NIST Update: What's new? What's going on?

Erica Romans
Peter Vallone, Becky Steffen, Lisa Borsuk, Kevin Kaseler, Sarah Riman, Katherine Gettings

What's new?

SRM 2391d: PCR-based DNA Profiling Standard

On sale as of July 9, 2019
2300 units were vialled

Developed as a successor to SRM 2391c
SRM 2391c is no longer available for sale at NIST
The expiration date was recently extended by 2 years to Feb. 3, 2022
We will not further extend the expiration date after this date

Characteristics of sequencing noise in DNA profiles

The characterization of sequencing noise and artifacts is critical for assessing thresholds for data interpretation
Developed a framework using statistical tools to systematically interpret the characteristics of single-source DNA profiles generated by targeted sequencing

Population Sample Sequencing

Illumina MiSeq FGx instrument, ForenSeq
- 27 autosomal STRs + 24 YSTR + 7 XSTR + Amel
- 94 HID-SNPs + 56 ancestry SNPs + 22 phenotype SNPs
- 1036 Samples
- Sequenced in batches of 24 or 32
- 41 total sequencing runs in 2016
Population Sample Sequencing

2018 27 auSTR
2018 SE33
2020 7 X-STR

24 Y-STR in progress

Sequencing of mitochondrial genomes

In collaboration with the Armed Forces DNA Identification Laboratory 3,327 mitochondrial genomes were sequenced

656 NIST U.S. population samples were sequenced at NIST on the Illumina MiSeq platform

Protocols optimized to reduce error and allow for ‘platinum quality’ complete mtGenomes

Dataset # of Samples Attempted # of Samples in Finalized Data Passing Removed
COAF 124 112 112 0 82 2 0
COCN 119 112 112 0 52 00
COHS 113 109 109 0 13 00
NTAF 258 256 251 5 1 1 00
NTCN 262 260 258 2 1 00 1
NTHS 139 138 138 0 0 0 1
DSAS 175 169 165 4 3 3 0 0
DSNA 175 171 158 13 1 3 0 0
Total 1365 1327 1303 24 20 14 2 2

FY20 Highlights
Principal coordinate analyses of the eight datasets

Data was submitted to EMPOP

Additional Things Going on…

• Forensics@NIST talks are available:
  - https://www.nist.gov/news-events/events/2020/11/forensics@nist-2020
• Katherine is presenting in a workshop at AAFS February 2021
  - STR Wars: The Rise of Sequencing
• NIST Validation Symposium
  - More information coming soon!
• POC: Robert Ramotowski
What else is going on?

Examination of “Front End” methods in DNA typing

Examination of Front End Methods in DNA Typing

- Problem: Amount of sample loss during the extraction process is unknown for commonly used extraction methods
  - Low extraction efficiency cold result in overall lower sample quantity
  - May fail to yield full STR profiles or minor components in mixtures

Methods for determining extraction efficiency and sample loss vary

DNA Extraction

DNA extraction is the first step after collection in the DNA typing process

- Cell lysate
- Cell lysis: Break open the cell to release the proteins that protect the DNA
- Removal of Lipids: By adding a detergent
- Removal of Proteins: By adding a protease
- DNA Binding: DNA binds to magnetic or silica beads and is washed
- Precipitation of DNA: DNA falls out of solution
- Rehydration of DNA: Adding desired buffer
- Elution of DNA: Into desired buffer

DNA extraction is the first step after collection in the DNA typing process

Purification methods are often used to try to eliminate the presence of additional proteins, lipids, and inhibitors

Relative Extraction Efficiency

- Measuring recovery of DNA compared to another method or technique
  - Comparing current method in a laboratory to a new method to bring online
- Reporting in the literature when comparing methods
  - Full vs. Partial profiles or allele/locus counting
    - Assumes that a full profile is equivalent to 100% efficiency
  - Total peak height measurement
    - Assumes summation of peak heights from known sample input to be the threshold of 100% efficiency

Measurement is made when comparing known samples in the genotyping phase between extraction protocols without accounting for the quantity of material obtained or lost
**Absolute Extraction Efficiency**

- Measures the ratio of the amount of DNA recovered (quantity) to the original amount of DNA (known) after extraction.
- Offers the ability to evaluate individual extraction processes and their efficiency independent of another method.

**Testing Absolute Extraction Efficiency**

Placing a **known amount** of DNA into the extraction process and determining the **amount recovered**.

**Limitations**

- **DNA Source**: Known amounts of DNA as a starting material.
- **Extraction Protocol**: Understanding if different protocols perform differently.

**DNA Sources**

- Phenol Chloroform (Organic)
  - Phenol Chloroform (Organic) is thought as the gold standard.
  - Proteinase K digestion of the cells.
  - Equal volumes of Phenol Chloroform added.
  - Phase lock gel tubes used for promoting separation.
  - DNA was precipitated with Ethanol and resolubilized with TE buffer.
### Extraction Methods: Qiagen Commercial Chemistries

- **QIAamp Spin Columns**
  - Manual method commonly used in forensic DNA laboratories
  - Silica columns for collection of DNA
  - Elution in proprietary buffer
  - Similar to TE

- **Qiagen EZ1 Advanced XL**
  - Robotic purification instrument
  - Cell lysis takes place on the benchtop in a thermonixer
  - Purification with paramagnetic bead collection
  - Elution in TE

### Measurement Method: Droplet Digital PCR

- dPCR is counting accessible amplifiable targets

\[ \lambda = -\ln \left( \frac{N_{\text{neg}}}{N_{\text{tot}}} \right) \]

\( \lambda \) = number of targets per partition

From here, concentration can be calculated.

\[ \lambda = -\ln \left( \frac{12731}{18946} \right) \]

\[ \lambda = 0.3976 \]

\[ \text{DNA in ng/\muL} = 3.301 \times (\lambda / (\text{Dilution} \times \text{Droplet Vol})) \]

\( \lambda = 0.3976 \)

DNA in ng/\muL = 3.301 x (0.3976 / (0.25 \times 0.07349))

DNA = 71.4 ng/\muL

*Source: ZL, Khoo MC, Bessou EL, Tenen B. Anal Bioanal Chem. 2018 May;410(12):2879-2887*
Four DNA input amounts were tested in replicates of 5 for each extraction method.

<table>
<thead>
<tr>
<th>DNA Source</th>
<th>Amount (ng)</th>
<th># of Cells</th>
<th>Uncertainty ± # Cells</th>
<th># of Replicates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracted DNA</td>
<td>50</td>
<td>8,333</td>
<td>833</td>
<td>5 per amount (50 per DNA Source)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>3,333</td>
<td>333</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>1,667</td>
<td>167</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>781</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Cells</td>
<td>38</td>
<td>6,200</td>
<td>620</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>3,125</td>
<td>312</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>1,563</td>
<td>156</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>781</td>
<td>78</td>
<td></td>
</tr>
<tr>
<td>Blood</td>
<td>276</td>
<td>48,000</td>
<td>4,800</td>
<td></td>
</tr>
<tr>
<td></td>
<td>138</td>
<td>23,000</td>
<td>2,300</td>
<td></td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>4,857</td>
<td>485</td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>2,333</td>
<td>233</td>
<td></td>
</tr>
</tbody>
</table>

Robotic purification (EZ1) had the greatest efficiency among the DNA sources.

QIAamp had the greatest variability across sources of DNA.

Organic method (PCI) had the lowest efficiency, but highest reproducibility.

Conclusions

- Different methods yield different efficiencies.
- Consistency across different DNA sources for each method.
- Increase in DNA starting amount showed a trend toward increased efficiency.
• Further experimental data to confirm repeatability of these measurements
• Optimization of extraction processes
  • Incubation time, purification method, addition of carrier RNA, microcon concentration (organic), varying input amounts, etc
• Examination of alternate extraction methods and operators
• Addition of a substrate to the extraction process

Questions?
erica.romos@nist.gov
1-301-975-5107

Acknowledgements
Margaret Kline
David Duewer
Steven Lund