

Identification of genomic regions associated with differences in fleece type in Huacaya and Suri alpacas (*Vicugna pacos*)

Abstract

The difference in fleece type is the distinguishing trait between the two types of alpacas (*Vicugna pacos*), Huacaya and Suri. The Suri fleece type has been found to be inherited dominantly over the Huacaya type, resulting in offspring with the Suri phenotype. The aim of our study was to map genomic regions associated with the two different fleece types. In this study, 91 alpacas (54 Huacayas and 37 Suris) from Germany and Switzerland were genotyped using the 76k alpaca SNP array. Only 59k chromosome-localised markers map to the alpaca reference assembly VicPac3.1, and after quality control 49 866 SNPs, were retained for population structure assessment and to conduct a genome-wide association study. Both principal component and neighbour-joining tree analysis showed that the two fleece-type cohorts overlapped rather than forming two distinct clusters. Genome-wide significantly associated markers were observed in the scaffold region of chromosome 16 (NW_021964192.1), which contains a cluster of keratin genes. A haplotype predominantly found in Suri alpacas has been identified which supports dominant inheritance. Variant filtering of nine whole-genome sequenced alpacas from both fleece types in the critical interval of 0.4 Mb did not reveal perfect segregation of either fleece type for specific variants. To our knowledge, this is the first study to use the recently developed species-specific SNP array to identify genomic regions associated with differences in fleece type in alpacas. There are still some limitations, such as the preliminary status of the reference assembly and the incomplete annotation of the alpaca genome.

Alpaca fibre has grown in popularity outside of South America in recent decades owing to its use in the production of high quality fleece. Alpaca fibre is preferred for its unique thermoregulatory, water repellent, flame retardant and hypoallergenic properties. The appearance of the fleece is the main trait that distinguishes the

two fleece types, Huacaya and Suri (Figure 1a) are the alpaca breeds of this domestic South American Camelid species (*Vicugna pacos*). The differences in texture and fibre characteristics between the two fleece types allow them to be used only for specific textile products. The elasticity of Huacaya fleece makes it suitable for knitted products. Suri fleece is too fine and lacks elasticity, making it more suitable for woven products (Frank et al., 2006). The Suri phenotype has been observed to be dominant over Huacaya (Ponzoni et al., 2003). Breeding results do not fully support a single-gene model of inheritance. According to Sponenberg (2010), as the main features, the excess of Huacaya offspring in Huacaya × Suri matings, and the rare production of Suri offspring from Huacaya × Huacaya matings is consistent with a single autosomal dominant gene that is subject to suppression by modifiers. Presciuttini et al. (2010) proposed based on statistical analysis of breeding results a genetic model with two linked loci which both need to be homozygous for recessive alleles in order to produce the Huacaya phenotype. Therefore, crossbred Huacaya and Suri alpacas might have an overall Suri phenotype but fleece type that does not fit specifically in the distinct fleece type classification of Suri or Huacaya when manufactured for textile products (Sponenberg, 2010).

The alpacas sampled in this study were phenotyped according to their fleece type from their pedigree data and pictures. Blood samples from 91 unrelated animals (54 Huacayas and 37 Suris) were collected in Germany ($n=29$) and Switzerland ($n=62$) (Table S3) for genotyping using the Illumina 76k alpaca SNP array, performed externally by NEOGEN (Ayr, Scotland). See Appendix S1 for details of genomic analysis.

A single genome-wide significant association signal was located on a scaffold (NW_021964192.1) on chromosome 16 (Figure 1b). The top-associated markers were located in a ~0.4 Mb-sized region containing a cluster of keratin genes (Figure 1c). These findings suggest an important influence of keratin on the difference between Huacaya and Suri fleece. The results from this study further confirm and refine the findings in a previous study in which *KRT15*, together with five other keratin genes

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Animal Genetics* published by John Wiley & Sons Ltd on behalf of Stichting International Foundation for Animal Genetics.

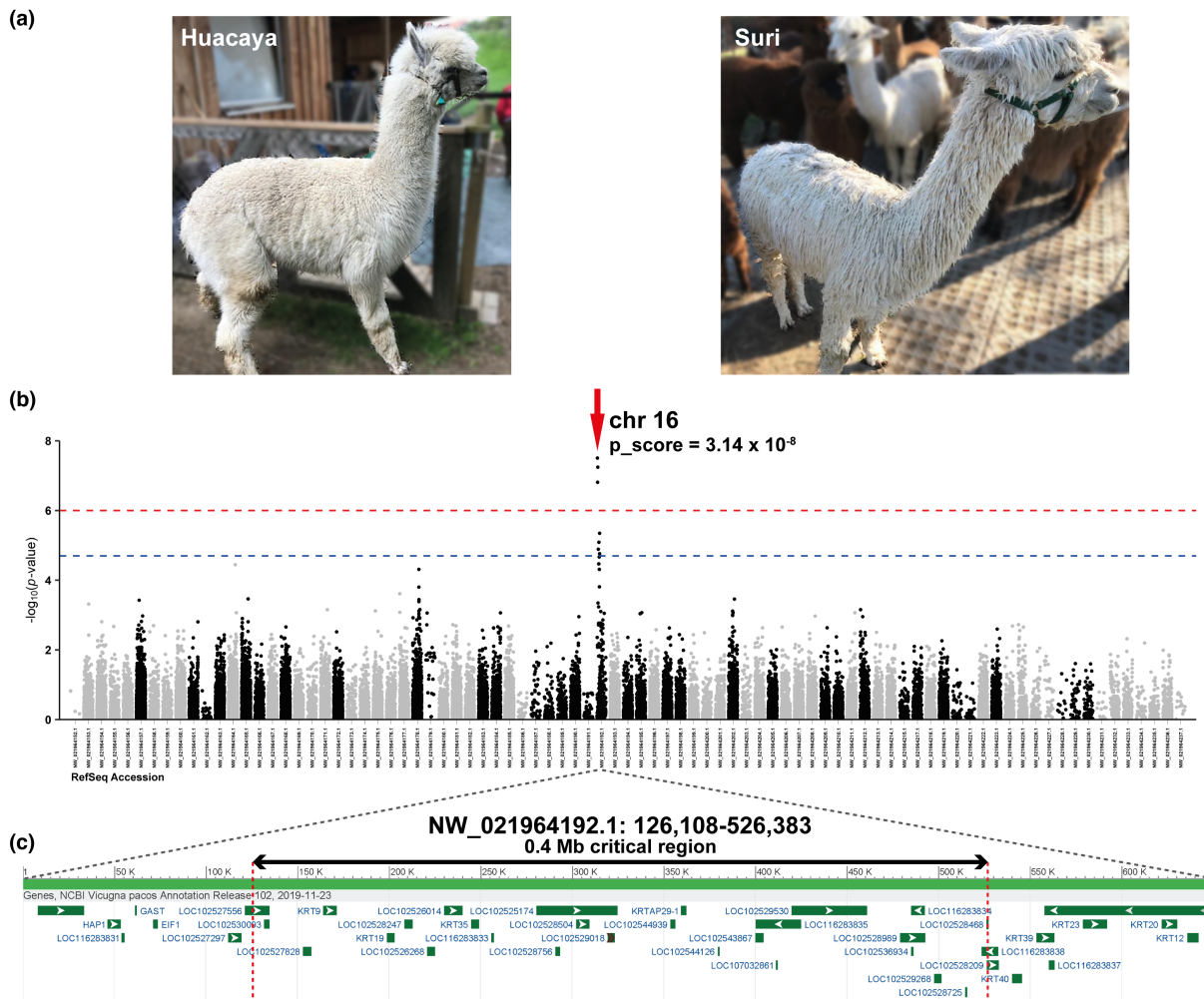


FIGURE 1 (a) Comparison between the two fleece types, Huacaya (left) and Suri (right). (b) Manhattan plot from the genome-wide association study of Huacaya vs. Suri alpacas. The red line represents Bonferroni corrected genome-wide significance threshold ($p = 1.0 \times 10^{-6}$), while the blue line corresponds to the suggestive threshold ($p = 2.0 \times 10^{-5}$). The x-axis corresponds to the scaffolds, in ascending order of chromosome number. Note that all adjacent scaffolds of one colour (grey or black) localise to the same chromosome. (c) Gene content in the 0.4 Mb critical region identified from the genome-wide association study.

No. of copies of the main associated haplotype	Two	One	Zero
Suri	11	21	5
Huacaya	1	0	53

TABLE 1 Genotypes for a Suri-associated haplotype in the mapped region on chromosome 16 in relation to fleece type in alpacas.

(*KRT31*, *KRT13*, *KRT9*, *KRT14* and *KRT16*) cytogenetically mapped to chromosome 16, making it the main target for identifying sequence variants associated with hair texture and growth (Mendoza et al., 2019). Keratins are the main protein components of hair, skin and nails. They affect the structural integrity and quality of animal fibres (Balmer et al., 2017). Variants affecting fleece type have been identified in keratin genes in other species, notably *KRT74* for fleece type in sheep (OMIA:002687-9940), *KRT71* for curly coat in dogs (OMIA:000245-9615) and cats (OMIA:001581-9685, OMIA:001712-9685), *KRT27* for karakul (tightly curled) coat type in cattle (OMIA:000246-9913) and *KRT25* for curly coat type in horses (OMIA:000245-9796).

The SNP genotypes from all 225 markers mapped to the NW_021964192.1 scaffold were phased using fastPHASE software (Scheet et al., 2006) and then ordered with respect to genome-wide significantly associated markers. This revealed the presence of a predominant haplotype in the mapped region on chromosome 16 in Suri animals (Table S1), supporting the assumption of a dominant mode of inheritance. Over 86% (32 out of 37) of the Suri alpacas carried at least one copy of this haplotype, whereas it was present in only one of the 54 Huacaya animals (Table 1). These findings affirm the genetic model of dominance proposed in an earlier study, where 588 Suri × Suri and 2126 Huacaya × Huacaya offspring were identified and

a single gene with two alleles model for an inheritance pattern of seven dominant Suri:1 recessive Huacaya was accepted based on segregation of the offspring (Renieri et al., 2008). A principal component analysis (PCA) plot and neighbour-joining (NJ) tree are presented in [Figures S1](#) and [S2](#). The results showed incomplete clustering of Huacayas and Suris according to the fleece type and country of origin. The partial overlap of the groups suggests genetic similarities between both fleece types (Presciuttini et al., 2010) and calls into question previous classification of the Huacaya and Suri as two separate breeds. Interestingly, two Suri alpacas with known Huacaya ancestry cluster with other Suris but not with Huacayas.

Since Huacaya has been identified as the wild-type variety that was domesticated from vicunas, it is likely that Suris were probably derived from Huacayas through mutation, resulting in a much thinner fleece type with reduced ability to withstand the harsh climate of high altitudes (Presciuttini et al., 2010). The clustering of samples could also be the result of different breeding practices of alpaca owners and breeders. Most breeders keep the populations separate, mainly for the purposes of producing a specific type of fleece for yarn, but there is a probability of occasional crossbreeding, especially by some owners keeping these animals only as pets. In addition, crossbreeding in previous generations, particularly of imported animals as was seen from the pedigree data of some of the sampled alpacas led to genetic introgression of Suri into Huacaya or vice versa. An example is the alpaca with a clear Suri phenotype ([Figure 1a](#)) which was found to descend in the fourth generation from a Suri dam who had two Huacaya parents. Such crossbred alpacas increase the difficulty of identifying specific variants that cause the difference in fleece type. Since Suri has been identified as the dominant fleece type, most animals that have been crossbred are expected to show the Suri phenotype, making it more difficult to determine their actual genotype.

However, the genomic region identified here associated with fleece type on alpaca chromosome 16, with functionally significant candidate genes, provides a first starting point for finding potentially causal functional variant(s) of interest, similar to what has been done for inherited body traits in other animal species (Paris et al., 2020). Therefore, exonic variant filtering was subsequently performed on nine whole-genome sequenced alpacas (four Suris and five Huacayas) from the critical interval on chromosome 16 ([Figure 1c](#)) to identify heterozygous and homozygous variants that might have a low (synonymous variant, splice region variant), moderate (missense variant, inframe indel) or high (frameshift variant, splice site variant) impact on the encoded protein. This resulted in a total of 331 variants (158 low, 149 moderate and 24 high impact) detected in the critical interval ([Table S2](#)). No consistent patterns of segregation were found for any of the variants with respect to

the different fleece types, neither for the recessive nor for the dominant mode of inheritance.

There are still certain limitations in the analysis of variants that cause the Huacaya and Suri fleece type. One important issue is the inheritance of the Suri phenotype, which seems not to be completely dominant (Presciuttini et al., 2010; Sponenberg, 2010). It is likely that at least two variants are controlling this trait (Presciuttini et al., 2010; Sponenberg, 2010), making their identification, e.g. via genome sequencing and variant filtering, much more difficult. The earlier example of the Suri animal ([Figure 1a](#)) that has Huacaya ancestry demonstrates the difficulty of grouping this animal as a 'purebred' Suri. According to the PCA plot and NJ tree, it groups more closely together with other Suris. In addition, the NJ tree revealed six Suris having a Huacaya-like genome and one Huacaya with a Suri-like genome, indicating a low frequency of the recessive Huacaya allele in the Suri population ([Figure S2](#)). The exclusion of Suris and Huacayas from the analysis with ancestors of the other phenotype may be helpful, but unfortunately we do not have such extensive or even any pedigree information for a large part of the sampled alpacas ([Table S3](#)).

In addition, SNP annotation on unplaced contigs and scaffolds, gaps in the reference assembly and incomplete annotation of the current assembly pose problems when trying to align sequences and identify specific variants. Identification of variants on keratin genes is particularly complicated, as most keratin genes for alpacas are still not well annotated and they cluster together owing to their proximity to each other and were therefore mapped as a single keratin cluster (Richardson et al., 2019). A similar situation was encountered when analysing keratin genes from dogs and horses (Balmer et al., 2017).

A recent paper examining the genome sequence data of two alpaca trios with Suri and Huacaya animals confirmed the dominant inheritance of the Suri fleece type postulated here (Pallotti et al., 2023). One of the two *KIT* missense variants (XM_031688000.1:c.35G>A) they presented as a candidate causal variant for the white coat phenotype was absent from the genomes of the nine whole-genome sequenced alpacas from our study, while the other variant (XM_031688000.1:c.982G>C) is absent in two completely white alpacas and heterozygous in two pigmented alpacas including a grey Huacaya that also carries the most likely causal missense variant (XM_031688000.1: c.391G>A) reported previously (Jones et al., 2019; Tan et al., 2022). Preliminary conclusions of Pallotti et al. (2023) are therefore invalidated by our data ([Table S4](#)).

Further whole-genome sequencing of well-phenotyped alpacas from both fleece types and different coat colours with deep pedigree information is required to unravel the causative variants. Nevertheless, the data presented strongly support for the first time that the keratin locus on chromosome 16 is a highly plausible candidate for influencing fleece type in alpacas.

KEYWORDS

fleece type, GWAS, keratin, SNP array, South American camelids, WGS

ACKNOWLEDGEMENTS

The authors would like to thank the alpaca owners for donating samples and information. Anne Kramer (Alpakavet, Seedorf BE) and Katja Voigt (Clinic for ruminants, LMU München) are acknowledged for their help with sampling. This study was financially supported by the 'Verein zur Förderung der Forschung im Gesundheitssektor von Lamas und Alpakas e.V.'

CONFLICT OF INTEREST STATEMENT










The authors declare no conflicts of interests.

DATA AVAILABILITY STATEMENT

All underlying raw data are freely available. SNP genotyping data can be retrieved at (<https://www.ebi.ac.uk/ena/>). Whole genome sequences associated with this study are deposited at the EBI/NCBI Sequence Read Archive under project PRJEB64162.

ETHICS STATEMENT

The alpacas in this study were privately owned and samples were collected with the consent of their owners. The collection of blood samples was approved by the 'Cantonal Committee For Animal Experiments' (Canton of Bern; permit 71/19; approval date 9 September 2019).

K. Tan¹ 
 O. O. Adeniyi¹ 
 A. Letko² 
 G. RuddGarces¹ 
 E. Manz³ 
 H. Wagner⁴ 
 P. Zanolari⁵ 
 C. Drögemüller² 
 G. Lühken¹ 

¹Institute of Animal Breeding and Genetics, Justus Liebig University Giessen, Giessen, Germany

²Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland

³Generatio GmbH, Heidelberg, Germany

⁴Animal Clinic for Reproduction and Neonatology, Justus Liebig University Giessen, Giessen, Germany

⁵Clinic for Ruminants, Vetsuisse Faculty, University of Bern, Bern, Switzerland

Correspondence

C. Drögemüller, Institute of Genetics, Vetsuisse Faculty, University of Bern, Bern, Switzerland.

Email: cord.droegemueller@unibe.ch

C. Drögemüller and G. Lühken contributed equally to this work.

ORCID

K. Tan  <https://orcid.org/0000-0003-0923-1924>
 O. O. Adeniyi  <https://orcid.org/0000-0001-6726-1262>
 A. Letko  <https://orcid.org/0000-0002-6521-1285>
 G. RuddGarces  <https://orcid.org/0000-0001-7043-3117>
 E. Manz  <https://orcid.org/0000-0003-1445-0486>
 H. Wagner  <https://orcid.org/0000-0001-8262-3751>
 P. Zanolari  <https://orcid.org/0000-0002-9549-2487>
 C. Drögemüller  <https://orcid.org/0000-0001-9773-522X>
 G. Lühken  <https://orcid.org/0000-0002-2032-0840>

REFERENCES

- Balmer, P., Bauer, A., Pujar, S., McGarvey, K.M., Welle, M., Galichet, A. et al. (2017) A curated catalog of canine and equine keratin genes. *PLoS One*, 12(8), e0180359. Available from: <https://doi.org/10.1371/journal.pone.0180359>
- Frank, E.N., Hick, M.V.H., Gauna, C.D., Lamas, H.E., Renieri, C. & Antonini, M. (2006) Phenotypic and genetic description of fibre traits in South American domestic camelids (llamas and alpacas). *Small Ruminant Research*, 61(2–3), 113–129. Available from: <https://doi.org/10.1016/j.smallrumres.2005.07.003>
- Jones, M., Sergeant, C., Richardson, M., Groth, D., Brooks, S. & Munyard, K. (2019) A non-synonymous SNP in exon 3 of the KIT gene is responsible for the classic grey phenotype in alpacas (*Vicugna pacos*). *Animal Genetics*, 50(5), 493–500. Available from: <https://doi.org/10.1111/age.12814>
- Mendoza, M.N., Raudsepp, T., Alshanbari, F., Gutiérrez, G. & De León, F.A.P. (2019) Chromosomal localization of candidate genes for fiber growth and color in alpaca (*Vicugna pacos*). *Frontiers in Genetics*, 10, 583. Available from: <https://doi.org/10.3389/fgene.2019.00583>
- Pallotti, S., Antonini, M., Napolioni, V. & Renieri, C. (2023) Whole-genome sequencing of alpaca revealed variants in KIT gene potentially associated with the white coat phenotype. *Animal Genetics*, 54, 816–817. Available from: <https://doi.org/10.1111/age.13366>
- Paris, J.M., Letko, A., Häfliger, I.M., Ammann, P. & Drögemüller, C. (2020) Ear type in sheep is associated with the MSRB3 locus. *Animal Genetics*, 51(6), 968–972. Available from: <https://doi.org/10.1111/age.12994>
- Ponzoni, R.W., Hubbard, D.J., Kenyon, R.V., Tuckwell, C.D., McGregor, B.A., Howse, A. et al. (2003) *Phenotypes resulting from Huacaya By Huacaya, Suri by huacaya and suri by suri alpaca crossings*.
- Presciuttini, S., Valbonesi, A., Apaza, N., Antonini, M., Huanca, T. & Renieri, C. (2010) *Fleece variation in alpaca (Vicugna pacos): a two-locus model for the Suri/Huacaya phenotype*. Available from: <http://www.biomedcentral.com/1471-2156/11/70>
- Renieri, C., Valbonesi, A., La Manna, V., Antonini, M. & Asparrin, M. (2008) Inheritance of Suri and Huacaya type of fleece in alpaca. *Italian Journal of Animal Science*, 8(1), 83–91. Available from: <https://doi.org/10.4081/ijas.2009.83>
- Richardson, M.F., Munyard, K., Croft, L.J., Allnut, T.R., Jackling, F., Alshanbari, F. et al. (2019) Chromosome-level alpaca reference genome VicPac3.1 improves genomic insight into the biology of new world camelids. *Frontiers in Genetics*, 10, 586. Available from: <https://doi.org/10.3389/fgene.2019.00586>
- Scheet, P., Stephens, M. & Scheet, M.P. (2006) A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. *The American Journal of Human Genetics*, 78, 629–644. Available from: www.ajhg.org
- Sponenberg, D.P. (2010) Suri and huacaya alpaca breeding results in North America. *Small Ruminant Research*, 93(2–3), 210–212.

Available from: <https://doi.org/10.1016/j.smallrumres.2010.05.004>

Tan, K., Roy, M., Manz, E., Wagner, H., Zanolari, P., Drögemüller, C. et al. (2022) The KIT:c.376G>A variant in German and Swiss alpacas (*Vicugna pacos*) with different coat colors. *Animal Genetics*, 53(5), 718–720. Available from: <https://doi.org/10.1111/age.13231>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.