OSAC RESEARCH NEEDS ASSESSMENT FORM

Title of research need: Methods in phenotype and ancestry inference

Describe the need:

Forensic phenotype inference is the estimation of externally visible traits using DNA, whereas, forensic ancestry inference is the estimation of the most likely population(s) of origin using DNA. Different subsets of single nucleotide polymorphisms (SNPs) are necessary to generate informative profiles for either application, with some markers overlapping. Phenotype and ancestry inference use a variety of statistical calculations, from probability to likelihood, with estimation systems ranging from regression to Bayesian statistics. Ancestry inference using small sets of ancestry informative markers (AIMS), which can be isolated SNPs or microhaplotypes if phased information is available. Utilizing PCA data reduction techniques with reference populations for comparison makes continental separation possible. However, subpopulation level ancestry or admixture inference can be extremely difficult using such techniques. With phenotype inference, it is possible to estimate some traits with higher accuracies than others (e.g., pigment). However, even more simple traits are still dependent on the number of markers and how well the inference models were built. Performance metrics and adequate comparisons are vital to improve systems and ensure models are built with a wide variety of variation (e.g., multiple populations and individuals with unique admixture proportions). Some traits are more difficult to estimate due to the influence of environmental and lifestyle variables (e.g., height). Further, there is a severe lack of fundamental genetic knowledge to accurately estimate other characteristics such as facial structure, yet there are claims in the field that they are possible. Overall, this application (particularly the reporting and interpretation of the results) is still new, and therefore there are many areas of research and characterization that are needed before forensic DNA inference for phenotype and ancestry can be considered for casework implementation.

Areas of improvement for both include: 1. Continued development and refinement of inference models and system with tools (e.g., bioinformatic pipelines and analysis software) to perform phenotype and/or ancestry inference with currently used traits (e.g., hair pigmentation). 2. Identification and characterization of genotyping methods and inference models for emerging traits beyond pigmentation that show promise but need further development. 3. Assessment of the application of phenotype and ancestry inferences to challenging sample types, such as low-input samples and deconvoluted profiles from mixtures. 4. A comparison of the applications of phenotype and ancestry inference using techniques beyond SNP genotyping such as epigenetics or proteomics.

Across all of these research needs, it is paramount that proper reporting language and guidelines are developed to highlight the investigative nature of these inferences and guide practitioners from multiple disciplines (i.e., Forensic DNA analysts, law enforcement, genealogists, etc.) in understanding what the inference does and does not mean, and considering ethical, privacy, and legal concerns. In addition, training of what is currently available and possible, including caveats and explanation of errors, is needed.

Keyword(s):	SNPs, Massively Parallel Sequencing, Microarray hybridization, bioinformatic pipelines, analysis software, prediction models and systems

Submitting subcommittee(s):	Human Biology	**Date Approved:**	05/16/2025

Background Information:

1. Does this research need address a gap(s) in a current or planned standard? (ex.: Field identification system for on scene opioid detection and confirmation)

Yes, in the use of Forensic DNA phenotyping (appearance and ancestry) for casework.

2. Are you aware of any ongoing research that may address this research need that has not yet been published (e.g., research presented in conference proceedings, studies that you or a colleague have participated in but have yet to be published)?

Yes.

3. Key bibliographic references relating to this research need: (ex.: Toll, L., Standifer, K. M., Massotte, D., eds. (2019). Current Topics in Opioid Research. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88963-180-3)

1. Schneider PM, Prainsack B, Kayser M. The Use of Forensic DNA Phenotyping in Predicting Appearance and Biogeographic Ancestry. *Dtsch Arztebl Int*. 2019;51-52(51-52):873-880.
2. McNevin, D. Forensic inference of biogeographical ancestry from genotype: The Genetic Ancestry Lab. *WIREs Forensic Sci*. 2020; 2:e1356.
3. Kenneth K. Kidd, William C. Speed, Andrew J. Pakstis, Manohar R. Furtado, Rixun Fang, Abeer Madbouly, Martin Maiers, Mridu Middha, Françoise R. Friedlaender, Judith R. Kidd. Progress toward an efficient panel of SNPs for ancestry inference. *Forensic Sci Int Genet*. 2014;10:23-32,
4. Christopher Phillips, Maria de la Puente; The analysis of ancestry with small-scale forensic panels of genetic markers. *Emerg Top Life Sci* 2021; ETLS20200327.
5. de la Puente M, Ruiz-Ramírez J, Ambroa-Conde A, et al. Broadening the Applicability of a Custom Multi-Platform Panel of Microhaplotypes: Bio-Geographical Ancestry Inference and Expanded Reference Data. *Front Genet*. 2020;11:581041.
6. Kayser, M, Branicki, W, Parson, W, & Phillips, C. Recent advances in Forensic DNA Phenotyping of appearance, ancestry and age. *Forensic Science International: Genetics*. 2023;65.
7. Ruiz-Ramírez, J, de la Puente, M, Xavier, C, Ambroa-Conde, A, Álvarez-Dios, J, Freire-Aradas, A, Mosquera-Miguel, A, Ralf, A, Amory, C, Katsara, MA, Khellaf, T, Nothnagel, M, Cheung, EYY, Gross, TE, Schneider, PM, Uacyisrael, J, Oliveira, S, Klautau-Guimarães, MdN, Carvalho-Gontijo, C, Phillips, C. Development and evaluations of the ancestry informative markers of the VISAGE Enhanced Tool for Appearance and Ancestry. *Forensic Science International: Genetics.* 2023;64.

4. Review the annual operational/research needs published by the National Institute of Justice (NIJ) at https://nij.ojp.gov/topics/articles/forensic-science-research-and-development-technology-working-group-operational#latest? Is your research need identified by NIJ?

No.

5. In what ways would the research results improve current laboratory capabilities?

It is important to compare and contrast commercial and non-commercial laboratory SNP genotyping methods used in generating data for forensic DNA phenotype and ancestry inference. Further, it is critical to evaluate the inference models and how the data used to build these models can impact results, along with developing and evaluating performance metrics of the models. Full disclosure on inference models, how they are built, utilized and their output must be mandatory for any approach and direct comparisons of the same samples (a control set) should ideally be implemented.

6. In what ways would the research results improve understanding of the scientific basis for the subcommittee(s)?

Better exploration of the minimum data (e.g., coming from low-input and/or degraded samples) requirements for accurate predictions, from the very first step in data generation, to the final stated inference are needed. Expansion into visual outcomes may benefit the field, especially the performances so that laboratories may decide when it may be time to implement this intelligence method in their cases.

7. In what ways would the research results improve services to the criminal justice system?

It will provide an overview of the current field in terms of commercial and non-commercial kits/systems and highlight what is possible and what is not. It can show the reliability of phenotype and ancestry inference, providing a firm foundation to aid cases with actionable intelligence answers, for past and future successful applications. This could be independent or alongside forensic investigative genetic genealogy to limit potential relationships to explore, suspect pools for law enforcement investigation, and additional STR typing. Finally, intelligence information generated through rapid DNA testing may be useful on the ground in time-sensitive cases.

8. Status assessment (I, II, III, or IV): II

	Major gap in current knowledge	Minor gap in current knowledge
No or limited current research is being conducted	I	III
Existing current research is being conducted	II	IV

This research need has been identified by one or more subcommittees of OSAC and is being provided as an informational resource to the community.