Scientific & Technical Review Panel Final Report for 2022-S-0011 Standards for Construction of Multilocus Databases

Organization of Scientific Area Committees (OSAC) for Forensic Science

STRP Final Report

2022-S-0011

Standards for Construction of Multilocus Databases

Organization of Scientific Area Committees (OSAC) for Forensics Science
April 25, 2022

Disclaimer:

This report was produced by an independent Scientific and Technical Review Panel (STRP). The views expressed in the report do not necessarily reflect the views or policies of the U.S. Government. Visit the OSAC website for more information on OSAC's STRP process.

Scientific & Technical Review Panel Members

- Alyse Bertenthal, Wake Forest University
- Kelly Carrothers, California Department of Fish and Wildlife
- Rachel Houston, Sam Houston State University
- Naomi Kaplan-Damary, The Hebrew University of Jerusalem/CSAFE
- Christina Lindquist, UC Davis School of Veterinary Medicine
- Laurence Mueller, University of California, Irvine
- Dyan Straughan, U.S. Fish and Wildlife Services

Report Summary:

The Scientific and Technical Review Panel (STRP) for “Standards for Construction of Multilocus Databases” is an independent panel appointed by the National Institute of Standards and Technology (NIST). A STRP is established with a range of experts to consider how well a standard meets the needs of the forensic science, law enforcement, and legal communities, and to recommend improvements to the standards under review. The STRP appreciates the efforts of Hope Draheim, Wildlife Forensic Biology Subcommittee member, while serving as the subcommittee liaison to this STRP during the review process.

The STRP began its review process with a kickoff meeting on November 19, 2021 and concluded with this STRP final report. The panel reviewed the draft standard and prepared comments for the Wildlife Forensic Biology Subcommittee.

Report Components:

The STRP reviewed this draft standard against OSAC’s *STRP Instructions for Review* which include the following content areas: scientific and technical merit, human factors, quality assurance, scope and purpose, terminology, method description and reporting results. The details below contain a brief description of each reviewed content area and the STRP’s assessment of how that content was addressed in the draft OSAC Proposed Standard.

1. **Scientific and Technical Merit:** OSAC-approved standards must have strong scientific foundations so that the methods practitioners employ are scientifically valid, and the resulting claims are trustworthy. In addition, standards for methods or interpretation of results must include the expression and communication of the uncertainties in measurements or other results.

    1.1 Consensus View – The STRP finds this standard to be scientifically and technically sound to establish minimum requirements to develop multilocus population genetic databases for wildlife forensics. While this standard includes criteria for the identification of samples; inclusion of associated biological information; choice and evaluation of genetic markers; standard statistical evaluation of the reference database; and evaluation and quality assurance protocols for databases, it is not intended to constitute a standard for what specific validation studies are necessary once databases are constructed and implemented in forensic casework. The STRP believes this standard will be beneficial to the wildlife forensics community.

1.2 Minority View – None.

2. **Human Factors:** All forensic science methods rely on human performance in acquiring, examining, reporting, and testifying to the results. In the examination phase, some standards rely heavily on human judgment, whereas others rely more on properly maintained and calibrated instruments and statistical analysis of data.

2.1. Consensus View – No significant human factor issues were identified by the STRP. Prior to the creation of the STRP, there was communication with human factors experts regarding this standard. They recommended that a population genetics expert serve on the STRP in lieu of a human factors expert. This recommendation was based on the observation that, aside from the need for additional detail and specific requirements to “raise the bar” for these databases, a population genetics expert could provide more focused recommendations for this standard. Ultimately the decision was made to include a population genetics expert.

2.2. Minority View – None

3. **Quality Assurance:** Quality assurance covers a broad range of topics. For example, a method must include quality assurance procedures to ensure that sufficiently similar results will be obtained when the methodology is properly followed by different users in different facilities.

3.1. Consensus View – The STRP finds the various recommendations for quality assurance in sections 4.1.1 through 4.7 to be satisfactory and appropriate for the construction of a multilocus database. The majority of the STRP believes that an important quality assurance consideration is to subject the database to external review by the scientific community. When constructing a multilocus database, the Forensic Service Provider (FSP) should consider the feasibility of making their datasets available; however, it may not be practicable for the FSP to allow public access to certain software and datasets due to law enforcement sensitive information.

We continue to support the recommendation that section 4.5 should read “*Once* *initially validated, databases augmented with new samples or subsetted shall be re-evaluated as in 4.2. The most current version of the database should then be made available for public review.*"

3.2. Minority View – The original suggestion by this panel to make these databases publicly available was declined with the explanation that harm can come since, "*publication of genetic data and metadata may harm the species of interest such as locations with an abundance of endangered and protected species*". The STRP does not view this as a barrier to making data public because the important information in these databases that should be checked or re-evaluated by other scientists is the genetic data. This data can be anonymized as human data often is. Locations of the samples are not needed and can be withheld in publicly available databases. If databases for multiple populations of the same species are collected they can be labelled so as to anonymize the population (e.g., "A", "B", etc.). The STRP suggests the subcommittee reconsider its blanket rejection of this recommendation, especially since the concern about harm to protected species is not implicated by databases that are constructed using populations that are not from endangered species. As elaborated in Appendix 1, the prevailing scientific consensus is that data must be subject to public review.

4. **Scope and Purpose:** Standards should have a short statement of their scope and purpose. They should list the topics that they address and the related topics that they do not address. Requirements, recommendations, or statements of what is permitted or prohibited do not belong in this section.

4.1. Consensus View – No extensive issues were identified by the STRP. The requirements to "have a procedure", absent details of what should be in the procedure is not as helpful or as specific as defining the contents. Since the subcommittee's view is that adding those specifics goes beyond the purview of the charge to write this particular standard, making the limitations of the standard explicit in the scope is prudent. The following wording is recommended to clarify the relationship of the requirements in this standard with other requirements for procedures and protocols:

"This standard specifies that a laboratory should have technical procedures for evaluation of software and databases, database archival and version control, and for promoting transparency through documentation, but it does not fully specify the content that those procedures must have to be fit for those purposes. This standard sets out only procedures for construction of multilocus databases and does not substitute for standards that will more specifically set forth minimum requirements or best practices for the other procedures mentioned here."

4.2. Minority View – None

5. **Terminology:** Standards should define terms that have specialized meanings. Only rarely should they give a highly restricted or specialized meaning to a term in common use among the general public.

5.1. Consensus View – The STRP finds that this draft standard does not adequately define all terminology. Specifically, two of the terms are not accurate as written.

5.1.1. <u>Coefficient of co-ancestry:</u> The current definition of coefficient of co ancestry is the definition of IBD (see OSAC lexicon). The word 'coefficient' means that this definition refers to a number or, in this case, a probability. Evett and Weir point out that the coefficient of co-ancestry is a probability and not simply the concept of IBD here where they say '...the coancestry coefficient θ_{XY}, defined as the *probability* [emphasis added] that two alleles, one taken at random from *X* and *Y*, are IBD' (Evett and Weir, 1998, p. 100). Furthermore, in the widely used textbook by Hartl and Clark (1989, *Principles of Population Genetics*, *Second Edition*, Sinauer, Sunderland, MA) the authors state that '... the coefficient of coancestry, defined as the *probability* [emphasis added] that any two randomly chosen alleles in a population (not necessarily in the same individual) are identical by descent' (Hartl and Clark, 1989, pg. 238).

A suggested definition of co-efficiency of co-ancestry is **'The probability that two alleles randomly sampled from different individuals are copies of the same ancestral allele without mutation, that is, the probability that they are identical by descent**'. Alternatively, the STRP advises that the definition could also be kept but changed to the definition of IBD and the new definition of coefficiency of co-ancestry be 'The probability that two alleles are identical by decent (IBD)'.

5.1.2. <u>Statistical power:</u> The definition is confusing and does not provide the true definition to the reader. The conclusion of the definition that 'the phenomenon exists' is confusing. Is this the phenomenon driving the null hypothesis or the phenomenon driving the alternative hypothesis? Since the 'phenomenon' is not defined, the reader cannot make sense of this definition. Additionally, the current definition does not take into account that the null hypothesis can be incorrectly rejected under a Type 1 error. A relatively straightforward discussion of the ideas presented above is given by Weir (1990, Genetic Data Analysis, Sinauer, Sunderland, MA). Weir states that 'Statistical tests of hypotheses are subject to two kinds of error: a true hypothesis may be rejected or a false hypothesis may not be rejected. The significance level measures the probability of the first kind of error, while POWER is one minus the probability of the second kind. … Powers are calculated for each alternative value of the parameter being tested. ' (Weir, 1990, p. 87). Here Weir points out that statistical power must be defined in relation to a specific alternative to the null hypothesis. Thus, a definition of statistical power must also refer to the existence of a specific alternative hypothesis in an explicit fashion.

A more acceptable definition would be '**The probability of rejecting a null hypothesis given that a specific alternative hypothesis is true.**'

5.2. Minority View – None

6. **Method Description:** There is no rule as to the necessary level of detail in the description of the method. Some parts of the method may be performed in alternative ways without affecting the quality and consistency of the results. Standards should focus on standardizing steps that must be performed consistently across organizations to ensure equivalent results. Alternatively, standards can define specific performance criteria that are required to be demonstrated and met rather than specifying the exact way a task must be done. For example, it may be enough to specify the lower limit for detecting a substance without specifying the equipment or method for achieving this limit of detection.

6.1. Consensus View – The STRP believes the level of detail in the description of the method for constructing multilocus databases is sufficient for promoting consistent application of the standard across wildlife forensic laboratories. The standard makes clear that it is up to each individual laboratory how they want to construct their database (4.1.1, 4.1.3, 4.1.4) and for what purpose (4.3).

6.2. Minority View – The criteria for inclusion of metadata (4.1.4) is overly strict. 4.1.4 should state that all known metadata shall be included in the multilocus database only if a separate reference specimen database is not used to account for the information since both methods are equally valid and how the database is used is at the discretion of the individual laboratory.

7. **Reporting Results:** Methods must not only be well described, scientifically sound, and comprehensive but also lead to reported results that are within the scope of the standard, appropriately caveated, and not overreaching.

7.1. Consensus View – This topic was not applicable to this draft

standard. 7.2. Minority View – None

Appendix 1

Additional applicable references and background

Background Related to Section 3. Quality Assurance

The importance of outside scientific review was recognized early in the history of forensic DNA typing when the 1992 National Academy of Science committee stated that “If scientific evidence is not yet ready for both scientific scrutiny and public re-evaluation by others, it is not yet ready for court." ("DNA Technology in Forensic Science", 1992, National Research Council, pg. 94”.

Additionally major scientific journals now have public availability of data as a primary requirement. The journal, *Science,* requires that “All data used in the analysis must be available to any researcher for purposes of reproducing or extending the analysis. Data must be available in the paper or deposited in a community special-purpose repository or a general-purpose repository such as Dryad (see Data and Code Deposition)” (https://www.science.org/content/page/science-journals-editorial-policies#data-and-materials after-publication). Likewise, *Nature* holds that “An inherent principle of publication is that others should be able to replicate and build upon the authors' published claims. A condition of publication in a Nature Portfolio journal is that authors are required to make materials, data, code, and associated protocols promptly available to readers without undue qualifications” (https://www.nature.com/nature/editorial-policies/reporting-standards).

These standards are also required by Federal Agencies which use taxpayer money to fund research. The National Institutes of Health, states that “The National Institutes of Health (NIH) *Policy for Data Management and Sharing* (herein referred to as the DMS Policy) reinforces NIH's longstanding commitment to making the results and outputs of NIH-funded research available to the public through effective and efficient data management and data sharing practices. **Data sharing enables researchers to rigorously test the validity of research findings,[5] strengthen analyses through combined datasets, reuse hard-to-generate data, and explore new frontiers of discovery** [emphasis added].” (https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-013.html). The emphasized text succinctly summarizes why data sharing is so important.

The National Science Foundation has a similar policy, “Investigators are expected to share with other researchers, at no more than incremental cost and within a reasonable time, the primary data, samples, physical collections and other supporting materials created or gathered in the course of work under NSF grants. Grantees are expected to encourage and facilitate such sharing” (https://www.nsf.gov/pubs/policydocs/pappg20_1/pappg_11.jsp#XID4).

These journals and funding agencies are representing the gold standard today for scientific research. The STRP urges the subcommittee to implement similar standards for construction of multilocus population genetic databases. As the requirements noted above indicate, subjecting data to public review and analysis is foundational to establishing data as a trustworthy source that can be used to make scientific inferences.