OSAC RESEARCH NEEDS ASSESSMENT FORM

Title of research need: Wildlife Forensics: Bioinformatics for Next Generation Sequencing (NGS)

Describe the need: Next generation sequencing (NGS), including massively parallel sequencing, genotype by sequencing, and whole genome sequencing, are revolutionizing the field of human population genetic analysis, as well as animal population genetics and conservation biology and ecology. These technological advances are accompanied by some substantial bioinformatics challenges, including scaling of algorithms normally used for Sanger sequence data, data storage issues (volume of raw and processed data, analytical space), training of personnel in data analysis methods, and estimation of statistical significance and error rates. Forensic use of NGS technologies will need to deal with additional challenges related to functional validation of genomic data, and standardization of data collection and analysis.

Keyword(s): Whole genome sequencing, SNP, genotype-by-sequencing, restriction-site associated DNA (RADSeq), DNA, mRNA, smallRNA, bioinformatics, standardization, validation, non-model species

Submitting subcommittee(s): Wildlife Forensics **Date Approved:** 10/7/2022

Background Information:

1. Does this research need address a gap(s) in a current or planned standard? (ex.: Field identification system for on scene opioid detection and confirmation)

ANSI ASB 019, Wildlife Forensics General Standards states "4.3.2 Analytical procedures shall be based on peer-reviewed studies and/or validated prior to use in casework. "

2. Are you aware of any ongoing research that may address this research need that has not yet been published (e.g., research presented in conference proceedings, studies that you or a colleague have participated in but have yet to be published)?

The current research is limited to that at the USFWS Conservation Genetics Laboratories and NOAA (Northwest Fisheries Science Center), which have ongoing research directed at mitogenome, genotyping by sequencing, and/or whole genome sequencing of aquatic species of interest to federal fish management, primarily focusing on wet lab applications, not bioinformatics pipelines.

3. Key bibliographic references relating to this research need: (ex.: Toll, L., Standifer, K. M., Massotte, D., eds. (2019). Current Topics in Opioid Research. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-180-3)

Dormontt, E. E., Jardine, D. I., van Dijk, K. J., Dunker, B. F., Dixon, R., Hipkins, V. D., Tobe, S., Linacre, A., & Lowe, A. J. (2020). Forensic validation of a SNP and INDEL panel for individualization of timber from bigleaf maple (*Acer macrophyllum* Pursch). Forensic Science International. Genetics, 46, 102252.

Magi A, Benelli M, Gozzini A, Girolami F, Torricelli F, Brandi ML. (2010). Bioinformatics for next generation sequencing data. Genes (Basel). 1(2):294-307.

Arulandhu AJ, Staats M, Hagelaar R, et al. (2017). Development and validation of a multi-locus DNA metabarcoding method to identify endangered species in complex samples. Gigascience. 6(10):1-18.

Ballard D, Winkler-Galicki J, Wesoly J. (2020). Massive parallel sequencing in forensics: advantages, issues, technicalities, and prospects. Int J Legal Med. 134, 1291-1303.

4. Review the annual operational/research needs published by the National Institute of Justice (NIJ) at https://nij.ojp.gov/topics/articles/forensic-science-research-and-development-technology-working-group-operational#latest? Is your research need identified by NIJ?

Forensic Biology/DNA, Scientific Research, Increased information about the discriminatory power and sensitivity of alternate biological analyses (e.g., proteomics, microbiome, plants, animals) to associate individuals with crime scene evidence.

5. In what ways would the research results improve current laboratory capabilities?

The research results would provide a robust process for the application of bioinformatics to NGS data generated for wildlife forensic investigations, as well as provide a foundation for statistical assessment of error rates.

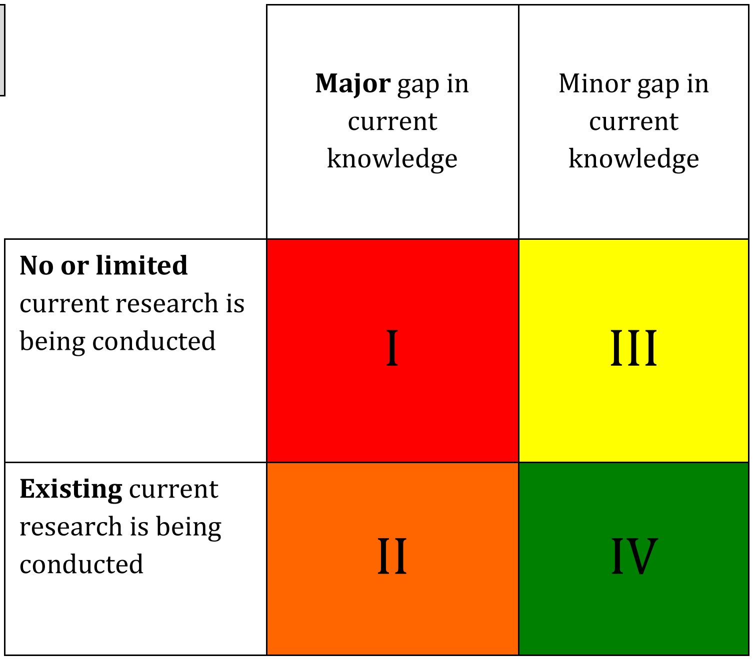
6. In what ways would the research results improve understanding of the scientific basis for the subcommittee(s)?

Not every member of the subcommittee understands the principles of underlying NGS analysis, particularly the data processing steps (pipelines) that identify metrics such as sequence coverage, sequence accuracy, and other traditional population genetic metrics. These research results would provide basic guidelines for developing standardized bioinformatics processes for wildlife species, and statistically sound methods for estimation of error rates.

7. In what ways would the research results improve services to the criminal justice system?

Accurate and rigorous data analysis is essential for presenting robust scientific information to the criminal justice system. NGS bioinformatics is a growing discipline, and currently many novice users of the technology are not well-versed in analysis of volumes of data exponentially scaled up from traditional Sanger sequence data. The research would provide the standard methods developed by trained specialists in bioinformatics.

8. Status assessment (I, II, III, or IV): I



	Major gap in current knowledge	Minor gap in current knowledge
No or limited current research is being conducted	I	III
Existing current research is being conducted	II	IV

This research need has been identified by one or more subcommittees of OSAC and is being provided as an informational resource to the community.