OSAC 2025-S-0012
Best Practice Recommendations for Publicly Sharing Short Tandem Repeat Data from Wildlife Panels

Wildlife Forensic Biology Subcommittee
Biology Scientific Area Committee (SAC)
Organization of Scientific Area Committees (OSAC) for Forensic Science

OSAC Proposed Standard

DRAFT OSAC 2025-S-0012
Best Practice Recommendations for Public Sharing Short Term Repeat Data from Wildlife Panels

Prepared by
Wildlife Forensic Biology Subcommittee
Version: 1.0
February 2025

Disclaimer:

This OSAC Proposed Standard was written by the Wildlife Forensic Biology Subcommittee of the Organization of Scientific Area Committees (OSAC) for Forensic Science following a process that includes an open comment period. This Proposed Standard will be submitted to a standard developing organization and is subject to change.

There may be references in an OSAC Proposed Standard to other publications under development by OSAC. The information in the Proposed Standard, and underlying concepts and methodologies, may be used by the forensic-science community before the completion of such companion publications.

Any identification of commercial equipment, instruments, or materials in the Proposed Standard is not a recommendation or endorsement by the U.S. Government and does not imply that the equipment, instruments, or materials are necessarily the best available for the purpose.

To be placed on the OSAC Registry, certain types of standards receive a Scientific and Technical Review (STR). The STR process is vital to OSAC's mission of generating and recognizing scientifically sound standards for producing and interpreting forensic science results. The STR shall provide critical and knowledgeable reviews of draft standards to ensure that the published methods that practitioners employ are scientifically valid, and the resulting claims are trustworthy.

The STR consists of an independent and diverse panel, which may include subject matter experts, human factors scientists, quality assurance personnel, and legal experts as applicable. The selected group is tasked with evaluating the proposed standard based on a defined list of scientific, administrative, and quality assurance based criteria.

For more information about this important process, please visit our website at: https://www.nist.gov/organization-scientific-area-committees-forensic-science/scientific-technical-review-str-process

Foreword

Wildlife forensic laboratories routinely utilize short tandem repeats (STRs) and associated allele frequencies in casework. Public sharing of allele frequencies and metadata is encouraged because it minimizes duplication of efforts, improves standardization across laboratories, and increases transparency. This Best Practice Recommendation document provides guidance on how to make publicly available the allele frequencies and associated metadata from STR panels used in wildlife forensic casework.

Keywords: *short tandem repeats, allele frequencies, publicly available, datasets, wildlife forensics*

Table of Contents

Best Practice Recommendation for Publicly Sharing Short Tandem Repeat Data from Wildlife Panels

1 Scope

This Best Practice Recommendation document provides guidance on how to make publicly available the allele frequencies and associated metadata from short tandem repeat (STR) panels used in wildlife forensic casework; this does not cover data generated from single nucleotide polymorphisms (SNPs) genotyping or Sanger sequencing. It is expected that individuals who are using this best practice document have a working understanding of STR typing, allele frequency calculations, underlying population genetic theory, and taxonomic considerations (e.g., species complexes, evolutionary significant units, and lack of consensus between legal and scientific population/naming/species delineations). Individuals should be cognizant of the species, subspecies, or population(s) from which the data were derived to ensure that they are being used appropriately for the probative question (e.g., California black bear allele frequencies would not be appropriate for individual matching of North Carolina black bears). This Best Practice Recommendation document is specific to organisms encountered in wildlife forensic casework only and does not outline how the publicly shared data can be used.

2 Normative References

The document contains no normative references. See Annex A, Bibliography for other references.

3 Terms and Definitions

For the purposes of this document, the following definitions apply.

3.1

relevant non-practitioners

organization or individual that could either generate or utilize data outside of forensic science services applications

NOTE 1 to entry: This could include academics, conservation agencies, and non-profit or for-profit organizations.

4 Recommendations

4.1 Forensic Science Service Providers (FSSP) and relevant non-practitioners should follow OSAC 2022-S-0011, *Standard for Construction of Multilocus Databases*, to determine which loci and individuals should be included in the database.

4.2 FSSP and relevant non-practitioners should establish the appropriate level and type of data to be shared. Options include but are not limited to:

4.2.1 Aggregate data, which could include allele frequencies and other metrics derived from individuals of the species, subspecies, or population(s) of interest.[1]

4.2.2 Individual data, which could include individual genotypes and other metrics derived from individuals of the species, subspecies, or population(s) of interest.

4.3 FSSP and relevant non-practitioners should consider providing the following accompanying information:

4.3.1 Panel and loci information. If this information is contained in a peer-reviewed scientific journal, the relevant citation can be provided. If not, this information should include but is not limited to:

4.3.1.1 Locus name, chromosome, and start/stop position.[2,3]

4.3.1.2 Primer sequences along with dye used.[4]

4.3.1.3 Repeat motif or structure.

4.3.1.4 PCR reaction and cycling conditions and expected interpretation thresholds.

4.3.1.5 If genotyping via capillary electrophoresis, the polymer and array length used for genotyping.[5]

4.3.1.6 Example electropherogram from a high-quality sample (e.g., sole-source sample from blood, buccal, or tissue).

4.3.1.7 Observed alleles and base pair size range.

4.3.1.8 Observed common artifacts for each locus (stutter, minusA, etc.), if applicable.

4.3.1.9 Unique attributes for certain loci or species/subspecies (e.g., in certain panels species-specific alleles might occur and be noteworthy).

4.3.1.10 Naming of alleles (i.e., based on size or number of repeats).

---

[1] Only one individual from either a known parent-offspring or full-sibling relationship should be included in aggregate where possible. The inclusion of such individuals may bias allele and genotype frequencies.

[2] For some species without a reference genome, identifying the chromosome and start/stop position will not be possible.

[3] Where possible, provide the NCBI accession number of the reference genome (typically referred to as "NCBI RefSeq assembly") or contig used.

[4] It can be informative to specify which species the primers were derived from and potential other related taxa that the primers may also work with.

[5] Allele size can vary for the same sample if separated via capillary electrophoresis under different conditions (e.g., polymer, array length, dye, internal lane standards, injection time and voltage).

4.3.1.11 Method used to generate the genotypes.[6]

4.3.2 Sample details and associated metadata for individuals used to generate allele frequencies. FSSP and non-practitioners should consult OSAC-2022-S-0011, *Standard for the Construction of Multilocus Databases,* Section 4.1.4, as it provides details of relevant metadata that should be documented for samples included in multilocus databases.[7]

4.4 FSSP and relevant non-practitioners are encouraged to make the dataset(s) utilized in casework publicly available via a public-facing website with minimal barriers for access (e.g., no username, no fees) or via peer-reviewed scientific journal. In instances where the underlying data are included in a non-open access, peer-reviewed scientific journal, data should also be separately shared via a public-facing website or repository (e.g., DRYAD, FigShare).[8]

4.5 Publicly shared data should be updated at the discretion of the FSSP and relevant non-practitioners. It is not suggested that this be completed at predefined time intervals (e.g., yearly), as FSSPs rarely update allele frequencies used in casework in this manner. It is suggested that version numbers (or date last updated) be linked to each dataset and provided in case notes. The rationale behind updating the publicly shared data should be annotated. Considerations of when to update publicly shared data include but are not limited to:

4.5.1 The FSSP updates the allele frequencies used for forensic calculations on case samples.[9]

4.5.2 Individuals that represent new or previously under-sampled locations or populations are acquired and genotyped.

4.5.3 A substantial increase (determined as a proportion of known individuals in that species, subspecies, or population(s)) in the number of reference or non-probative samples has been acquired.

4.5.4 The panel utilized for case samples is modified (e.g., removal or addition of loci, modification of primer sequences causing alterations to allele frequency data).

4.5.5 Reassignment of individuals to new or different species, population, sub-population or taxonomic units.

---

[6] Massively parallel sequencing (MPS) would allow for the detection of isoalleles and SNPs in flanking sequencing.

[7] In some instances, collectors, date of collection, and/or location of samples cannot be provided due to law enforcement sensitivities, policy, or organizational legal limitations.

[8] It is only expected that FSSP and relevant non-practitioners make datasets that are currently in use publicly available.

[9] FSSP should also publicly archive older allele frequencies, so that analyses could be repeated with the relevant dataset if required.

4.5.6 Allele frequency changes due to temporal variation (e.g., bottlenecks, translocation, increased migration, etc.).

4.5.7 Genotyping is completed on a different platform.

4.6 End users of publicly available datasets should ensure that they archive a copy of the accessed version.

Annex A

(informative)

This Best Practice Recommendation document is in alignment with the trend towards transparency in forensic science. Additionally, the open sharing of genetic marker data amongst FSSPs and relevant non-practitioners should both reduce the duplication of effort between laboratories that are performing similar genetic investigations and support the standardization of genetic marker panels in the wildlife forensic community.

The specific recommendations for the sharing of genetic marker information are largely reflective of the types of data presented in peer-reviewed marker characterization (e.g., primer notes) and developmental validation publications. Open access to such marker information is not intended to negate or replace the need for proper laboratory-specific validation studies, but rather to facilitate more effective information sharing and standardization across laboratories.

It is acknowledged that there are many different approaches that could be taken to make genetic marker data publicly available. Individual laboratories and agencies may have existing internal policies about making their data publicly accessible with varying levels of professional freedom for selecting the venue for data sharing (e.g., Github, Dryad, laboratory and agency websites). Hence this document is a “Best Practice Recommendation” and thus not overly prescriptive nor required by FSSP and relevant non-practitioners who conduct STR genetic analysis in wildlife forensics.

Annex B

(informative)

Bibliography

[1] Chin, J.M, Ribero, G., Raiden, A. “Open forensic science.” *Journal of Law and the Biosciences*, 2019; 255-288.

[2] Chin, J.M., Ibaviosa, C.M. “Beyond CSI: Calibrating public beliefs about the reliability of forensic science through openness and transparency.” *Science and Justice,* 2022. 62: 272-281.

[3] Hamlin, B.C., Meredith, E.P., Rodzen, J., Strand, J.M. “OdoPlex: An STR multiplex panel optimized and validated for forensic identification and sex determination of North American mule deer (*Odocoileus hemionus*) and white-tailed deer (*Odocoileus virginianus*).” *Forensic Science International: Animals and Environments,* 2021. 1: 100026.

[4] Meredith, E.P., Adkins, J.K., Rodzen, J.A. “UrsaPlex: An STR multiplex for forensic identification of North American black bear (*Ursus americanus*).” *Forensic Science International: Genetics,* 2020; 44: 102161.

[5] Wostenberg, D.J., Burnham-Curtis, M.K. “The development of multiplex STR panels for the identification of bald eagles (*Haliaeetus leucocephalus*) and golden eagles (*Aquila chrysaetos*).” *Forensic Science International: Animals and Environments,* 2023; 3: 100062.