OSAC RESEARCH NEEDS ASSESSMENT FORM

Title of research need: Characterization, optimization, and comparison of DNA sequencing methods

Describe the need: DNA sequencing is a considered a major technological advancement in the field of forensic DNA profiling. As opposed to Whole Genome Sequencing (WGS), the forensics field tends to focus on select regions of the genome to avoid potential ethical pitfalls in generating too much data, in addition to dealing with the significantly lower DNA quality and quantity that is often observed in casework. To date, there are two commercial companies that have produced benchtop sequencers that fit the need of most forensic laboratories in terms of cost and throughput. With the advent of Massive Parallel sequencing (MPS), genetic data being generated using commercial sequencing kits and commercial software for Short Tandem Repeats (STRs), in addition to intelligence-based methodologies such as ancestry/phenotype prediction and genetic genealogy that require SNP data input, there is a need to 1. compare the performance of data generated from STR CE analyses and SNP arrays versus commercial sequencing platforms using MPS (SNP and/or STR). 2. Compare the use of MPS pipelines, both commercial and non-commercial, for their efficacy in alignment and genotype calling between platforms/methods. 3. In terms of the physical chemistry and newly developing methods/approaches to sequencing, (i.e., hybridization, amplicon sequencing etc.), a comparison of these methodologies for low input DNA and mixture samples in particular. Integrating MPS sequencing into casework will be an extreme benefit to the field of DNA profiling and intelligence, however a full assessment of what's available/possible including performances is necessary.

Keyword(s): SNPs, STRs, Massive Parallel Sequencing, pipelines, software

Submitting subcommittee(s): Human Biology **Date Approved:** 10/05/2021

(If SAC review identifies additional subcommittees, add them to the box above.)

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)

Yes, in the use of MPS for 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)?

Yes, but not enough in detail.

3. Key bibliographic references relating to this research need:

1) Alonso, A., Barrio, P.A., Müller, P., Köcher, S., Berger, B., Martin, P., Bodner, M., Willuweit, S., Parson, W., Roewer, L. and Budowle, B. Current state-of-art of STR sequencing in forensic genetics. *Electrophoresis*. 2018;39: 2655-2668.
2) Peter de Knijff. From next generation sequencing to now generation sequencing in forensics. *Forensic Sci Int Genet*. 2019;38:175-180
3) Ballard, D., Winkler-Galicki, J. & Wesoły, J. Massive parallel sequencing in forensics: advantages, issues, technicalities, and prospects. *Int J Legal Med.* 2020; 134:1291–1303.
4) Tytgat, O.; Gansemans, Y.; Weymaere, J.; Rubben, K.; Deforce, D.; Van Nieuwerburgh, F. Nanopore Sequencing of a Forensic STR Multiplex Reveals Loci Suitable for Single-Contributor STR Profiling. *Genes.* 2020;*11*:381.
5) Wang, D., Tao, R., Li, Z. *et al.* STRsearch: a new pipeline for targeted profiling of short tandem repeats in massively parallel sequencing data. *Hereditas. 2020;* 157: 8.
6) Jonathan L. King, August E. Woerner, Sammed N. Mandape, Kapema Bupe Kapema, Rodrigo Soares Moura-Neto, Rosane Silva, Bruce Budowle. STRait Razor Online: An enhanced user interface to facilitate interpretation of MPS data. *Forensic Sci Int Genet*. 2021; 52:102463.
7) Nam Nhut Phan, Amrita Chattopadhyay, Tsui-Ting Lee, Hsiang-I Yin, Tzu-Pin Lu, Liang-Chuan Lai, Hsiao-Lin Hwa, Mong-Hsun Tsai, Eric Y Chuang. High-performance deep learning pipeline predicts individuals in mixtures of DNA using sequencing data. *Briefings in Bioinformatics*. 2021.
8) Erin M. Gorden, Kimberly Sturk-Andreaggi, Charla Marshall. Capture enrichment and massively parallel sequencing for human identification. *Forensic Sci Int Genet*. 2021;102496.
9) Michael Hofreiter, Jiri Sneberger, Martin Pospisek, Daniel Vanek. Progress in forensic bone DNA analysis:Lessons learned from ancient DNA. *Forensic Sci Int Genet*. 2021;54.

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?

Yes, "Better ways to enrich or target genomic areas of forensic DNA interest as opposed to a traditional PCR-based approach".

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

Must compare and contrast commercial and non-commercial methods/software/lab techniques used in sequencing particular regions of the genome for STR/SNP genotype calling. Need to generate comparison information on this data, min read count, algorithms used etc. so that future standards may be implemented. At present consolidating all types of chemistries, pipelines, sample preparation, sample performance and concordance etc. used for sequencing protocols in forensic casework (in particular low input and mixtures) will be extremely valuable, especially with regards STR data for identification and SNP data for intelligence purposes.

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

Better exploration of the caveats of bad samples/data and minimum requirements for accurate genotyping calls due to their importance in downstream identifying and intelligence information.

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

It will provide an overview of the current climate in terms of commercial entities, types of methods, kit performance, including new approaches (i.e. hybridization v PCR) pre sequencing, and more importantly a

significant exploration of the algorithms utilized for variant calling, from STRs to SNPs. This step is vital for correct input of data into statistical and modelling tools/programs downstream. Sequencing alignment methodologies and thresholds for genotyping calling need exploration and comparison within commercial and non-commercial software/tools.

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

	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.