**Minutes for the Interagency Workshop on the use of Next-Generation DNA Sequencing for Human Identification and Characterization (January 31, 2012)**

**National Institute of Standards and Technology, Biochemical Science Division, Gaithersburg, MD 20899**

---

Agenda (download 01)

Introductions (over 40 attendees DoD, NIJ, DHS, NIST, ATF, FBI, others)

Opening remarks (Ken Kroupa - DoD) (download 02)

Objectives: Identify common issues across agencies for next generation sequencing (NGS) technologies. Identify focus areas, need for focused groups on governance, build on past success of ANDE project as a model/example. In the climate of limited funding it is important to work together when goals align. This is not a requirements meeting.

Questions/comments for future discussion:
How would NGS be applied to forensics/biometrics (would it be similar)?
What would the accreditation aspect involve?

---

Speaker Peter M Vallone (NIST) – "Expanding Upon STR Typing for Human Identification"

Talk slides (download 03)

Discussion of current STR typing technologies and how deeper sequencing of STR loci can expand the information for human identity purposes. Examples of Sanger and Mass Spectrometry were presented. Application of NGS for STRs and SNPs were discussed (applications: kinship, biogeographical ancestry, and externally visible traits). Challenges of current NGS typing compared to traditional STR typing were discussed.

Questions/comments for future discussion:
Could microarrays play a role in the SNP typing for the forensic community?
Possibly, but STR loci could not be included.

---

Speaker James Harper (MIT/LL) – "Next Gen Sequencing For Human DNA ID"

Discussion of how the large excess of information content available from NGS when compared to current STR typing could be used to enable a variety of valuable forensics applications. NGS is not for rapid runs and field applications today, but technology development continues to be rapid and laying the scientific foundation now is important to enabling these applications in the future. Discussion of a range of potential applications enabled by high NGS data volumes including: increasing analysis throughput, forensic mixture analysis, kinship analysis, biogeographical estimation, forensic DNA phenotyping (externally visible characteristics like

eye color, hair color, etc), sample authentication, and correlation of individual activity/history epigenetic and metagenomic profiles.

---

Speaker Eric Schwoebel (MIT/LL) – "Technology Review for STR Profile Sequencing"

Discussion of NGS platforms and read length capabilities. Longer read lengths are needed for STRs (if we are to include STR loci for database legacy considerations – there were no objections to this). Discussion of acceptable NGS accuracy (<1%). Described the workflow for 454, Illumina, and Ion-Torrent platforms. An example cost analysis of running various platforms was discussed. Sample multiplexing was encouraged.

Comments:
We (the community) should decide on the information we require (STR, SNPs, etc) and not reduce our expectations based on current read lengths for each platform.

---

Speaker Darrell Ricke (MIT/LL) – "Sequence Analysis"

Discussion of moving from sizing STRs to sequencing STRs and SNPs. Current NGS sequencers have less than 1% base calling errors. Higher coverage of alleles overcomes the NGS errors through consensus base calling. NGS sequence data with errors and allele calling was presented for the vWA locus that was generated as part of a multiple individual multiplex experiment of multiple loci.

---

Speakers Marc Salit and Justin Zook (NIST) – "NGS biases, systematic sequencing errors, and accuracy"

Talk slides (download 04 ,05 06)

Discussion of platform and bioinformatic bias in NGS technologies. A discussion of systematic sequencing errors (SSEs) was presented. The source of these SSEs can be challenging to identify. A discussion of how reference materials can help calibrate data sets was presented.

---

Speaker Rebecca Just (AFDIL) – "NGS applied to Missing Persons work"

Talk slides: please email Rebecca Just (rebecca.s.just@us.army.mil) for slides

The use of NGS for mitochondrial genome and STR sequencing. The Illumina platform was used to sequence 90 unique mitochondrial genomes (multiplexed with bar coding in two flow cell lanes) using amplification and hybridization enrichment for mtDNA. A solution-based hybridization enrichment applied to highly degraded templates, followed by Illumina sequencing, produced complete mtGenome haplotypes with high sequence coverage where previous Sanger sequencing failed. Initial studies with sequencing miniSTRs was discussed. AFDIL presently has a 454 Jr., Illumina MiSeq, and a Pacific Bio system in house. Bioinformatic support is an issue in terms of available expertise.

---

Speaker Todd Bille (ATF) – “ATF Laboratory”

Talk slides (download 07)

The ATF works with primarily touch evidence (90% of cases). Mixtures are a primary concern. Funding is an issues so there are no immediate plans to use NGS, but the interest is there if mixtures can be addressed.

---

Speaker Tom Callaghan (FBI) – “FBI and NGS”

Tom addressed the 6 questions posed to the attendees. The FBI would be interested in having STR loci typed along with other markers by NGS. The FBI would like to see reference materials early in the process. This would assist with validation (platform and software). The FBI plans are for small NGS platform in 2012. There are no concerns with sequencing CODIS loci (STRs, mito), but non-CODIS loci would need further discussion.

---

Final Discussion led by Ed Wack (MIT/LL)

The discussion of NGS applications for forensics/human ID was useful. Many of the applications were cross-cutting across agencies.

The agencies were asked to think about the information they were presented with and focus on needs for their specific agency (applications, barriers to adoption, possibility of a steering committee, operation scenarios). In approximately 1.5 months a ‘survey’ or ‘worksheet’ would be sent out.

The idea of a NGS steering committee was met with agreement.

For the time being we will set aside ‘technology limitations’ and focus on our own (specific agency) needs.

There was a general consensus that we would like to see the NGS data compatible with STR (core database loci). The idea that we could obtain a ‘full set’ of marker information on a reference sample and possibly a ‘sub set’ of the full maker set on a casework sample was put forth.

We also follow the work of the clinical/medical NGS community as they will be driving the NGS technology forward.

NIJ handed out a list of their current NGS funded projects for attendees (download 08).

USACIL discussed (download 09) a NGS meeting at the end of May that would involve a panel discussion related to NGS and human identification.