American Genetic Association

Journal of Heredity, 2014, 1–4
doi:10.1093/jhered/esv031
Brief Communication

OXFORD

Brief Communication

Molecular Genetics of Sex Identification, Breed Ancestry and Polydactyly in the Norwegian Lundehund Breed

Regina Kropatsch, Claudia Melis, Astrid V. Stronen, Henrik Jensen, and Joerg T. Epplen

From the Department of Human Genetics, Ruhr University, Universitätsstr. 150, 44801 Bochum, Germany (Kropatsch and Epplen); the Department of Biology, Centre for Biodiversity Dynamics, Norwegian University of Science and Technology, Trondheim, Norway (Melis and Jensen); the Department of Chemistry and Bioscience, Aalborg University, Fredrik Bajers Vej 7H, Aalborg Øst, Denmark (Stronen); and the Faculty of Health, University Witten/Herdecke, Witten, Germany (Epplen).

Address correspondence to Regina Kropatsch, Department of Human Genetics, Ruhr University, Universitätsstrasse 150, 44801 Bochum, Germany, or e-mail: regina.kropatsch@rub.de.

Data deposited at Dryad: http://dx.doi.org/doi:10.5061/dryad.67sj1

Received November 7, 2014; First decision December 16, 2014; Accepted April 13, 2015.

Corresponding Editor: Elaine Ostrander

Abstract

The Norwegian Lundehund breed of dog has undergone a severe loss of genetic diversity as a result of inbreeding and epizootics of canine distemper. As a consequence, the breed is extremely homogeneous and accurate sex identification is not always possible by standard screening of X-chromosomal loci. To improve our genetic understanding of the breed we genotyped 17 individuals using a genome-wide array of 170 000 single nucleotide polymorphisms (SNPs). Standard analyses based on expected homozygosity of X-chromosomal loci failed in assigning individuals to the correct sex, as determined initially by physical examination and confirmed with the Y-chromosomal marker, amelogenin. This demonstrates that identification of sex using standard SNP assays can be erroneous in highly inbred individuals.

Subject area: Genomics and gene mapping
Key words: ancestors, mtDNA, Norwegian Lundehund, polydactyly, sex identification, Y chromosome

Accurate sex identification from noninvasive hair and fecal samples of wild and free-ranging animals is often uncertain or impossible as samples often provide incomplete information. Although poor quality data obtained by noninvasive sampling can result in low amplification success (Santini et al. 2007), information on individual sex from molecular markers provides important data for research and conservation (Mumma et al. 2014). Sample identity problems can often be revealed by checking the reported sex of each individual against the one predicted by genetic data (Turner et al. 2011). Importantly, however, this requires reliable sex assignment. Despite genome-wide coverage from high-quality samples this may not be guaranteed.

Mitochondrial DNA (mtDNA) mainly used for phylogenetic population studies in the domestic dog (*Canis lupus familiaris*) regarding their geographic and temporal origin (Savolainen et al. 2002; Boyko et al. 2009; Pang et al. 2009; Vonholdt et al. 2010) as well as their evolutionary history (Tsuda et al. 1997; Vilà et al. 2005) also provides information regarding maternal gene flow and phylogenetic relationships within and among purebred dog breeds because of its maternal

pattern of inheritance. In addition, the paternally inherited Y chromosome allows conclusions about evolutionary events in paternal lineages of mammals. Thus, mtDNA and Y-chromosomal analyses yield information about sex-specific contributions at the time of the last bottleneck event in the history of the breed.

The Norwegian Lundehund is an endangered small Spitz breed which has gone through at least 2 known severe genetic bottlenecks (Melis et al. 2013). It was developed to hunt puffins (*Fratercula arctica*) on steep cliffs in Northern Norway 5 centuries ago (Espelien 2012), and is characterized by unique breed-defining traits such as a great flexibility of joints and neck as well as the presence of extra toes in the fore and hind limbs (polydactyly; Park et al. 2008; Melis et al. 2013). These traits were probably under selective pressure as they may have given an advantage when hunting puffins (Galis et al. 2001) by preventing the dogs from slipping off the rocks. Preaxial polydactyly (PPD) is a common congenital anomaly of the limb with abnormal number of digits caused by alterations in the antero-posterior axis of limb development (for a review see Al-Qattan 2013). Human and murine PPDs are known to be associated with the same highly conserved gene *LMBR1* (limb development membrane protein 1; Clark et al. 2000) or more precisely with substitutions in an intronic regulatory element called the zone of polarizing activity regulatory sequence (ZRS; Lettice et al. 2003). Canine homologous PPD has been mapped to CFA16 (Park et al. 2004), and it is caused by substitutions in a conserved intronic region of the *LMBR1* gene with limb-specific enhancer activity upstream of human ZRS (preZRS) as identified in Korean and Western dog breeds (Park et al. 2008).

Our research had 3 separate objectives. At first, we use the endangered Lundehund breed, where male and female assignment was certified through independent observation and genetic analyses, to illustrate how theoretical assumptions in statistical genetics can produce erroneous sex estimates. Furthermore, to estimate the minimum number of male and female ancestors in this dog breed, we investigate genetic relationships in a cohort of this breed based on both mtDNA and Y-chromosomal markers. Finally, to identify the cause of PPD in the Lundehund breed, we examine the known mutations in the conserved intronic sequence of *LMBR1* gene identified in Korean and Western dog breeds.

Materials and Methods

Two independent data sets were used, a "Norwegian" (investigated in Norway) for sex validation analyses and a "German" (studied in Germany) for studying ancestry and PPD mutations. For sex validation, 52 Lundehund individuals were DNA-sampled by buccal swabs at an international breed meeting in Norway (2013). DNA was extracted using the Isohelix DDK-50 isolation kit (Cell Projects Ltd, Sutton Valence, UK). A subset of 17 individuals of the "Norwegian" study cohort (5 males and 12 females) was chosen by avoiding siblings and parents and was genotyped with the Canine Illumina High Density Beadchip (including more than 170 000 single nucleotide polymorphisms [SNPs]). Sex assignment was obtained using GENOME STUDIO V 2011.1 software (Illumina, San Diego, CA, USA) which estimates the sex of each sample using X-chromosomal SNPs. Similarly, the option "Sex check" in the software PLINK (Purcell et al. 2007) uses X chromosome data to determine sex (based on heterozygosity rates) and identifies individuals as males if the inbreeding coefficient F is higher than 0.8 and as females if F is lower than 0.2. Subsequently, polymerase chain reaction amplification with the amelogenin microsatellite marker for sex identification on all 52 samples was performed as described in the study by Kekkonen et al. (2011), and alleles were scored using GENEMAPPER 4.0 software (Applied Biosystems, Darmstadt, Germany). The percentage of missing alleles, observed (H_O) and expected heterozygosity (H_E) and relatedness (identity by descent) were calculated in PLINK based on 165 293 SNPs.

For the estimation of patri- and matrilineages, the "German" study population comprising 57 to the largest extent unrelated Lundehunds (23 males and 34 females) of different Norwegian stud book lines was characterized for Y-chromosomal markers and mtDNA. Blood and buccal swab samples were collected in cooperation with the Norwegian Lundehund Club. We isolated genomic DNA from peripheral blood cells according to the standard protocols (Miller et al. 1988) and from epithelial mucosal cells with QIAamp DNA Mini kit (Quiagen, Hilden, Germany) according to the manufacturer's instructions. In order to estimate the number of male and female ancestors of the current Lundehund population we genotyped each individual at canine Y-chromosomal markers (5 SNPs and 4 microsatellite loci) and sequenced a 1947 bp-sized portion of canine mitochondrial genome (NCBI accession number U96639) as described in a previous study (Kropatsch et al. 2011).

As the breed standard for Lundehunds implies PPD, we investigated for PPD mutation analysis a subset of male and female dogs (n = 20) of the "German" study cohort as well as 4 healthy Schapendoes dogs without PPD as controls. A 783 bp-sized fragment of intron 5 of the *LMBR1* gene including preZRS with both PPD mutations (DC-1 and DC-2) for Korean and Western dog breeds (Park et al. 2008) was amplified using M13-tailed primers 5′-GCAAATGTATCACAGACATTGAC-3′ and 5′-GATTGAGAAA TAAGATCAATTTGATAAACA-3′ and sequenced directly as described previously (Kropatsch et al. 2011) on an automated capillary DNA sequencer (ABI3500, Applied Biosystems).

In fulfillment of data archiving guidelines (Baker 2013), we have deposited in Dryad the primary data underlying the analyses reported in this paper (doi:10.5061/dryad.67sj1).

Results

For sex validation, we obtained identical sex assignments by both GENOME STUDIO and PLINK, and these did not match those expected from sampling, because all 17 individuals of the "Norwegian" data set were identified as males. At first we considered that the samples might have been contaminated, but the amelogenin typing results were consistent with the sex assessed at sample collection for all individuals. For these 17 individuals, the average percentage of missing SNP genotypes was 0.71% and mean relatedness 0.899 (range 0.842–0.954). Across all SNPs, H_O was 0.038 (SE 0.00031) and H_E was 0.035 (SE 0.00028). For the X chromosome (4982 SNPs), H_O was 0.036 (SE 0.00141) and H_E was 0.038 (SE 0.00149).

The analysis of the Y-chromosomal markers revealed that all 9 markers were monomorphic in 16 of 23 males of the "German" data set. For the remaining 7 males, Y-chromosomal marker analyses failed because of poor DNA quality probably caused by protein contamination. We identified 1 Y-chromosomal haplotype or patrilineage (Table 1). In addition, 1 mtDNA haplotype or matrilineage (Table 2) was identified in all 57 individuals. Accordingly, stud book analyses revealed paternal gene flow from 1 known male ancestor called Kvikk 2 and 1 maternal lineage from 1 known female ancestor named Mosti (Figure 1).

PPD sequence analyses revealed for the DC-1 mutation of Korean dog breeds the wild-type allele in homozygous state (G/G) for all investigated individuals including the Schapendoes controls, whereas the DC-2 mutation was identified in homozygous state (A/A) for all 20 Lundehund individuals of the "German" study population. This was in contrast to all controls that showed the wild-type allele in homozygous state (G/G).

Table 1. SNPs and microsatellite markers on the Y chromosome used for minimal male ancestor estimation and Y-chromosomal haplotype in Norwegian Lundehund individuals

Y-chromosomal marker		EMBL accession number	Exchange position	Wild-type allele	Simple repeat motif	Identified haplotype
SNPs	Ydog28	DQ973642	873	G	—	G
	Ydog21	DQ973639	417	G	—	G
	YdogG	DQ973680	219	T	—	T
	YdogN	DQ973627	608	C	—	GC
	Ydog20	DQ973692	35	A	—	A
Microsatellites	MS41B	AF192268	—	224	$(CA)_4TAG(AC)_{19}$	224
	MS18	DQ973643	—	366	mixed repeat	370
	MS98	AY466398	—	151	$(CA)_{13}$	151
	MS97	AY466397	—	268	$(CA)_{14}$	268

EMBL, European Molecular Biology Laboratory.

Table 2. Mitochondrial DNA haplotype as identified based on variable sites in the *ATP synthase subunit 6* gene and a 3′ portion of *ATP synthase subunit 8* gene as well as a portion of D-loop sequence of Norwegian Lundehund individuals

Mitochondrial DNA amplicons	Exchange position	Reference sequence U96639	Identified haplotype
ATPase6 + ATPase 8 genes	8368	C	T
D-loop sequence	15 625	T	C
	15 814	C	T

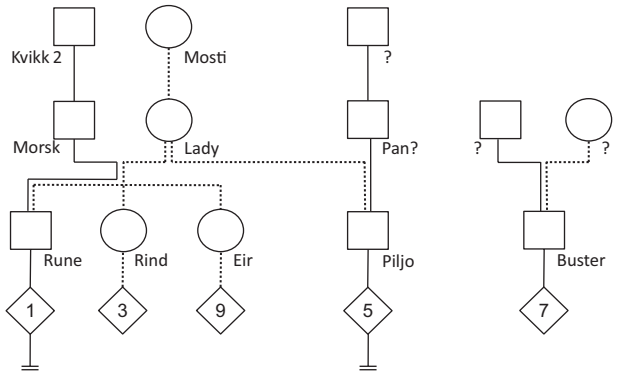

Figure 1. The ancestry of the Norwegian Lundehund breed based on stud book information. Squares and circles symbolize male and female dogs, respectively. Rhombuses correspond to descendants of unknown sex. Solid lines represent paternal thus Y-chromosomal lineages, whereas dotted lines indicate maternal or mitochondrial DNA lineages. Question marks denote doubtful or unknown descent.

Discussion

As expected from high-density genome-wide SNP results with good coverage (>99%), we identified in the "Norwegian" study cohort almost the same low level of heterozygosity and high degree of inbreeding as found in previous studies based on autosomal microsatellite markers (Melis et al. 2013; Pfahler and Distl 2014). This extremely reduced level of genetic variation appears to result in male sex assignment for all dogs when using GENOME STUDIO or PLINK, despite 71% being females. In current molecular genetic research approaches it is common to generate high-density genome-wide SNP data of the investigated individuals and simply infer each individual's sex based on these data. Thus, in general, no further investigations are necessary. Yet, because sex estimates based on theoretical expectations in statistical genetics can be erroneous for highly inbred individuals and populations, it is important to consider the level of inbreeding when using programs like GENOME STUDIO or PLINK. This is relevant especially in cases where sex identification is important for breeding, conservation, or management purposes. In such cases, it may be necessary to perform independent tests such as amelogenin gene amplification which could be quite helpful also for poor quality DNA samples with low success in high-density SNP genotyping.

Y-chromosomal and mtDNA results of the "German" data set suggest that the current Lundehund population in Norway has 1 male and 1 female ancestor. Potentially, this observation only provides the minimum estimated number of male and female ancestors. It is for example possible that several Lundehund males with identical Y-chromosomal haplotypes contributed. In addition, we only investigated parts of the mt genome which could lead to an underestimation of the number of female founders as described (Pang et al. 2009). Previously immense haplotype diversity has been observed in Y-chromosomal and mtDNA analyses in 27 other dog breeds (Kropatsch et al. 2011). Although only small numbers of individuals had been investigated per breed, we identified enormous mtDNA and Y-chromosomal haplotype variability per breed (*op. cit.*), a result contrasting with our ancestor estimations for the Lundehund despite analyzing a larger cohort of different lines for this breed. For example, Weimaraners, hunting dogs with high genetic diversity based on autosomal markers (Streitberger et al. 2012), comprise 4 Y-chromosomal- and 3 mtDNA haplotypes in the current German population, whereas the close-bred Saarloos Wolfhound, with high degrees of homozygous genome regions (Kropatsch et al., unpublished data), also show very low mtDNA and Y-chromosomal haplotype diversity (Kropatsch et al. 2011).Thus, our findings are consistent with the known Lundehund history, meaning that the current global population originates from 5 dogs who were close relatives (Frimann-Clausen and Laane 1968; Melis et al. 2013). Based on our results, we propose that perhaps only 4 autosomal-, 3 X-chromosomal-, and 1 Y-chromosomal chromatids were provided by 1 male and 1 female ancestor, and hence represented in the Lundehund breed population.

Here, the extremely low genetic variation in the Lundehund was confirmed according to previous findings where autosomal microsatellite markers were employed (Melis et al. 2013; Pfahler and Distl 2014). Albeit variation in neither the Y chromosome nor the mtDNA has been identified within the "German" study population, the age of the most recent mutations in the Lundehund breed can be estimated based on the number of generations since the last genetic bottleneck event. In addition, any hints for introgression of genetic variation from other breeds (Fossum 1973) like Norwegian Buhund in the breed of Lundehund were not evident from our analyses. MtDNA and Y-chromosomal investigations in a single Norwegian Buhund from

Germany support our previous assumption as the identified mtDNA and Y-chromosomal haplotypes differ from these of the Lundehund population remarkably (Kropatsch et al., unpublished data).

According to our results, the Lundehund breed is an exceptional animal model for monogenic diseases such as polydactyly, which appears to be caused by the same intronic mutation in the preZRS of *LMBR1* gene as in the Western dog breeds Beagle, Cocker Spaniel, Malinois, Rottweiler, Shetland Sheepdog, Standard Poodle, Standard Schnauzer, Shih Tzu, and Yorkshire Terrier (Park et al. 2008). Mutation sharing or rather sharing identical haplotypes in different breeds is commonly observed (Park et al. 2008; Kropatsch et al. 2011), yielding evidence for a common or shared origin (Bannasch et al. 2005), which is reflected in their breeding history (Sutter and Ostrander 2004). Although illustrating the common origin with other Western breeds, this mutation led to functional phalanges exclusively in the breed of Lundehunds.

Funding

Norwegian Genetic Resource Center and private funds of the Department of Human Genetics, Ruhr University Bochum; Danish Council for Independent Research (DFF—grant no. 1337-00007 to A.V.S.).

Acknowledgments

We would like to thank the Norwegian Lundehund Club and all the dog owners and breeders who helped with data collection. We are also grateful to A. Billing for assistance with lab analyses.

References

Al-Qattan MM. 2013. Preaxial polydactyly of the upper limb viewed as a spectrum of severity of embryonic events. *Ann Plast Surg*. 71:118–124.

Baker CS. 2013. Journal of heredity adopts joint data archiving policy. *J Hered*. 104:1.

Bannasch DL, Bannasch MJ, Ryun JR, Famula TR, Pedersen NC. 2005. Y chromosome haplotype analysis in purebred dogs. *Mamm Genome*. 16:273–280.

Boyko AR, Boyko RH, Boyko CM, Parker HG, Castelhano M, Corey L, Degenhardt JD, Auton A, Hedimbi M, Kityo R, *et al*. 2009. Complex population structure in African village dogs and its implications for inferring dog domestication history. *Proc Natl Acad Sci USA*. 106:13903–13908.

Clark RM, Marker PC, Kingsley DM. 2000. A novel candidate gene for mouse and human preaxial polydactyly with altered expression in limbs of Hemimelic extra-toes mutant mice. *Genomics*. 67:19–27.

Espelien IS. 2012. *Lundehundboka*. Bergen: Vigmostad Bjørke.

Fossum R. 1973. Quo vadis lundehundrase? *Hundesport*. 5:158.

Frimann-Clausen C, Laane M. 1968. Cyto-genetical and anatomical studies in the Lundehound, a deviating member of the Canis familiaris palustris group. *Nytt Mag Zool*. 16:72–78.

Galis F, van Alphen J, Metz J. 2001. Why five fingers? Evolutionary constraints on digit numbers. *Trends Ecol Evol*. 16:637–646.

Kekkonen J, Hanski I, Jensen H. 2011. Increased genetic differentiation in house sparrows after a strong population decline: from panmixia towards structure in a common bird. *Biol Conserv*. 144:2931–2940.

Kropatsch R, Streitberger K, Schulte-Middelmann T, Dekomien G, Epplen JT. 2011. On ancestors of dog breeds with focus on Weimaraner hunting dogs. *J Anim Breed Genet*. 128:64–72.

Lettice LA, Heaney SJ, Purdie LA, Li L, de Beer P, Oostra BA, Goode D, Elgar G, Hill RE, de Graaff E. 2003. A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. *Hum Mol Genet*. 12:1725–1735.

Melis C, Borg ÅA, Espelien IS, Jensen H. 2013. Low neutral genetic variability in a specialist puffin hunter: the Norwegian Lundehund. *Anim Genet*. 44:348–351.

Miller SA, Dykes DD, Polesky HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res*. 16:1215.

Mumma MA, Soulliere CE, Mahoney SP, Waits LP. 2014. Enhanced understanding of predator-prey relationships using molecular methods to identify predator species, individual and sex. *Mol Ecol Resour*. 14:100–108.

Pang JF, Kluetsch C, Zou XJ, Zhang AB, Luo LY, Angleby H, Ardalan A, Ekström C, Sköllermo A, Lundeberg J, *et al*. 2009. mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol Biol Evol*. 26:2849–2864.

Park K, Kang J, Park S, Ha J, Park C. 2004. Linkage of the locus for canine dewclaw to chromosome 16. *Genomics*. 83:216–224.

Park K, Kang J, Subedi KP, Ha JH, Park C. 2008. Canine polydactyl mutations with heterogeneous origin in the conserved intronic sequence of LMBR1. *Genetics*. 179:2163–2172.

Pfahler S, Distl O. 2014. A massive reduction of the genetic diversity in the Lundehund. *Anim Genet*. 45:154.

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ, *et al*. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 81:559–575.

Santini A, Lucchini V, Fabbri E, Randi E. 2007. Ageing and environmental factors affect PCR success in wolf (Canis lupus) excremental DNA samples. *Mol Ecol Notes*. 7:955–961.

Savolainen P, Zhang YP, Luo J, Lundeberg J, Leitner T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science*. 298:1610–1613.

Streitberger K, Schweizer M, Kropatsch R, Dekomien G, Distl O, Fischer MS, Epplen JT, Hertwig ST. 2012. Rapid genetic diversification within dog breeds as evidenced by a case study on Schnauzers. *Anim Genet*. 43:577–586.

Sutter NB, Ostrander EA. 2004. Dog star rising: the canine genetic system. *Nat Rev Genet*. 5:900–910.

Tsuda K, Kikkawa Y, Yonekawa H, Tanabe Y. 1997. Extensive interbreeding occurred among multiple matriarchal ancestors during the domestication of dogs: evidence from inter- and intraspecies polymorphisms in the D-loop region of mitochondrial DNA between dogs and wolves. *Genes Genet Syst*. 72:229–238.

Turner S, Armstrong LL, Bradford Y, Carlson CS, Crawford DC, Crenshaw AT, de Andrade M, Doheny KF, Haines JL, Hayes G, *et al*. 2011. Quality control procedures for genome-wide association studies. *Curr Protoc Hum Genet*. Chapter 1:Unit1.19.

Vilà C, Seddon J, Ellegren H. 2005. Genes of domestic mammals augmented by backcrossing with wild ancestors. *Trends Genet*. 21:214–218.

Vonholdt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P, Degenhardt JD, Boyko AR, Earl DA, Auton A, *et al*. 2010. Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature*. 464:898–902.