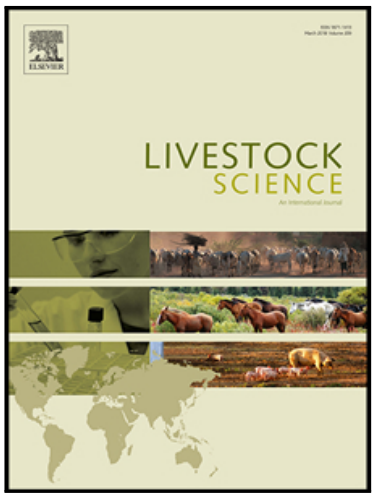
View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Archivio istituzionale della ricerca - Università di Palermo

# Accepted Manuscript




A combined genome-wide approach identifies a new potential candidate marker associated with the coat color sidedness in cattle

Salvatore Mastrangelo , Gianluca Sottile , Maria Teresa Sardina , Anna Maria Sutera , Marco Tolone , Rosalia Di Gerlando , Baldassare Portolano

PII: S1871-1413(18)30543-2
DOI: https://doi.org/10.1016/j.livsci.2019.05.009
Reference: LIVSCI 3709

To appear in: *Livestock Science*

Received date: 22 October 2018
Revised date: 14 March 2019
Accepted date: 10 May 2019

Please cite this article as: Salvatore Mastrangelo , Gianluca Sottile , Maria Teresa Sardina , Anna Maria Sutera , Marco Tolone , Rosalia Di Gerlando , Baldassare Portolano , A combined genome-wide approach identifies a new potential candidate marker associated with the coat color sidedness in cattle, *Livestock Science* (2019), doi: https://doi.org/10.1016/j.livsci.2019.05.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

## Highlights

- Coat color is one of the most important phenotypic features in livestock breeds.
- The comparison among GWAS and $F_{ST}$ analysis revealed a single SNP, significantly associated with color sidedness.
- The local breeds represent an important resource and model to study the genetic basis affecting peculiar traits.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

# A combined genome-wide approach identifies a new potential candidate marker associated with the coat color sidedness in cattle

Salvatore Mastrangelo[1], Gianluca Sottile[2], Maria Teresa Sardina[1], Anna Maria Sutera[1], Marco Tolone[1], Rosalia Di Gerlando[1], Baldassare Portolano[1]

[1] Dipartimento Scienze Agrarie, Alimentari e Forestali, University of Palermo, 90128 Palermo, Italy

[2] Dipartimento Scienze Economiche, Aziendali e Statistiche, University of Palermo, 90128 Palermo, Italy

Corresponding author

Salvatore Mastrangelo

Tel: +39-09123896069

salvatore.mastrangelo@unipa.it

**running title:** A new candidate marker associated with the coat colour in cattle

ACCEPTED MANUSCRIPT

**ABSTRACT**

Coat color is one of the most important phenotypic features in livestock breeds. Cinisara is a local cattle breed generally of uniform black color which occasionally presents a particular phenotype, with animals typically display a white band along their spine, from the head to the tail, and on the ventral line (color sidedness). Therefore, this breed provides an ideal model to study the genetic components underlying phenotypic variation in coat color. A total of 63 animals, ten with sidedness phenotype and 53 with uniform black color were genotyped with Illumina Bovine 50K. The comparison among genome-wide association study and $F_{ST}$ analysis revealed a single nucleotide polymorphism (SNP), ARS-BFGL-NGS-55928, significantly associated with the trait. Only one gene (*PLK2*) was annotated near the associated SNP in a window of ±200 kb. The protein encoded by this gene is a member of the polo-like kinases, the same family of several known coat-color candidate genes. Based on the reported results, we draw the possible conclusion that the identified marker is potentially associated with the coat color sidedness in Cinisara. The local breeds with their genetic variability represent an important resource and model to study the genetic basis affecting peculiar traits. Future studies would be particularly relevant to refine these results and to better understand the genetic basis for this phenotype.

*Keywords*: Coat color, GWAS, $F_{ST}$, cattle, *PLK2* gene

## 1. Introduction

Mammalian coat color is determined by the pigment melanin synthesized within melanosomes, specific organelles in the melanocytes. There are two types of melanin: the black to brown eumelanin and the red to yellow pheomelanin. The quantity and ratio of eumelanin to pheomelanin determine the color of mammalian skin, hair and eyes (Barsh, 1996). Coat color genetics has been the subject of a large number of studies that, in mice, led to the identification of



ACCEPTED MANUSCRIPT

more than 300 loci affecting pigmentation (Montoliu et al., 2009). In other species, a small number of loci (fewer than 10) had been recognized as being responsible for coat color (Mohanty et al., 2008; Hayes et al., 2010). As visual characteristics of animals, pigmentation traits are used in breed recognition for cattle and other livestock species, and therefore they represent an important phenotype of interest for breeding and research (Fan et al., 2014). The three most common coat colors observed in cattle are solid black, brown and red (Adalsteinsson et al., 1995). Other phenotypes are basically produced by modifications of these three main colors. Color sidedness is a dominantly inherited phenotype of cattle characterized by the polarization of pigmented sectors on the flanks, snout and ear tips (Olson, 1999). This phenotype has been described in several cattle breeds, such as Belgian Blue, Brown Swiss (Durkin et al., 2012) and African Nguni cattle (Szczerbal et al., 2017).

A substantial proportion of todays endangered cattle breeds displays distinct color phenotypes (Klungland et al., 2000). Cinisara is a local cattle breed reared in Sicily; it is inscribed as ‘at risk’ in the Food and Agricultural Organization (FAO) list of endangered breeds. The breed is included in the Podolic cattle group and it is well adapted to marginal, harsh, wind-swept and high-summer temperature climates and challenging endemic disease-rich environments (Mastrangelo et al., 2014). The coat, generally uniform black color, can occasionally present the color sidedness phenotype called *Agghio* in the local dialect (Fig. 1). Therefore, this breed provides an ideal model to study the genetic components underlying phenotypic variation in coat color. In fact, so far, the genetic mechanism is unclear and the potential causative variants for this phenotype have not been identified in this local breed.

The availability of genome-wide SNP panels has dramatically increased the power of genome-wide association studies (GWAS) to identify new markers associated with different phenotypes in domestic animals (Kijas et al., 2013). In cattle, GWAS can be used to localize the genomic regions that contribute to natural genetic variation in any phenotypic trait (Matukumalli et al.,



ACCEPTED MANUSCRIPT

2009). Application of the SNP arrays through genome-wide approaches has also elucidated the markers responsible of coat color in several livestock species (Hayes et al., 2010; Kijas et al., 2013; Mészáros et al., 2015; Kim et al., 2017; Nazari-Ghadikolaei et al., 2018). Therefore, a combined genome-wide approach (GWAS and $F_{ST}$) has been performed, comparing the two coat color phenotypes within Cinisara breed, to investigate the peculiar genetic features of this breed, and to uncover the genomic regions that are likely responsible for color sidedness.

## 2. Materials and methods

### *2.1 Samples, genotyping and data filtering*

Blood samples were collected from cattle by trained veterinarians. All the procedures were approved by "Organismo Preposto al Benessere Animale" (O.P.B.A) of University of Palermo in agreement with the recommendations of European Union (EU) Directive 2010/63/EU to ensure an appropriate animal care.

About 10 ml of blood was collected from caudal vein using tubes with EDTA as anticoagulant. A total of 63 unrelated animals from ten farms have been collected. In particular, ten color sidedness (case) (one per herd) (Figure 1) and 53 uniform black (control) (about 5 per herd) individuals were genotyped for 54,609 single nucleotide polymorphisms (SNPs), using the Illumina Bovine SNP50K v2 BeadChip. For these animals pedigree data were not available. SNPs were mapped using the Bos taurus UMD 3.1.1 genome assembly. Data filtering was performed using PLINK 1.7 (Purcell et al., 2007). The dataset was filtered to remove animals with more than 2% of missing genotypes, non-autosomal markers, SNPs with a call rate < 90%, with a minor allele frequency (MAF) < 0.01 and Hardy-Weinberg equilibrium P-value < 0.001.

### *2.2 Statistical analysis*

Pairwise genetic relationships were estimated to evaluate population substructure within cases and controls using identity-by-state (IBS) genetic distances calculated by PLINK 1.7 (Purcell et al., 2007) and graphically represented by multidimensional scaling (MDS) analysis with the statistical software R-3.5.1 (https://cran.r-project.org/doc/FAQ/R-FAQ.html#Citing-R). The same software was used to visualize the IBS matrices using the Heatmap function.

We performed the genome-wide association study (GWAS) using the univariate case/control model of the SNPassoc R package (Gonzalez et al., 2007). We used Bonferroni-corrected thresholds to determine genome-wide significant and suggestive thresholds, which were defined as 0.05/N and 0.10/N, respectively (N is the number of tested SNPs). After Bonferroni correction, significant thresholds were $P$<1.11E-06 for genome-wide ($P$<0.05) and $P$<2.21x10$^{-6}$ for suggestive, respectively. Genomic inflation factor (λ) and quantile-quantile (Q-Q) plot were obtained with GenABEL (Aulchenko et al., 2007) in order to assess potential systematic bias due to population structure or analytical approach. The genome-wide $F_{ST}$ case-control analysis was also performed using PLINK 1.9 (Chang et al., 2015). The top 0.9999 SNPs of the percentile distribution were considered as the most divergent across the comparison. A Manhattan plot of the results was generated using R in which $P$- and $F_{ST}$ values of each SNP were plotted as a function of its position along each autosomal chromosome within the genome. Moreover, with the aim to identify private copy number variants (CNVs) for sidedness animals, we used the results of Di Gerlando et al. (2018) which analyzed the same animals here considered to call CNVs in Cinisara breed. In particular, we compared the results of CNV calls between the two groups (ten color sidedness *vs.* 53 uniform black individuals), to identify CNVs carried only by sidedness and not by uniform black individuals. For more information on CNVs identification methodology, see Di Gerlando et al. (2018).

## 3. Results



ACCEPTED MANUSCRIPT

After filtering for quality, the final number of retained SNPs for the analysis was 45,246. All 63 animals had high quality genotyping and were included in the analysis. To examine the relationships between all individuals, MDS was performed. The obtained results showed some structure not well defined and noteworthy, the plot did not separate the animals on the basis of the different phenotypes (Figure S1). The heatmap of genetic similarity corroborated the findings obtained with MDS, showing the absence of substructure between individuals with different phenotypes that clustered independently of pigmentation (Figure S2).

We performed the quantile-quantile (Q-Q) plot to evaluate the reliability of GWAS data. An average inflation factor ($\lambda$) of 1.08 indicated that the GWAS was not inflated by population structure (Figure S3). In the GWAS, at the $P$<0.05 Bonferroni corrected ((-log10(P) = 5.95; $P_{\text{nominal value}}$ = 1.11E-06), we identified a single strongly associated SNP, ARS-BFGL-NGS-55928 (rs110452481) ((-log10(P) = 6.37; $P_{\text{nominal value}}$ = 4.237775e-07) at the position 21,048,672 bp on bovine chromosome (BTA) 20. No other SNP reached the Bonferroni corrected suggestive significant threshold ($P$<0.10) ((-log10(P) = 5.65). Fig. 2a reports the Manhattan plot obtained in this GWAS. To further support the association, a genome-wide $F_{ST}$ case-control analysis was also performed. A total of five SNPs were above the selected threshold ($F_{ST}$ = 0.432) (Table 1): three on BTA1, one on BTA6 and one on BTA20. The Manhattan plot of the $F_{ST}$ analysis is shown in Fig. 2b. The marker with the highest $F_{ST}$ value ($F_{ST}$ = 0.559) which differentiated between the groups overlapped with the significant SNP identified in the GWAS. This SNP showed markedly different genotypic frequencies between the two groups. The uniform black had higher frequency of GG homozygote genotype compared with color sidedness individuals, which were mostly heterozygous (Table 2). Only one gene (*PLK2*) was annotated near the identified marker.

The comparison of CNVs data between the two groups revealed six private CNVs (five deletions and one duplication) on chromosomes BTA12, 13, 26 and 29 (Table S1) carried only by few color sidedness individuals.

ACCEPTED MANUSCRIPT

## 4. Discussion

Potentially, there is much unrecognized beneficial genetic variation in local autochthonous breeds and populations, which supposes important reservoirs of non-exploited resources. In this study, genome-wide analysis has been performed for the color sidedness phenotype in Cinisara cattle. Due to its phenotypic variability, this local breed provides a model for investigating the genetic bases of color sidedness. To the best of our knowledge, no study with combined genome-wide approach (GWAS and $F_{ST}$) has been reported for color sidedness in cattle.

To overcome some limitation of this study, such as low number of tested animals and the use of low-density array, a combined analytical approach has been used (Kijas et al., 2013; Mastrangelo et al., 2018). The identification of candidate genomic regions by more than one methodology may be seen as strong evidence of the activity on a particular phenotype. This also allowed to exclude other regions that reached or were close to the defined thresholds derived by several factors that could not be better managed (i.e. genetic drift, population structure, ascertain bias of the SNP chip tool) (Schiavo et al., 2018).

In this study, to gain insights into the molecular basis of color sidedness, ten cases and 53 controls were used and several analyses were conducted. The comparison among GWAS and $F_{ST}$ revealed a single strong association signal, ARS-BFGL-NGS-55928, to be potentially associated with the investigated phenotype in Cinisara breed. No other region in the genome exhibited a significant differentiation. In a GWAS case-control for pigmentation traits in Holstein, several SNPs associated with the phenotypes were identified (Fan et al., 2014). Li et al. (2014) conducted a GWAS to identify SNPs associated with white and non-white coats in Finn sheep and found 35 SNPs related with those phenotypes. In this study, only one SNP was found associated with the investigated trait and GWAS results that are characterized by a single outlier marker, in the absence of other associated peaks, are biologically relevant (Kijas et al., 2013). Moreover, the

results of genomic inflation factor and Q-Q plot suggest that there is no population stratification in our data, although the Cinisara is a local breed with a reduced population size (Mastrangelo et al., 2014).

*Polo Like Kinase 2* (*PLK2*) was the closest gene annotated near the peak SNP (in a window of ±200 kb) located approximately 188 Kb downstream at position 20,854,607–20,860,584 bp. PLK2 protein is a member of the polo-like kinase family that includes serine-threonine kinases and that plays a crucial role in cell cycle progression, mitotic division and cytokinesis in response to DNA damage and to hypoxic conditions (Archambault & Glover, 2009; Valenti et al., 2011). Gene Ontology (GO) annotations related to this gene include transferase activity, transferring phosphorus-containing groups and protein tyrosine kinase activity. Coat color in cattle is controlled by several genes and their corresponding alleles, such as: *ASIP* (agouti), *TYR* (albinism), *TYRP1* (brown), *KIT* (color sided, dominant white), *KITLG* (roan), *PMEL* (dilution), *MC1R* (extension) and *MITF* (white spotting) (Brenig et al., 2013). Some of these genes belong to the protein kinase family, as well as the *PLK2* here identified. For example the *KIT* gene, which encodes tyrosine kinase receptor, has been already described as the major determinant of the differences in coat color phenotypes (Fontanesi et al., 2010). Another gene, the *KITLG* (tyrosine-protein kinase ligand), encodes a tyrosine kinase receptor ligand (Hui et al., 2006). In cattle, the mutations at this gene are responsible for white coat color (Seitz et al., 1999). Fan et al. (2014), in a GWAS for pigmentation traits in Chinese Holstein population reported several novel candidate markers for coat color traits. Therefore, the results of the two genome-wide approaches together with the biological function of the gene, suggested that the *PLK2* is a potential candidate gene in color sidedness phenotype in Cinisara cattle breed.

Durkin et al. (2012) showed that color sidedness in cattle is determined by two CNVs that duplicate a small part of the BTA6 and BTA29. More precisely, the authors have shown that a duplication of the *KIT* gene located on BTA6 and its aberrant insertion on BTA29 ($Cs_{29}$) resulted

in the white coat color phenotype 'color sided' in Belgian Blue and other cattle breeds. Moreover, the authors scanned the genome of color sided animals and identified a shared haplotype which encompassed *LUZP2* gene, not known to be involved in pigmentation. Our results on the identification of the private CNVs for case group, revealed a duplication of ~400 kb on BTA29 within the *LUZP2* gene. However, private CNVs are of particular interest if they are fixed or with high frequency in the breed or sub-group, since they could be involved in breed specific characteristics (Molin et al., 2014). Unfortunately, no fixed private CNVs were detected and this duplication presented low frequency within color sidedness individuals (Table S1). Moreover, no CNVs have been identified close to the candidate marker here associated with the phenotype.

The vast majority of the coat color phenotype related studies are based on candidate gene-approaches focusing on commercial cattle breeds. The patterns of coat color observed in modern cattle breeds are quite variable and it is possible to hypothesize that the variation in coat color patterns is determined by different genes and mutations (Edea et al., 2017). Consistent with our results, several studies in mammalian species, such as cat (Gandolfi et al., 2013) and dog (Cadieu et al., 2009), as well as in livestock species, through GWAS case/control for coat color, reported peaks/signals within or close to potential candidate genes. In fact, Kijas et al. (2013), in a study for coat pigmentation in sheep, reported only one strongly associated SNP close to *MCR1* gene. Similar results have been also reported by Menzi et al. (2016) and by Kumar et al. (2018), which in studies for coat color in goat, clearly demonstrated that a single SNP controls the phenotype. Matukumalli et al. (2009) exploited the black/red coat color phenotype to demonstrate the utility of the Illumina BovineSNP50 assay for genome-wide association studies in cattle by localizing known variation for the phenotype. Moreover, different studies showed the power of medium density chip to detect association using a low number of animals (Becker et al., 2010; Gandolfi et al., 2013; Kijas et al., 2013; Muniz et al., 2016). Based on the reported results, together with what

is shown in other studies, we draw the possible conclusion that the identified marker is potentially associated with the coat color sidedness in Cinisara.

Coat color is a qualitative trait and an indicator of adaptability to heat stress (Fadare et al., 2013). The variation in the size of pigmented sectors is a phenotype that might have an important impact in the future, because it is expected that animals with a higher proportion of white coat absorb less solar radiation and therefore they are better buffered against heat stress (Reinsch et al., 1999). In an intricate interactive fashion with environment, some mechanisms that maintain genetic diversity such as crossing-over, independent assortment and sexual reproduction, generate an astonishing phenotypic variation which provides the substrate for adaptive mechanisms (Qanbari and Simianer, 2014). Therefore, these results also showed and confirmed as the local animal genetic resources can be used to acquire information on genetic factors that might have potential impacts in the future breeding programs for other cosmopolitan breeds.

## 5. Conclusions

Coat color is one of the most important qualitative traits in livestock breeds. This study contributes to better understanding the genetic architecture of coat color in Cinisara cattle. Determination of the genomic regions with the potential candidate gene for this characteristic will help to protect these animal genetic resources. Once again, the local breeds with their genetic variability represent an important resource and a model to study the genetic basis affecting peculiar traits. Our results showed that this phenotype in Cinisara breed is associated with a marker close to a potential candidate gene. Therefore, this work constitutes a preliminary report to further genomic research on coat color sidedness. It will be important to refine and validate these results. The candidate gene here identified can be considered in further studies based on the comparison of sequence data in other breeds displaying this phenotype. Moreover, combined genome-wide approaches using different populations, the high-density SNP chip and an increase



ACCEPTED MANUSCRIPT

in the number of genotyped individuals would be particularly relevant to further enhance the power to identify additional SNPs and genes and to better understand the genetic basis for this phenotype in cattle.

**Acknowledgements**

This research was financed by the PON01_02249 project, funded by MIUR.

**Conflict of interest statement**

The authors have no conflict of interest to declare.

**References**

Adalsteinsson, S., Bjarnadottir, S., Vage, D.I., Jonmundsson, J. V., 1995. Brown coat color in Icelandic cattle produced by the loci Extension and Agouti. J. Hered. 86, 395-398.

Archambault, V., Glover D. M., 2009. Polo-like kinases: conservation and divergence in their functions and regulation. Nat. Rev. Mol. Cell Biol. 10, 265–275.

Aulchenko, Y.S., Ripke, S., Isaacs, A., Van Duijn, C.M., 2007. GenABEL: an R library for genome-wide association analysis. Bioinformatics 23(10), 1294-1296.

Barsh G.S., 1996. The genetics of pigmentation: from fancy genes to complex traits. Trends Genet. 12, 299–305.

Becker, D., Tetens, J., Brunner, A., Bürstel, D., Ganter, M., Kijas, J., International Sheep Genomics Consortium, Drögemuller, C., 2010. Microphthalmia in Texel sheep is associated with a missense mutation in the paired-like homeodomain 3 (PITX3) gene. PLoS One 5(1), e8689.

Brenig, B., Beck, J., Floren, C., Bornemann-Kolatzki, K., Wiedemann, I., Hennecke, S., Swalve, E., Schütz, E., 2013. Molecular genetics of coat colour variations in White Galloway and White Park cattle. Anim. Genet. 44, 450-453.



ACCEPTED MANUSCRIPT

Cadieu, E., Neff, M.W., Quignon, P., Walsh, K., Chase, K., Parker, H. G., VonHoldt, B.M., Rhue, A., Boyko, A., Byers, A., Wong, A., Mosher, D.S., Elkahloun, A.G., Spady, T.C., André, C., Lark, K.G., Cargill, M., Bustamante, C. D., Wayne, R. K., Ostrander, E.A., 2009. Coat variation in the domestic dog is governed by variants in three genes. Science 326(5949), 150-153.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 4, 7.

Di Gerlando, R., Sardina, M.T., Tolone, M., Sutera, A.M., Mastrangelo, S., Portolano, B., 2018. Genome-wide detection of copy-number variations in local cattle breeds. Anim. Prod. Sci. doi.org/10.1071/AN17603.

Durkin, K., Coppieters, W., Drögemüller, C., Ahariz, N., Cambisano, N., Druet, T., Fasquelle, C., Haile, A., Horin, P., Huang, L., Kamatani, Y., Karim, L., Lathrop, M., Moser, S., Oldenbroek, K., Rieder, S., Sartelet, A., Sölkner, J., Stålhammar, H., Zelenika, D., Zhang, Z., Leeb, T., Georges, M., Charlier, C., 2012. Serial translocation by means of circular intermediates underlies colour sidedness in cattle. Nature 482(7383), 81.

Edea, Z., Dadi, H., Dessie, T., Kim, I.H., Kim, K.S., 2017. Association of MITF loci with coat color spotting patterns in Ethiopian cattle. Genes Genom. 39(3), 285-293.

Fadare, A. O., Peters, S.O., Yakubu, A., Sonibare, A.O., Adeleke, M.A., Ozoje, M.O., Imumorin, I.G., 2012. Physiological and haematological indices suggest superior heat tolerance of white-coloured West African Dwarf sheep in the hot humid tropics. Trop. Anim. Health Prod. 45(1), 157-165.

Fan, Y., Wang, P., Fu, W., Dong, T., Qi, C., Liu, L., Guo, G., Li, C., Ciu, X., Zhang, S., Zhang, Q., Zhang, Y., Sun, D., 2014. Genome-wide association study for pigmentation traits in Chinese Holstein population. Anim. Genet. 45(5), 740-744.

ACCEPTED MANUSCRIPT

Fontanesi, L., Tazzoli, M., Russo, V., Beever, J., 2010. Genetic heterogeneity at the bovine KIT gene in cattle breeds carrying different putative alleles at the spotting locus. Anim. Genet. 41, 295-303.

Gandolfi, B., Alhaddad, H., Joslin, S.E., Khan, R., Filler, S., Brem, G., Lyons, L.A., 2013. A splice variant in KRT71 is associated with curly coat phenotype of Selkirk Rex cats. Sci. Rep. 3, 2000.

Gonzalez, J.R., Armengol, L., Sole, X., Guino, E., Mercader, J.M., Estivill, X., Moreno V., 2007. SNPassoc: an R package to perform whole genome association studies. Bioinformatics 23, 654-655.

Hayes, B.J., Pryce, J., Chamberlain, A.J., Bowman, P.J., Goddard M.E., 2010. Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in Holstein cattle as contrasting model traits. PLoS Genet. 6(9), e1001139.

Hui, E.S., Udofa, E.A., Soto, J., Vanderhoof, V.H., Zachman, K., Tong, Z.B., Nelson, L.M., 2006. Investigation of the human stem cell factor KIT ligand gene, KITLG, in women with 46, XX spontaneous premature ovarian failure. Fertil. Steril. 85, 1502-1507.

Kijas, J.W., Serrano, M., McCulloch, R., Li, Y., Salces Ortiz, J., Calvo, J. H., Pérez-Guzmán, M.D., The International Sheep Genomics Consortium. 2013. Genome wide association for a dominant pigmentation gene in sheep. J. Anim. Breed. Genet. 130(6), 468-475.

Kim, N.Y., Bhuiyan, M.S.A., Chae, H.S., Baek, K.S., Son, J.K., Shin, S.M., Woo, J.H., Park, S.H., Lee, S.H., 2017. Genome-wide association study for tobiano spotting coat color in Korean Jeju× Thoroughbred horse population. Anim Genet. 48(6), 728-729.

Klungland, H., Olsen, H.G., Hassanane, M.S., Mahrous, K., Våge, D.I., 2000. Coat colour genes in diversity studies. J. Anim. Breed. Genet. 117, 217-224.

Kumar, C., Song, S., Dewani, P., Kumar, M., Parkash, O., Ma, Y., Malhi, K.K., Yang, N., Mwacharo, J.M., He, X., Jiang, L., 2018. Population structure, genetic diversity and selection signatures within seven indigenous Pakistani goat populations. Anim. Genet. 49(6), 592-604.

Li, M.H., Tiirikka, T., Kantanen, J., 2014. A genome-wide scan study identifies a single nucleotide substitution in ASIP associated with white versus non-white coat-colour variation in sheep (Ovis aries). Heredity 112(2), 122.

Mastrangelo, S., Saura, M., Tolone, M., Salces-Ortiz, J., Di Gerlando, R., Bertolini, F., Fontanesi, L., Sardina, M.T., Serrano, M., Portolano, B., 2014. The genome-wide structure of two economically important indigenous Sicilian cattle breeds. J. Anim. Sci. 92(11), 4833-4842.

Mastrangelo, S., Sottile, G., Sutera, A.M., Di Gerlando, R., Tolone, M., Moscarelli, A., Sardina, M.T., Portolano, B., 2018. Genome-wide association study reveals the locus responsible for microtia in Valle del Belice sheep breed. Anim. Genet. 49(6), 636-640.

Matukumalli, L.K., Lawley, C.T., Schnabel, R.D., Taylor, J.F., Allan, M.F., Heaton, M.P., O'Connell, J., Moore, S.S., Smith, T.P.L., Sonstegard, T.S., Van Tassell, C.P., 2009. Development and characterization of a high density SNP genotyping assay for cattle. PloS One 4:e5350.

Menzi, F., Keller, I., Reber, I., Beck, J., Brenig, B., Schütz, E., Tosso, L., Drögemüller C., 2016. Genomic amplification of the caprine EDNRA locus might lead to a dose dependent loss of pigmentation. Sci. Rep. 6, 28438.

Mészáros, G., Petautschnig, E., Schwarzenbacher, H., Sölkner, J., 2015. Genomic regions influencing coat color saturation and facial markings in Fleckvieh cattle. Anim. Genet. 46(1), 65-68.

Mohanty, T.R., Seo, K.S., Park, K.M., Choi T.J., Choe, H.S., Baik, D.H., Hwang, I.H., 2008. Molecular variation in pigmentation genes contributing to coat colour in native Korean Hanwoo cattle. Anim. Genet. 39, 550-553.

Molin, A.M., Berglund, J., Webster, M.T., Lindblad-Toh, K., 2014. Genome-wide copy number variant discovery in dogs using the CanineHD genotyping array. BMC Genom. 15, 210.

Montoliu, L., Oetting W.S., Bennett, D.C., 2009. European Society for Pigment Cell Research. World Wide Web (URL:http://www.espcr.org/micemut).

Muniz, M.M.M., Caetano, A.R., McManus, C., Cavalcanti, L.C.G., Façanha, D.A.E., Leite, J.H. G.M., Facò, O., Paiva, S.R., 2016. Application of genomic data to assist a community-based breeding program: A preliminary study of coat color genetics in Morada Nova sheep. Livest. Sci. 190, 89-93.

Nazari-Ghadikolaei, A., Mehrabani-Yeganeh, H., Miarei-Aashtiani, S.R., Staiger, E.A., Rashidi, A., Huson, H.J., 2018. Genome-Wide Association Studies Identify Candidate Genes for Coat Color and Mohair Traits in the Iranian Markhoz Goat. Front. Genet. 9, 105.

Olson, T., 1999 . The Genetics of Cattle (CAB International, 1999).

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I.W., 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(3), 559-575.

Qanbari, S., Simianer, H., 2014. Mapping signatures of positive selection in the genome of livestock. Liv. Sci. 166: 133-143.

R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. 2011. http://www.R-project.org.

Reinsch, N., Thomsen, H., Xu, N., Brink, M., Looft, C., Kalm, E., Brockmann, G.A., Grupe, S., Kuhn, C., Schwerin, M., Leyhe, B., Hiendleder, S., Erhardt, G., Medjugorac, I., Russ, I., Förster, M., Reents, R., Averdunk, G., 1999. A QTL for the degree of spotting in cattle shows synteny with the KIT locus on chromosome 6. J. Hered. 90(6), 629-634.

Schiavo, G., Bertolini, F., Utzeri, V.J., Ribani, A., Geraci, C., Santoro, L., Ovilo, C., Fernandez, A.I., Gallo, M., Fontanesi L., 2018. Taking advantage from phenotype variability in a local



ACCEPTED MANUSCRIPT

animal genetic resource: identification of genomic regions associated with the hairless phenotype in Casertana pigs. Anim. Genet. 49(4), 321-325.

Seitz, J.J., Schmutz, S.M., Thue, T. D., Buchanan F.C., 1999. A missense mutation in the bo-Pleiotropic effects of coat colour genes on animal health 17 vine MGF gene is associated with the roan phenotype in Belgian Blue and Shorthorn cattle. Mamm. Genome 10, 710-712.

Szczerbal, I., Robinson, T.J., Duran, A., Davey, S.C., Andersson, M., Switonski, M., 2017. Ectopic position of duplicated KIT gene in African Nguni cattle, associated with color sidedness, confirms its shared ancestry with the Bos taurus lineage. Anim. Genet. 48(1), 122-123.

Valenti, F., Fausti, F., Biagioni, F., Shay, T., Fontemaggi, G., Domany, E., Domany, E., Yaffe, M.B., Strano, S., Blandino, G., Di Agostino, S., 2011. Mutant p53 oncogenic functions are sustained by Plk2 kinase through an autoregulatory feedback loop. Cell Cycle 10(24), 4330-4340.

ACCEPTED MANUSCRIPT

**Table 1.** Significant ($P$<0.05) SNPs obtained in the GWAS and the top 0.9999 detected in the $F_{ST}$ analysis associated with color sidedness phenotype in Cinisara cattle breed.

| **BTA**[a] | **SNP**[b] | **Position (bp)** | **GWAS**[c] | **$F_{ST}$ value** |
|---|---|---|---|---|
| 1 | ARS-BFGL-BAC-19439 | 66,252,625 | | 0.468 |
| 1 | ARS-BFGL-NGS-100745 | 69,045,500 | | 0.445 |
| 1 | BTA-105039-no-rs | 116,569,378 | | 0.454 |
| 6 | Hapmap43679-BTA-76147 | 48,423,262 | | 0.434 |
| 20 | ARS-BFGL-NGS-55928 | 21,048,672 | 4.237775e-07 | 0.559 |

[a]BTA: *Bos Taurus* chromosome number. [b]SNP: single nucleotide polymorphisms. [c]GWAS: genome-wide association study.

**Table 2.** Genotypic frequencies of the best associated SNP, ARS-BFGL-NGS-55928, in case and control groups.

| **Group** | GG | GA | AA |
|---|---|---|---|
| Case (color sidedness) | 0.30 | 0.70 | - |
| Control (uniform black) | 0.98 | 0.02 | - |

ACCEPTED MANUSCRIPT

**FIGURES**

**Figure 1.** Example of color sidedness Cinisara cattle.

ACCEPTED MANUSCRIPT

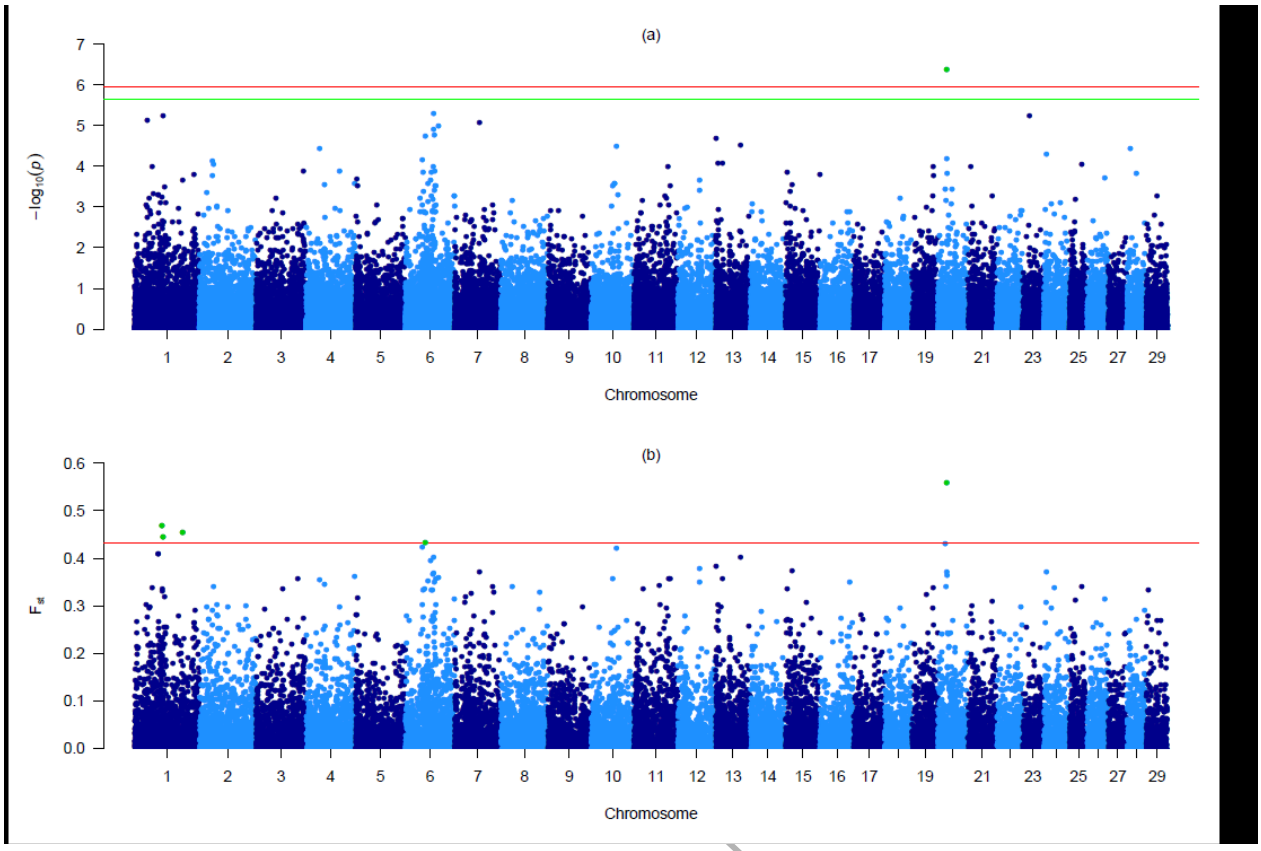



**Figure 2. a**) Manhattan plot of the *P*-values in the genome-wide association study (GWAS). The horizontal lines represent the genome-wide significance (red; $P<0.05$) and suggestively significant (green; $P<0.10$) single nucleotide polymorphisms (SNPs). (**b**) Manhattan plot of the $F_{ST}$ showing the top 0.9999 SNPs above the red line ($F_{ST} = 0.432$).

ACCEPTED MANUSCRIPT