assessment of genetic structure among eastern North Pacific gray whales on their feeding grounds. *Marine Mammal Science* (early view).

Lang A.R., Weller D.W., LeDuc R.G., Burdin A.M., Brownell R.L.J. 2010. Delineating patterns of male reproductive success in the western gray whale (*Eschrichtius robustus*) population. Paper SC/62/BRG10 presented to the International Whaling Commission Scientific Committee. 22 pp.

Lang A.R., Weller D.W., LeDuc R.G., Burdin A.M., Pease V.L., Litovka D., Burkanov V.,

Brownell R.L.J. 2011. Genetic analysis of stock structure and movements of gray whales in the eastern and western North Pacific. Paper SC/63/BRG10 presented to the International Whaling Commission Scientific Committee. 20 pp.

LeDuc R.G., Weller D.W., Hyde J., Burdin A.M., Rosel P.E., Brownell R.L.J., Wursig B., Dizon A.E. 2002. Genetic differences between western and eastern gray whales (*Eschrichtius robustus*). *Journal of Cetacean Research and Management* 4:1-5.

Mate B., Bradford A., Tsidulko G., Vertyankin V., Ilyashenko V. 2011. Late feeding season movements of a western North Pacific gray whales off Sakhalin Island, Russia and subsequent migration into the eastern North Pacific. Paper SC/63/BRG23 presented to the International Whaling Commission Scientific Committe*e*. 7 pp.

Morin P.A., McCarthy M. 2007. Highly accurate SNP genotyping from historical and low-quality samples. *Molecular Ecology Resources* 7:937-946.

Urbán R. J., Weller D., Tyurneva O., Swartz S., Bradford A., Yakovlev Y., Sychenko O., Rosales N. H., Martínez A. S., Burdin A.M., Gómez-Gallardo U. A. 2013. Report on the photographic comparison of the Sakhalin Island and Kamchatka Peninsula with the Mexican gray whale catalogues. *Paper SC/65a/BRG04 presented to the International Whaling Commission Scientific Committee*.

Weller D.W., Klimek A., Bradford A.L., Calambokidis J., Lang A.R., Gisborne B., Burdin A.M., Szaniszlo W., Urban J., Gomez-Gallardo Unzueta A., Swartz S., Brownell R.L.J. 2012. Movements of gray whales between the western and eastern North Pacific. *Endangered Species Research* 18:193-199.