

1 **Identification of a novel missense variant in *SLC45A2* associated with dilute snowdrop**
2 **phenotype in Gypsy horses**

3

4 D. Bisbee*, M. L. Carpenter*, K. Hoefs-Martin*, S. A. Brooks*†, and C. Lafayette*

5 *Etalon Inc, Menlo Park, CA 94025, USA. †Department of Animal Sciences, University of Florida,

6 Gainesville, FL 32607, USA.

7

8 **SUMMARY**

9 *SLC45A2* mutations are responsible for several dilution phenotypes in horses, including *cream*,
10 *pearl*, and *sunshine*. We sequenced the *SLC45A2* gene in a horse that possessed a diluted coat
11 color but tested negative for any known dilution genotypes. We identified a novel homozygous
12 missense mutation in *ECA21: SLC45A2:c.305G>A; p.(Arg102Gln)* which we have named
13 *snowdrop* (C^{sno}). The *snowdrop* dilution is autosomal recessive and dilutes both red and black
14 pigment in the homozygous state, creating a phenotype similar to homozygous *cream*. We also
15 identified this mutation in the heterozygous state in several relatives, including the sire and dam
16 of the affected horse, where it has no visible effect on phenotype. Based on these data, the
17 *snowdrop* variant produces a recessive dilution similar to *cream*, and its discovery confirms that
18 multiple *SLC45A2* alleles cause dilute phenotypes in horses.

19 -----

20

21 In many animals, *SLC45A2* encodes a transmembrane protein in melanocytes that is
22 hypothesized to transport molecules necessary for proper melanosome function. Coat and skin
23 color dilution phenotypes caused by variants in *SLC45A2* are documented in multiple vertebrate

24 species¹⁻⁴. In equines, mutations in the *SLC45A2* gene, located on chromosome 21, are
25 responsible for the *cream*, *pearl*, and *sunshine* dilutions^{5,6}.

26 In this study, we focused on a Gypsy breed mare that appeared phenotypically cremello
27 (light creamy white) but tested negative for any known dilution or spotting alleles, including
28 sabino, W20, tobiano, PATN1, and SW1-4 (Figure 1). We obtained hair samples from 15 Gypsy
29 horses, including the affected mare, her dam and sire, and three half siblings. DNA was
30 extracted according to a previously reported protocol⁷, and genotyping of known alleles was
31 performed using an Agena MassArray assay. PCR and Sanger sequencing of the seven exons in
32 the *SLC45A2* gene in the affected mare were performed as previously described⁶. Non-
33 synonymous variants were analyzed using the PROVEAN web server⁸. Ensembl protein
34 sequences for orthologs of *SLC45A2* in 106 mammals were downloaded from OrthoDB and
35 aligned using clustal omega to assess the conservation of each amino acid change^{9,10}.

36 Sanger sequencing identified three missense mutations in *SLC45A2* in the affected mare.
37 Two variants, *SLC45A2:c.32C>A* and *SLC45A2:c.872C>A* (numbering based on transcript
38 ENSECAT00000026240.1 in the Ensembl EquCab2.0 assembly), had positive PROVEAN
39 scores but were in a location not conserved across vertebrates and were not investigated further.
40 Variant *SLC45A2:c.305G>A; p.(Arg102Gln)* in exon 1 was detected in the homozygous state in
41 the affected mare, and in the heterozygous state in the sire, dam, and a half-sibling of the mare
42 (Table 1), and was termed *snowdrop* (C^{sno}). This mutation was not identified in the other two
43 half-siblings (Table 1) or in any of the additional nine Gypsy horses tested. Only the
44 homozygous mare displayed a diluted phenotype. The sequence containing this variant was
45 deposited in Genbank under accession number MN704281.

46 This C^{sno} mutation at arginine 102 is predicted to be deleterious, with a PROVEAN score
47 of -3.661 . In an alignment of the *SLC45A2* protein sequence from 106 mammalian species, the
48 horse reference allele at this position was conserved across 100% of species analyzed, also
49 suggesting that this variant likely has an impact on protein function. The topography of the
50 *SLC45A2* transmembrane protein is shown in Fig. S1 as predicted with PHOBIUS protein
51 prediction software¹¹; the location of the *snowdrop* allele as well as several other *SLC45A2*
52 dilutions are shown. The C^{sno} mutation replaces the highly conserved positively charged arginine
53 residue with a polar glutamine. Cytoplasmic regions of transmembrane proteins tend to be
54 positively charged, a phenomenon termed the “positive-inside rule”, and disruptions in this
55 positive charge can affect the topology and assembly of these proteins^{12,13}. Thus, the *snowdrop*
56 mutation is likely to impact the function of the *SLC45A2* protein as a membrane-associated
57 transporter.

58 It is unclear whether the *snowdrop* mutation affects eumelanin or pheomelanin, but it is
59 suspected to dilute eumelanin in the homozygous form. Homozygous C^{sno}/C^{sno} appears to dilute
60 black pigment in both the mane and tail, and homozygous horses have pink around the nose and
61 white eyelashes (Figure 1). Heterozygous C^{sno} appears to have no effect on phenotype; thus, C^{sno}
62 appears to be an autosomal recessive mutation. At this time, it is unknown if the *snowdrop*
63 dilution has any deleterious effects. Testing across additional Gypsy horses could possibly
64 determine this new variant's origin and allele frequency.

65

66 **Conflict of Interest**

67 All of the authors are affiliated with Etalon Diagnostics, which offers testing of the C^{sno} variant.

68

69 Acknowledgements

70 The authors would like to thank Clare Brennan, Kari Newman, and the other owners who
71 provided samples for this study. Photos were provided by Clare Brennan.

72

73 References

- 74 1 Cho, Y. S. *et al.* The tiger genome and comparative analysis with lion and snow leopard
75 genomes. *Nat Commun* **4**, 2433, doi:10.1038/ncomms3433 (2013).
- 76 2 Du, J. & Fisher, D. E. Identification of Aim-1 as the underwhite mouse mutant and its
77 transcriptional regulation by MITF. *J Biol Chem* **277**, 402-406,
78 doi:10.1074/jbc.M110229200 (2002).
- 79 3 Gunnarsson, U. *et al.* Mutations in SLC45A2 cause plumage color variation in chicken
80 and Japanese quail. *Genetics* **175**, 867-877, doi:10.1534/genetics.106.063107 (2007).
- 81 4 Prado-Martinez, J. *et al.* The genome sequencing of an albino Western lowland gorilla
82 reveals inbreeding in the wild. *BMC Genomics* **14**, 363, doi:10.1186/1471-2164-14-363
83 (2013).
- 84 5 Mariat, D., Taourit, S. & Guerin, G. A mutation in the MATP gene causes the cream coat
85 colour in the horse. *Genet Sel Evol* **35**, 119-133, doi:10.1051/gse:2002039 (2003).
- 86 6 Holl, H. M. *et al.* A candidate gene approach identifies variants in SLC45A2 that explain
87 dilute phenotypes, pearl and sunshine, in compound heterozygote horses. *Anim Genet* **50**,
88 271-274, doi:10.1111/age.12790 (2019).
- 89 7 Cook, D., Gallagher, P. C. & Bailey, E. Genetics of swayback in American Saddlebred
90 horses. *Anim Genet* **41 Suppl 2**, 64-71, doi:10.1111/j.1365-2052.2010.02108.x (2010).
- 91 8 Choi, Y. & Chan, A. P. PROVEAN web server: a tool to predict the functional effect of
92 amino acid substitutions and indels. *Bioinformatics* **31**, 2745-2747,
93 doi:10.1093/bioinformatics/btv195 (2015).
- 94 9 Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence
95 alignments using Clustal Omega. *Mol Syst Biol* **7**, 539, doi:10.1038/msb.2011.75 (2011).
- 96 10 Zdobnov, E. M. *et al.* OrthoDB v9.1: cataloging evolutionary and functional annotations
97 for animal, fungal, plant, archaeal, bacterial and viral orthologs. *Nucleic Acids Res* **45**,
98 D744-D749, doi:10.1093/nar/gkw1119 (2017).

- 99 11 Kall, L., Krogh, A. & Sonnhammer, E. L. Advantages of combined transmembrane
100 topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res* **35**,
101 W429-432, doi:10.1093/nar/gkm256 (2007).
- 102 12 Baker, J. A., Wong, W. C., Eisenhaber, B., Warwicker, J. & Eisenhaber, F. Erratum to:
103 Charged residues next to transmembrane regions revisited: "Positive-inside rule" is
104 complemented by the "negative inside depletion/outside enrichment rule". *BMC Biol* **15**,
105 72, doi:10.1186/s12915-017-0410-6 (2017).
- 106 13 von Heijne, G. Control of topology and mode of assembly of a polytopic membrane
107 protein by positively charged residues. *Nature* **341**, 456-458, doi:10.1038/341456a0
108 (1989).
- 109
- 110
- 111
- 112
- 113
- 114
- 115

116

117 **Table 1: Genotypes and phenotypes of horses tested in this study**

Horse	Genotypic coat color	Sno genotype	Hypothetical coat color	Phenotypic coat color
Affected mare	aa Ee W20/n	C ^{sno} /C ^{sno}	Black w/ white markings	Light creamy white
Sire	Aa EE W20/W20	C ^{sno} /n	Bay w/ white markings	Bay
Dam	Aa Ee Gg W20/n	C ^{sno} /n	Grey w/ white markings	Grey
Half-sibling 1	aa Ee W20/n TO/n	C ^{sno} /n	Black w/ white markings	Black Tobiano
Half-sibling 2	Aa Ee CR/n W20/n TO/n	n/n	Buckskin w/ white markings	Buckskin Tobiano
Half-sibling 3	aa Ee W20/n	n/n	Black w/ white markings	Black

118

119



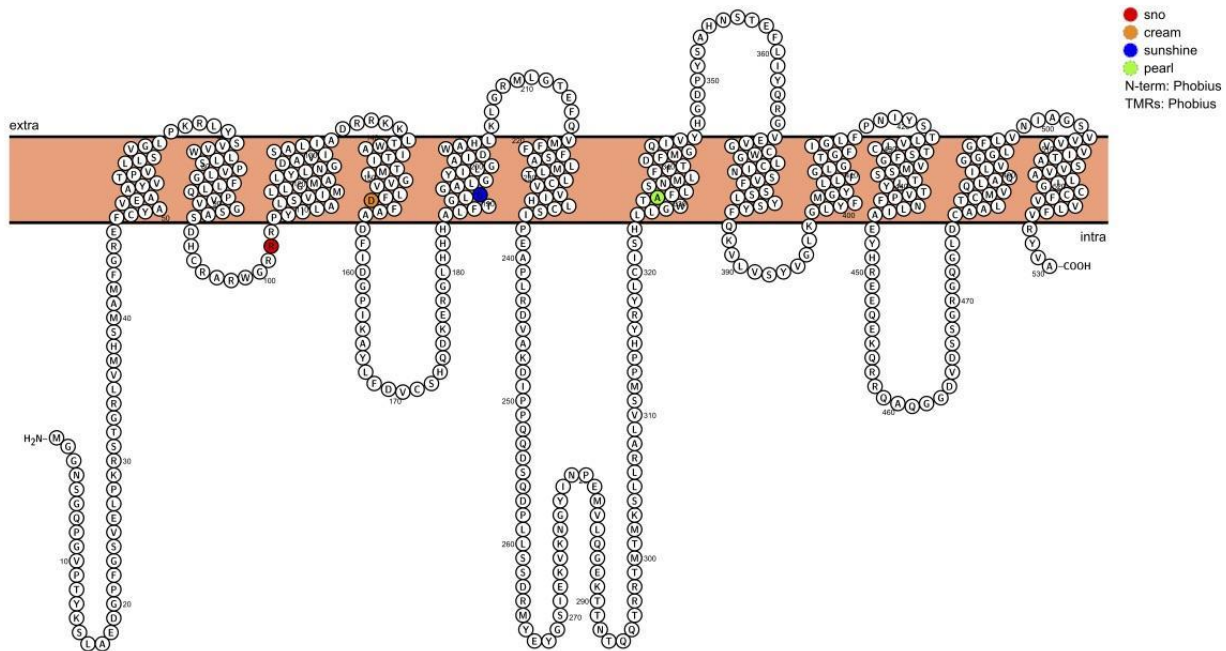
120

121 **Figure 1.** Phenotypes & genotypes of different C^{sno} positive horses. A and B are the dam and
122 sire, respectively, of C, which is homozygous for the C^{sno} mutation. B is also the sire of D and
123 therefore half-sibling to C.

124

125

126



127

128

129 **Figure S1:** Topography of equine SLC45A2 with transmembrane regions determined using

130 PHOBIUS protein prediction software. The *snowdrop* allele (Red) replaces a positively charged

131 arginine residue with a neutral polar glutamine residue. The locations of previously identified

132 mutations resulting in the *cream* (orange), *sunshine* (blue), and *pearl* (green) dilutions are shown

133 for reference.

134