

## **The Breed Makes the Difference:** Pedigree Composition, Canine STR Allele Frequencies, and VI **Random Match Probabilities**



MEDICAL UNIVERSITY

INNSBRUCK

<u>Burkhard Berger<sup>1</sup></u>, Cordula Berger<sup>1</sup>, Josephin Heinrich<sup>1</sup>, Werner Hecht<sup>2</sup>, Andreas Hellmann<sup>3</sup>, Uwe Schleenbecker<sup>3</sup>, Udo Rohleder<sup>3</sup>, Nadja Morf<sup>4</sup>, the CaDNAP Group, Walther Parson<sup>1,5</sup>

<sup>1</sup> Institute of Legal Medicine, Medical University of Innsbruck, Austria; <sup>2</sup> Institute of Veterinary Pathology, Justus-Liebig-University Giessen, Germany; <sup>3</sup> Bundeskriminalamt, Kriminaltechnisches Institut, Wiesbaden, Germany; <sup>4</sup> Institute of Forensic Medicine, University of Zurich, Switzerland <sup>5</sup> Forensic Science Program, The Pennsylvania State University, PA, USA

burkhard.berger@i-med.ac.at



Domestic dogs are popular and as integral part in everyday life they can become relevant to forensic DNA analysis, which is typically conducted by canine STR

analysis followed by assessing the weight of evidence of a matching DNA profile, e.g. by random match probability (RMP) calculations.

Noteworthy in this context is the **outstanding diversity** of dogs. Breeders have exerted selective pressure throughout a long period creating the extraordinary variety of modern dog breeds, unparalleled in any other animal species.

This diversity applies also to forensically used markers - such as the CaDNAP 13-STR panel [1] which was applied in this study – and is demonstrated here by STRUCTURE analysis of 312 pedigree dogs from 26 breeds (Figure 1) and Principal Component Analysis of 30 individuals from 3 breeds, which were used for RMP calculations (Figure 2).

The breed specificity of STR-genotypes makes the breed composition a key factor for a dog population sample (DPS) featuring a reliable source of allele frequencies and correction factors (*Theta, Fst*) applicable for conservative RMP calculations suitable for use in court.



Figure 1: STRUCTURE analyses [4] of 13-STR-genotypes from DPS-312. The samples were grouped by the 26 breeds.

Figure 3: Breed composition of two dog population samples (DPS): DPS-**1184** consists of 166 dog breeds. The proportions of the breeds mirror the proportions determined in the dog population of Germany, Austria and Switzerland [2]. Mongrels (18%), German Shepherd (7%) and Labrador Retriever (3%) account for the largest shares. **DPS-312** consists of 26 breeds comprising 12 individuals each (4%).



In [2] we recently proposed to approximate the breed composition of a dog population sample as closely as possible to that of the actual dog population (in this case of Germany, Austria, and Switzerland), which means that popular breeds such as the German Shepherd account for a much larger proportion than rare breeds such as the Husky (Figure 3, DPS-1184: 13-STR-genotypes from 1,184 dogs).

To test the robustness of this approach, we compiled a sub-set of DPS-1184 with a balanced breed composition (Figure 3, DPS-312: genotypes of 312 dogs). Allele frequencies and Fst values between all breeds (DPS-1184: 0.16; DPS-312: 0.15) for Theta correction were the basis for RMP calculations applying NRCII formulas 4.10 [3] to account for breed-based population stratification.

RMP values for members of 3 selected breeds are shown in Figure 4. The calculations comprised dogs from the popular German Shepherd (n=10), the medium frequent **Belgian Shepherd** (n=7), and the rare **Siberian Husky** (n=13). The latter breed was not included in DPS-312.



Figure 2: Principal Component Analysis of 13-STR-genotypes from 10 German Shepherds, 7 Belgian Shepherds, and 13 Siberian Huskies. The substructure analysis was performed using the software STRAF [5].

**Figure 4: Random Match Probabilities** of 13-STR-genotypes from 30 dogs and 3 breeds: The -log of the RMP values is shown; the solid bars correspond to calculations based on DPS-1184, the patterned bars to DPS-312.

## **Results and Conclusions**

- **RMP values ranged** between 2.6E-15 and 3.5E-12 (based on DPS-1184) and 7.7E-16 and 5.7E-13 (based on DPS-312).
- The RMPs showed no substantial differences between the breeds, but a tendency towards higher RMP values for the German Shepherd.
- The simplified breed composition and the reduced sample number of DPS-312 had a clear impact on RMP values. In all 30 cases DPS-312 based RMPs are lower than those based on DPS-1184 (higher –log values, Figure 4). This difference is most pronounced for the German Shepherd.
- For conservative RMP calculations a comprehensive population sample compiled with realistic breed proportions proved to be the approach of **choice** as suggested earlier [2].

References: [1] Berger et al. (2014), Forensic Sci Int Genet 8, 90-100 [2] Berger et al. (2019), Forensic Sci Int Genet 42, 90-98 [3] National Research Council, The Evaluation of Forensic DNA Evidence. Washington (DC), 1996. [4] Pritchard, Stephens and Donnelly (2000), Genetics 155, 945-959 [5] Gouy and Zieger (2017), Forensic Sci Int Genet 30, 148-151 Acknowledgements: We extend our warmest thanks to all dog owners and breeders enabling the present study through their interest in supporting canine forensic STR analysis.