Report: ISFG Peter M. Schneider Short Term Fellowship 2024
Fellow: Dr. Catherine M. Grgicak, Rutgers University
Host: Prof. Klaas Slooten, NFI
Date: June 9-13, 2025
Topic: **Evaluations of single-cell data and ground truth knowledge**

Recently, Dr. Grgicak, Prof. Slooten and collaborators completed work demonstrating three architecturally distinct models, generally, produced calibrated WoE for 996 individuated single-cell electropherograms (scEPGs) [1]. Catalyzed by these results, Rutgers advanced single-cell interpretive capacity by developing a single-cell evaluator that automatically clusters scEPGs by virtue of their similarity, assigns WoE for persons of interest to each cluster, and averages those weights. In sharing Rutgers findings with NFI, one aim of the collaboration was to define a cogent evaluative strategy capable of buttressing declarations that single-cell evaluations are legitimate to the forensic domain. To do this we: i) confirmed a broad (enough) mixture test set was constructed; ii) confirmed that the experimental design applied permitted ground truth knowledge be retained throughout the experiment; iii) defined at what level of mixture evaluation ground truth knowledge applies; iv) discussed what diagnostics best demonstrate single-cell data evaluations are ready for translation from research to legal settings.

To accomplish the first, we reviewed the factor and sample space covered. It was agreed that 336 admixtures composed of 2-5 donors, in and out of balance, with minor contributions as low as 2-cells is a reasonable test set. In the second place, we confirmed that a valued experimental design was one that mixed scEPGs from cells that were isolated while the donor was still single source. In this way we have a set of experiments for which we have both a true/false outcome and an assigned WoE. In the third place we asked at what level are true false labels applicable. To establish this, we considered a thought experiment and supposed we collected many cells from known donors. From this set we construct admixtures by sampling a sub-set of *m* scEPGs, which can, therefore, be from any number of donors. To evaluate the mixture of scEPGs we cluster the *m* scEPGs into *r* groups. Given two clustering errors might occur: i) that scEPGs of a donor reside in more than one cluster; and ii) that scEPGs of more than one donor reside in a single cluster, we ask if ground truth knowledge is retained at the cluster or at the admixture level. With a cluster, potentially, holding scEPGs from more than two donors, we have that true/false assignments retain their meaning only at the level of admixture. It is for this reason that in the fourth place we diagnose the potential of single-cell data by way of normalizedWoE (NWoE), which is $\log LR_{avg}$ across clusters divided by $\log(1/RMP_{POI})$. The distribution plot showing $P(NWoE \leq x)$ exhibits that 80% of the time we received 90% of the maximal amount of evidence that could have been obtained for all true donors tested.

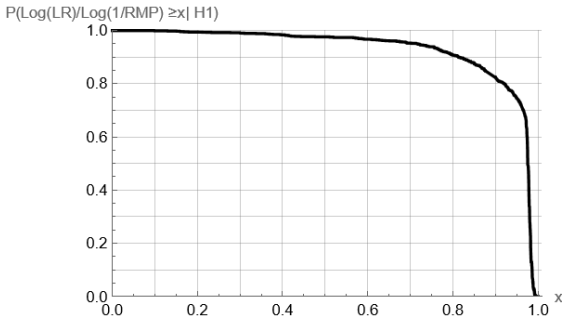


**Figure 1.** Proportion of NWoE returning at least a value of x. For example, the proportion of true donors giving at least 90% of log(1/RMP) is 80%.

We also discussed whether LRs supporting non-contribution should be reported and if so, if this should be accompanied by guidance on further interpretation and consequences for the case. In so doing, we located single-cell datasets that will serve to illuminate guidance on the topic.

Finally, using scDNA as an exemplar, Dr Grgicak presented views to NFI's Biological Traces Department on what features support translation from research to operations by sharing her experience in working within an SLC (Salience, Legitimacy and Credibility) framework [2], and articulated that clear indications of each are advantageous when deciding at what point a novelty should translate from research to operations.

**References**

[1] C. M. Grgicak, K. Slooten, R. G. Cowell, Q. Bhembe, and D. S. Lun, "The (in)dependence of single-cell data inferences on model constructs," (in eng), *Forensic Sci Int Genet,* vol. 76, p. 103220, Jan 3 2025, doi: 10.1016/j.fsigen.2024.103220.

[2] D. Cash, W. Clark, F. Alcock, N. Dickson, N. Selin, and J. Jager, *Salience, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making*. 2002.