

geneGIS: Computational tools for spatial analyses of DNA profiles and associated photo-identification records of whales and dolphins

Dori Dick¹, Dawn Wright^{1,2}, John Calambokidis³, Jason Holmberg⁴, Erin Falcone³, Tomas Follett⁵, Debbie Steel⁵, Beth Slikas⁵, C. Scott Baker⁵

¹Oregon State University, Corvallis, OR, ²Esri, Redlands, CA, ³Cascadia Research Collective, Olympia, WA, ⁴Shepherd Project, Portland, OR, ⁵Marine Mammal Institute, Oregon State University, Newport, OR

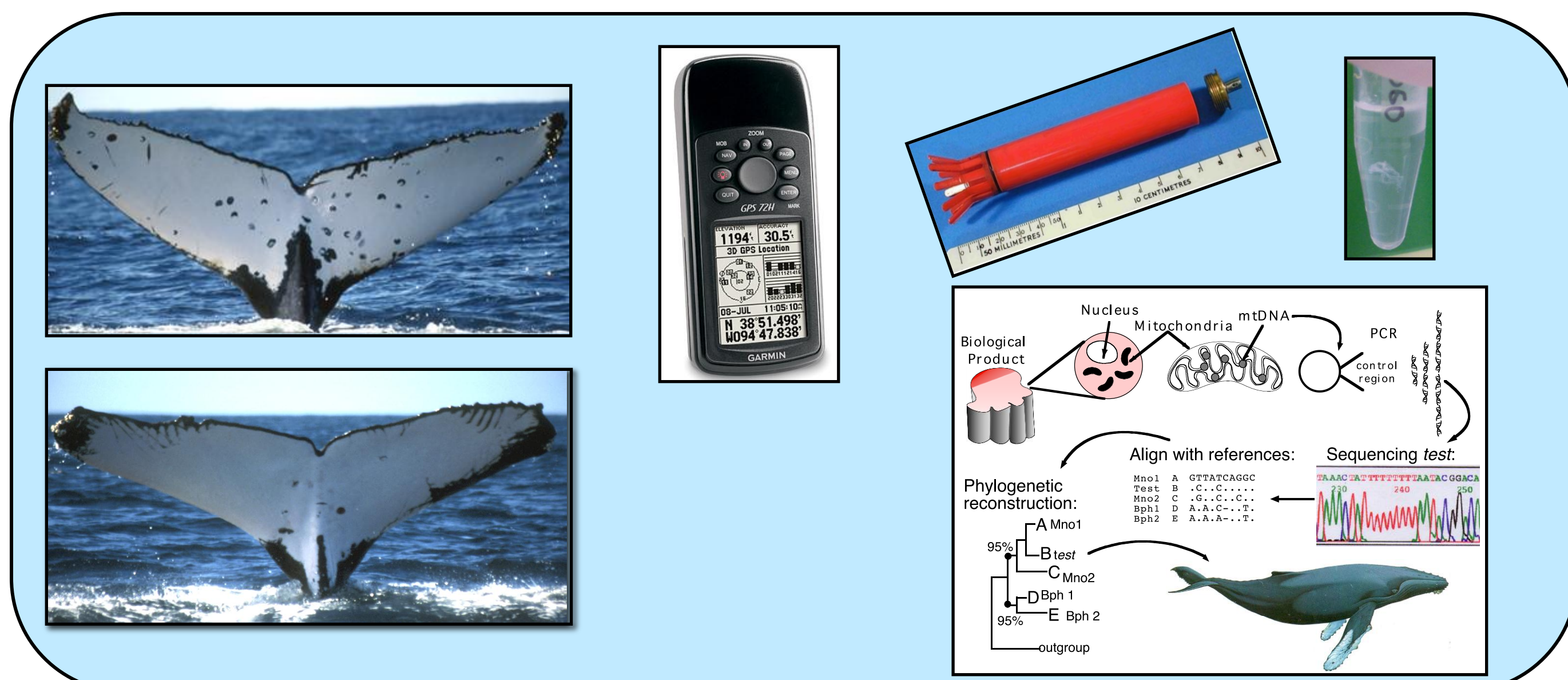


Contact: dickd@geo.oregonstate.edu



BACKGROUND

- More large-scale studies of long-lived species (e.g., whales and dolphins) are now collecting spatially explicit records linked through individual identification to genetic samples and photo-identification.



- Such spatio-temporal records are used to track individual migration and life history parameters, to estimate the abundance and trends of populations and, in the case of genetic markers, to infer close kinship (e.g., parent/offspring relationships) and define management units, or Distinct Population Segments.
- There is a conspicuous absence of computational tools for the integration and spatial exploration of individual records, particularly the potential for linking photo-identification to genetic information (e.g., DNA profiles) and for extending genetic identity to include close kinship.

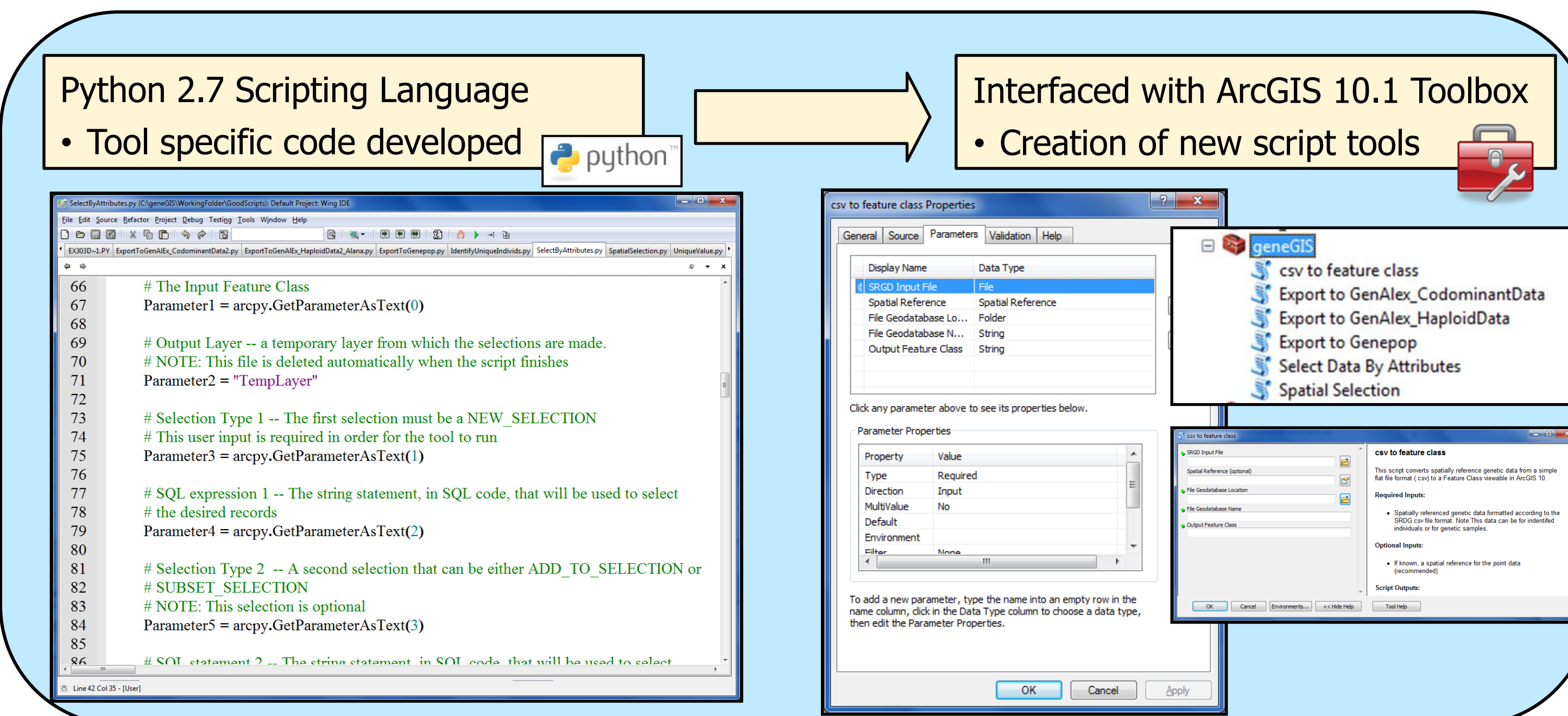
OBJECTIVES

- To enhance understanding of cetacean population structure, ecosystem relationships, and predictive models of human impact we are developing and implementing *geneGIS* – a geographic information system architecture and toolbox to facilitate visual exploration and spatial analyses of individual-based records from DNA profiles and photo-identification.
- Specifically:
 - Data visualization
 - Spatial exploration and selection
 - Spatial analyses
 - Data export to formats required by other genetic analyses software
- We demonstrate the utility of *geneGIS* using SPLASH, a collated database of humpback whale (*Megaptera novaeangliae*) records in the North Pacific (Barlow et al. 2011) with more than 8,000 naturally marked individuals and 2,700 associated DNA profiles including 10 microsatellites, mtDNA haplotypes, and sex.
- Tools for *geneGIS* are under development in two different environments:
 - Python 2.7 and ArcGIS 10.1 via ArcToolbox
 - Java-based, web interface supported by the Shepherd Project (see Jason Holmberg, 8:50am on 10/15: "The Shepherd Project: a software framework for mark-recapture and molecular ecology")



PROGRESS TO DATE

I. Tool Development



II. Data Input

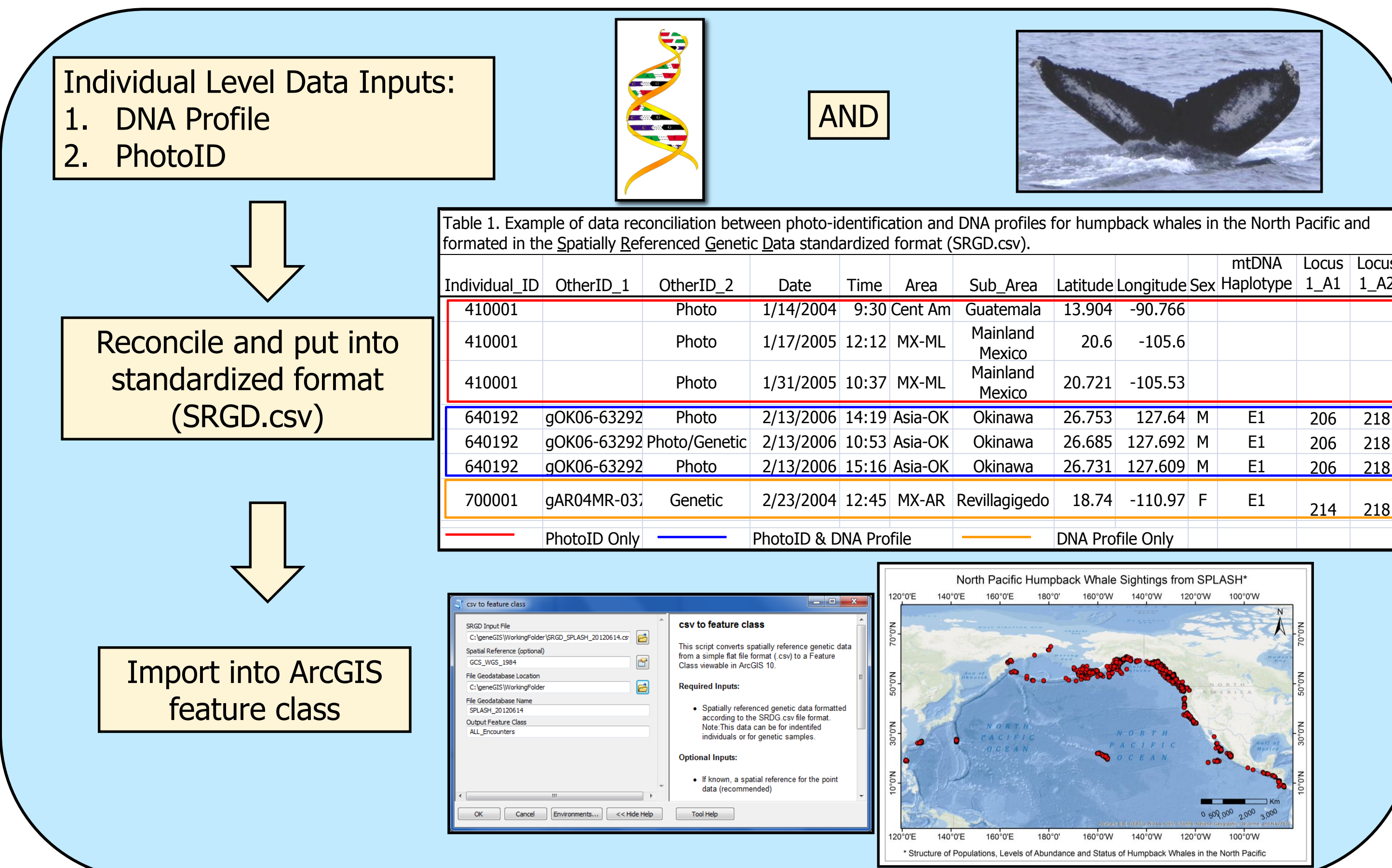
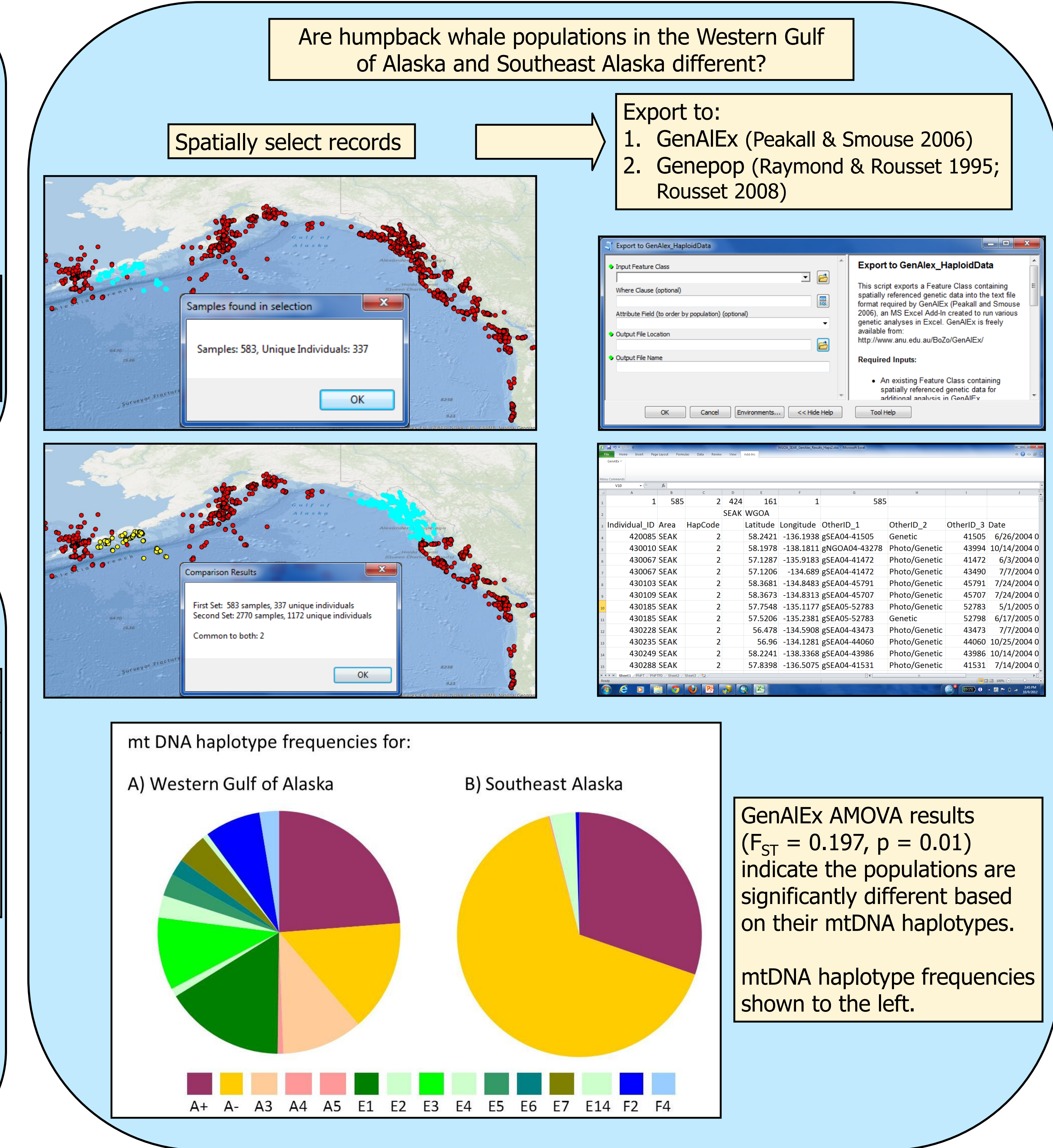


Table 1. Example of data reconciliation between photo-identification and DNA profiles for humpback whales in the North Pacific and formatted in the Spatially Referenced Genetic Data standardized format (SRGD.csv).													
Individual_ID	OtherID_1	OtherID_2	Date	Time	Area	Sub_Area	Latitude	Longitude	Sex	Haplotype	mtDNA	Locus	Locus
410001		Photo	1/14/2004	9:30	Cent Am	Guatemala	13.904	-90.766					
410001		Photo	1/17/2005	12:12	MX-ML	Mainland Mexico	20.6	-105.6					
410001		Photo	1/31/2005	10:37	MX-ML	Mainland Mexico	20.721	-105.53					
640192	gOK06-63292	Photo	2/13/2006	14:19	Asia-OK	Okinawa	26.753	127.64	M	E1	206	218	
640192	gOK06-63292	Photo/Genetic	2/13/2006	10:53	Asia-OK	Okinawa	26.685	127.692	M	E1	206	218	
640192	gOK06-63292	Photo	2/13/2006	15:16	Asia-OK	Okinawa	26.731	127.609	M	E1	206	218	
700001	gAR04MR-03	Genetic	2/23/2004	12:45	MX-AR	Revillagigedo	18.74	-110.97	F	E1	214	218	
PhotoID Only PhotoID & DNA Profile DNA Profile Only													

III. Application Example



SIGNIFICANCE OF RESEARCH

- geneGIS* provides a quantitative approach to conservation priorities by mapping spatial data and relevant attributes (e.g., sex, mtDNA haplotype, microsatellite alleles), improving our ability to study the influence of the seascape on population structure.
 - This is critical to understanding, minimizing, and predicting the impacts of anthropogenic seascape alteration and global climate change on marine biodiversity.
 - geneGIS* is accessible to molecular ecologists who want to visualize and discover spatial patterns in their data.
 - Although *geneGIS* was developed to enhance our understanding of the spatial relationships between marine mammals and their environment, we envision the functionality of *geneGIS* to extend to other taxa in terrestrial and aquatic environments.
 - Future developments include tools to identify and connect individuals between regions and to extract values from input environmental layers (e.g., bathymetry, sea surface temperature).
- ***We are seeking access to additional individual-based genetic and photoID datasets from other species and populations to test tool functionality. Please contact Dori Dick for more information.***

REFERENCES

- Barlow, J. et al. 2011. Humpback whale abundance in the North Pacific estimated by photographic capture-recapture with bias correction from simulation studies. *Marine Mammal Science*. 27:793-818.
- Peakall, R. and P. E. Smouse. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 6:288-295.
- Raymond, M. and F. Rousset. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*. 86:248-249.
- Rousset, F. 2008. GenePop'007: a complete reimplementation of the GenePop software for Windows and Linux. *Molecular Ecology Resources*. 8:103-106.

ACKNOWLEDGEMENTS

We would like to thank Shaun Walbridge of Esri for his help with tool development in ArcGIS 10.1. Thanks to all the researchers involved with SPLASH, Cascadia Research Collective for database maintenance, and OSU Cetacean Conservation Genetics Lab for genetic analyses. This research is funded by ONR contract N0270A to Oregon State University.

