GROUPED AND UNGROUPED SINGLE-CELL ELECTROPHEROGRAMS ENABLE PRECISION DNA INTERPRETATION: RELEVANCY AND LEGITIMACY OF SINGLE-CELL FORENSICS

Madison Mulcahey, Leah O'Donnell, Nidhi Sheth, Desmond S. Lun PhD, Ken R. Duffy PhD, Catherine M. Grigac PhD

+Rutgers University, Camden, US; Maynooth University, Ireland

Forensic DNA pipelines are complex systems, where each step influences outcomes of subsequent steps:
- The E in the likelihood ratio, \( LR = \frac{Pr(E|H,N,I)}{Pr(E|R,N,I)} \), is influenced by processing, threshold and high-pass artifact filtering decisions.
- The \( N \) indicates number of contributors (NOC), and is assigned
  - The \( I \) is context, and is assigned

In Single-cell analysis, \( n = 1 \), and \( I = I_{\text{none}} \), simplifying interpretation

Model Based Clustering

- These scatter plots give an overview of the quality of the data, or level of genetic information, inherent in each of the 634 single-cell electrophorograms used to construct 630 single-cell admixtures ranging from 2- to 5- contributors and containing up to 75 cells, where the proportion of the smallest contributor was as low as 3.5%.
- The scatter plots show Log\(_{10}\) LR for each cell against the sum of the peak height (ΣPHt) across the single cell electropherogram. When each cell was run against a true contributor, only 6 of 643 resulted in Log\(_{10}\) LR < 0. These were generally the result of overexpressed stutter. When each sample was run against a false contributor the Log\(_{10}\) LR was always < -40.
- Overall, scEPGs carry sufficient data for single-cell analysis.

Conclusion

1. Model Based Clustering is a credible means by which to cluster scEPGs.
2. Single cell pipelines have forensic relevance since Log\(_{10}\)LR(C|H\(_{\text{true}}\)) ≥ 25 for most clusters, and stays at that value regardless of admixture complexity.
3. Log\(_{10}\)LR(C|H\(_{\text{true}}\)) are robust to overclustering, showing the legitimacy of single cell genetics to the forensic domain.

References


Funding
This work was partially supported by NIDDK R21-DK-069548 and NINDS R01-NS-075143.