JABG Journal of Animal Breeding & Genomics

Journal of Animal Breeding and Genomics, 2026 June, 10(2)43-50
https://doi.org/10.12972/jabng.2026.10.2.2

Print ISSN 1226-5543
Online ISSN 2586-4297

Research Article

Genomic Analysis of Post-Mortem Feline Tissue Samples for Forensic Identification and Kinship Verification

Yun Kim[1,2,†], Seo-Yoon Kang[3,†], Seung-Wan Woo[1,2], Dayeon Kang[1,2], Sung-Lim Lee[3], Jaemin Kim[1,2], Bok Kyung Ku[4], Jae-Won Byun[4,*]

[1]Division of Applied Life Science (BK21 Four), Gyeongsang National University, Jinju 52828, Republic of Korea

[2]Institute of Agriculture and Life Sciences, Gyeongsang National University, Jinju 52828, Republic of Korea

[3]College of Veterinary Medicine, Gyeongsang National University, Jinju 52828, Republic of Korea

[4]Animal Disease Diagnostic Division, Animal and Plant Quarantine Agency, Gimcheon 39660, Republic of Korea

[†] These authors contributed equally to this manuscript.
[*]Corresponding author: jaewon8911@korea.kr

ABSTRACT

The application of whole genome sequencing (WGS) for precise individual identification and kinship verification in veterinary forensics is not yet fully established. To evaluate the forensic utility of genomic profiling, we analyzed WGS data from post-mortem feline tissues suspected of animal abuse, alongside domestic cat reference genomes, identifying over 4.7 million high-quality single nucleotide polymorphisms (SNPs) from approximately 34 million raw variants. First, we established clear genomic evidence for individual identity and familial relatedness among the fragmented tissues. Through comparative population genomic analysis incorporating diverse pre-existing cat breeds, we evaluated the genetic structure and lineage of the specimens. Furthermore, targeted genetic screening provided molecular evidence for specific morphological traits, such as coat color phenotypes, which reflects the physical characteristics of the unidentified animal remains. Collectively, these findings highlight the efficacy of WGS as a robust and comprehensive analytical tool in veterinary forensics, offering critical evidence for resolving complex animal cruelty cases.

Keywords: Forensic genetics, Individual identification, Kinship verification, Single nucleotide polymorphism, Whole genome sequencing

INTRODUCTION

In veterinary forensics, the genetic investigation of biological evidence—such as post-mortem tissues—plays a crucial role in uncovering the circumstances surrounding suspected animal abuse cases (Mores et al., 2025). Determining whether fragmented tissue samples originate from a single individual or multiple genetically related individuals provides critical information for reconstructing the events of a crime and identifying potential victims.

To achieve precise individual identification and kinship verification, advanced genomic approaches such as whole genome sequencing (WGS) are increasingly being utilized (Cardinali et al., 2023). WGS offers a comprehensive genomic profile that goes beyond basic identification, enabling the simultaneous analysis of genetic relatedness, biological sex, population structure, and morphological traits from a single robust dataset.

In this study, we analyzed WGS data from three post-mortem feline tissues recovered from a suspected animal cruelty case to establish genetic identity, familial relatedness, and biological sex. To evaluate population structure and potential lineage, we integrated a reference cohort of 48 domestic cats representing 20 breeds. Additionally, targeted screening of the *ASIP* gene was performed to infer prospective coat color. Ultimately,

Received June 02, 2026 **Revised** June 16, 2026 **Accepted** June 19, 2026

this approach demonstrates the robust utility of comprehensive genomic profiling in providing actionable evidence for veterinary forensic investigations.

MATERIALS AND METHODS

DNA Extraction and Whole Genome Sequencing

Three post-mortem feline tissue samples, designated as 26QO035, 26QO036, and 26QO046, were recovered from an environment suspected of animal cruelty (Figure 1). Genomic DNA extraction, library preparation, and high-throughput sequencing were conducted by Macrogen Inc. (Seoul, South Korea). Sequencing libraries were constructed using the TruSeq DNA PCR Free kit (Illumina, San Diego, CA, USA) with an insert size of approximately 350 bp. Sequencing was performed on an Illumina platform to generate 151-bp paired-end raw reads in FASTQ format.

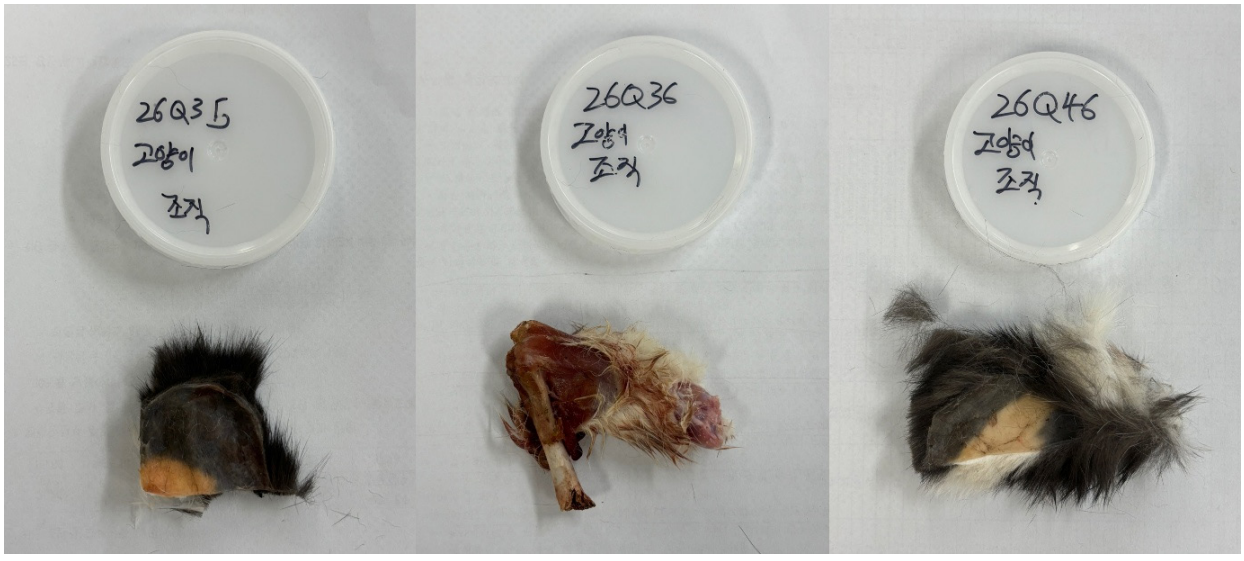

Figure 1. Photographic documentation of the three post-mortem feline tissue specimens (26QO035, 26QO036, and 26QO046) recovered for forensic genomic analysis.

Quality Control and Reference Data Integration

Raw sequence quality was evaluated using FastQC v0.11.9 (Andrews, 2010), confirming high-quality sequences with an average GC content of approximately 42-43% across all samples. To contextualize the genetic background of the forensic samples, genomic data from 48 reference individuals were obtained from a previously curated dataset (Kim et al., 2023). Because these historical reference datasets were already aligned to the older FelCat 8.0 assembly in the previous study, their genomic coordinates were converted to the F.catus_Fca126_mat1.0 assembly using the LiftoverVcf utility in PicardTools v2.21.1 (https://github.com/broadinstitute/picard).

Read Alignment and Variant Calling Pipeline

The bioinformatics workflow for variant discovery strictly followed the Genome Analysis Toolkit GATK v4.1.4.0 Best Practices guidelines (McKenna et al., 2010) (Figure 2). Raw sequence reads from both the forensic samples and reference controls were mapped to the F.catus_Fca126_mat1.0 reference genome using the BWA-MEM algorithm v0.7.17-r1188. To minimize PCR-induced artifacts and computational

bias, duplicate reads were identified and removed via Picard MarkDuplicates v2.21.1 (Li and Durbin, 2009). Base Quality Score Recalibration (BQSR) was subsequently applied using the standard Ensembl feline variation database (felis_catus.vcf.gz) as a known-sites resource to statistically adjust empirical base quality scores and enhance variant calling precision. Initial variant extraction was performed for each individual sample using GATK HaplotypeCaller. To create a unified dataset for comparative analysis, the individual variant call format (VCF) files of the three forensic samples and the five reference controls were subsequently directly merged. To account for any missing genotypes resulting from this direct merging process, a highly stringent zero-missingness filter was applied in the subsequent step.

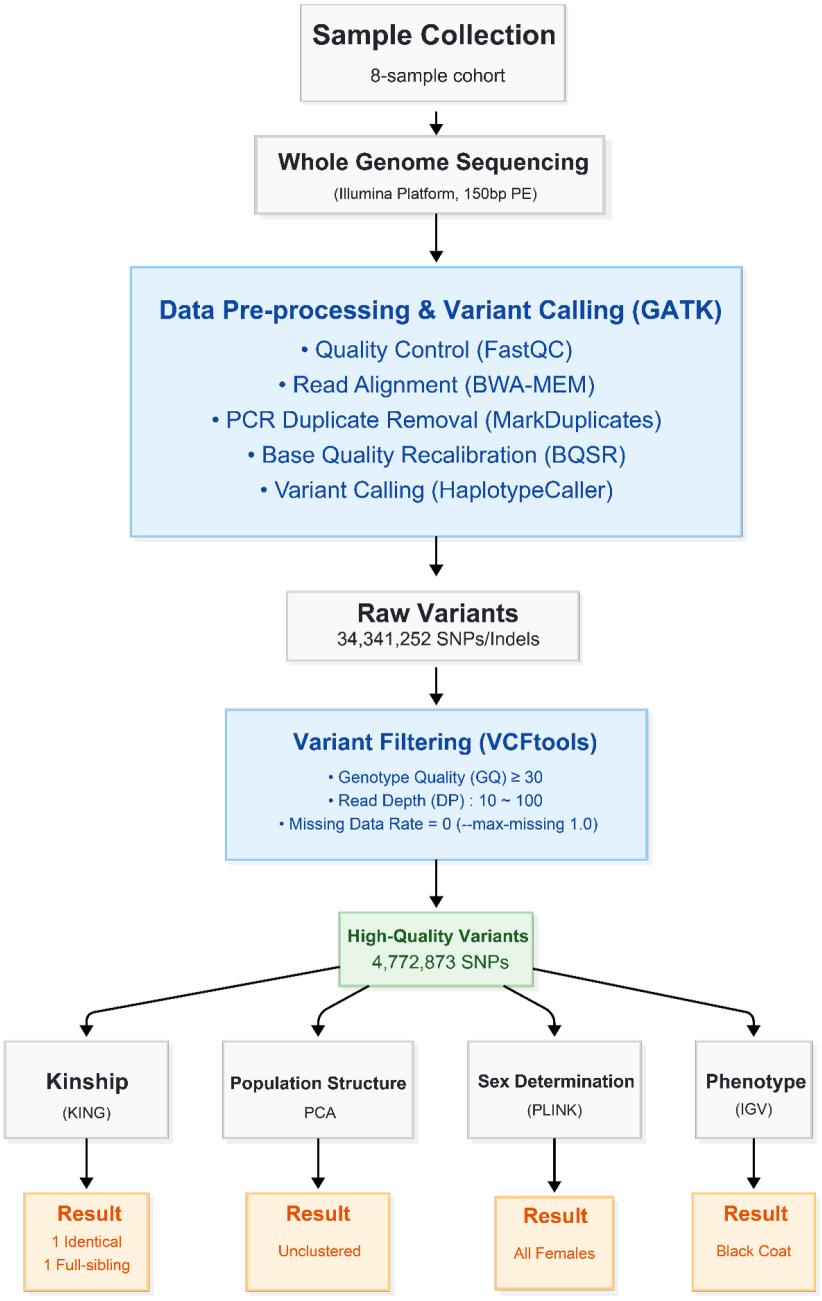

Figure 2. Flowchart outlining the overall study design and bioinformatic analytical approach for the forensic genomic profiling of the feline samples.

Quality Filtering and Dataset Construction

To ensure statistical stringency for kinship verification, five domestic cats were strategically selected from the 48-reference panel to serve as an unrelated negative control cohort. The selection was strictly based on high-quality alignment metrics, requiring a sequencing coverage exceeding 42X and a mapping rate greater than 95%. Because our pipeline applied an absolute zero-missingness threshold (--max-missing 1.0), restricting the negative control cohort to these five optimal genomes was strategically crucial to maximize the retention of high-confidence SNPs. Furthermore, this 5-sample cohort provides 15 independent pairwise comparisons against the forensic subjects, which is statistically robust for establishing a definitive baseline of background unrelatedness (negative PHI values) without introducing missing data bias from lower-coverage samples. These five control genomes were directly merged with the three forensic samples. Rigorous quality control filtering was applied to this combined 8-sample dataset using VCFtools v0.1.13 (Danecek et al., 2011). Variants were filtered based on a minimum Genotype Quality (GQ) of 30, a Read Depth (DP) of 10-100, and a strict zero-missingness threshold (--max-missing 1). This stringent filtering pipeline successfully retained 4,772,873 high-quality single nucleotide polymorphisms (SNPs) for the final high-resolution dataset.

Kinship and Individual Identification Analysis

To evaluate genetic identity and familial relationships, Identity-by-Descent (IBD) sharing proportions and Kinship coefficients (PHI) were calculated across the 8-sample dataset using the KING-robust estimator, which is implemented via the --relatedness2 function within VCFtools (Manichaikul et al., 2010). The inclusion of the five external domestic cats functioned as a baseline to validate zero-relatedness and prevent statistical inflation. Biological sex was determined by computing the X-chromosome inbreeding coefficient (F-value) using PLINK v1.90b6.24 (Chang et al., 2015), classifying individuals with a low F-value (< 0.2) as females.

Population Structure Analysis and Phenotypic Inference

To assess population genetic structure and potential breed lineages, a combined dataset of 51 individuals (3 forensic and 48 reference samples) was generated. Chromosome nomenclatures were harmonized, and biallelic SNPs with no missing genotypes were extracted and merged using bcftools. Principal Component Analysis (PCA) was subsequently performed on this highly filtered dataset using PLINK v1.90b6.24 to calculate the top principal components. Additionally, to infer prospective coat color phenotypes, aligned BAM files were visually inspected using Integrative Genomics Viewer (IGV) v2.7.5 (Robinson et al., 2011) to screen for known structural variants within the *ASIP* gene (Eizirik et al., 2003).

RESULTS

Genomic Data Quality and Variant Discovery

High-throughput WGS of the forensic tissues yielded high-quality sequence reads with a stable GC content (42-43%). Through the direct merger of the variant call format (VCF) files of the three forensic subjects alongside the five domestic reference controls, an initial set of 34,341,252 raw variants was produced across the complete 8-sample cohort. Following the stringent quality filtration—specifically excluding any loci with missing data—a high-resolution dataset of 4,772,873 high-confidence SNPs was established. This robust dataset provided a statistically stringent foundation for all subsequent downstream evaluations.

Individual Identification and Kinship Verification

Pairwise genetic comparisons across the 8-sample cohort revealed a PHI coefficient of 0.4997 between samples 26QO036 and 26QO046 (Table 1). The near-zero count of opposite homozygous loci (N_AAaa = 1) definitively confirmed they originated from the exact same biological individual. In contrast, sample 26QO035 exhibited a PHI of 0.2287 against both identical samples, which strictly falls within the standard KING threshold for first-degree relatives (0.177–0.354; Manichaikul et al., 2010). While full siblings and parent-offspring pairs both fall within this specific PHI range, they can be statistically distinguished by their proportion of zero identity-by-state (IBS0) sharing. The elevated opposite homozygous counts (N_AAaa = 133,111 and 133,041), which directly represent high IBS0 proportions, effectively excluded a parent-offspring arrangement. Because an offspring obligatorily inherits one allele from each parent, true parent-offspring pairs theoretically exhibit zero opposite homozygous sites (e.g., one being AA and the other aa). This elevated count, therefore, definitively classified 26QO035 as a full sibling. Crucially, all comparisons between the forensic tissues and the five control cats yielded negative PHI values (-0.10 to -0.27; Table 1 and Table S1), validating the absence of background relatedness and the statistical robustness of the pipeline. Consequently, the three forensic tissues were decisively determined to originate from two distinct full siblings.

Table 1. Pairwise kinship coefficients (PHI) and opposite homozygous loci (N_AAaa) for individual identification and relationship verification among the forensic samples and external domestic controls.

INDV1	INDV2	N_AAaa	PHI
26QO035	26QO036	133,111	0.2287
26QO036	26QO046	1	0.4997
26QO046	26QO035	133,041	0.2288
26QO035	SRR8092630	593,586	-0.1374
26QO036	SRR8092630	609,375	-0.1479
26QO046	SRR8092630	609,239	-0.1478

Table S1. Pairwise kinship coefficients (PHI) and opposite homozygous locus counts (N_AAaa) between the three forensic specimens and the remaining four external domestic reference controls

INDV1	INDV2	N_AAaa	PHI
26QO035	SRR11392564	516,859	-0.1027
26QO035	SRR11392531	571,378	-0.1635
26QO035	SRR11392632	654,534	-0.1964
26QO035	SRR11392454	681,536	-0.2334
26QO036	SRR11392564	517,071	-0.1039
26QO036	SRR11392531	568,239	-0.1631
26QO036	SRR11392632	667,391	-0.2050
26QO036	SRR11392454	690,580	-0.2376
26QO046	SRR11392564	516,906	-0.1038
26QO046	SRR11392531	568,086	-0.1630
26QO046	SRR11392632	667,181	-0.2049
26QO046	SRR11392454	690,346	-0.2374

Sex Determination

The biological sex of the subjects was inferred by analyzing the inbreeding coefficient (F-value) based on X-chromosome heterozygosity rates. Because mammalian females (XX) possess two homologous X chromosomes, they typically maintain a significantly higher level of X-chromosome-specific heterozygosity compared to hemizygous males (XY). The F-value analysis for all three post-mortem samples yielded low negative values (Table 2): -0.2650 for 26QO035, -0.4005 for 26QO036, and -0.4018 for 26QO046. These negative values reflect

a high proportion of heterozygous loci relative to expectation, confirming the presence of two functional X chromosomes. Accordingly, all investigated samples were decisively identified as female individuals.

Table 2. Biological sex determination of the three forensic samples based on X-chromosome inbreeding coefficients (F-values).

Sample ID	SEX	F
26QO035	Female	-0.2650
26QO036	Female	-0.4005
26QO046	Female	-0.4018

Population Genetic Structure and Breed Affiliation

PCA integrating the forensic samples with 48 reference individuals across 20 breeds revealed that the subjects formed a tight, independent cluster on the PC1-PC2 projection (Figure 3). This distinct genetic isolation from all established purebred clades indicates that the individuals are derived from a localized, random-bred domestic lineage rather than a specialized breed.

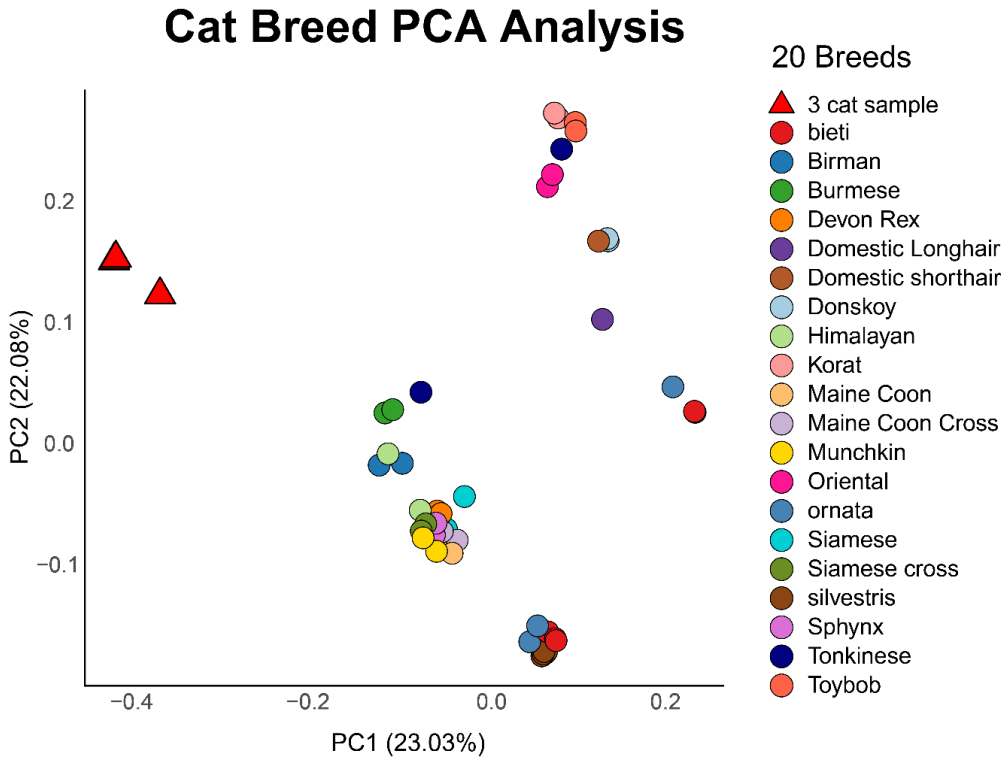

Figure 3. Principal component analysis of population genetic structure between the three forensic samples and 20 purebred feline reference clades.

Targeted Phenotypic Inference

To reconstruct the potential physical appearance of the specimens, genomic screening was focused on the *ASIP* (Agouti Signaling Protein) gene via visual inspection of the Binary Alignment Map (BAM) files using the Integrative Genomics Viewer (IGV). Structural variant analysis revealed a distinctive homozygous deletion within the *ASIP* gene region across the samples. This specific 2-bp deletion

(A3:24831083_24831084del) represents a well-characterized loss-of-function mutation that disrupts normal agouti signaling, preventing the alternating yellow-black banding of hair shafts. While feline coloration is inherently complex and governed by epistatic interactions among multiple loci, the presence of this homozygous non-agouti mutation allows for the prospective inference that the subjects displayed a solid black coat phenotype.

DISCUSSION

While traditional veterinary forensic investigations have primarily relied on short tandem repeat (STR) markers for basic individual identification (Linacre, 2021), this study demonstrates the superior utility of whole genome sequencing (WGS). WGS enables the simultaneous extraction of identity, complex kinship, population lineage, and morphological traits from a single sample. Crucially, by quantifying opposite homozygous loci across millions of SNPs and utilizing external domestic controls, we definitively resolved full-sibling and identical DNA relationships. This high-density genomic profiling overcomes the statistical limitations of traditional STR panels, offering precise kinship resolution essential for linking multiple specimens in animal cruelty cases.

Furthermore, population structure analysis revealed that the experimental subjects formed an independent cluster distinctly separated from the 20 established purebred clades. This isolation reflects a structural bias in public feline genomic repositories, which overwhelmingly favor Western purebred lines (Buckley et al., 2020). Since many investigations involve random-bred domestic cats whose genetic signatures are underrepresented, this highlights the critical necessity of localized reference panels for accurate lineage assignment.

Despite these advantages, phenotypic inferences carry inherent biological limitations. While the identified homozygous *ASIP* deletion provides a molecular basis for a non-agouti (solid black) coat, mammalian pigmentation is polygenic and subject to epistatic interactions (Kaelin and Barsh, 2013). Alleles such as dominant white (W locus) or extensive white spotting (S locus) could completely mask the *ASIP* mutation, meaning morphological reconstructions represent prospective estimations rather than absolute determinations.

However, despite the unparalleled data resolution offered by WGS, its routine implementation in veterinary forensics currently faces significant practical limitations compared to traditional STR-based profiling. WGS workflows inherently entail substantially higher operational costs and longer computational and sequencing turnaround times, which may restrict scalability in time-sensitive or resource-constrained investigations. Furthermore, unlike well-established forensic STR marker sets, WGS currently lacks universally standardized validation frameworks required for seamless courtroom admissibility. Acknowledging these trade-offs is essential; while WGS serves as a powerful tool for comprehensive genomic discovery, establishing standardized consensus guidelines remains a critical prerequisite for its routine legal application. Ultimately, this robust, WGS-based analytical framework provides the comprehensive, objective, and legally defensible evidence required to effectively resolve complex animal abuse investigations.

CONFLICT OF INTERESTS

The authors declare no potential conflicts of interest relevant to this article.

ACKNOWLEDGEMENTS

This study was supported by the research fund by the Animal and Plant Quarantine Agency (APQA), Republic of Korea (No. N-1543069-2025-29-01).

REFERENCES

Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

Buckley RM, Davis BW, Brashear WA, Farias FH, Kuroki K, Graves T, Hillier LW, Kremitzki M, Li G and Middleton RP. 2020. A new domestic cat genome assembly based on long sequence reads empowers feline genomic medicine and identifies a novel gene for dwarfism. PLoS Genetics. 16: e1008926.

Cardinali I, Tancredi D and Lancioni H. 2023. The revolution of animal genomics in forensic sciences. International Journal of Molecular Sciences. 24: 8821.

Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM and Lee JJ. 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 4: s13742–13015–10047–13748.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT and Sherry ST. 2011. The variant call format and VCFtools. bioinformatics. 27: 2156–2158.

Eizirik E, Yuhki N, Johnson WE, Menotti-Raymond M, Hannah SS and O'Brien SJ. 2003. Molecular genetics and evolution of melanism in the cat family. Current biology. 13: 448–453.

Kaelin CB and Barsh GS. 2013. Genetics of pigmentation in dogs and cats. Annu. Rev. Anim. Biosci. 1: 125–156.

Kim T, Go M, Kang D and Kim J. 2023. Genetic Differentiation between Domestic Cats and Wildcats. Journal of Intellectual Property. 7: 9–15.

Li H and Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. bioinformatics. 25: 1754–1760.

Linacre A. 2021. Animal forensic genetics. Genes. 12: 515.

Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M and Chen W-M. 2010. Robust relationship inference in genome-wide association studies. bioinformatics. 26: 2867–2873.

McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S and Daly M. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome research. 20: 1297.

Mores CM, Castro MBd and Melo CBd. 2025. Forensic examinations in the investigations of crimes involving animals: a review. Ciência Rural. 55: e20240071.

Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G and Mesirov JP. 2011. Integrative genomics viewer. Nature biotechnology. 29: 24–26.