Short Communication

A splice site variant in the *SUV39H2* gene in Greyhounds with nasal parakeratosis

Anina Bauer[1,2], Judith Nimmo[3], Ross Newman[4], Magdalena Brunner[2,5], Monika M. Welle[2,5], Vidhya Jagannathan[1,2], Tosso Leeb[1,2]

[1] Institute of Genetics, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland
[2] DermFocus, University of Bern, 3001 Bern, Switzerland
[3] ASAP Laboratory, Mulgrave, VIC 3170, Australia
[4] Mobile Vet Services and Supplies, Warwick Qld 4370, Australia
[5] Institute of Animal Pathology, Vetsuisse Faculty, University of Bern, 3001 Bern, Switzerland

Running title: Canine *SUV39H2* splice site variant

Address for correspondence

Tosso Leeb
Institute of Genetics
Vetsuisse Faculty
University of Bern
Bremgartenstrasse 109a
3001 Bern
Switzerland

Phone: +41-31-6312326
Fax: +41-31-6312640
E-mail: Tosso.Leeb@vetsuisse.unibe.ch

Summary

Hereditary nasal parakeratosis (HNPK) described in the Labrador Retriever breed is a monogenic autosomal recessive disorder that causes crusts and fissures on the nasal planum of otherwise healthy dogs. Our group previously showed that this genodermatosis may be caused by a missense variant located in the *SUV39H2* gene encoding a histone 3 lysine 9 methyltransferase, a chromatin modifying enzyme with a potential role in keratinocyte differentiation. In the present study, we investigated a litter of Greyhounds, in which six out of eight puppies were affected with parakeratotic lesions restricted to the nasal planum. Clinically and histologically, the lesions were comparable to HNPK in Labrador Retrievers. Whole genome sequencing of one affected Greyhound revealed a 4 bp deletion at the 5'-end of intron 4 of the *SUV39H2* gene that was absent in 188 control dog and three wolf genomes. The variant was predicted to disrupt the 5'-splice site with subsequent loss of SUV39H2 function. The six affected puppies were homozygous for the variant, while the two non-affected littermates were heterozygous. Genotyping of a larger cohort of Greyhounds revealed that the variant is segregating in the breed and that this breed might benefit from genetic testing to avoid carrier x carrier matings.

Keywords: skin, whole genome sequencing, dog, Canis lupus familiaris, keratinocyte, differentiation, genodermatosis, HNPK, epigenetics, chromatin

57 The epidermis, the outermost layer of the skin, undergoes a continuous process of self-renewal
58 throughout lifetime. Keratinocytes migrate from the basal layer to the stratum corneum, as they
59 differentiate. Terminally differentiated keratinocytes are present as corneocytes without nuclei
60 and organelles that form the tightly sealing outermost layer of the epidermis before being
61 sloughed from the skin surface (Candi et al. 2005; Fuchs et al 2007). The terminal
62 differentiation of keratinocytes is tightly controlled and there is increasing evidence that
63 epigenetic processes play an important role in keratinocyte differentiation (Botchkarev et al.
64 2012). Variants in genes encoding structural components, cell cycle regulators and adhesion
65 molecules involved in this process can lead to heritable skin diseases, so called
66 genodermatoses.
67 Our group previously identified the SUV39H2:XM_535179.5:c.972T>G [p.(Asn324Lys)]
68 missense variant to be causative for hereditary nasal parakeratosis (HNPK) in Labrador
69 Retrievers (Jagannathan et al. 2013). This monogenic autosomal recessive skin disorder leads
70 to crusts and fissuring of the nasal planum and has so far only been described in the Labrador
71 Retriever breed (Pagé et al. 2003; Peters et al. 2003). *SUV39H2* encodes the "suppressor of
72 variegation 3-9 homolog 2 (Drosophila)", a histone 3 lysine 9 (H3K9) methyltransferase which
73 mediates chromatin silencing (O'Carroll et al. 2000; Jenuwein & Allis 2001; Peters et al. 2001).
74 The *SUV39H2* missense variant detected in Labrador Retriever HNPK affects the catalytic site
75 and abolishes enzyme function (Schuhmacher et al. 2015). Loss of *SUV39H2* function may
76 result in delayed terminal keratinocyte differentiation, which seems a plausible
77 pathomechanism to explain the HNPK phenotype in Labrador Retrievers (Jagannathan et al.
78 2013).
79 In the present study, we investigated a Greyhound litter with eight puppies. Six puppies were
80 affected with nasal parakeratosis to varying degrees of severity (Figure 1). Only the nasal
81 planum was affected, the dogs' other skin appeared normal. Topical treatment with zinc cream
82 led to improvement of the lesions in one case but not in others. Clinically and histologically,
83 the lesions were comparable to the changes observed in Labrador Retrievers with HNPK. The
84 parents were not affected and genetic testing of the mother revealed that she was not a carrier

for the SUV39H2:c.972T>G variant causing HNPK in Labrador Retrievers (Orivet Genetics, Australia). Based on these data, we hypothesized that a novel genetic variant, inherited in a monogenic autosomal recessive mode, was responsible for the nasal parakeratosis in the investigated Greyhounds. We isolated DNA from EDTA blood from all eight puppies and submitted them for genotyping on the canine illumina HD 220 k SNP chip (Neogen/GeneSeek). The SNP genotypes are publicly available at https://www.animalgenome.org/repository/pub/BERN2017.1102/. We used these genotype data and plink version 1.07 (Purcell et al. 2007) to search for extended regions of homozygosity with allele sharing ≥ 1 Mb present in all six cases. Markers located on the sex chromosomes were excluded. Using default settings, this resulted in 48 homozygous regions totalling 122 Mb or 5% of the canine genome. The largest homozygous segment spanned ~22 Mb, was located on chromosome 2 and harboured the *SUV39H2* gene (Table S1).

We prepared a PCR-free DNA library of an affected Greyhound, collected 2 x 150 bp reads on an illumina HiSeq 3000 instrument, and re-sequenced the genome at 37x coverage (ENA project accession PRJEB16012, sample accession SAMEA104125118). Variants were called with respect to the CanFam 3.1 reference genome assembly and compared to 188 control dog and three wolf genomes as described previously (Table S2; Bauer et al. 2017). We filtered for variants that were present in homozygous state in the affected Greyhound and absent in the control genomes, assuming that the causative variant was only present in the Greyhound breed. We detected 28 variants with high, moderate or low predicted impact on protein function (Table S3). Eight were located within a shared homozygous region in the cases including a splice site variant in the *SUV39H2* gene, our primary functional candidate gene for nasal parakeratosis. This variant was a 4 bp deletion within the 5'-splice site of intron 4, XM_535179.6:c.996+3_996+6delAAGT or Chr2:21,731,812_21,731,815delACTT (CanFam 3.1).

The mutant sequence at the 5'-splice site retained the strictly conserved GT-dinucleotide at the first two bases of the intron. However, it differed in six out of nine nucleotides of the entire 5'-splice site (Figure 2). To assess the likelihood of a splice defect, we analysed the frequency

of the wildtype and mutant sequence motifs in a compilation of 186,630 human 5'-splice sites (http://katahdin.mssm.edu/splice/viewsplicemotifgraphform.cgi?database=spliceNew; Sheth et al. 2006). The canine wildtype sequence AGTgtaagt was identical to the sequence of 288 human 5'-splice sites, while the mutant sequence motif AGTgtgata did not occur in human 5'-splice sites. A splice site prediction software also clearly recognized the wildtype sequence, but not the mutant sequence (http://www.fruitfly.org/seq_tools/splice.html; Reese et al. 1997). Thus, both *in silico* analyses suggested that the mutant 5'-splice site in Greyhounds with nasal parakeratosis is non-functional.

As we did not have access to RNA from an affected dog, we experimentally assessed the functional consequence on splicing by an RNA-seq experiment on skin RNA from a heterozygous Greyhound in comparison to a homozygous wildtype Greyhound (project accession PRJEB21761, sample accessions SAMEA104393648 and SAMEA104393651). This experiment demonstrated retention of intron 4 in transcripts originating from the mutant allele in the heterozygous dog (Figure S1).

We genotyped the eight puppies for this variant by Sanger sequencing and found a perfect association with the phenotype: The six affected puppies were homozygous for the deletion and the two non-affected littermates were both heterozygous carriers. We then genotyped a larger cohort of Greyhounds as well as 483 control dogs of different breeds for this variant (Table S4). As expected, the variant was absent from all tested breeds other than the Greyhounds. In our cohort of 420 Greyhounds, which were not closely related to the investigated litter, we did not detect any homozygous mutant dogs, but found eight additional heterozygous dogs. These data indicate a carrier frequency of roughly 2% in the Greyhound population.

In light of the previous knowledge on the SUV39H2:c.972C>T variant in Labrador Retrievers with HNPK, we think that our data strongly suggest that SUV39H2:c.996+3_996+6delAAGT is the causative genetic variant underlying the observed nasal parakeratosis in Greyhounds. The Greyhound breed might benefit from genetic testing to avoid future carrier x carrier matings.

Acknowledgements

The authors would like to thank the dog owners for donating samples and pictures and for sharing information of their dogs. The authors also wish to thank Nathalie Besuchet Schmutz, Muriel Fragnière, and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform and the Interfaculty Bioinformatics Unit of the University of Bern are acknowledged for performing the whole genome re-sequencing experiments and providing high performance computing infrastructure. We acknowledge collaborators of the Dog Biomedical Variant Database Consortium (DBVDC), Gus Aguirre, Catherine André, Danika Bannasch, Doreen Becker, Cord Drögemüller, Kari Ekenstedt, Kiterie Faller, Oliver Forman, Steve Friedenberg, Eva Furrow, Urs Giger, Christophe Hitte, Marjo Hytönen, Vidhya Jagannathan, Tosso Leeb, Hannes Lohi, Cathryn Mellersh, Jim Mickelson, Leonardo Murgiano, Anita Oberbauer, Sheila Schmutz, Jeffrey Schoenebeck, Kim Summers, Frank van Steenbeck, Claire Wade for sharing dog genome sequence data from control dogs and wolves. This study was supported by grants from the Swiss National Science Foundation (CRSII3_160738 / 1) and the Albert-Heim Foundation (no. 105).

References

Bauer A., Waluk D.P., Galichet A., Timm K., Jagannathan V., Sayar B.S., Wiener D.J., Dietschi E., Müller E.J., Roosje P., Welle M.M. & Leeb T. (2017) A *de novo* variant in the *ASPRV1* gene in a dog with ichthyosis. PLoS Genet 13, e1006651.

Botchkarev V.A., Gdula M.R., Mardaryev A.N., Sharov A.A. & Fessing M.Y. (2012) Epigenetic regulation of gene expression in keratinocytes. J Invest Dermatol 132, 2505-21.

Candi E., Schmidt R. & Melino G. (2005) The cornified envelope: a model of cell death in the skin. Nat Rev Mol Cell Biol 6, 328-40.

Fuchs E. (2007) Scratching the surface of skin development. Nature 445, 834-42.

Jagannathan V., Bannoehr J., Plattet P., Hauswirth R., Drögemüller C., Drögemüller M., Wiener D.J., Doherr M., Owczarek-Lipska M., Galichet A., Welle M.M., Tengvall K., Bergvall K., Lohi H., Rüfenacht S., Linek M., Paradis M., Muller E.J., Roosje P. & Leeb T. (2013) A

mutation in the *SUV39H2* gene in Labrador Retrievers with hereditary nasal parakeratosis (HNPK) provides insights into the epigenetics of keratinocyte differentiation. PLoS Genet 9, e1003848.

Jenuwein T., Allis C.D. (2001) Translating the histone code. Science 293, 1074-80.

O'Carroll D., Scherthan H., Peters A.H., Opravil S., Haynes A.R., Laible G., Rea S., Schmid M., Lebersorger A., Jerratsch M., Sattler L., Mattei M.G., Denny P., Brown S.D., Schweizer D. & Jenuwein T. (2000) Isolation and characterization of Suv39h2, a second histone H3 methyltransferase gene that displays testis-specific expression. Mol Cell Biol 20, 9423-33.

Pagé N., Paradis M., Lapointe J.M. & Dunstan R.W. (2003) Hereditary nasal parakeratosis in Labrador Retrievers. Vet Dermatol 14, 103-10.

Peters A.H., O'Carroll D., Scherthan H., Mechtler K., Sauer S., Schofer C., Weipoltshammer K., Pagani M., Lachner M., Kohlmaier A., Opravil S., Doyle M., Sibilia M. & Jenuwein T. (2001) Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107, 323-37.

Peters J., Scott D.W., Erb H.N. & Miller W.H. (2003) Hereditary nasal parakeratosis in Labrador retrievers: 11 new cases and a retrospective study on the presence of accumulations of serum ('serum lakes') in the epidermis of parakeratotic dermatoses and inflamed nasal plana of dogs. Vet Dermatol 14, 197-203.

Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M. A, Bender D., Maller J., P. Sklar, de Bakker P. I., Daly M. J., & Sham P.C. (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 81, 559-575.

Reese M. G., Eeckman F. H., Kulp D. & Haussler D. (1997) Improved splice site detection in Genie. J Comp Biol 4, 311-23.

Schuhmacher M.K., Kudithipudi S., Kusevic D., Weirich S. & Jeltsch A. (2015) Activity and specificity of the human SUV39H2 protein lysine methyltransferase. Biochim Biophys Acta 1849, 55-63.

194 Sheth N., Roca X., Hastings M.L., Roeder T., Krainer A.R. & Sachidanandam R. (2006)
195 Comprehensive splice-site analysis using comparative genomics. Nucleic Acids Res 34,
196 3955-67.

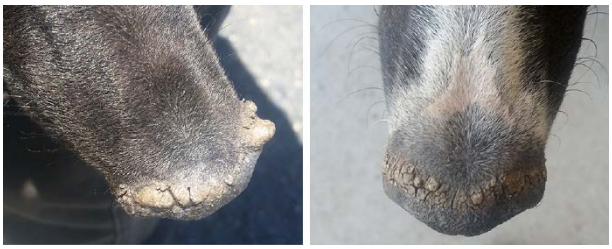

Figure 1. Nasal parakeratosis in two male Greyhounds of the same litter. Note the varying severity of the lesions affecting exclusively the nasal planum. The affected dogs did show any other clinical signs than the changes of the nose.

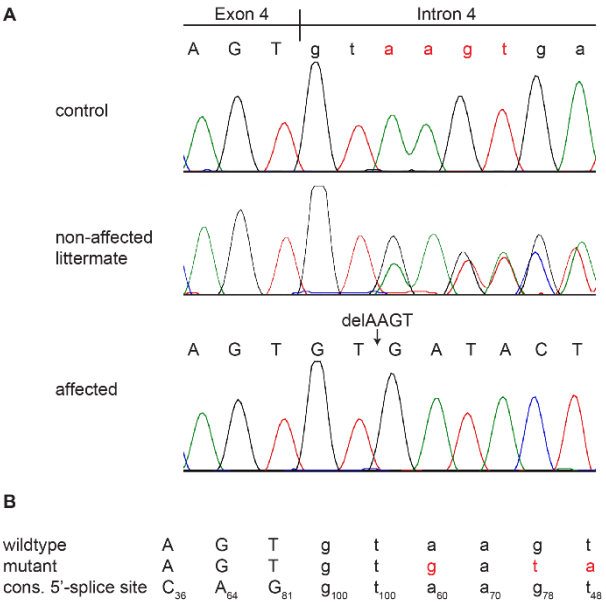

Figure 2. Electropherograms showing the *SUV39H2*:c.996+3_996+6delAAGT variant and its impact on the splice donor motif. (A) Electropherograms showing the wildtype splice site in a control dog, a non-affected littermate with a wt/del genotype and an affected Greyhound homozygous for the deletion. The four nucleotides deleted in the mutant allele are shown in red in the wildtype sequence. (B) Wildtype and mutant allele compared to the consensus sequence for the human U2 GT-AG type 5'-splice sites (Sheth et al. 2006). Subscript numbers in the consensus sequence indicate the percentage of the respective conserved nucleotide in 183,682 investigated human 5'-splice site motifs of the U2 GT-AG type. Note that the wildtype sequence already deviates from the perfect consensus sequence at 3 of the 9 positions. The mutant sequence differs at 6 of the 9 nucleotides in the U1 spliceosomal RNA recognition site. The additional differences to the optimal consensus are highlighted in red.

Supplementary Material

Figure S1. *SUV39H2* splice defect.

Table S1. Homozygous intervals ≥ 1 Mb shared between 6 cases.

Table S2. Information on 192 dog/wolf genome sequences.

Table S3. Private variants detected by whole genome sequencing.

Table S4. Genotypes of 420 Greyhounds and 483 control dogs from 65 various other breeds.