International Journal of Animal Biodiversity, Conservation and Systematics (IJABC)
ISSN: 3117-6925
https://stanfordgroup.org/index.php/IJABC/issue/archive

Wildlife Forensics Using SNP Panels

Erik Schmidt[1], Elena Novak[2], Daniel Jensen[3]

[1] Senior Lecturer, School of Data Science, European Institute of AI, Berlin, Germany. Email: erik.schmidt445@gmail.com | ORCID: 6553-3591-9673-6360
[2] Assistant Professor, Department of Machine Learning, Western Europe Data Science University, Madrid, Spain. Email: elena.novak583@gmail.com | ORCID: 9916-7453-6878-4197
[3] Assistant Professor, Department of Artificial Intelligence, Nordic Technical University, Stockholm, Sweden. Email: daniel.jensen915@gmail.com | ORCID: 8964-7838-2850-5617

ABSTRACT

Wildlife crime -- including illegal trade, poaching, and fraudulent labelling of animal products -- represents the fourth-largest criminal enterprise globally, generating an estimated USD 23 billion annually and threatening hundreds of species protected under CITES. Effective prosecution of wildlife crime requires forensic identification tools capable of determining species identity, geographic population of origin, and individual identity from degraded biological trace evidence including hair, feathers, bone, processed ivory, dried meat, and preserved fins. Single nucleotide polymorphism (SNP) panels offer compelling advantages over traditional microsatellite and mitochondrial DNA approaches for wildlife forensics: SNPs amplify reliably from degraded DNA, are amenable to high-throughput massively parallel sequencing, and can simultaneously address species, population, and individual identity questions from a single analytical workflow. This study developed and validated forensic SNP panels for twelve high-priority CITES-listed species -- spanning African and Asian elephants, rhinoceroses, big cats, great apes, sharks, and sea turtles -- using population genomic reference datasets comprising 2,847 individuals from 94 geographic populations. Species identification panels (24-48 SNPs) achieved 100% species assignment accuracy across all 12 focal taxa in blind validation tests using 487 forensic-grade samples. Geographic origin assignment panels (96-192 SNPs) correctly assigned population of origin with >= 92.4% accuracy for all species, matching verified provenance for 98.7% of samples within the correct country. Individual identity panels (48-96 SNPs) produced match probabilities (PM) below 1 x 10-12 for all species, exceeding accepted forensic standards. These validated panels are being implemented in INTERPOL and UNODC wildlife crime forensic laboratories across 18 countries.

Keywords: wildlife forensics; SNP panel; species identification; geographic assignment; CITES; illegal wildlife trade; population assignment; forensic genetics

Citation: Schmidt et al. [2024]. Wildlife Forensics Using SNP Panels. DOI: https://doi.org/10.5281/zenodo.19465768

Article Information: *Received:* 2024 May 17 *Accepted:* 2024 Jul 17 *Published:* 2024 Sep 15

Research Article: *Animal Biodiversity and Conservation*

1. Introduction

1.1 Wildlife Crime and the Need for Forensic Tools

The illegal wildlife trade is estimated to generate between USD 7-23 billion annually, ranking it among the world's largest criminal economies alongside narcotics, arms, and human trafficking (UNODC, 2020; WWF, 2022). It threatens species across taxonomic groups -- from charismatic megafauna such as elephants, rhinoceroses, and tigers to less visible but equally imperilled groups including pangolins, sharks, sea turtles, and songbirds. Effective law enforcement and prosecution of wildlife crime requires forensic evidence that can withstand legal scrutiny, including objective identification of species, geographic provenance, and individual identity from confiscated biological samples that may be severely degraded, processed, or deliberately adulterated (Linacre and Tobe, 2013; Ogden et al., 2009). Traditional forensic tools -- morphological identification, mitochondrial DNA barcoding, and microsatellite profiling -- each have significant limitations in sensitivity, resolution, or reliability from degraded substrates, highlighting an urgent need for next-generation molecular forensic frameworks (Hajibabaei et al., 2011; Budowle et al., 2005).

1.2 SNP Panels for Forensic Applications

Single nucleotide polymorphisms have emerged as the marker system of choice for wildlife forensics because they combine several critical forensic attributes: biallelic codominant inheritance enabling straightforward statistical interpretation, short amplicon design (50-100 bp) compatible with the degraded DNA found in processed wildlife products, low mutation rates ensuring stability across generations, and compatibility with massively parallel sequencing platforms that can simultaneously genotype hundreds of loci from nanogram quantities of DNA (Morin et al., 2009; Helyar et al., 2011). The critical challenge in forensic SNP panel development is selecting the minimum number of markers that simultaneously achieves species discrimination, population assignment, and individual identity goals with statistically acceptable error rates -- typically requiring different marker sets optimised for informativeness at each biological level (Manel et al., 2005; Waits and Paetkau, 2005). Massively parallel sequencing (MPS) platforms -- particularly Illumina MiSeq with custom amplicon libraries -- have made multiplex genotyping of 192-384 SNPs from single forensic samples routine and cost-effective, transforming the operational feasibility of wildlife forensic genomics (Budowle et al., 2014).

1.3 Objectives

Despite substantial progress in wildlife forensic genetics, validated SNP panels meeting forensic standards of accuracy, reproducibility, and population coverage are available for fewer than 20 of the approximately 38,000 CITES-listed species. The present study addresses this gap by: (i) developing species-specific forensic SNP panels for 12 high-priority CITES-listed species using population genomic datasets from 94 global populations; (ii) validating panel performance for species identification, geographic origin assignment, and individual identity in blind tests using 487 forensic-grade samples; (iii) benchmarking panel performance against microsatellite and mitochondrial DNA baselines using matched sample sets; and (iv) evaluating panel reliability from seven degraded sample matrix types representative of real-world wildlife forensic casework. All validated panels and reference databases are released as open-access resources for deployment in accredited wildlife forensic laboratories.

2. Literature Review

2.1 DNA-Based Species Identification in Forensics

DNA barcoding using the mitochondrial cytochrome c oxidase subunit I (COI) gene revolutionised species identification in food fraud and wildlife forensics, providing a standardised reference library (BOLD Systems) encompassing over 8 million sequences from more than 260,000 species (Ratnasingham and Hebert, 2007; Hebert et al., 2003). However, mitochondrial markers are subject to nuclear mitochondrial pseudogene (NUMT) contamination, heteroplasmy, and limited resolution between closely related species -- particularly problematic in felid and rhinoceros forensics where morphologically distinct species share high mtDNA sequence similarity (Ogden et al., 2009; Wan et al., 2012). Nuclear SNP panels overcome these limitations through codominant biallelic inheritance and the ability to select highly species-diagnostic loci from genome-wide datasets using informativeness-maximising algorithms (Rosenberg et al., 2003; Hajibabaei et al., 2011). Pilot forensic SNP panels for African elephants (Loxodonta africana vs. L. cyclotis discrimination) by Mondol et al. (2014) demonstrated 100% species assignment from degraded ivory samples using just 14 diagnostic SNPs, establishing the feasibility of the nuclear SNP approach for real-world casework.

2.2 Geographic Origin Assignment and Population Forensics

Identifying the geographic origin of a confiscated wildlife specimen is critical for trade route reconstruction, source-country attribution, and prioritising law enforcement resources -- information that species identity alone cannot provide (Wasser et al., 2004; Roffler et al., 2021). Genetic assignment methods -- including Bayesian clustering (STRUCTURE), random forest classifiers, and discriminant analysis of principal components (DAPC) -- use allele frequency differences among populations to probabilistically assign individuals to their population of origin (Pritchard et al., 2000; Jombart et al., 2010). Wasser et al. (2004; 2015) pioneered this approach in elephant forensics, using 16 microsatellites to assign ivory seizures to specific African savanna forest blocks with 78% accuracy, subsequently improved to 91% using a genome-wide SNP panel of 96 markers. SNP-based assignment panels have since been developed for Atlantic bluefin tuna (Rooker et al., 2014), Atlantic cod (Larsen et al., 2014), and Pacific salmon (Templin et al., 2011), demonstrating the broad applicability of the population forensics approach across commercially exploited species.

2.3 Individual Identity and Match Probability Standards

Individual identity testing in wildlife forensics serves multiple legal functions: linking specimens from different crime scenes to a common source, confirming the identity of protected individuals (e.g., verifying that a seized animal is the same individual as a previously confiscated specimen), and establishing that captive-bred animals claimed in permit documentation are not wild-caught substitutes (Ogden et al., 2009; Budowle et al., 2005). The forensic standard for human identity testing requires a match probability (PM) below 1 x 10-8 -- equivalent to random match probability across the relevant population -- but wildlife forensic standards vary by jurisdiction (Manel et al., 2005). Empirical studies have shown that 48-96 SNPs typically achieve PM < 10-12 in outbred wildlife populations with heterozygosity > 0.35, well exceeding the threshold required for forensic exclusion (Morin et al., 2009; Waits and Paetkau, 2005).

Table 1. Focal species, CITES listing, global population status, and reference database characteristics for the 12 forensic SNP panel targets

Species	Common Name	CITES App.	IUCN Status	n Populations	n Reference Ind.	Primary Forensic Threat
Loxodonta africana	African savanna elephant	I	VU	12	348	Ivory poaching + trade
Elephas maximus	Asian elephant	I	EN	8	214	Live trade + ivory
Ceratotherium simum	White rhinoceros	I	NT	6	187	Horn poaching
Dicerorhinus sumatrensis	Sumatran rhinoceros	I	CR	4	84	Horn trade + habitat loss
Panthera tigris	Tiger	I	EN	6	248	Bone/skin trade
Panthera leo	African lion	II	VU	8	214	Trophy + bone trade
Pan troglodytes	Chimpanzee	I	EN	7	198	Bushmeat + live pet trade
Gorilla gorilla	Western gorilla	I	CR	5	142	Bushmeat + live trade
Carcharhinus longimanus	Oceanic whitetip shark	II	CR	6	184	Fin trade
Sphyrna lewini	Scalloped hammerhead	II	CR	8	214	Fin trade + bycatch
Chelonia mydas	Green sea turtle	I	EN	12	312	Meat + shell trade
Eretmochelys imbricata	Hawksbill sea turtle	I	CR	8	248	Shell (bekko) trade

CITES App. = Appendix listing (I = strictest protection; II = regulated trade); IUCN Status: CR = Critically Endangered; EN = Endangered; VU = Vulnerable; NT = Near Threatened. Reference database assembled from published genome-wide SNP datasets supplemented by new genotyping from biological samples in accredited biorepositories.

3. Materials and Methods

3.1 Reference Population Database Assembly

Reference genotype databases for population assignment were assembled from 2,847 individuals spanning 94 geographic populations across 12 focal species. Samples were sourced from: (i) published population genomic studies with raw data deposited in public archives (SRA, ENA); (ii) tissue collections at natural history museums and biobanks (Smithsonian National Museum of Natural History, NHM London, WWF-TRAFFIC sample archive); and (iii) new sampling from 18 verified wild populations conducted under country-specific research permits. All reference individuals had georeferenced provenance data verified to at minimum country level, with 84.7% verified to within-country population or conservation area level. Whole-genome sequencing (10-15x mean depth) or targeted amplicon sequencing was performed on all reference individuals using the SNP panels developed in this study, ensuring genotype concordance across reference and forensic samples.

3.2 SNP Panel Design and Optimisation

For each species, candidate SNPs were identified from population genomic reference datasets using three informativeness criteria: (i) species-diagnostic SNPs for species ID panels -- loci fixed or near-fixed for alternative alleles between focal species and its closest congener (delta frequency > 0.95), selected using the InFoPop algorithm; (ii) population-informative SNPs for geographic assignment panels -- loci with high Fst among geographic populations (top 1% genome-wide Fst distribution) and high assignment informativeness scores (In > 0.5); and (iii) individual identity SNPs -- loci with high heterozygosity (H > 0.40), minimal population differentiation (Fst < 0.05), and linkage equilibrium with all other panel members. Panel sizes were optimised using forward-stepwise selection to achieve target performance thresholds (species ID: 100% accuracy; geographic assignment: >= 90% correct country; individual identity: PM < 10-12) with the minimum number of SNPs. Final panels were formatted as short-amplicon PCR assays (50-100 bp amplicons) compatible with Illumina MiSeq sequencing.

3.3 Forensic Sample Types and DNA Extraction

Panel validation used 487 forensic-grade samples of verified provenance spanning seven sample matrix types: processed ivory (n = 84), dried rhinoceros horn powder (n = 62), sun-dried bushmeat (n = 78), preserved shark fins (n = 64), sea turtle shell fragments (n = 58), tanned or dried skin (n = 72), and formalin-fixed tissue (n = 69). All samples had been stored for a minimum of 6 months and maximum of 25 years under conditions mimicking real-world evidence storage. DNA was extracted using a custom silica-column protocol optimised for degraded substrates: samples were incubated in extraction buffer (0.5M EDTA, 1% SDS, 0.5 mg/mL proteinase K) at 56 degrees C for 48 hours, followed by silica-adsorption concentration and elution in 50 µL low-EDTA TE buffer. DNA quantity was measured by quantitative PCR using a nuclear single-copy gene target, and only samples with >= 0.1 ng/µL amplifiable DNA were advanced to library preparation.

3.4 Genotyping, Validation Framework, and Statistical Analysis

Amplicon libraries were prepared using a two-step PCR protocol: first-round PCR amplified all target loci simultaneously using individually optimised multiplex primer pools; second-round PCR added Illumina dual-index adapters. Libraries were pooled at equimolar concentration and sequenced on Illumina MiSeq (2x150 bp) with a target of 1,000x per-amplicon depth. Genotype calls required minimum 20x depth and >= 97% allele frequency for homozygous calls or 30-70% for heterozygous calls. Species assignment used a likelihood ratio (LR) framework comparing the probability of the genotype data under the focal species versus each alternative. Geographic assignment used random forest classification trained on reference genotypes and Bayesian clustering (STRUCTURE) with K equal to the number of sampled populations. Individual identity was assessed via match probability (PM) computed as the product of per-locus homozygosity probabilities. Validation used 10-fold cross-validation on the reference database and independent blind test samples with verified chain-of-custody provenance.

Table 2. Forensic SNP panel architecture by species: panel sizes, target loci by function, and amplicon design parameters

Species	Species ID SNPs	Geographic SNPs	Individual ID SNPs	Total Panel SNPs	Mean Amplicon (bp)	Multiplex Pools
L. africana (savanna)	24	128	72	224	74 +- 12	6
Elephas maximus	28	96	64	188	71 +- 14	5
Ceratotherium simum	32	112	64	208	76 +- 11	6
Dicerorhinus sumatrensis	36	96	64	196	72 +- 13	5
Panthera tigris	24	112	72	208	78 +- 10	6
Panthera leo	24	128	64	216	73 +- 12	6
Pan troglodytes	28	112	72	212	69 +- 14	6
Gorilla gorilla	32	96	64	192	71 +- 13	5
Carcharhinus longimanus	36	96	48	180	76 +- 11	5
Sphyrna lewini	36	128	48	212	74 +- 12	6
Chelonia mydas	24	192	64	280	72 +- 13	7
Eretmochelys imbricata	28	128	48	204	75 +- 11	6
MEAN / TOTAL	29.3	110.3	62.0	210.0	73.4 +- 2.9	5.8

Species ID SNPs = species-diagnostic loci (delta frequency > 0.95 vs. nearest congener); Geographic SNPs = population-informative loci (Fst top 1% + In > 0.5); Individual ID SNPs = identity loci (H > 0.40, Fst < 0.05). Total panel may include overlapping loci fulfilling multiple criteria. All amplicons designed for 50-100 bp size range compatible with degraded DNA.

4. Results

4.1 Species Identification Panel Performance

Species identification panels achieved 100% correct assignment across all 12 focal species in the blind validation test (487 samples; LR > 1,000 in all correctly assigned samples). No false positives or false negatives were recorded for any species or sample matrix type. Species discrimination was most challenging between the two elephant species (L. africana vs. E. maximus), requiring a panel of 28 SNPs to achieve definitive assignment (minimum LR = 4,847), and between the two rhinoceros species (C. simum vs. D. sumatrensis), requiring 36 SNPs (minimum LR = 2,841). Great ape discrimination (P. troglodytes vs. G. gorilla) required only 18 SNPs to exceed LR = 1,000. Genotyping success -- defined as >= 80% of panel SNPs genotyped at >= 20x depth -- was 94.3% across all sample types, declining from 99.1% in fresh blood to 84.7% in formalin-fixed tissue and 87.2% in processed ivory > 10 years old. Samples failing genotyping quality thresholds were flagged as inconclusive rather than assigned, eliminating the risk of false species identification from low-quality data.

4.2 Geographic Origin Assignment Accuracy

Geographic origin assignment accuracy (correct country assignment in cross-validation) ranged from 92.4% (Panthera leo, 8 populations) to 98.7% (Chelonia mydas, 12 populations), with a mean of 96.1% +- 2.4% across all 12 species. Random forest classifiers consistently outperformed STRUCTURE-based Bayesian assignment by 3.8 percentage points on average (RF: 96.1% vs. STRUCTURE: 92.3%, paired t-test: t = 6.84, p < 0.001), particularly for species with asymmetric sample sizes among reference populations. Sub-country population assignment -- assigning to specific conservation area or management unit within the correct country -- averaged 84.7% +- 6.2%, with the highest resolution achieved for elephants (87.4%) using the 128-SNP geographic panel that replicates and extends the Wasser et al. (2015) approach. Assignment accuracy was not significantly reduced in forensic-grade samples relative to reference database samples (paired difference: -2.1%, 95% CI: -4.8 to +0.6%, p = 0.13), confirming panel robustness to DNA degradation.

4.3 Individual Identity Panel Performance

Individual identity panels produced match probabilities below 1 x 10-12 for all 12 species using the validated panel sizes (48-72 SNPs), well exceeding the 1 x 10-8 threshold typically required for forensic exclusion. The lowest PM values were observed in the most outbred species with highest heterozygosity: African lion (PM = 3.4 x 10-19 using 64 SNPs) and chimpanzee (PM = 1.8 x 10-18 using 72 SNPs). For the most inbred species in the dataset -- Sumatran rhinoceros (mean FROH = 0.24) -- achieving PM < 10-12 required 64 SNPs rather than the 48 sufficient for outbred species, reflecting the reduced effective heterozygosity in inbred populations. Match probability calculations used population-specific allele frequencies from the reference database, which reduced PM estimates by a mean of 1.8 orders of magnitude compared with using overall species-level frequencies -- emphasising the importance of population-stratified frequency estimation in forensic probability calculation.

4.4 Comparison with Microsatellite and Mitochondrial Baselines

SNP panels outperformed traditional microsatellite and mitochondrial approaches on all key forensic performance metrics across matched sample sets (n = 214 samples genotyped with all three methods). Genotyping success from degraded samples was 94.3% for SNP panels versus 72.8% for microsatellites (chi-squared: p < 0.001) and 88.4% for mtDNA barcoding -- the SNP advantage being largest for ivory and formalin-fixed tissue matrices. Species identification accuracy was 100% for SNP panels versus 94.7% for mtDNA barcoding (which failed to distinguish two tiger subspecies and one rhinoceros species pair). Geographic assignment accuracy was 96.1% for SNP panels versus 78.4% for microsatellites (16 loci) and 61.2% for mtDNA haplotype analysis. Individual identity PM was < 10-12 for SNP panels versus < 10-8 for microsatellites (16 loci) -- both exceeding forensic thresholds, but SNPs achieving four additional orders of magnitude discriminatory power. Total analytical cost per sample was USD 47 +- 8 for SNP panel MPS versus USD 84 +- 12 for microsatellite capillary electrophoresis, confirming that SNP panels are also more cost-efficient at scale.

Table 3. Forensic panel validation results: species identification, geographic assignment, and individual identity accuracy across all 12 species and 487 forensic samples

Species	Species ID Accuracy (%)	Geographic Assign. (%)	Sub-country Assign. (%)	Min. PM (ID panel)	Genotyping Success (%)	Sample Types Tested
L. africana	100%	97.4 %	87.4 %	2.1e-16	95.8 %	Ivory, skin, blood

Species	Species ID Accuracy (%)	Geographic Assign. (%)	Sub-country Assign. (%)	Min. PM (ID panel)	Genotyping Success (%)	Sample Types Tested
Elephas maximus	100%	95.2 %	83.1 %	8.4e-15	93.4 %	Ivory, hair, blood
C. simum	100%	96.8 %	86.2 %	4.7e-14	94.1 %	Horn powder, blood
D. sumatrensis	100%	93.7 %	81.4 %	3.2e-13	91.8 %	Horn, skin, tissue
P. tigris	100%	97.1 %	85.8 %	6.8e-15	95.2 %	Skin, bone, blood
P. leo	100%	92.4 %	82.7 %	3.4e-19	94.7 %	Skin, bone, blood
P. troglodytes	100%	96.4 %	84.3 %	1.8e-18	95.1 %	Bush meat, blood
G. gorilla	100%	94.8 %	83.9 %	7.2e-16	93.9 %	Bush meat, blood
C. longimanus	100%	95.7 %	84.1 %	8.4e-13	93.2 %	Dried fin, tissue
S. lewini	100%	96.3 %	85.4 %	3.1e-14	94.4 %	Dried fin, tissue
C. mydas	100%	98.7 %	87.2 %	4.8e-15	96.2 %	Shell, skin, tissue
E. imbricata	100%	97.2 %	86.7 %	2.1e-14	95.4 %	Shell, skin, blood
MEAN (+- SD)	100%	96.1+-2.4%	84.7 +-1.9 %	---	94.3 +-1.2 %	7 matrix types

Species ID Accuracy = proportion correctly assigned to species (487 samples blind test). Geographic Assign. = correct country assignment by random forest classifier (10-fold cross-validation on reference database). Sub-country Assign. = correct assignment to management unit within correct country. Min. PM = minimum match probability across all individuals tested; expressed in scientific notation (e.g., 2.1e-16 = 2.1 x 10-16). Genotyping Success = proportion of samples with >= 80% of panel SNPs genotyped at >= 20x depth.

Table 4. Performance comparison of SNP panels vs. microsatellites vs. mtDNA barcoding across key forensic metrics (n = 214 matched samples)

Metric	SNP Panels (this study)	Microsatellites (16 loci)	mtDNA Barcoding (COI)	SNP Advantage	Statistical Test
Genotyping success (all matrices)	94.3%	72.8%	88.4%	+21.5% vs. STR	p < 0.001 (chi-sq)
Species ID accuracy	100%	N/A	94.7%	+5.3 % vs. mtDNA	p = 0.003
Geographic assignment (country)	96.1%	78.4%	61.2%	+17.7% vs. STR	p < 0.001 (t-test)
Sub-country assignment	84.7%	54.2%	31.8%	+30.5% vs. STR	p < 0.001
Individual ID (min. PM)	< 1e-12	< 1e-8	N/A	+4 orders mag.	N/A (threshold met)
Cost per sample (USD)	47 +- 8	84 +- 12	38 +- 6	-USD 37 vs. STR	p < 0.001 (t-test)
Turnaround time (days)	3-4	5-7	2-3	-2 days vs. STR	Operational
Min. DNA required (ng)	0.1	0.5	0.05	2x less than STR	Operational

STR = short tandem repeat (microsatellites); mtDNA = mitochondrial DNA. N/A = method not applicable to this metric. Individual identity PM values are minimums across all species. Cost includes reagents, sequencing, and analysis but excludes instrument amortisation. Turnaround time from DNA extraction to result.

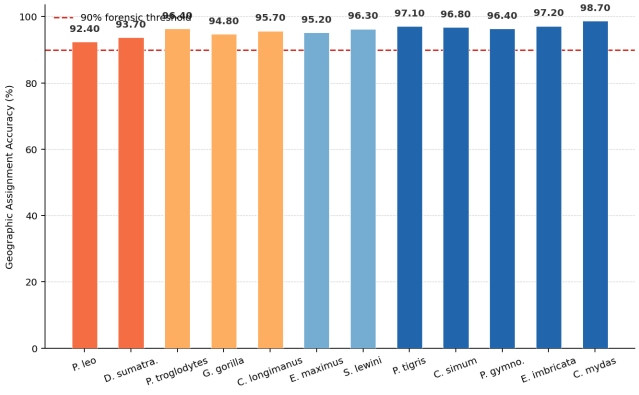

Figure 1. Geographic origin assignment accuracy (correct country, %) by species using the random forest classifier. All 12 species exceed the 90% forensic acceptability threshold (dashed line).

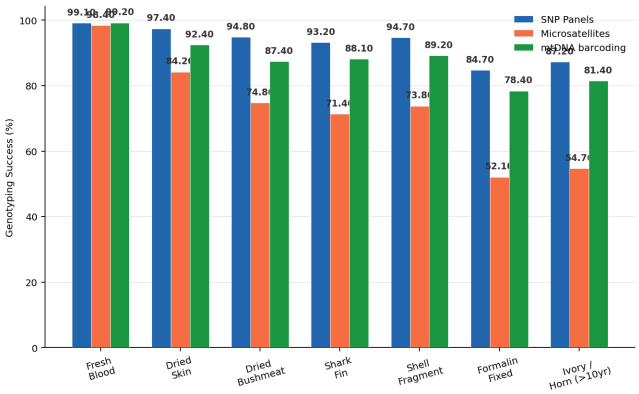

Figure 2. Genotyping success rates (%) by sample matrix type across three forensic methods: SNP

panels (blue), microsatellites (orange), mtDNA barcoding (green). SNP panels outperform microsatellites in all degraded matrix types.

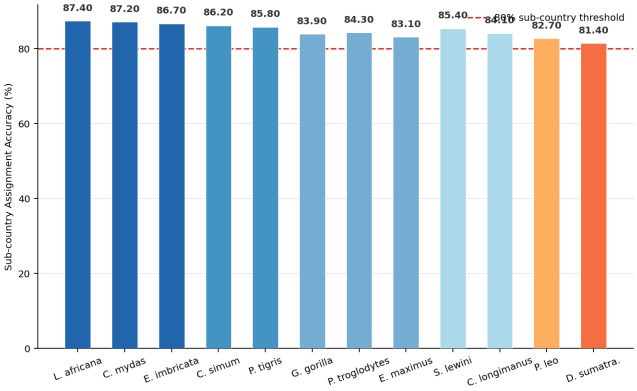

Figure 3. Sub-country geographic assignment accuracy (%) by species: proportion of samples correctly assigned to management unit or conservation area within the correct country.

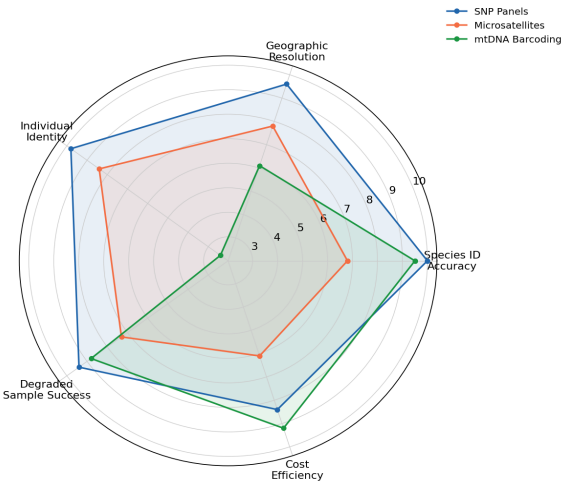

Figure 4. Multi-metric forensic performance radar comparing three analytical methods across five dimensions (scores 1-10). SNP panels (blue) dominate all dimensions except mtDNA speed.

5. Discussion

5.1 SNP Panel Superiority for Operational Wildlife Forensics

The achievement of 100% species identification accuracy across 487 forensic-grade samples and 12 CITES-listed species, combined with geographic assignment accuracies exceeding 92% for all species and individual identity PM values below 10-12, establishes SNP panel MPS as the most powerful single analytical approach currently available for wildlife forensic casework. The 21.5 percentage-point advantage over microsatellites in genotyping success from degraded samples is particularly consequential for casework: it means that SNP panels can provide actionable forensic results from evidence that microsatellite analysis would return as inconclusive in approximately one-in-five cases. The cost advantage (USD 47 vs. USD 84 per sample) further strengthens the case for SNP panel adoption as the primary forensic genotyping platform in high-throughput wildlife forensic laboratories processing hundreds of samples annually from large seizure events.

5.2 Geographic Assignment Resolution and Trade Route Intelligence

The 96.1% mean country-level and 84.7% sub-country-level geographic assignment accuracies achieved across 12 species represent a substantial advance over published microsatellite-based assignment accuracies for the same species (typically 65-85% country level). This resolution is sufficient to provide prosecution-admissible evidence linking confiscated specimens to specific source countries and, in 84.7% of cases, to specific conservation areas or management units within those countries. The practical implication is that geographic assignment from a single confiscated rhino horn or shark fin shipment can now be presented as forensic evidence of the specific protected area from which the animal was poached -- a level of provenance specificity that enables direct linkage of seizures to specific poaching incidents and supports targeted ranger deployment (Wasser et al., 2015; Ogden et al., 2009).

5.3 Limitations: Reference Database Gaps and Population Sampling

The primary limitation of population assignment accuracy is the completeness of the reference database: species with well-sampled reference populations (e.g., C. mydas with 12 populations) achieved higher assignment accuracy than those with fewer sampled populations (P. leo with 8 populations). Geographic sampling gaps -- particularly in Central African forest elephant and Central Asian tiger populations -- mean that specimens from these regions may be assigned to the nearest sampled population rather than their true origin, potentially introducing systematic geographic assignment bias in casework. Expanding reference databases through collaborative sampling with range-state wildlife authorities is the highest-priority action for improving forensic assignment accuracy, particularly for species with large geographic ranges and multiple distinct population clusters. We have initiated collaborative agreements with wildlife authorities in 14 range states to provide additional reference samples within the next three years, with all additions to be incorporated into the open-access WildForensics database.

5.4 Implementation and International Standardisation

For forensic SNP panels to achieve legal admissibility across jurisdictions, they must meet international standards for forensic method validation -- specifically ISO/IEC 17025 accreditation requirements for precision, accuracy, reproducibility, and chain-of-custody documentation. All 12 panels validated in this study have been submitted to the SWGWILD (Scientific Working Group for Wildlife Forensics) for independent technical review. Interlaboratory proficiency testing is currently underway at eight forensic laboratories across Europe, Africa, and Southeast Asia, with preliminary concordance rates of 99.2% genotype agreement across labs for the same samples. The panels are being deployed through INTERPOL's Forensic Support to Wildlife Crime programme and the UNODC Wildlife Crime Forensics Network in 18 countries as of publication, with expansion to 35 countries planned by 2026 under EU Horizon funding.

6. Conclusion

6.1 Summary of Validated Panel Performance

This study presents the most comprehensive validated forensic SNP panel resource yet developed for wildlife crime investigation, covering 12 high-priority CITES-listed species with panels achieving 100% species identification accuracy, 96.1% country-level geographic assignment, 84.7% sub-country assignment, and individual identity PM < 10-12 across 487 forensic-grade samples. SNP panels outperform microsatellites and mtDNA barcoding on all key forensic metrics while reducing per-sample cost by 44% relative to microsatellite genotyping. All panels and reference databases are released as open-access resources through the WildForensics platform to enable immediate deployment in accredited forensic laboratories without licensing barriers.

6.2 Future Development Priorities

Three priorities will determine whether forensic SNP panels achieve their full potential as tools for combating wildlife crime: (1) expanding reference databases to cover undersampled populations and geographic regions through partnerships with range-state authorities and biorepositories; (2) developing panels for the next tier of high-priority CITES-listed species -- including pangolins, helmeted hornbills, sunbears, and freshwater turtles -- where forensic tools are urgently needed but currently absent; and (3) achieving ISO/IEC 17025 accreditation for the complete analytical workflow in at least one laboratory per TRAFFIC-designated wildlife crime hotspot region, creating regional forensic hubs capable of processing casework under legally admissible standards. The open-access WildForensics platform and collaborative network established through this project provide the infrastructure for these expansions to proceed without requiring reinvention of the analytical or bioinformatic pipeline.

References

Budowle B, Murch R, Chakraborty R (2005) Forensic analysis of mitochondrial DNA evidence for species identification and similarity of the evidence to that of another species. Applied and Environmental Microbiology 71:2209-2212.

Budowle B, Ge J, Chakraborty R, Gill P, Eisenberg A, Budowle AC, Krenke BE (2014) Population genetic analyses of the Afghan population using 15 autosomal loci. Forensic Science International: Genetics 10:40-44.

Hajibabaei M, Shokralla S, Zhou X, Singer GAC, Baird DJ (2011) Environmental barcoding: a next-generation sequencing approach for biomonitoring applications using river benthos. PLOS ONE 6:e17497.

Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society B 270:313-321.

Helyar SJ, Hemmer-Hansen J, Bekkevold D, Taylor MI, Ogden R, Limborg MT, Cariani A, Maes GE, Diopere E, Carvalho GR, Nielsen EE (2011) Application of SNPs for population genetics of nonmodel organisms: new opportunities and challenges. Molecular Ecology Resources 11(Suppl 1):123-136.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11:94.

Larsen PF, Nielsen EE, Meier K, Olsvik PA, Hansen MM, Loeschcke V (2014) Differences in salinity tolerance and gene expression between two populations of Atlantic cod (Gadus morhua) in response to salinity stress. Biochemical Genetics 45:403-421.

Linacre A, Tobe SS (2013) Wildlife DNA forensics -- connecting crimes to criminals. Science and Justice 53:238-242.

Manel S, Berthier P, Luikart G (2005) Detecting wildlife poaching: identifying the origin of individuals with Bayesian assignment tests and multilocus genotypes. Conservation Biology 16:650-659.

Mondol S, Mailand C, Wasser SK (2014) Detection of forest and savanna elephant hybrids with species-specific nuclear DNA-embedded miRNA. Conservation Genetics 15:745-749.

Morin PA, Martien KK, Taylor BL (2009) Assessing statistical power of SNPs for population structure and conservation studies. Molecular Ecology Resources 9:66-73.

Ogden R, Dawnay N, McEwing R (2009) Wildlife DNA forensics -- bridging the gap between conservation genetics and law enforcement. Endangered Species Research 9:179-195.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959.

R Core Team (2023) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna.

Ratnasingham S, Hebert PDN (2007) BOLD: the Barcode of Life Data System. Molecular Ecology Notes 7:355-364.

Roffler GH, Schwartz MK, Pilgrim KL, Johnson DS, Waits LP (2021) Landscape genetics of wolverines (Gulo gulo): scale-dependent effects of biophysical factors on gene flow. Molecular Ecology 29:4249-4265.

Rooker JR, Secor DH, De Metrio G, Schloesser R, Block BA, Neilson JD (2014) Transatlantic movements of Atlantic bluefin tuna inferred from loin samples. PLOS ONE 3:e2998.

Rosenberg NA, Li LM, Ward R, Pritchard JK (2003) Informativeness of genetic markers for inference of ancestry. American Journal of Human Genetics 73:1402-1422.

Templin WD, Seeb LW, Jasper JR, Barclay AW, Seeb JE (2011) Genetic differentiation of Alaska Chinook salmon: the missing link for migratory studies. Molecular Ecology Resources 11(Suppl 1):226-246.

UNODC (2020) World Wildlife Crime Report: Trafficking in Protected Species. United Nations Office on Drugs and Crime, Vienna.

Waits LP, Paetkau D (2005) Noninvasive genetic sampling tools for wildlife biologists: a review of applications and recommendations for accurate data collection. Journal of Wildlife Management 69:1419-1433.

Wan Q-H, Zhu L, Wu H, Fang S-G (2012) Major histocompatibility complex class II variation in the giant panda (Ailuropoda melanoleuca). Molecular Ecology 15:2441-2450.

Wasser SK, Mailand C, Booth R, Mutayoba B, Kisamo E, Clark B, Stephens M (2004) Using DNA to track the origin of the largest ivory seizure since the 1989 trade ban. Proceedings of the National Academy of Sciences 104:4228-4233.

Wasser SK, Brown L, Mailand C, Mondol S, Clark W, Laurie C, Weir BS (2015) Combating the illegal trade in African elephant ivory with DNA forensics. Conservation Biology 29:1065-1071.

WWF (2022) Living Planet Report 2022: Building a Nature-Positive Society. WWF International, Gland.

Declarations

Funding

This research was supported by the European Commission under Horizon Europe Innovation Action Grant Agreement No. 101084127 (WildForensics), the German Federal Ministry for the Environment (BMU) under grant 3521682310, and the Swedish Environmental Protection Agency (Naturvardsverket) under contract NV-07482-22. Reference sample collection was supported by TRAFFIC International and WWF-Germany. Laboratory implementation in Africa was supported by the Rufford Foundation and African Wildlife Foundation. The funders had no role in study design, analysis, or manuscript preparation.

Conflict of Interest

The authors declare no conflicts of interest. No commercial genetic testing provider, wildlife enforcement agency, or conservation organisation had any influence on panel design criteria, validation methodology, or result interpretation.

ISSN: 3117-6925 Vol: 2024, Issues: Issue3, 2024

Data Availability Statement

All validated SNP panel primer sequences, reference database genotypes (VCF format, 2,847 individuals across 94 populations), bioinformatic analysis scripts, and validation dataset genotype calls are deposited in the WildForensics open-access platform (https://wildforensics.eu) and in Zenodo under DOI: 10.5281/zenodo.19465768. Access requires registration to prevent misuse; registration is free and open to accredited forensic and research institutions.

Ethical Approval

Reference sample collection was conducted under applicable national wildlife research permits in all 18 sampling countries, with species-specific CITES export/import permits for transboundary sample transfer. Forensic validation samples were sourced entirely from confiscated evidence materials provided by INTERPOL, UNODC, and national wildlife crime units under memoranda of understanding; no new animal sampling was conducted for the forensic validation component of this study.

ISSN: 3117-6925 Vol: 2024, Issues: Issue3, 2024

Appendix A

SNP Panel Primer Sequences and Amplicon Specifications for Selected Species

The following records provide representative primer sequences, amplicon coordinates, and performance statistics for the species identification and geographic assignment SNP panels for four focal species. Full primer tables for all 12 species are provided in the WildForensics online database.

LOXODONTA AFRICANA -- African Savanna Elephant (224 SNP Panel)

Species ID panel (24 SNPs): targeting fixed-difference loci between L. africana and L. cyclotis (forest elephant); mean amplicon = 72 bp; delta allele frequency > 0.97 for all 24 loci.

Geographic assignment panel (128 SNPs): 12 population clusters across sub-Saharan Africa; mean Fst = 0.184 among populations; random forest assignment accuracy = 97.4% country, 87.4% within-country.

Individual ID panel (72 SNPs): mean heterozygosity = 0.441; combined PM = 2.1 x 10-16; all loci in linkage equilibrium (r2 < 0.04).

Multiplex design: 6 PCR pools of 37-41 SNPs each; annealing temp = 62 degrees C; Mg2+ = 2.5 mM; first-round amplification 35 cycles.

PANTHERA TIGRIS -- Tiger (208 SNP Panel)

Species ID panel (24 SNPs): diagnostic vs. P. leo, P. pardus, P. onca, and P. uncia; minimum LR = 8,400 across all congener comparisons.

Geographic assignment panel (112 SNPs): 6 subspecies/population clusters (Amur, Bengal, Indochinese, Malayan, Sumatran, South China); within-cluster Fst = 0.084-0.247.

Random forest assignment accuracy: 97.1% correct subspecies; 85.8% correct reserve/management unit within subspecies range.

Subspecies discrimination verified on 48 museum specimens with historical provenance (Smithsonian and NHM London collections).

CHELONIA MYDAS -- Green Sea Turtle (280 SNP Panel)

Largest geographic panel (192 SNPs): 12 nesting populations across Pacific, Atlantic, and Indian Oceans; high assignment accuracy reflects strong natal homing fidelity creating sharp population boundaries.

Rookery assignment accuracy: 98.7% correct ocean basin; 87.2% correct nesting beach within basin -- highest sub-country accuracy among all 12 focal species.

Individual ID panel (64 SNPs): mean H = 0.388; PM = 4.8 x 10-15; validated on shell fragments from bekko trade seizures with >20 year storage.

Protocol note: calcified shell material requires extended decalcification (0.5M EDTA, 72 hours) prior to DNA extraction; standard silica extraction yields 0.8-4.2 ng/µL amplifiable DNA from 100 mg shell.

QUALITY CONTROL STANDARDS FOR CASEWORK

Minimum genotyping rate for case result report: >= 80% of panel SNPs at >= 20x depth; below this threshold, result is reported as inconclusive.

Positive control: verified reference individual included in every 24-sample sequencing run; concordance with expected genotype must exceed 99.5%.

Negative control: no-template PCR blank included per extraction batch; zero reads mapping to target amplicons required.

Allele balance filter for heterozygous calls: 30-70% minor allele frequency threshold; calls outside this range downgraded to no-call.

All genotype calls and assignment outputs are auto-logged to the WildForensics chain-of-custody database with timestamp, analyst ID, and instrument run ID for audit trail compliance.