SELECTIVELY ANALYZING AND INTERPRETING DNA FROM MULTIPLE DONORS WITH A FULL SINGLE-CELL STRATEGY

Nidhi C. Sheth
Rutgers University
Camden, NJ
NIJ Forensic Science R&D Symposium
To develop a method by which we can compute the weight of evidence for an admixture using a single-cell pipeline

I. Implement a suspect agonistic clustering method.

II. Combine information from single cells that is relevant to forensic science.
**Single-Cell Vs. Bulk Mixture Pipeline**

**Extraction**

\[ V_{\text{ext}} = V_{\text{PCR}} \]

**Electrophoresis**

- Allele 1
- Allele 2

**PCR**

\[ V_{\text{ext}} \geq V_{\text{PCR}} \]

**Not Sampled**
PRELIMINARY WORK: FOUR DIRECT-TO-PCR EXTRACTION METHODS

Towards developing forensically relevant single-cell pipelines by incorporating direct-to-PCR extraction: compatibility, signal quality, and allele detection

Nidhi Sheth1, Harish Swaminathan2, Amanda J. Gonzalez3, Ken R. Duffy4, Catherine M. Gregor1,2,3

Received: 17 August 2020 / Accepted: 4 January 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH, DE part of Springer Nature 2021

Direct-To-PCR Extraction Treatment

<table>
<thead>
<tr>
<th></th>
<th>DP</th>
<th>FG</th>
<th>LP</th>
<th>PP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Viagen</td>
<td>Direct PCR Lysis Solution</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>forensicGEM® Saliva</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEPArray™ LysePrep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arcturus® PicoPure®</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Freq. Alleles Detected /Cell**

- DP: 0.70
- FG: 0.65
- LP: 0.85
- PP: 0.90

P≤ 10^-4

**β-Value/Cell**

- DP: -0.007
- FG: -0.004
- LP: -0.005
- PP: -0.002

P≤ 10^-4

**Avg. Peak Height /Cell**

- DP: 320
- FG: 306
- LP: 350
- PP: 509
Extraction Treatment Comparison Using DEPAArray

**Extraction Treatment Variations**

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>Higher Pro K and PBS (1X)</td>
</tr>
<tr>
<td>PP2</td>
<td>Lower Pro K and PBS (1X)</td>
</tr>
<tr>
<td>PP3</td>
<td>Lower Pro K and PBS (0.5X)</td>
</tr>
<tr>
<td>PP4</td>
<td>Lower Pro K and PBS (0.25X)</td>
</tr>
</tbody>
</table>

High-quality data from a forensically relevant single-cell pipeline enabled by low PBS and proteinase K concentrations

Nidhi Sheth MS | Ken R. Duffy PhD | Catherine M. Grgicak PhD

![Graphs showing data comparison for different extraction treatments.](image)
High-quality data from a forensically relevant single-cell pipeline enabled by low PBS and proteinase K concentrations

Nidhi Sheth MS | Ken R. Duffy PhD | Catherine M. Grgicak PhD

**Extraction Treatment Variations**

<table>
<thead>
<tr>
<th>Extraction Treatment</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>PP1</td>
<td>Higher Pro K and PBS (1X)</td>
</tr>
<tr>
<td>PP2</td>
<td>Lower Pro K and PBS (1X)</td>
</tr>
<tr>
<td><strong>PP3</strong></td>
<td>Lower Pro K and PBS (0.5X)</td>
</tr>
<tr>
<td>PP4</td>
<td>Lower Pro K and PBS (0.25X)</td>
</tr>
</tbody>
</table>

**Pie charts.** Percentage of profiles exhibiting full (100%), most (50-99%), few (1-49%), and no alleles across heterozygous loci where the known alleles are at least two STR units apart.
**Experimental Conditions**

<table>
<thead>
<tr>
<th>Extraction</th>
<th>DDT Conc. for Sperm Extraction: 1.5mM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amplification, Separation, Analysis</td>
<td>Global Filer™ (30 Cycles), 25 sec injection on 3500 Genetic analyzer, and Osiris(10.3.1) Analysis</td>
</tr>
</tbody>
</table>

- Lower peak height indicates decreased $\beta$-value for the epithelial cell, which could result from DNA damage.
- Moving forward, 1X PBS and Lower ProK will be implemented in the experimental pipeline.
SINGLE-CELL CLUSTERING AND INTERPRETATION STRATEGY

Admixture  Single-cell per tube  Single-cell EPGs  Clustering scEPGs based on similarity

Determining the LR for admixture by averaging LR from individual clusters

\[ \text{WoE} = \log \text{LR}_{avg}(A, s_{true}) \]
**Alternative Clustering Techniques Comparison: Leukocyte**

**Correct Cluster**
- L
- H

**Over Cluster**
- L
- H

**Mis Cluster**
- L
- H

**Model-Based**
- Number of Contributors: 2, 3, 4, 5
- Proportion: 0, 0.5, 1

**Probabilistic**
- Number of Contributors: 2, 3, 4, 5
- Proportion: 0, 0.5, 1

**K-mean**
- Number of Contributors: 2, 3, 4, 5
- Proportion: 0, 0.5, 1

Smallest Contributor: L < 0.2
Contributor: H ≥ 0.2

Model Based Clustering performs better, no misclusters
Model Based Clustering was performed using “mclust” as implemented in R.
Computing The Likelihood Ratio Using EESCIIt (Evidentiary Evaluation of Single Cells)

Snapshot of the EESCIIt software that computes Likelihood Ratio (LR) for single-cells - Each cluster was run in EESCIIt against the true contributor.
Heat maps show that most admixtures resulted in $\log_{10}\text{LR}(C,s_{\text{true}})$ of at least 25, regardless of sample complexity. $C=$cluster

**RESULTS FOR EACH CLUSTER**

<table>
<thead>
<tr>
<th>Leukocyte</th>
<th>Epithelial</th>
<th>Mixed</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="LogLR(C,s) heatmap" /></td>
<td><img src="image" alt="LogLR(C,s) heatmap" /></td>
<td><img src="image" alt="LogLR(C,s) heatmap" /></td>
</tr>
</tbody>
</table>

- **Smallest Contributor** $L < 0.2$
- **Contributor** $H \geq 0.2$

<table>
<thead>
<tr>
<th>Number of Contributors</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Leukocyte</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Epithelial</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mixed</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **Prop. Admixtures** 0.1% 35%

---

**Legend**

- L = Low
- H = High
Model-Based Clustering: Overall LogLR

Heat maps show that most admixtures resulted in $\text{LogLR}_{\text{avg}}(A, s_{\text{true}}) > 5$, regardless of sample complexity and clustering performance. Notably, the highest proportion of $\text{LogLR}_{\text{avg}}(A, s_{\text{true}})$ falls within $[25, 30]$.
ACKNOWLEDGMENTS

This work was partially supported by NIJ2020-R2-CX-0032 awarded by the National Institute of Justice, Office of Justice Programs

The opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not reflect those of the Department of Justice

Collaborators: Dr. Catherine Grgicak
Dr. Desmond S. Lun
Dr. Ken R. Duffy

Students: Dr. Harish Swaminathan
Leah O’Donnell
Madison Mulcahy
Qhawe Bhembe
Amanda Gonzalez

nidhi.sheth@rutgers.edu
www.lftdi.com