Short Communication

Whole genome sequencing reveals a novel deletion variant in the *KIT* gene in horses with white spotted coat colour phenotypes

Nicole Dürig[1,2], Rony Jude[1,2,3], Heather Holl[4,5], Samantha A. Brooks[4], Christa Lafayette[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] RJC, 53919 Weilerswist, Germany

[4] Department of Animal Sciences, University of Florida, Gainesville, FL 32611-0910, USA

[5] Etalon Inc., Menlo Park, CA 94025, USA

Running title: Equine *KIT* deletion

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

White spotting phenotypes in horses can range in severity from the common white markings up to completely white horses. *EDNRB*, *KIT*, *MITF*, *PAX3,* and *TRPM1* represent known candidate genes for such phenotypes in horses. For the present study, we re-investigated a large horse family segregating a variable white spotting phenotype, for which conventional Sanger sequencing of the candidate genes' individual exons had failed to reveal the causative variant. We obtained whole genome sequence data from an affected horse and specifically searched for structural variants in the known candidate genes. This analysis revealed a heterozygous ~1.9 kb deletion spanning exons 10-13 of the *KIT* gene (chr3:77,740,239_77,742,136del1898insTATAT). In continuity with previously named equine *KIT* variants we propose to designate the newly identified deletion variant *W22*. We had access to 21 horses carrying the *W22* allele. Four of them were compound heterozygous *W20/W22* and had a completely white phenotype. Our data suggest that *W22* represents a true null allele of the *KIT* gene, while the previously identified *W20* leads to a partial loss of function. These findings will enable more precise genetic testing for depigmentation phenotypes in horses.

Keywords: Equus caballus; melanocyte; pigmentation; coat colour; structural variant; *KIT*; heterogeneity

Many domesticated animals show white spotting phenotypes, which can range from tiny unpigmented spots up to completely white animals. These phenotypes are the result of an altered embryonic development of the neural crest derived melanocyte lineage ("leucism"). Candidate genes for such phenotypes include *EDNRB*, *KIT*, *MITF, PAX3*, and *TRPM1* (OMIA 000629, 000209, 001688, 000214, 001341).

Prior to the elucidation of the molecular mechanisms it has been recognized that semi-dominant genetic variants can lead to white horses in heterozygous state and embryonic lethality in homozygous state (Pulos & Hutt, 1969). Among domesticated animals the horse currently represents the species with the largest number of molecularly defined white spotting alleles. A missense variant in the *EDNRB* gene causes the frame overo white spotting pattern or lethal white foal syndrome, if present in heterozygous or homozygous state, respectively (Santschi *et al.* 1998). Variants in *MITF* and *PAX3* cause the so-called splashed white phenotype, which closely resembles human Waardenburg syndrome (Hauswirth *et al.* 2012; Hauswirth *et al.* 2013). A variant at the equine *TRPM1* gene causes the so-called leopard complex spotting (Bellone *et al.* 2013). Furthermore, according to our knowledge, 23 different functional alleles at the equine *KIT* gene have been described so far. These include the alleles for sabino-1 and tobiano spotting, and an allelic series termed *W1* - *W21* (Brooks & Bailey 2005; Brooks *et al.* 2007; Haase *et al.* 2007; Haase *et al.* 2008; Haase *et al.* 2009; Holl *et al.* 2010; Haase *et al.* 2011; Hauswirth *et al.* 2013; Haase *et al.* 2016; Table S1).

Our earlier analyses of horses with white spotting phenotypes involved the individual PCR amplification of candidate genes' exons followed by subsequent Sanger sequencing to identify potentially causative genetic variants. During these analyses, we encountered a family of horses segregating a dominant white spotting phenotype that resembled the phenotypes of other horses with *KIT* variants. However, our Sanger sequencing-based analysis of the *KIT* gene failed to identify the causative variant (Haase *et al.* 2007).

New variant in the *KIT* gene

We re-sequenced the genome of a white-spotted Thoroughbred horse at 32x coverage using 2 x 150 bp reads on an Illumina HiSeq 3000 instrument. Sequencing and read mapping to the EquCab 2 reference assembly was performed as previously described (Murgiano et al. 2016). Data were deposited at the European Nucleotide Archive (study accession PRJEB14779, sample ERS1451611). This analysis again failed to reveal any single nucleotide or small indel variants in the candidate genes *EDNRB*, *KIT*, *MITF, PAX3*, and *TRPM1*. In order to search for large structural variants we visually inspected the read alignments for these genes in the Integrative Genomics Viewer (IGV; Robinson *et al.* 2011).

This experiment revealed a large heterozygous deletion spanning exons 10 – 13 of the *KIT* gene in the white spotted horse. Specifically, 1,898 bp were missing and replaced by 5 bp not present in the reference genome (chr3:77,740,239_77,742,136del1898insTATAT). Assuming that the remaining exons are correctly spliced, the deletion allele was predicted to maintain the reading frame and give rise to a transcript lacking 450 nucleotides encoding 150 amino acids including the transmembrane domain (c.1529_1978del; p.Glu510_Gly659del). We designed primers flanking the deletion (W22_F: CACCTGCGTTCTGAGCACTA, W22_R: CCAAGGCAGGAGTTTTGTTG) and confirmed the breakpoints of the deletion by PCR and Sanger sequencing. A PCR with these primers and SequalPrep long range polymerase (ThermoFisher) was used as diagnostic assay to genotype additional horses for the deletion. This PCR yielded two bands of 349 bp and 2,242 bp in horses that were heterozygous for the deletion, whereas only the longer band was amplified in horses with the wildtype genotype (Figure 1). In line with previous nomenclature, we propose to term this new *KIT* allele as *W22*. We did not find any other obvious structural variants in the *EDNRB*, *MITF, PAX3*, and *TRPM1* genes.

Interaction of *W22* with *W20*

We identified a total of 21 horses carrying the *W22* allele. Their phenotypes were quite variable and ranged from a solid-coloured body with a broad blaze and high white legs up to completely white horses (Figure 2). The quantitative degree of white spotting may be seen as a trait with complex inheritance. Previous work has shown that numerous coding and non-coding variants at several genes have an influence on depigmentation. The *W22* allele probably represents a null allele at the *KIT* gene. Thus, horses carrying one copy of the *W22* allele are particularly well suited to study the genotype-phenotype correlation of their remaining second copy of the *KIT* gene. We previously described the *W20* allele (*KIT*:p.Arg682His), which probably leads to only a minor reduction in KIT function (Hauswirth *et al.* 2013). *W20* is a very common allele segregating in many diverse horse breeds. The *W22* variant was found on a haplotype with the *W20* allele (Figure S1). We determined the *W20* genotypes in all 21 horses carrying the *W22* variant (Table S2). Four of them were homozygous for *W20* whereas 17 did not carry the *W20* variant on their second (functional) *KIT* haplotype. All four horses with the *W20/W22* genotype were completely white. The other 17 horses showed variable phenotypes ranging from ~15% up to 100% depigmented body surface (Figure 2). The proportion of depigmentation was significantly different between *W22/+* horses (mean 62%) and *W20/W22* horses (mean 100%, $p = 0.0003$). Thus, our data confirm the earlier reported “white increasing” effect of the *W20* allele while at the same time suggesting that other unknown genetic variants with similar effects may explain the pronounced white spotting phenotypes of some horses

with the *W22/+* genotype. Another independent study very recently also replicated the association of *W20* with white facial markings and total white markings in Spanish horses (Negro *et al.* 2017).

In conclusion, our study revealed a new structural variant at the *KIT* gene and added new evidence for the subtle effect of the *KIT W20* allele on pigmentation. The quantitative extent of depigmentation is a complex trait influenced by multiple genetic and possibly also stochastic/environmental factors.

Acknowledgements

The authors would like to thank all involved horse owners for donating samples and pictures and for sharing pedigree information of their horses. The authors also wish to thank Nathalie Besuchet, Muriel Fragnière, and Sabrina Schenk for expert technical assistance. The Next Generation Sequencing Platform of the University of Bern is acknowledged for performing the whole genome re-sequencing experiments and the Vital-IT high-performance computing center of the Swiss Institute of Bioinformatics for performing computationally-intensive tasks (http://www.vital-it.ch/). This study was supported by a grant from the Swiss National Science Foundation. Conflicts of interest: Rony Jude is an equine consultant. Heather Holl and Christa Lafayette are affiliated with a genetic testing laboratory offering tests for white spotting in horses.

References

Adzhubei I.A. Schmidt S., Peshkin L., Ramensky V.E., Gerasimova A., Bork P., Kondrashov A.S. & Sunyaev S.R. (2010) A method and server for predicting damaging missense mutations. *Nature Methods* **7**, 248-9.

Bellone RR, Holl H, Setaluri V, Devi S, Maddodi N, Archer S, Sandmeyer L, Ludwig A, Foerster D, Pruvost M, Reissmann M, Bortfeldt R, Adelson DL, Lim SL, Nelson J, Haase B, Engensteiner M, Leeb T, Forsyth G, Mienaltowski MJ, Mahadevan P, Hofreiter M, Paijmans JL, Gonzalez-Fortes G, Grahn B, Brooks SA (2013) Evidence for a retroviral insertion in *TRPM1* as the cause of congenital stationary night blindness and leopard complex spotting in the horse. *PLoS One* **8**, e78280.

Brooks S. A. & Bailey E. (2005) Exon skipping in the *KIT* gene causes a Sabino spotting pattern in horses. *Mammalian Genome* **16**, 893-902.

Brooks S.A., Lear T.L., Adelson D.L. & Bailey E. (2007) A chromosome inversion near the KIT gene and the Tobiano spotting pattern in horses. *Cytogenetic & Genome Research* **119**, 225-30.

Ferrer-Costa C., Gelpí J.L., Zamakola L., Parraga I., de la Cruz X. & Orozco M. (2005) PMUT: a web-based tool for the annotation of pathological mutations on proteins. *Bioinformatics* **21**, 3176-8.

Haase B., Brooks S.A., Schlumbaum A., Azor P.J., Bailey E., Alaeddine F., Mevissen M., Burger D., Poncet P.-A., Rieder S. & Leeb T. (2007) Allelic heterogeneity at the equine *KIT* locus in dominant white (*W*) horses. *PLoS Genet* **3**, e195.

Haase B., Jude R., Brooks S.A. & Leeb T. (2008) An equine chromosome 3 inversion is associated with the tobiano spotting pattern in German horse breeds. *Animal Genetics* **39**, 306-9.

Haase B., Brooks S.A., Tozaki T., Burger D., Poncet P.-A., Rieder S., Hasegawa T., Penedo C. & Leeb T. (2009) Seven novel *KIT* mutations in horses with white coat colour phenotypes. *Animal Genetics* **40**, 623-9.

Haase B., Rieder S., Tozaki T., Hasegawa T., Penedo M.C., Jude R. & Leeb T. (2011) Five novel *KIT* mutations in horses with white coat colour phenotypes. *Animal Genetics* **42**, 337-9.

Hauswirth R., Haase B., Blatter M., Brooks S.A., Burger D., Drögemüller C., Gerber V., Henke D., Janda J., Jude R., Magdesian K.G., Matthews J.M., Poncet P.-A., Svansson V., Tozaki T., Wilkinson-White L., Penedo M.C., Rieder S. & Leeb T. (2012) Mutations in *MITF* and *PAX3* cause "splashed white" and other white spotting phenotypes in horses. *PLoS Genetics* **8**, e1002653.

Hauswirth R., Jude R., Haase B., Bellone R.R., Archer S., Holl H., Brooks S.A., Tozaki T., Penedo M.C.T., Rieder S. & Leeb T. (2013) Novel variants in the *KIT* and *PAX3* genes in horses with white-spotted coat colour phenotypes. *Animal Genetics* **44**, 763-5.

Holl H., Brooks S. & Bailey E. (2010) *De novo* mutation of *KIT* discovered as a result of a non-hereditary white coat colour pattern *Animal Genetics* **41 Suppl 2**, 196-198.

Negro S., Imsland F., Valera M., Molina A., Solé M. & Andersson L. (2017) Association analysis of *KIT*, *MITF*, and *PAX3* variants with white markings in Spanish horses. *Animal Genetics* doi: 10.1111/age.12528.

Murgiano L., Waluk D.P., Towers R., Wiedemar N., Dietrich J., Jagannathan V., Drögemüller M., Balmer P., Druet T., Galichet A., Penedo M.C., Müller E.J., Roosje P., Welle M.M. & Leeb T. (2016) An intronic MBTPS2 variant results in a splicing defect in horses with brindle coat texture. *G3 (Bethesda)* **6**, 2963-70.

Pingault V., Ente D., Dastot-Le Moal F., Goossens M., Marlin S. & Bondurand N. (2010) Review and update of mutations causing Waardenburg syndrome. *Human Mutation* **31**, 391-406.

Pulos W.L. & Hutt F.B. (1969) Lethal dominant white in horses. *Journal of Heredity* **60**, 59-63.

Robinson J.T., Thorvaldsdóttir H., Winckler W., Guttman M., Lander E.S., Getz G. & Mesirov J.P. (2011) Integrative genomics viewer. *Nature Biotechnology* **29**, 24-6.

Santschi E.M., Purdy A.K., Valberg S.J., Vrotsos P.D., Kaese H. & Mickelson J.R. (1998) Endothelin receptor B polymorphism associated with lethal white foal syndrome in horses. *Mammalian Genome* **9**, 306-9.

Thomas A.J. & Erickson C.A. (2008) The making of a melanocyte: the specification of melanoblasts from the neural crest. *Pigment Cell & Melanoma Research* **21**, 598-610.

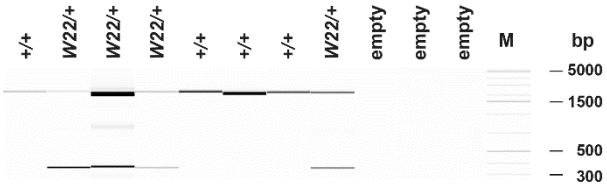

Figure 1 Long-range PCR to genotype the *W22* deletion. Genomic DNA samples from 8 horses were amplified with primers W22_F and W22_R. The PCR products were separated on a FragmentAnalyzer™ capillary gel electrophoresis instrument. The wildtype allele gives a band of 2,242 bp, the *W22* allele gives a band of 349 bp. The genotypes of the eight tested horses are indicated above the respective lanes.

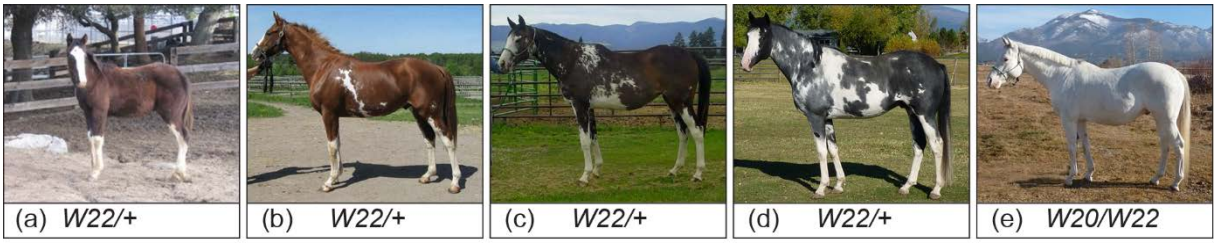

Figure 2 Coat colour phenotypes of horses carrying the *W22* allele. (a-e) represent horses with an increasing proportion of depigmented skin areas. Note that the horse shown in (e) is compound heterozygous *W20/W22* at the *KIT* gene. All *W20/W22* horses in our study showed a completely white coat colour phenotype whereas the degree of depigmentation in animals with the *W22/+* genotype varied between ~15% and 100%.

Supplementary Material:

Figure S1 Illustration of the variants and haplotypes.

Table S1 Compilation of variants with an influence on depigmentation.

Table S2 Horses, phenotypes and their genotypes.