UNIVERSIDADE ESTADUAL PAULISTA
"JÚLIO DE MESQUITA FILHO"
FACULDADE DE MEDICINA VETERINÁRIA
CÂMPUS ARAÇATUBA

BEATRIZ BATISTA TRIGO

Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

Araçatuba
2022

BEATRIZ BATISTA TRIGO

Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

Tese apresentada à Faculdade de Medicina Veterinária de Araçatuba da Universidade Estadual Paulista "Júlio de Mesquita Filho" – UNESP, como parte dos requisitos para a obtenção do título de Doutora em Ciência Animal (Área de Medicina Veterinária Preventiva e Produção Animal).

Orientador: Professor Dr. Yuri Tani Utsunomiya
Coorientadora: Pesquisadora Dra. Flávia Lombardi Lopes

Araçatuba
2022

T828i

Trigo, Beatriz Batista
Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno / Beatriz Batista Trigo. -- Araçatuba, 2022
102 p. : il., tabs.

Tese (doutorado) - Universidade Estadual Paulista (Unesp), Faculdade de Medicina Veterinária, Araçatuba
Orientador: Yuri Tani Utsunomiya
Coorientadora: Flávia Lombardi Lopes

1. Bovinos. 2. ASIP. 3. Genética animal. I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca da Faculdade de Medicina Veterinária, Araçatuba. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

AUTORA: BEATRIZ BATISTA TRIGO
ORIENTADOR: YURI TANI UTSUNOMIYA
COORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Prof. Dr. YURI TANI UTSUNOMIYA (Participaçao Virtual)
Programa de Pós-Graduação em Ciência Animal da Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. FABIANO ANTONIO CADIOLI (Participaçao Virtual)
Departamento de Clínica e Cirurgia Veterinária / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. GUILHERME DE PAULA NOGUEIRA (Participaçao Virtual)
Departamento de Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Profa. Dra. ELIANE PATRÍCIA CERVELATTI (Participaçao Virtual)
Curso de Ciencias Biológicas / Centro Universitário Católico Salesiano Auxilium / UniSALESIANO - Araçatuba

Dr. ANDRÉ VIEIRA DO NASCIMENTO (Participaçao Virtual)
Doutor em Genética e Melhoramento Animal pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/UNESP

Araçatuba, 22 de fevereiro de 2022.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

AUTORA: BEATRIZ BATISTA TRIGO
ORIENTADOR: YURI TANI UTSUNOMIYA
COORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Prof. Dr. YURI TANI UTSUNOMIYA (Participaçao Virtual)
Programa de Pós-Graduação em Ciência Animal da Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. FABIANO ANTONIO CADIOLI (Participaçao Virtual)
Departamento de Clínica e Cirurgia Veterinária / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. GUILHERME DE PAULA NOGUEIRA (Participaçao Virtual)
Departamento de Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Profa. Dra. ELIANE PATRÍCIA CERVELATTI (Participaçao Virtual)
Curso de Ciencias Biológicas / Centro Universitário Católico Salesiano Auxilium / UniSALESIANO - Araçatuba

Dr. ANDRÉ VIEIRA DO NASCIMENTO (Participaçao Virtual)
Doutor em Genética e Melhoramento Animal pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/UNESP

Araçatuba, 22 de fevereiro de 2022.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

AUTORA: BEATRIZ BATISTA TRIGO
ORIENTADOR: YURI TANI UTSUNOMIYA
COORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Prof. Dr. YURI TANI UTSUNOMIYA (Participaçao Virtual)
Programa de Pós-Graduação em Ciência Animal da Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. FABIANO ANTONIO CADIOLI (Participaçao Virtual)
Departamento de Clínica e Cirurgia Veterinária / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. GUILHERME DE PAULA NOGUEIRA (Participaçao Virtual)
Departamento de Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Profa. Dra. ELIANE PATRÍCIA CERVELATTI (Participaçao Virtual)
Curso de Ciencias Biológicas / Centro Universitário Católico Salesiano Auxilium / UniSALESIANO - Araçatuba

Dr. ANDRÉ VIEIRA DO NASCIMENTO (Participaçao Virtual)
Doutor em Genética e Melhoramento Animal pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/UNESP

Araçatuba, 22 de fevereiro de 2022.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

AUTORA: BEATRIZ BATISTA TRIGO
ORIENTADOR: YURI TANI UTSUNOMIYA
COORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Prof. Dr. YURI TANI UTSUNOMIYA (Participaçao Virtual)
Programa de Pós-Graduação em Ciência Animal da Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. FABIANO ANTONIO CADIOLI (Participaçao Virtual)
Departamento de Clínica e Cirurgia Veterinária / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. GUILHERME DE PAULA NOGUEIRA (Participaçao Virtual)
Departamento de Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Eliane Patrícia Cervelatti

Profa. Dra. ELIANE PATRICIA CERVELATTI (Participaçao Virtual)
Curso de Ciencias Biológicas / Centro Universitário Católico Salesiano Auxilium / UniSALESIANO - Araçatuba

Dr. ANDRÉ VIEIRA DO NASCIMENTO (Participaçao Virtual)
Doutor em Genética e Melhoramento Animal pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/UNESP

Araçatuba, 22 de fevereiro de 2022.

UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

Título: Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno

AUTORA: BEATRIZ BATISTA TRIGO
ORIENTADOR: YURI TANI UTSUNOMIYA
COORIENTADORA: FLÁVIA LOMBARDI LOPES

Aprovada como parte das exigências para obtenção do Título de Doutora em CIÊNCIA ANIMAL, área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Prof. Dr. YURI TANI UTSUNOMIYA (Participaçao Virtual)
Programa de Pós-Graduação em Ciência Animal da Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. FABIANO ANTONIO CADIOLI (Participaçao Virtual)
Departamento de Clínica e Cirurgia Veterinária / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Prof. Dr. GUILHERME DE PAULA NOGUEIRA (Participaçao Virtual)
Departamento de Produção e Saúde Animal / Faculdade de Medicina Veterinária - Câmpus de Araçatuba/UNESP

Profa. Dra. ELIANE PATRÍCIA CERVELATTI (Participaçao Virtual)
Curso de Ciencias Biológicas / Centro Universitário Católico Salesiano Auxilium / UniSALESIANO - Araçatuba

André Vieira do Nascimento
Dr. ANDRÉ VIEIRA DO NASCIMENTO (Participaçao Virtual)
Doutor em Genética e Melhoramento Animal pela Faculdade de Ciências Agrárias e Veterinárias - Câmpus de Jaboticabal/UNESP

Araçatuba, 22 de fevereiro de 2022.

À Deus e Nossa Senhora Aparecida por todas as bênçãos e aos meus pais Nilton e Isabel por todo o apoio, carinho, paciência, amor e por sempre acreditarem em minha trajetória.

AGRADECIMENTOS

À Deus e Nossa Senhora Aparecida por sempre me abençoarem e protegerem durante toda a minha vida e sempre atenderem às minhas preces me protegendo e iluminando.

Aos meus pais Nilton Godoy Trigo e Maria Isabel Marques Batista Trigo que são minha base e foram fundamentais para que eu chegasse até aqui sempre me apoiando com muito amor, carinho e dedicação, ouvindo minhas angústias e incertezas sempre com palavras de apoio.

Ao meu noivo Demetrius Diego Peres dos Santos por todo o amor, carinho e paciência, sempre me dando força para continuar com fé e dedicação. Obrigada por acreditar em mim e nos meus sonhos.

Ao meu orientador Professor Doutor Yuri Tani Utsunomiya pela oportunidade, orientando sempre com muita competência e paciência, enfrentando comigo os desafios que apareceram no caminho sempre com muito pé no chão e força, sendo além de tudo um amigo.

À Pesquisadora Doutora Flávia Lombardi Lopes por aceitar ser minha coorientadora, nos auxiliando sempre que necessário com muita competência e sendo sempre atenciosa e solicita.

À Dra. Sarita Priscila Gobbo Ferrari por todo o apoio, carinho e oportunidades, sendo uma grande amiga e mentora.

Aos amigos e colegas que convivi no Laboratório de Bioquímica e Biologia Molecular durante esse período: obrigada por todo apoio e convívio.

Às amigas que a pós me deu: Maria Fernanda Lopes, Natália Scaramelle, Juliana Felix, Mariana Cordeiro, Jéssica Troiano e Amanda Furlan, por estarem comigo nesta caminhada e tornarem o Laboratório de Epigenômica a minha segunda casa na FMVA.

À aluna de IC Natália Fonseca por todo o apoio e ajuda prestada durante todo o projeto, trabalhando sempre com muita dedicação e alegria.

Aos funcionários da empresa AgroPartners Consulting que me auxiliaram na bioinformática durante essa longa caminhada: minha gratidão por toda ajuda e disponibilidade.

À Professora Dra. Eliane Cervelatti e Margareth Caminhas pela oportunidade de aprendizado durante o Estágio de Docência realizado em suas

disciplinas e à todos os professores que passaram por minha vida contribuindo para meu crescimento pessoal e profissional, minha eterna gratidão.

Aos pesquisadores Dra. Danika Bannasch, Dr. Tosso Leeb, Dr. Abdulfatai Tijjani e toda sua equipe de pesquisa por todo o auxílio prestado durante a execução e elaboração do nosso trabalho.

À todos os coautores dos artigos publicados que foram resultado do meu Doutorado: obrigada pela ajuda com todas as ideias, amostras, testes, análises, reagentes e resultados fazendo que a junção de toda a ajuda se tornassem artigos para publicações relevantes em ótimas revistas.

Aos Professores Dr. Guilherme de Paula Nogueira, Dra. Alexeia Barufatti, Dr. Fabiano Antônio Cadioli, Dra. Eliane Patrícia Cervelatti e Dr. André Vieira do Nascimento por tantas considerações e sugestões importantes e pertinentes durante minha banca de qualificação e de defesa de tese às quais contribuíram imensamente para a continuidade e finalização do nosso trabalho.

À Universidade Estadual Paulista “Júlio de Mesquita Filho”; à Faculdade de Medicina Veterinária de Araçatuba; ao programa de Pós-Graduação em Ciência Animal e à CAPES pela bolsa de Doutorado.

“O êxito da vida não se mede pelo caminho que você conquistou,
mas sim pelas dificuldades que superou no caminho.”
Abraham Lincoln

TRIGO, B.B. **Identificação e caracterização de variantes genéticas associadas à pelagem em gado zebuíno.** 2022. 102 f. Tese (Doutorado) – Faculdade de Medicina Veterinária, Universidade Estadual Paulista, Araçatuba, 2022.

RESUMO

O gado zebuíno da raça Nelore (*Bos indicus*) é conhecido por sua fácil adaptação nos ambientes de climas tropicais e subtropicais. Uma das características fenotípicas relacionadas à essa adaptação e principalmente à tolerância à radiação ultra-violeta é o padrão de coloração da pele e da pelagem apresentado pela raça. Outras raças zebuínas também apresentam padrão de pelagem semelhante ao Nelore. Dada a potencial contribuição da variação na pelagem destes animais na adaptação dos mesmos nos climas tropicais e sub-tropicais, objetivou-se mapear variantes genéticas candidatas estruturais ou funcionais que estivessem associadas ao escurecimento da pelagem nos zebuínos. Realizando análise de associação genômica ampla (*genome-wide association study – GWAS*) e posterior análise de dados de sequenciamento completo foi possível encontrar uma complexa variante estrutural nas proximidades do gene da proteína sinalizadora agouti (*agouti signaling protein -ASIP*) a qual foi denominada de *ASIP*-SV1, sendo fortemente correlacionada ao escurecimento da pelagem em regiões específicas no corpo dos bovinos das raças Nelore e Brahman. A variante estrutural *ASIP*-SV1 também foi predita como sendo de origem *B. indicus*, para confirmar tal predição foram analisadas sequencias da região da variante estrutural de animais zebuínos e não-zebuínos para elucidar a segregação desta variante. Nas análises realizadas 45,3% dos zebuínos analisados apresentaram a variante estrutural enquanto que dos taurinos somente 6.3% continham a variante no gene *ASIP* sendo que esses taurinos eram de raças conhecidas por apresentarem introgressão zebuína em seu genoma. Concluiu-se que a variante *ASIP*-SV1 causa o escurecimento na pigmentação da pelagem em regiões específicas do corpo destes animais causando um aumento na produção de eumelanina, a variante se mostrou robusta e sugeriu-se que, em conjunto a outros genes, essa variação possa contribuir com a adaptação dos bovinos aos climas tropicais e sub-tropicais.

Palavras-chave: Bovinos. *ASIP*. Genética animal.

TRIGO, B.B. **Identification and characterization of genetic variants associated in zebu cattle's coat.** 2022. 102 f. Tese (Doutorado) – Faculdade de Medicina Veterinária, Universidade Estadual Paulista, Araçatuba, 2022.

ABSTRACT

Indicine Nellore cattle (*Bos indicus*) had as main characteristic their adaptation in tropical and subtropical climate. The pigmentation of skin and coat color are both phenotype characteristics related to its climate adaptation and heat tolerance. Several indicine breeds also had the same coat color pattern as shown by Nellore cattle. Given the potential contribution of coat color variation on their adaptation in tropical and subtropical environments the aim of the study was to map candidate genetic structure or functional variants associated with darkness of hair coat in zebu cattle. Using genome-wide association study (GWAS) and latter analysis of whole-genome sequence it was possible to identity a complex structural variant in vicinity of agouti signaling protein (ASIP) gene named as ASIP-SV1, being strongly correlated to darkness of hair coat in specific body regions of Nellore and Brahman cattle. ASIP-SV1 was also predicted to be B. indicus origin. To confirm this prediction, it was analyzed sequences from ASIP-SV1 region from indicine and non-indicine cattle to elucidate this variant segregation. It was possible to identify 45.3% from zebu animals analyzed showing ASIP-SV1 while only 6.3% from taurine animals analyzed had this structural variant. The taurine cattle with ASIP-SV1 in ASIP gene is known to had zebu introgression on their genome. It was possible to conclude that ASIP-SV1 variant cause the darkness of hair coat on specifics regions of the body of these animals increasing eumelanin production, the variant shows to be robust and results suggest that ASIP-SV1, in a set of other genes, should play a role on cattle tropical and subtropical climate adaptation.

Keywords: Bovine. *ASIP*. Animal genetics.

SUMÁRIO

1 INTRODUÇÃO GERAL

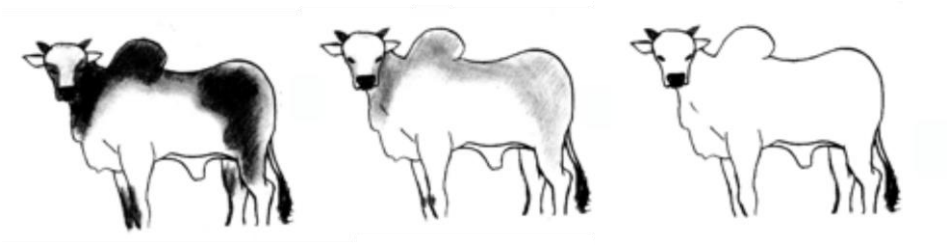
1.1 A bovinocultura no Brasil, características fenotípicas e a adaptação do gado Nelore na região tropical

Atualmente, o Brasil é o maior exportador, o segundo maior produtor e o terceiro maior consumidor de carne bovina no mundo com um rebanho de aproximadamente 187,7 milhões de cabeças de gado (ABIEC, 2022). De todo o rebanho de corte brasileiro, a raça Nelore (*Bos indicus*) representa 80% da população (ABCZ, 2022).

A raça Nelore é originária da região de Nellore-Ongole no leste da costa indiana. Também conhecida como Ongole, essa raça zebuína foi introduzida na costa indiana há mais de 4.000 anos e começou a ser importada para vários países tropicais no século XIX (AJMONE-MARSAN; GARCIA; LENSTRA, 2010). A motivação para a importação de animais desta raça para os países tropicais é devida à fácil adaptação do gado Nelore às condições extremas dos climas tropicais e subtropicais, a resistência da raça a diversas doenças e ectoparasitas, sobrevivência em pastagem de baixa qualidade nutritiva e principalmente devido à tolerância ao calor (KARTHICKEYAN et al., 2008).

Os bovinos da raça Nelore possuem, de forma geral, a pele preta. Enquanto as vacas possuem pelagem branca, os touros apresentam uma pelagem na qual a coloração pode variar de cinza claro a preto em regiões específicas do corpo como a cabeça, pescoço e barbela, traseiro e joelhos (FIGURA 1). As regiões que possuem escurecimento da pelagem na idade adulta do bovino geralmente apresentam coloração avermelhada desde o nascimento até aproximadamente um ano de idade. Um padrão interessante e curioso é que podem ser observados também touros que apresentem como padrão de coloração a pelagem branca, semelhante ao padrão observado nas fêmeas (FIGURA 1).

Figura 1 – Variação no padrão de coloração de pelagem do gado Nelore



Fonte: Adaptado de Trigo et al. 2021

Fadare et al. (2013) e Leite et al. (2018), analisando tolerância térmica em ovelhas, hipotetizaram que a tolerância ao calor pode ser contribuída por algumas características fenotípicas como por exemplo o tamanho do pelo, a pigmentação da pele e a coloração da pelagem deste animal.

Da Silva et al. (2003) descreveu que a mistura de pelos de pigmentação branca, cinza e escura, que fossem curtos (comprimento médio de 5 a 8 mm), grossos (> 50 mm), densos (> 1.400 pelos/cm^2) e que estivessem contra uma pele escura, assim como é o padrão de pelagem apresentado pelo Nelore, refletiria mais radiação de onda curta do que o padrão de pelagem apresentado pelas raças européias de gado *Bos taurus*. A radiação de onda curta está bastante presente nas regiões tropicais e subtropicais, e inclui as radiações ultra-violeta A (UVA) e B (UVB), todo o espectro de luz visível e o infravermelho, sendo este último geralmente percebido como calor pelos termorreceptores da pele. Assim, todos os fatos apresentados sugerem que o padrão de pelagem apresentado pelo Nelore contribui para a tolerância ao calor observada na raça.

Outra característica da adaptação do gado Nelore é a resistência a ectoparasitas e tal resistência foi também relacionada à coloração da pelagem em bovinos. A resistência contra ectoparasitas foi estudada em diferentes raças, e os autores concluíram que animais que apresentavam a pelagem clara se mostraram mais resistentes quando comparados com animais de pelagem escura (ANDRADE et al., 1998; CARDOSO et al., 2014).

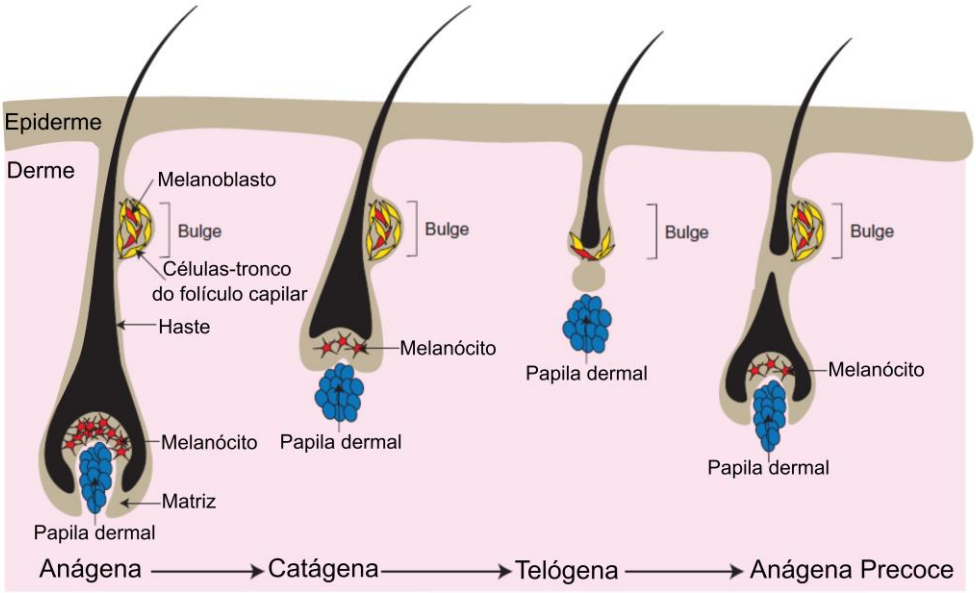
1.2 Características da pigmentação do pelo e suas vias funcionais

Como o padrão de pelagem é um dos fatores de interesse na adaptação dos zebuínos da raça Nelore nas regiões tropicais e subtropicais, é importante compreender o ciclo do pelo, as proteínas que fazem parte das vias funcionais e como elas afetam a dinâmica da pigmentação do pelo.

Os pelos são hastes queratinizadas que estão inseridas na epiderme e invaginadas na derme, onde encontra-se o folículo piloso (ERDOĞAN, 2017). O ciclo do pelo é dividido basicamente em 3 fases: (i) A fase anágena, na qual o pelo está em crescimento pela produção das células queratinizadas na matriz ou bulbo piloso; (ii) a fase catágena, na qual se observa uma redução na atividade celular e apoptose das células que formam a haste; e (iii) a fase telógena, na qual ocorre a soltura do

pelo do folículo e subsequente queda do pelo. O ciclo então se reinicia com a fase anágena ou anágena precoce (FIGURA 2) (LIM; NG; CLAVEL, 2019).

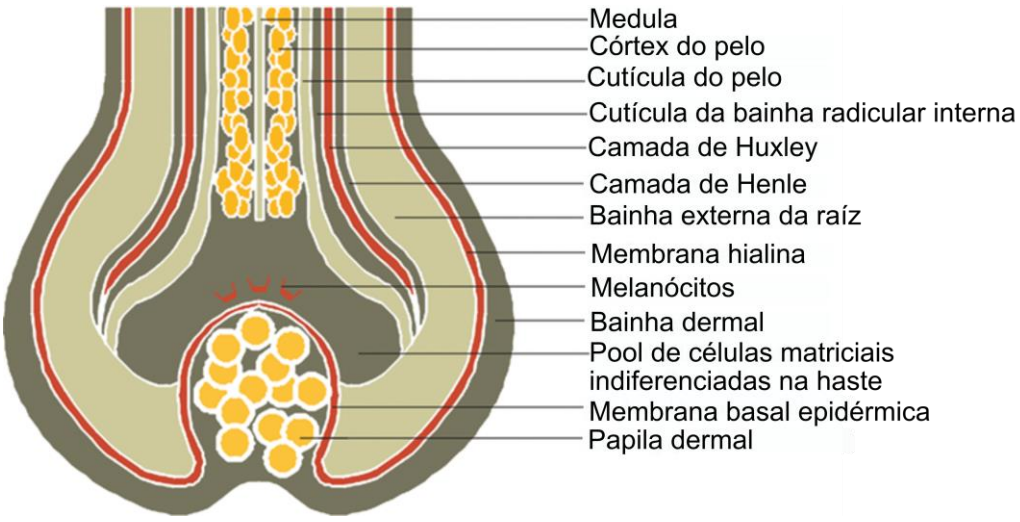
Figura 2 – Ciclo de crescimento do pelo



Fonte: Adaptado de LIM; NG; CLAVEL, 2019.

Na formação dos pelos, as células se originam na matriz e se queratinizam formando a haste pilar que é dividida em medula, córtex e cutícula (FIGURA 3). A medula é a porção central do pelo, a cutícula é a região externa do pelo e é na região intermediária do pelo, o córtex, onde é determinado o pigmento por conta da produção de melanócitos na matriz (ERDOĞAN, 2017).

Figura 3 – Estruturas presentes no pelo



Fonte: Adaptado de Erdoğan, 2017.

Os melanócitos, como dito anteriormente, são encontrados em sua maioria na matriz do pelo. Porém, a produção da melanina ocorre somente na fase anágena em unidades pigmentares específicas do folículo piloso, ocorrendo a síntese da melanina nos melanossomas nos melanócitos. A pigmentação da haste capilar ocorre no momento em que esses melanossomos sintetizados nos melanócitos são transferidos da matriz para os queratinócitos da haste (ERDOĞAN, 2017; STENN; PAUS, 2001). A melanina é dividida em dois tipos: a eumelanina, de coloração preta, e a feomelanina, de coloração vermelha ou amarela. A eumelanina é produzida a partir da reação da L-tirosina na presença da enzima Tirosinase (TYR), enquanto a feomelanina só é formada quando a L-tirosina está na presença da cisteína combinada com a falta da TYR ativada (SHARMA; WAGH; GOVINDARAJAN, 2002).

A pigmentação é uma característica de via complexa controlada por centenas de genes que são capazes de afetar a coloração e o padrão da pelagem nos mamíferos (CIESLAK et al., 2011; ROCHUS et al., 2019; SCIENSKI, 2018).

De modo geral, as colorações da pelagem são definidas de acordo com a relação entre eumelanina e feomelanina, a qual é controlada principalmente pelos agonistas e antagonistas do receptor de melanocortina 1 (*melanocortin 1 receptor* – MC1R) (CIESLAK et al., 2011; LIM; NG; CLAVEL, 2019). O gene *MC1R* produz a proteína transmembrana MC1R nos melanócitos, a qual é capaz de regular a produção de melanina.

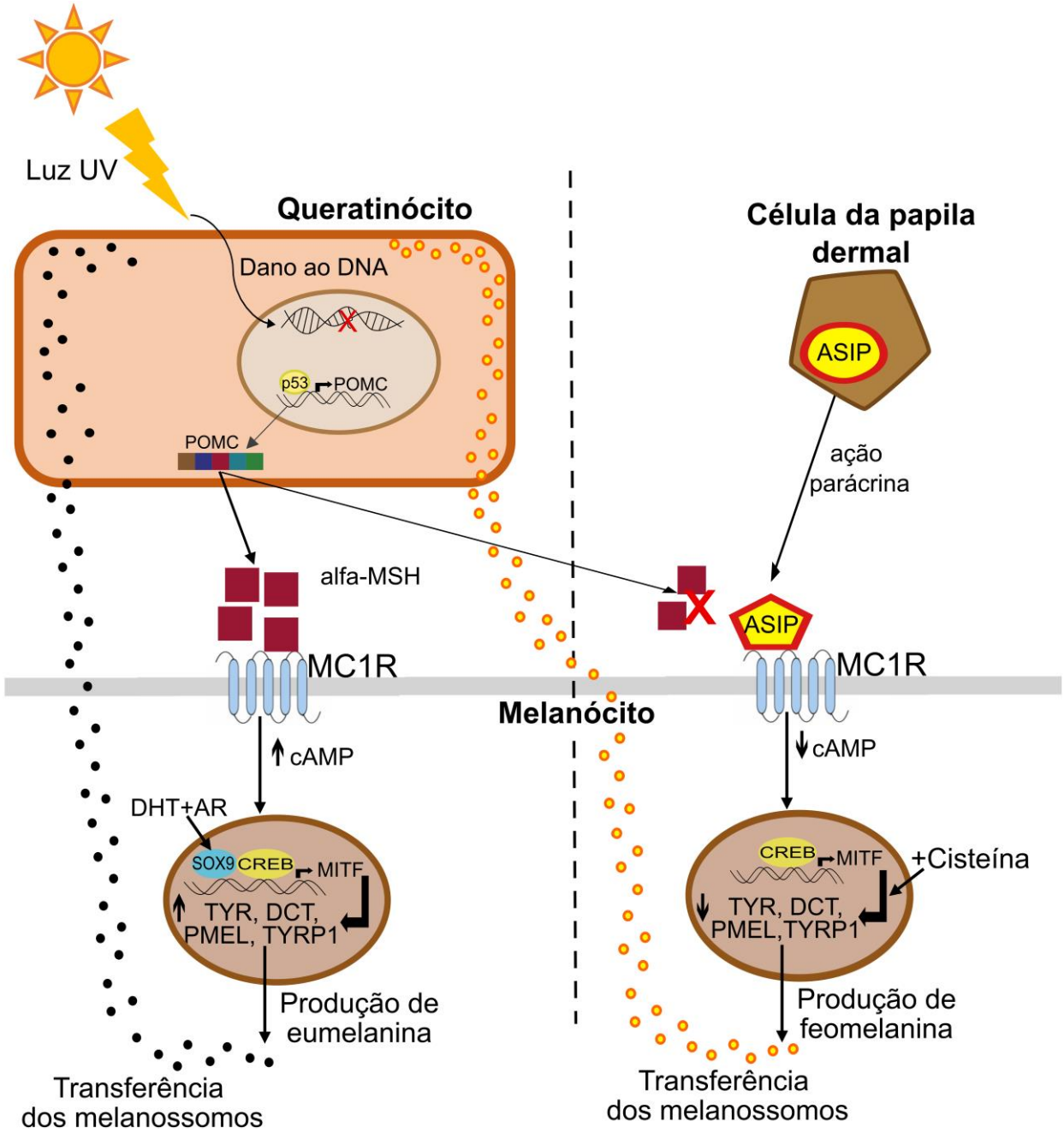
Radiações ultra-violetas (UV) causam danos no DNA, que por sua vez estabilizam a proteína supressora de tumor p53, reguladora central do ciclo celular e reparadora do dano ao DNA, que promove a ativação transcricional do gene pro-opiomelanocortina (*POMC*). O gene *POMC* codifica a produção e secreção de alguns peptídeos, incluindo a alfa-MSH (ligante primário do MC1R) que se liga no MC1R presente na membrana dos melanócitos e ativa a cascata de cAMP (do inglês, *cyclic adenosine monophosphate*) - CREB (do inglês, *cAMP response element- binding*) - MITF (do inglês, microphthalmia-associated transcription factor). Essa cascata eleva os níveis de TYR e outras enzimas biossintéticas da melanina. A eumelanina que foi sintetizada nos melanócitos é então transferida e acumulada nos queratinócitos por intermédio dos melanossomos (D'ORAZIO et al., 2013; NGUYEN; FISHER, 2019) (FIGURA 4).

Adicionalmente, participa da regulação da pigmentação a SOX9 (do inglês, *SRY sex-determining region Y-box 9*) que é expressa nos melanócitos e supra-regulada também pela exposição aos raios UVB. Juntamente com o CREB, o SOX9 é capaz de controlar o promotor MITF e TYRP, aumentando a expressão das proteínas-chave melanogênicas como a TYR e estimulando o aumento da pigmentação (ARORA; SIDDIQUI; MEHAN, 2021; PASSERON et al., 2007) (FIGURA 4).

Quando o inibidor competitivo ASIP é produzido pelas células da papila dermal e atua de forma parácrina nos melanócitos, ele evita com que o alfa-MSH se ligue ao receptor do MC1R, diminuindo assim a atividade da cAMP e alterando toda a cascata que viria a seguir diminuindo os níveis de TYR ativada e das outras enzimas biossintéticas. Além da diminuição da TYR, ocorre a inserção de cisteína na cascata resultando na síntese da feomelanina (melanossomo), a qual também será transferida e acumulada nos queratinócitos (CAVALCANTI, 2015; SHARMA; WAGH; GOVINDARAJAN, 2002), conforme observado na figura 4. Além de ser um antagonista para o MC1R, Passeron et al. (2007) demonstraram que o ASIP também é capaz de infra regular a expressão do SOX9 na cascata de pigmentação.

Controlando a proporção dos pigmentos produzidos, o gene *ASIP* age localmente como um modificador de cor extracelular através da proteína ASIP, influenciando na distribuição dos pigmentos pelo corpo (BARSH et al., 2000; CIESLAK et al., 2011; LETKO et al., 2019; SUZUKI, 2013).

Figura 4 – Via de regulação do MC1R



Fonte: Elaborado pela autora

1.3 Possíveis causas para a variação no escurecimento da pelagem do Nelore e o escurecimento da pelagem em outras raças zebuínas

O gado Nelore tem como característica fenotípica um padrão de pelagem variável que pode ser relacionado com a idade, as estações do ano, o sexo e a genética do animal.

Uma das características observadas que podem ser relacionadas à idade é o escurecimento da pelagem de avermelhado nos bezerros até a escala de cinza na idade adulta, parecendo ainda que a pelagem continua escurecendo conforme o passar da idade dos animais. A mudança na pigmentação de avermelhado para cinza pode sugerir um processo de maturação dos melanócitos relacionados à puberdade do animal e o escurecimento conforme o envelhecimento do animal é especulado por ser regulado por hormônios esteroides (FONSECA et al., 2019). O hormônio dihidrotestosterona (DHT) é a principal forma de testosterona que é sintetizada perifericamente e atua exclusivamente através do receptor de andrógeno (AR), que é um fator de transcrição intracelular presente nos queratinócitos foliculares e epidérmicos, células das papila-dermais, entre outras presentes na estrutura dérmica (ZOUBOULIS et al., 2007). O SOX9 foi apontado pelo estudo de Chen et al. (2006) por potencializar a esteroidogênese, além de ter sido detectado na pele. Diante dos fatos apresentados, é possível sugerir que a variação na intensidade de cor entre os machos e a variação na tonalidade da cor com o passar do tempo é devido à interação do DHT com o AR influenciando na regulação da expressão do SOX9 na cascata de pigmentação. Este fato também explicaria porque fêmeas Nelore não apresentam escurecimento do pelo como observado em machos.

Além dessas variações na intensidade de pigmentação, são observadas diferentes tonalidades ao longo do ano que pode ser explicada pela exposição do animal aos raios UVA e UVB que são capazes de supra-regular os fatores de transcrição envolvidos na melanogênese (D'ORAZIO et al., 2013; NGUYEN; FISHER, 2019).

Outra variação observada na pelagem é relacionada ao padrão na coloração da pelagem que difere nas fêmeas, sendo em um primeiro momento relacionada a uma variação sexual, inclusive pela interação do DHT com o AR se ligando ao SOX9 na cascata de pigmentação descrita anteriormente. No entanto, alguns touros também podem apresentar uma pelagem branca semelhante ao observado nas fêmeas, independentemente de seus níveis de DHT, idade ou estação do ano, sugerindo que a variação neste escurecimento da pelagem em machos é também determinada geneticamente.

De qualquer forma, os fatores genéticos/epigenéticos e mecanismos moleculares observados no dimorfismo sexual, a variação do escurecimento na

pelagem dos touros, e o escurecimento da pelagem dos machos relacionados à idade e à luminosidade permanecem descaracterizados no gado Nelore.

O padrão de escurecimento na pelagem em regiões específicas no corpo do gado Nelore não é exclusivo da raça e muitas outras raças zebuínas exibem padrão de coloração da pelagem semelhante. As raças indo-paquistanesas Bhagnari, Dajal, Guzerat, Hariana, Tharparkar e Ongole são exemplos com o padrão de pelagem similar, assim como outras raças bovinas presentes nas Américas e na Oceania que são derivadas da raça Ongole como o Tabapuã Brasileiro, o Peranakan Ongole Indonésio, e o Brahman Australiano e Americano.

Além dessas, outras raças taurinas oriundas da Europa Central e da Europa Oriental também apresentam o mesmo padrão de pelagem. O interessante é que algumas dessas raças já foram descritas como portadoras de haplótipos ancestrais de *B. indicus* (DECKER et al., 2014; BARBATO et al., 2020), como é o caso das raças Chianina, Corsa, Croatian Podolian, Garfagnina, Gascon, Guelmoise, Grey Húngaro, Podolian Italiano, Marchigiana, Maremmana, Piedmontese, Romagnola, Grey Turco, Grey Tirolês e raças Grey Ucranianas, entre outras.

Sabendo que a variação na pigmentação da pelagem é um potencial fator na adaptação dos bovinos em climas tropicais e subtropicais, considerando a complexidade do fenótipo de coloração de pelagem e uma possível existência de diversas vias moleculares capazes de determinar o escurecimento relacionado à idade em machos e o observado no dimorfismo sexual, supõe-se que as diferentes características seriam melhores avaliadas caso analisadas de forma individualizada.

Diante de todas as informações, é provável que esse padrão de coloração da pelagem seja explicado por variantes genéticas comuns nas raças as quais tiveram sua ancestralidade compartilhada.

1.4 Objetivo

Mapear variantes genéticas posicionais e funcionais associadas ao escurecimento da pelagem do gado zebuíno.

2 CAPÍTULO 1 – VARIANTS AT THE *ASIP* LOCUS CONTRIBUTE TO COAT COLOR DARKENING IN NELLORE CATTLE

Artigo publicado na Revista *Genomics, Selection and Evolution*, DOI: 10.1186/s12711-021-00633-2.

Beatriz B Trigo[1,2], Adam T H Utsunomiya[2,3], Alvaro A A D Fortunato[1,2,4], Marco Milanesi[1,2,3], Rafaela B P Torrecilha[2,3], Harrison Lamb[5], Loan Nguyen[5]; Elizabeth M Ross[5]; Ben Hayes[5], Rômulo C M Padula[6], Thayla S Sussai[6], Ludmilla B Zavarez[2], Rafael S Cipriano[6], Maria M T Caminhas[1], Flavia L Lopes[1], Cassiano Pelle[7], Tosso Leeb[8,9], Danika Bannasch[8,10], Derek Bickhart[11], Timothy P L Smith[12], Tad S Sonstegard[13], José F Garcia[1,2,3,14], Yuri T Utsunomiya[1,2,3]*

[1]São Paulo State University (Unesp), School of Veterinary Medicine, Araçatuba, Department of Production and Animal Health, São Paulo, Brazil

[2]International Atomic Energy Agency (IAEA) Collaborating Centre on Animal Genomics and Bioinformatics, Araçatuba, São Paulo, Brazil

[3]AgroPartners Consulting, R. Floriano Peixoto, 120 - Sala 43A - Centro, Araçatuba - SP, 16010-220

[4]Personal-PEC, R. Sebastião Lima, 1336 - Centro, Campo Grande - MS, 79004-600, Brazil

[5]Centre for Animal Science, Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Australia

[6]Centro Universitário Católico Salesiano, Araçatuba, São Paulo, Brazil

[7]CRV Lagoa, Sertãozinho, São Paulo, Brazil

[8]Institute of Genetics, Vetsuisse-Faculty, University of Bern, Bremgartenstrasse 109A, 3012, Bern, Switzerland

[9]Dermfocus, University of Bern, Bremgartenstrasse 109A, 3012, Bern, Switzerland

[10]Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA 95616, USA

[11]Dairy Forage Research Center, USDA-ARS, 1925 Linden Drive, Madison, WI, 53706, USA

[12]US. Meat Animal Research Center, USDA-ARS, 844 Road 313, Clay Center, NE 68933, USA

[13]Recombinetics Inc., St. Paul, MN, USA
[14]São Paulo State University (Unesp), School of Agriculture and Veterinarian Sciences, Jaboticabal, Department of Preventive Veterinary Medicine and Animal Reproduction, São Paulo, Brazil
*Corresponding autor

2.1 Resumo

O gado Nelore (*Bos indicus*) é conhecido pela sua adaptação em ambientes quentes e úmidos. O tamanho da pelo e a coloração da pelagem podem contribuir para a tolerância ao calor. A raça Nelore foi fortemente selecionada para a pelagem branca, mas os touros geralmente apresentam uma pelagem com uma coloração escura em escala que vai de cinza claro a preto nas regiões da cabeça, pescoço, traseiro e joelhos. Dado a potencial contribuição da variação na coloração de pelagem na adaptação dos gados em regiões tropicais e sub-tropicais, nosso objetivo foi mapear variantes genéticas candidatas posicionais e funcionais associadas ao escurecimento da pelagem em touros Nelore. Realizamos um estudo de associação genômica ampla (do inglês, GWAS) para o escurecimento de pelagem utilizando dados de 432 touros Nelore que foram genotipados para mais de 777 mil marcadores de polimorfismos de nucleotídeos únicos (SNP). Um único sinal principal foi detectado próximo do gene *proteína de sinalização aguti* (*ASIP*). A análise de dados do sequenciamento genômico completo de 21 touros revelou variantes funcionais associadas ao escurecimento de pelagem, incluindo um rearranjo estrutural envolvendo o *ASIP* (ASIP-SV1). Caracterizamos a variante estrutural utilizando dados de sequenciamento Oxford Nanopore de 13 novilhas Brahman australianas, as quais compartilham ancestralidade com o gado Nelore. Encontramos que essa variante originou uma deleção de 1155-pb seguida de uma inserção de elemento transponível de mais de 150 pb que podem impactar no recrutamento de exons não codantes do gene ASIP. Nossos resultados indicaram que a variante na sequência do ASIP causa escurecimento na pigmentação da pelagem e regiões específicas do corpo, provavelmente através da diminuição na expressão do ASIP e consequente aumento na produção da eumelanina.

Palavras-Chave: *Bos indicus.* GWAS. CNV. Deleção. Isoformas.

2.2 Abstract

Nellore cattle (*Bos indicus*) are well-known for their adaptation to warm and humid environments. Hair length and coat color may impact heat tolerance. The Nellore breed has been strongly selected for a white coat, but bulls generally exhibit darker hair ranging from light grey to black on the head, neck, hump, and knees. Given the potential contribution of coat color variation to the adaptation of cattle populations to tropical and sub-tropical environments, our aim was to map positional and functional candidate genetic variants associated with the darkness of hair coat (DHC) in Nellore bulls. We performed a genome-wide association study (GWAS) for DHC using data from 432 Nellore bulls that were genotyped for more than 777k single nucleotide polymorphism (SNP) markers. A single major association signal was detected in the vicinity of the *agouti signaling protein* gene (*ASIP*). The analysis of whole-genome sequence (WGS) data from 21 bulls revealed functional variants that are associated with DHC, including a structural rearrangement involving *ASIP* (*ASIP*-SV1). We further characterized this structural variant using Oxford Nanopore sequencing data from 13 Australian Brahman heifers, which share ancestry with Nellore cattle. We found that this variant originates from a 1155-bp deletion followed by an insertion of a transposable element of more than 150 bp that may impact the recruitment of *ASIP* non-coding exons. Our results indicate that the variant *ASIP* sequence causes darker coat pigmentation on specific parts of the body, most likely through a decreased expression of *ASIP* and consequently an increased production of eumelanin.

Keywords: *Bos indicus.* GWAS. CNV. Deletion. Isoforms.

2.3 Background [1]

Brazil is the largest exporter, second-largest producer, and third largest consumer of beef in the world [1]. Approximately 80% of the beef cattle in Brazil are Nellore. This *Bos indicus* cattle breed is native to the Nellore-Ongole region in the east coast of India and has been imported by many tropical countries since the late

[1] Artigo publicado na Revista *Genomics, Selection and Evolution*, DOI: 10.1186/s12711-021-00633-2. Normas de publicação da revista (ANEXO A)

19th century [2]. These imports were motivated by the adaptation of Nellore cattle to the challenging conditions of tropical and sub-tropical climates, since the breed is resistant to several diseases, survives on low-quality forage, and tolerates heat [3].

Length of the hair coat, skin pigmentation and coat color are traits that are often hypothesized to contribute to heat tolerance in mammals [4, 5]. In fact, the mixture of white/grey and dark hair that is short (5 to 8 mm average length), thick (> 50 mm), dense (> 1400 hair/cm^{-2}) and placed against a black skin, provides higher reflectance at shorter light wavelengths (of particular interest in tropical regions) in Nellore than in European *Bos taurus* cattle breeds [6]. However, Nellore cattle exhibit variation in coat color patterns that are associated with sex, age, and genetic background.

Nellore cattle have black skin with cows presenting a near-white hair coat and bulls generally exhibiting darker hair ranging from light grey to black, especially on the head, neck, hump, and knees. This observation points to sex as a first source of phenotypic variation. Curiously, some bulls also have a near-white coat that resembles that of females, which is further evidence that variation in the darkness of hair coat (DHC) in males is genetically determined. The parts of the body of the adult animal that present darker hair are generally reddish from birth to yearling, which suggests a melanocyte maturation process linked to puberty. Hair coat color seems to continue to darken throughout the adult life, an aging-related phenomenon that is speculated to be regulated by steroid hormones and the prolactin receptor in hair follicles [7]. However, the molecular mechanisms and genetic/epigenetic factors that underlie the observed sexual dimorphism, the variation in DHC between bulls, and the age-related darkening of hair coat in males remain uncharacterized in the Nellore breed. Given the potential complexity of coat color patterns in Nellore cattle and the putative existence of different molecular pathways governing each one of the three mentioned sources of variation, we hypothesize that they are better understood as separate, intermediate traits.

Several *B. indicus* breeds exhibit coat color patterns that are similar to those observed in Nellore cattle. Examples include the Indian-Pakistani breeds Tharparkar, Bhagnari, Dajal, Hariana, Guzerat and Ongole (from which Nellore are derived). Similar patterns are also seen in other cattle populations derived from Ongole that are reared in the Americas and Oceania, such as the Brazilian Tabapuã,

the Indonesian Peranakan Ongole, and the Australian and American Brahman. In addition, several Central and Eastern European *B. taurus* breeds also display white/grey hair-coat, some of which have been reported to carry haplotypes of *B. indicus* ancestry [8, 9]. This is the case of the Chianina, Corsa, Croatian Podolian, Garfagnina, Gascon, Guelmoise, Hungarian Grey, Italian Podolian, Marchigiana, Maremmana, Piedmontese, Romagnola, Turkish Grey, Tyrolean Grey and Ukrainian Grey breeds, among others. Therefore, it is likely that common genetic variants explain this coat color pattern across breeds with shared ancestry.

Senczuk and colleagues [10] performed a genetic divergence analysis in which the above-mentioned *B. taurus* breeds were contrasted with four non-grey northwestern cattle populations (Angus, Charolais, Limousin, and Holstein), and suggested that three loci may be linked to white/grey hair-coat: CHR2:6,510,630-7,010,630, CHR14:22,531,305-25,722,332 and CHR26:22,789,524-23,289,524 based on the ARS-UCD1.2 bovine genome assembly [11]. Holland [12] also mapped white/grey coat color in a Nellore-Angus crossbred population to chromosome 6, within a segment that harbors genes known to impact coat color such as *v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog* (*KIT*).

Since coat color variation may contribute to the adaptation of cattle populations to tropical and sub-tropical environments, and given the interesting complexity of the coat color phenotype in the Nellore breed, our aim was to map positional and functional candidate genetic variants associated with DHC in Nellore bulls. As argued above, the age-related darkening of hair coat in males and the sexual dimorphism component should be considered as separate traits for the sake of simplicity, and thus are the focus of separate ongoing studies.

2.4 Methods

2.4.1 Phenotypes

In total, 432 bulls were included in our genome-wide association study (GWAS). Photographs of the adult bulls, which were taken while they were on artificial insemination stations for semen sample collection, were inspected by six evaluators. Each animal was assigned a visual score based on a subjective scale ranging from 0 to 3. The lowest score (0) corresponded to completely white animals and the highest score (3) to animals with extremely dark hair on the head, neck,

hump, and knees. The scores of the six evaluators were averaged to achieve a final score that was adopted as the response variable in our GWAS screening. The distribution of the resulting phenotypes is in Figure S1 [see Additional file 1 Figure S1].

2.4.2 Genotypes

Illumina® BovineHD BeadChip (777k chip) genotypes from previous studies [13, 14] were available for all the bulls. The pre-commercial single nucleotide polymorphism (SNP) panel included 786,799 markers. Prior to the association analysis, the PLINK v1.9b4.6 [15, 16] software was used to exclude SNPs with a call rate lower than 90% or a minor allele frequency lower than 2%. All bulls had a minimum genotype call rate of 90%.

2.4.3 Genome-wide association study

Phenotypes were regressed onto genotypes using the mixed linear model analysis with the leave-one-chromosome-out (MLMA-LOCO) procedure in the GCTA v.1.90.2 beta software [17, 18]. This procedure accounts for population structure using the genomic relationship matrix between individuals. The inflation factor of the squared test statistics was measured as the slope of a linear regression between observed and theoretical quantiles in R version v3.6.2 [19]. The GWAS was performed a second time including the fixed effect of the top scoring SNP to test for allelic and gametic phase heterogeneity, i.e., for association signals that are driven by multiple underlying causal variants that are either independent or only partially correlated with each other.

2.4.4 Test for the presence of a dominance effect

The top scoring SNP in the GWAS screening was tested for the presence of a dominance effect. Following Falconer and Mackay [20], the model below was fitted to the data using ordinary least squares regression:

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{m}\alpha + \mathbf{w}d + \mathbf{e},$$

where α is the allele substitution effect, d is the dominance effect, and $\mathbf{m}$ and $\mathbf{w}$ are vectors relating $\mathbf{y}$ to α and d, respectively. For vector $\mathbf{m}$, genotypes 0, 1 and 2 were re-coded as $0-2p$, $1-2p$ and $2-2p$, respectively, where p is the

frequency of the counted allele. Conversely, vector $\mathbf{w}$ assumed values $-2p^2$, $2pq$ and $-2q^2$ for genotypes 0, 1 and 2, respectively, where $q = 1-p$. In this setting, α represented the average effect of extra copies of the counted allele, which is a function of both additive and dominance genetic effects, i.e. $\alpha = a + (q-p)d$. This model is convenient because $COV[\mathbf{m}, \mathbf{w}] = 0$, thus avoiding co-linearity in the simultaneous estimation of additive and dominance effects [21]. The coefficient of determination (R^2) of this regression model was adopted as a proxy for the proportion of phenotypic variance explained by the tested SNP. This model was fitted to the data in R version v3.6.2 [19].

2.4.5 Analysis of haplotype diversity

To assess haplotype diversity at candidate loci, chromosomes that presented evidence of association with DHC were subjected to phase inference with the software Eagle v2.4.1 [22]. Then, we used the GHap v2.0.0 R package [23] to extract haplotype alleles within the chromosomal segments that presented peak associations. The distribution of phenotypes conditional on haplotypes was then inspected using boxplots in R version v3.6.2 [19].

2.4.6 Whole-genome sequences of Nellore bulls

Whole-genome sequencing (WGS) data were available for 17 Nellore bulls from a previous study [14]. These animals had been sequenced on an Illumina HiSeq 2000 instrument at an average coverage of ~9x using paired-end reads of 100 bp. Since none of these sequenced bulls had an average phenotype score of zero, the data was complemented with additional sequences from four white bulls that were not included in the GWAS. DNA samples from these animals were processed with the TruSeq Nano library preparation kit (Illumina) and then sequenced at ~10x coverage on the Illumina Novaseq6000 platform using paired-end reads of 150 bp. All 21 bulls had their paired-end reads aligned against the ARS-UCD1.2 bovine genome assembly [11] with the Burrows-Wheeler Alignment (BWA) mem algorithm [24]. Optical and PCR duplicates were marked with the PicardTools v1.119 software (available at: http://broadinstitute.github.io/picard/). Single nucleotide variants (SNVs) and small insertions-deletions (INDEL) were extracted from aligned reads using the

mpileup algorithm from SAMtools v1.9 and BCFtools v1.10.2 [25]. Variant effects were predicted and annotated with the Ensembl Variant Effect Predictor tool (VEP) [26]. Structural variants (SV) were inferred with the Gaussian Mixture Model implemented in the CNVcaller toolkit [27].

2.4.7 Identification of putative causal variants

Sequence variants within each candidate region were further tested for associations with phenotypes using ordinary least squares regression in R version v3.6.2 [19]. Because of the limited sample size (21 animals), the resulting *p*-values were used only as auxiliary indicators for the location of putative causal variants. The Integrative Genomics Viewer (IGV) software [28] was used to visually confirm the existence of candidate variants and manually curate genotype calls made by CNVcaller, and to reveal additional structural variants that might have remained undetected by the described bioinformatics pipeline.

2.4.8 Simulation of candidate structural variants

We performed a series of trial-and-error simulations of short-read data and of mutant genomes based on the ARS-UCD1.2 bovine assembly. For the simulation of short reads, the suspected structural arrangements were created in FASTA files by manually editing the ARS-UCD1.2 sequence. The resulting files were used to generate simulated paired-end reads with the wgsim v0.3.1-r13 program from SAMtools [25]. In the second set of simulations, the modified FASTA files were used as the reference genome, and real sequence reads were aligned against the simulated mutant genome. Different structural arrangements were tested until the pattern of read alignments in the simulations became indistinguishable from that of the empirical data. All alignments were performed with the BWA mem algorithm [24].

2.4.9 Nanopore sequences of Brahman heifers

To evaluate independently candidate structural variants identified in the Nellore sequence data, we produced Oxford Nanopore Technologies (ONT) sequences from DNA extracted from the tail hair of 13 Australian Brahman heifers. Since the Brahman breed shares ancestry with Nellore cattle, the ONT Brahman sequence set not only served as a validation set but also helped improve the

resolution of structural variants via the use of long-read data. The ligation sequencing library preparation kit (SQK-LSK109) was used with a single R9.4.1 flow cell per animal on a MinION sequencer. Each flow cell was run for 96 hours with two to three nuclease flushes to increase the flow cell yield. An average coverage of ~ 8.73× was obtained across all samples with a maximum of 13.6× and minimum of 6.21×, and the average read length was 5.7 kb. Bases were called from fast5 signals using the Guppy software (v4.0.11, released August 2020, Oxford Nanopore Technologies). Long reads were aligned against the ARS-UCD1.2 bovine assembly using the Minimap2 software v2.14 [29] with the default ONT sequencing settings and the alignments were visualized with IGV [28].

2.4.10 Annotation of transcripts

We used cDNA and annotation data from previous reports [30-33] to improve the annotation of functional candidate genes in the ARS-UCD1.2 bovine assembly. When coordinates for transcript elements (e.g., exons, introns, and UTR) were not available, cDNA sequences were aligned to the ARS-UCD1.2 bovine genome assembly to obtain approximate coordinates using BLASTN in Ensembl (available at: https://www.ensembl.org/Bos_taurus/Tools/Blast) and BLAST in NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi).

2.4.11 Estimation of coalescence time

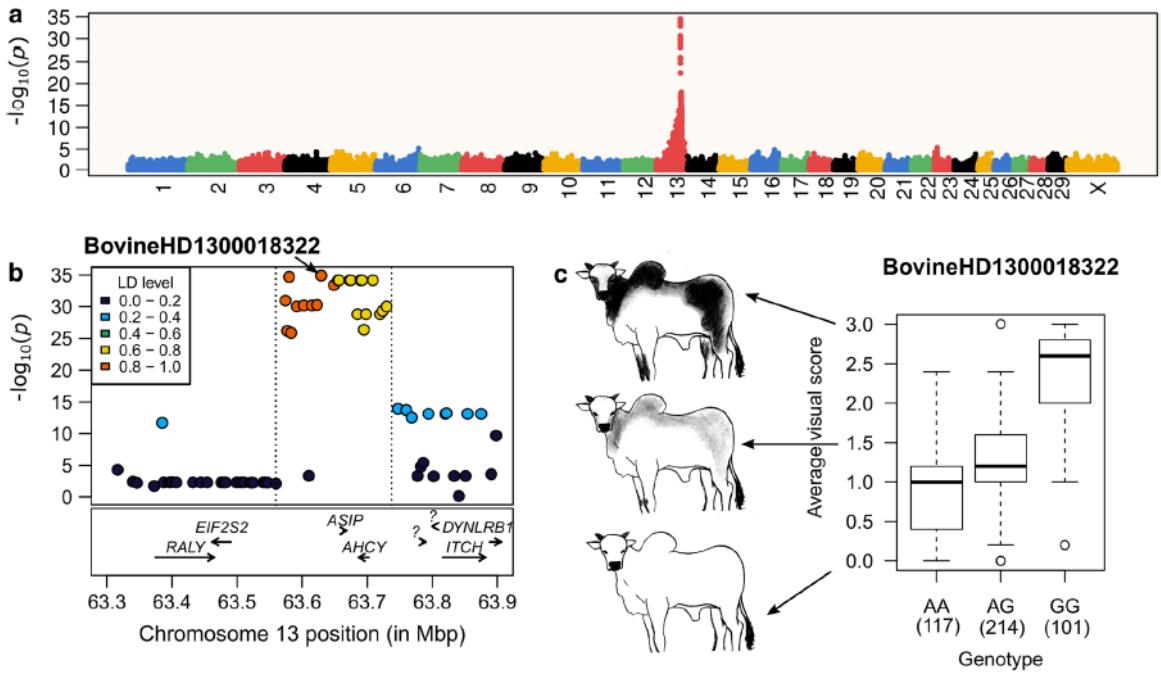
We used the Genealogical Estimation of Variant Age (GEVA) method [34] to estimate the time of coalescence (in generation units) between carrier and non-carrier haplotypes of putative causal variants. This analysis was based on the 21 Nellore whole-genome sequences. Prior to the analysis, targeted chromosomes were filtered for bi-allelic variants presenting a minimum quality score of 60, a genotype quality score of 20, a call rate of 95% and a minimum allele count of 3, and then phased with the Eagle v2.4.1 software [22]. The effective population size, a parameter required by the GEVA method, was obtained from previously reported chromosome-specific estimates [35].

2.5 Results

2.5.1 GWAS maps DHC to chromosome 13 in Nellore cattle

After filtering, 541,919 SNPs were screened for associations with DHC. The inflation factor was equal to 1.055, which indicates that the GWAS was properly corrected for relatedness and population substructure. A single GWAS hit was found on chromosome 13 (Fig. 1a). The most significant SNP, namely g.13:63,629,244A>G (rs109334889 or BovineHD1300018322, p = 1.27 x 10^{-35}), had an alternative allele frequency of 48.1% in our sample, and was located approximately 33.5 kb upstream of the *ASIP* gene (Fig. 1b). This GWAS hit disappeared when we repeated the analysis by including this SNP as a fixed effect, which indicated that the signal was most likely driven by a single underlying causal variant or haplotype [see Additional file 2 Figure S2]. Testing the marker for the presence of a dominance effect revealed estimates of α = 0.702±0.040 (p =1.42×10^{-52}) and d = -0.467±0.057 (p = 1.84×10^{-15}). This result suggests that DHC is associated with the alternative *G* allele in an additive pattern in Nellore cattle, but also that the reference *A* allele is related to a dominance deviation towards lighter coats than what would be expected for heterozygous animals on average (Fig. 1c). Furthermore, the leading SNP alone explained 46.6% of the variance in DHC.

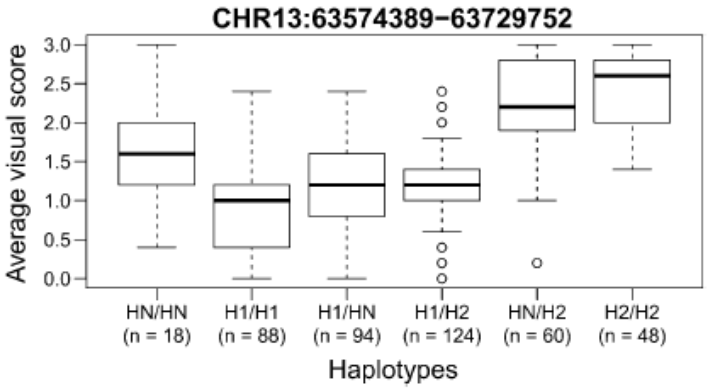
Figure 1 - Genome-wide association analysis for DHC in Nellore cattle.



(**a**) A single major locus mapping to chromosome 13 was detected (highest associated SNP, BovineHD1300018322; p = 1.27×10^{-35}). (**b**) Peak associations were found in a region (dashed lines) spanning the *ASIP* (CHR13:63,662,796-63,668,123) and *AHCY* (CHR13:63,686,723-63,702,437) genes. (**c**) The alternative G allele was correlated with increased darkness of hair coat in an additive pattern (p = 1.42×10^{-52}). Heterozygotes had their median phenotypic value skewed towards the median of the AA genotypic class, revealing a potential dominance effect associated with the reference A allele (p = 1.84×10^{-15}).

We analyzed the haplotypic diversity in the chromosome 13 region that presented peak associations (CHR13:63,574,389- 63,729,752) and included 24 SNPs (delimited by dashed lines in Fig. 1b). Eighteen haplotype alleles were identified, among which only two had a frequency higher than 5%. As expected, these two haplotypes, referred to as H1 (frequency of 45.6%) and H2 (frequency of 32.4%), had contrasting phenotypic distributions, with the H2 carriers tending to have a darker hair coat (Fig. 2). These results also support the hypothesis of a single underlying causal variant or haplotype driving the association signal at the candidate locus.

Figure 2 - Phenotypic distribution conditional on haplotypes at the DHC association signal.



Haplotype alleles were called on a block of 24 SNPs mapping to the association region on chromosome 13. Only two alleles had a frequency of at least 5%, namely H1 = TTGTATGTAACAATTGAAGGCCAA (frequency of 45.6%) and H2 = GCACGCGCGGTGGCCAGGAAATGG (frequency of 32.4%). The remaining 16 alleles were grouped in a single cluster (HN) for clarity. Alleles H1 and H2 exhibited contrasting phenotypic distributions, with H2 being involved with darker hair coat.

2.5.2 Analysis of WGS data of Nellore bulls indicate that the causative variant is likely to have a regulatory effect

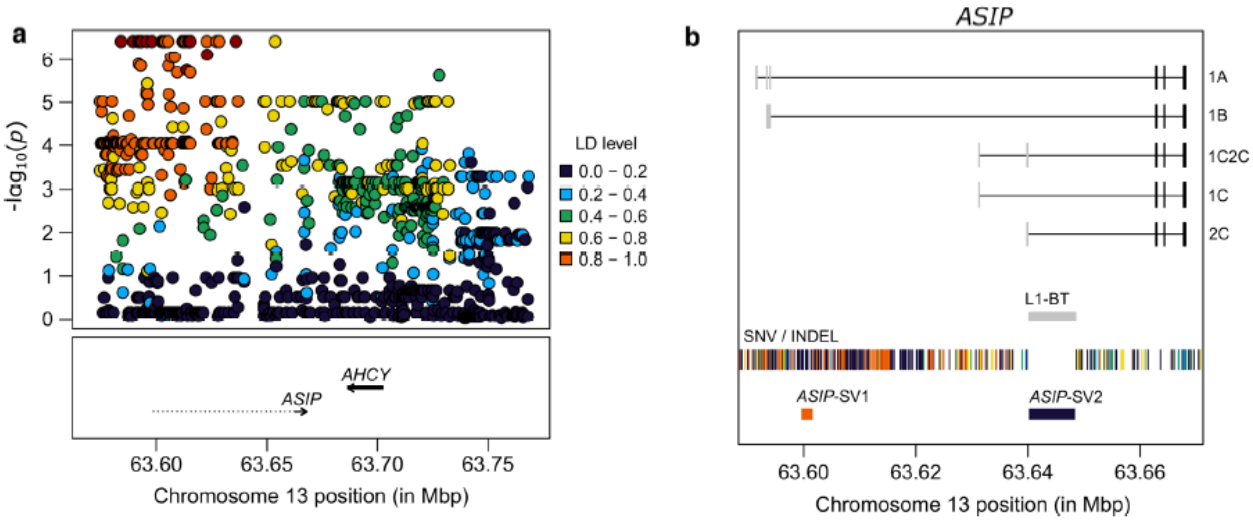
In total, 1,098 sequence variants that mapped to the region with a peak association (delimited by dashed lines in Fig. 1b) were extracted from WGS data of Nellore bulls. These included 932 intergenic, 65 intronic, 60 downstream, 32 upstream, five 3'-UTR, three synonymous, and one in-frame deletion variants in the *ASIP* gene [see Additional file 3 Table S1]. In addition, two structural variants (SV) were found within the same chromosomal window. Using the phenotypes from the 21 sequenced bulls, we found that the variants located upstream of the *ASIP* gene were more strongly associated with DHC (Fig. 3a) than the other variants, which indicates that the underlying causal variant is less likely to be located in the coding regions of the gene. However, we found that an in-frame deletion that deleted the CGGACC sequence from the last exon of *ASIP* (rs519457228, located at 13:63,667,797-63,667,802) was a promising functional candidate. Unfortunately, it was poorly genotyped in our study, with a call rate of ~67%, suggesting alignment or sequencing issues. Further inspection of read alignments with IGV revealed that the coverage of the 17 WGS obtained from a previous study [14] was poor for that exon, which is GC-

rich. This low coverage was most likely caused by PCR bias during library preparation, since these sequences were generated using legacy library preparation kits that are prone to amplification bias [36]. The sequence coverage drop was not observed in the four WGS generated with the TruSeq Nano kit for the current study, which supports our hypothesis of PCR bias. Thus, although rs519457228 was not included in the region with the strongest associations for DHC, we did not discard it as a candidate causal variant.

2.5.3 Improved annotation of *ASIP* reveals alternative transcripts that recruit non-coding exons

Because we found strong associations that mapped upstream of the *ASIP* gene, we explored the presence of putative regulatory elements and additional exons and introns that could be missing in the current gene annotation. By manually curating previously described bovine *ASIP* transcripts [30-33], we found that the 5'-end of the *ASIP* gene is at least 71 kb longer than in the current annotation of the ARS-UCD1.2 bovine genome assembly (Ensembl release 102). Analysis of this additional sequence revealed that the top scoring upstream variants overlapped two bovine transcripts that recruit non-coding exons of *ASIP*, historically termed 1A and 1B (Fig. 3b). Due to limitations in predicting the functional impact of the candidate variants on these transcripts, we could not prioritize any particular SNV or small INDEL for further analysis, but they were retained as plausible positional candidates [see Additional file 3 Table S1]. However, we did attempt to improve the resolution of the larger structural variants spanning the region, since they are more likely to impact gene expression and transcript diversity.

Figure 3 - **Regional association plots for DHC in 21 whole genome-sequenced Nellore bulls.**



(**a**) Peak associations were clustered upstream of the *ASIP* gene. The dashed line that extends upstream of the *ASIP* gene spans non-coding exons and introns that are not currently annotated in the ARS-UCD1.2 bovine genome assembly (Ensembl release 102). Linkage disequilibrium (LD) levels were calculated against the most significant variant in the region. (**b**) Details of the *ASIP* transcripts that differ in non-coding exons [33-35]. Single nucleotide variants (SNV) and small insertion/deletions (INDEL) are displayed as vertical bars, whereas structural variants (SV) are shown as rectangles. The long interspersed nuclear element (LINE) marked as L1-BT (*ASIP*-SV2) is responsible for a non-coding exon that is recruited by transcripts 2C and 1C2C. The first rectangle in the SV track (*ASIP*-SV1) is a 1,155 bp deletion significantly associated with dark hair in the 21 sequenced animals (p = 9.12×10^{-5}).

2.5.4 A 1155-bp deletion that overlaps with alternative *ASIP* transcripts is associated with dark hair

One of the detected structural variants (hereafter referred to as *ASIP*-SV1) overlapped with the transcripts 1A and 1B and comprised a 1155-bp deletion spanning the region CHR13:63,599,803-63,600,957. The deletion allele had an estimated effect of 0.900 ± 0.182 (p = 9.12×10^{-5}) in the regression analysis, and thus correlated with darker hair. The second structural variant (hereafter referred to as *ASIP*-SV2) was also a deletion with respect to the reference genome, which spanned the L1-BT repeat (CHR13:63,639,817-63,648,206) known to recruit an additional non-coding exon in bovine transcripts 2C and 1C2C [31, 32]. However, the 21 Nellore

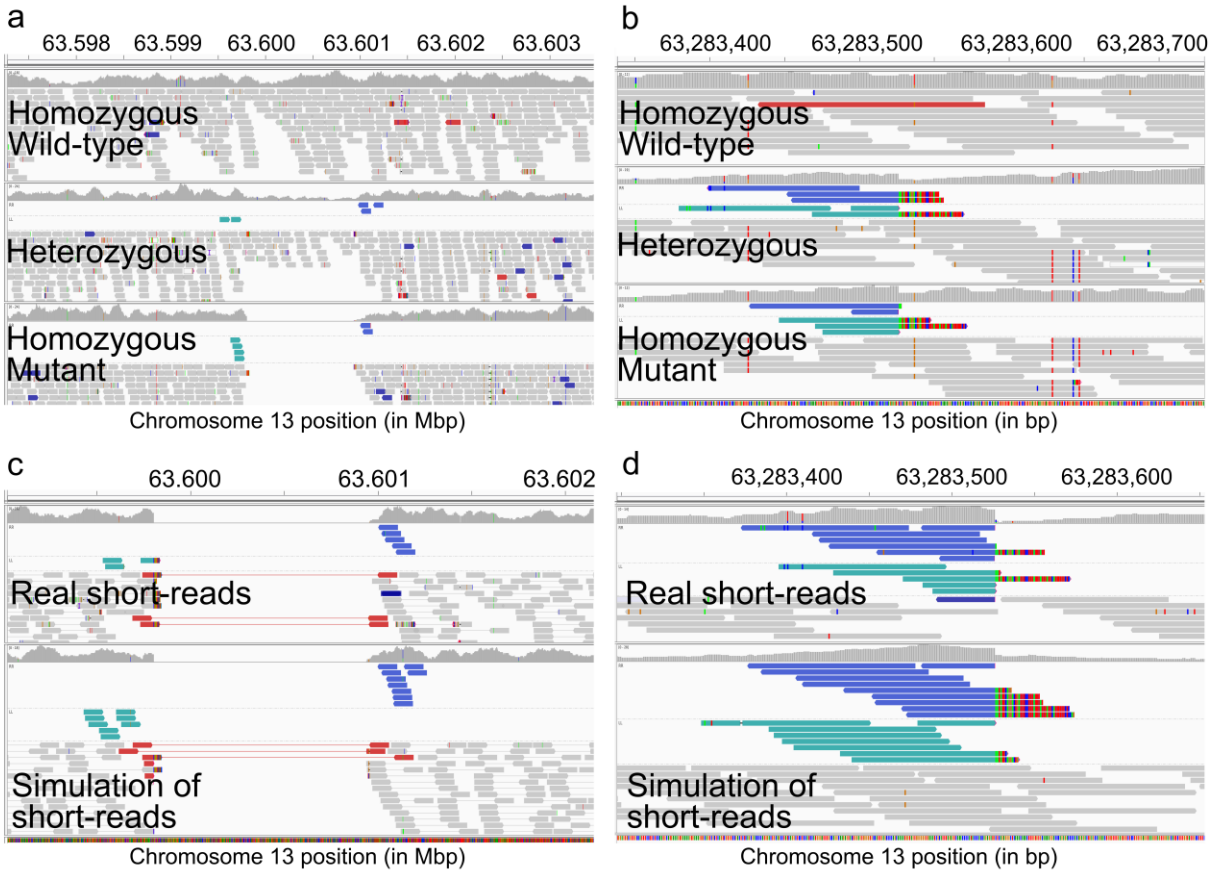
bulls analyzed here lacked the L1-BT insertion. Therefore, *ASIP*-SV2 was less likely to affect DHC in Nellore cattle and consequently not further analyzed.

2.5.5 The 1155-bp deletion serves as an insertion site for a duplicated inversion

Close inspection of *ASIP*-SV1 with IGV revealed the presence of read pairs that had an unexpected orientation (Fig. 4a). Whereas regular reads should be oriented inwards with respect to their inserts (RL orientation), for some bulls the read pairs had a single orientation, either pointing to the 3' (RR orientation) or 5' (LL orientation) end of the chromosome. In addition, the inserts for these single-orientation reads were typically longer than 316 kb, with paired-end reads mapping towards position CHR13:63,283,374 (Fig. 4b). Chimeric reads, as well as soft- and hard-clipped reads, were observed near the position CHR13:63,283,374. A BLAST analysis of the chimeric reads against the reference genome further confirmed that the foreign sequences belonged to the *ASIP*-SV1 region.

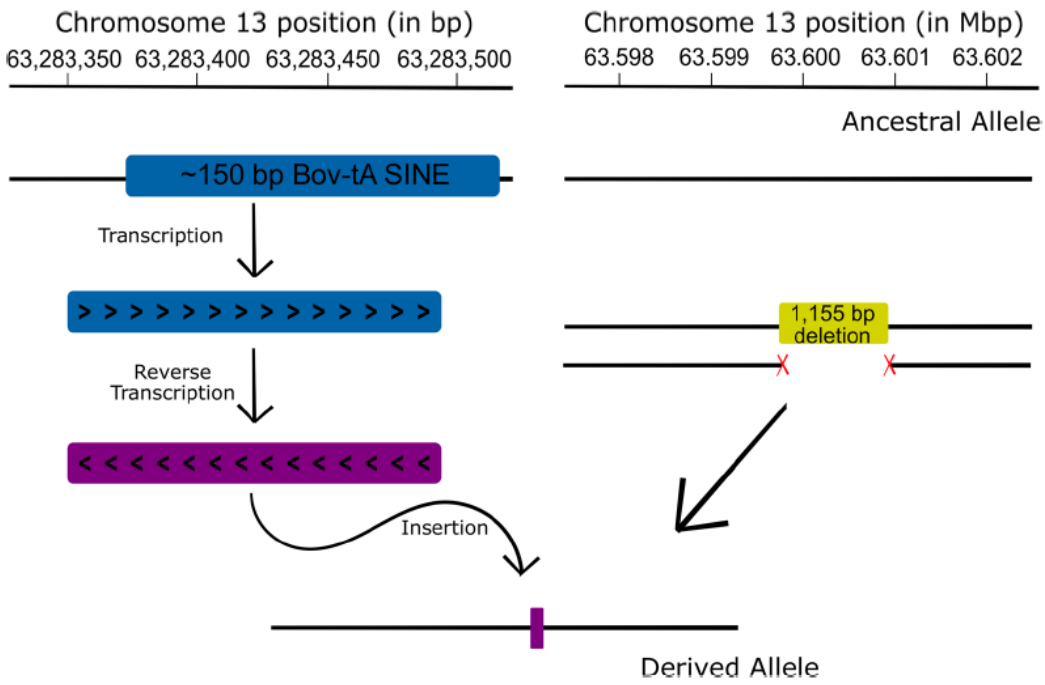
Taken together, our results suggest that the *ASIP*-SV1 region serves as an insertion site for an inverted duplication coming from CHR13:63,283,374. We performed several trial-and-error simulations of paired-end reads to find a structural rearrangement that could produce a good fit to the observed data. The best fitting simulation was a duplication of the CHR13:63,283,374-63,283,523 region, which had its reverse complement inserted at position CHR13:63,599,803 (Fig. 4c and d). We found that the inserted reverse complement had a match of 105 nucleotides with the Bov-tA SINE, which suggests that the inserted sequence was likely acquired through a retrotransposition event of an interspersed repeat (Fig. 5).

Figure 4 - IGV screenshots for Illumina short-read alignments spanning a structural variant (*ASIP*-SV1) associated with DHC in Nellore cattle.



(**a**) Presents an overview of the 1155 bp deletion at CHR13:63,599,803-63,600,957. Carriers presented read pairs with RR (cyan) and LL (blue) orientation flanking the deletion. (**b**) Shows a close inspection of RR and LL paired-end reads at the chromosome 13 position 13:63,283,374, revealing chimeric and soft- and hard-clipped reads. Simulations presented in panels (**c**) and (**d**) show that these alignments are consistent with the ~1-kb deletion being an anchoring point for the insertion of a reverse complement of the CHR13:63,283,374-63,283,523 sequence, which shares high similarity with the Bov-tA SINE. The wild type and mutant alleles were determined based on comparative genomics analyses presented in Additional file 5.

Figure 5 - Schematic representation of the hypothesized mutation event leading to the structural variant (*ASIP*-SV1) associated with DHC in Nellore cattle.



An expressed Bov-tA SINE was likely retrotranscribed and inserted in replacement of the 1,155 bp sequence at CHR13:63,599,803-63,600,957.

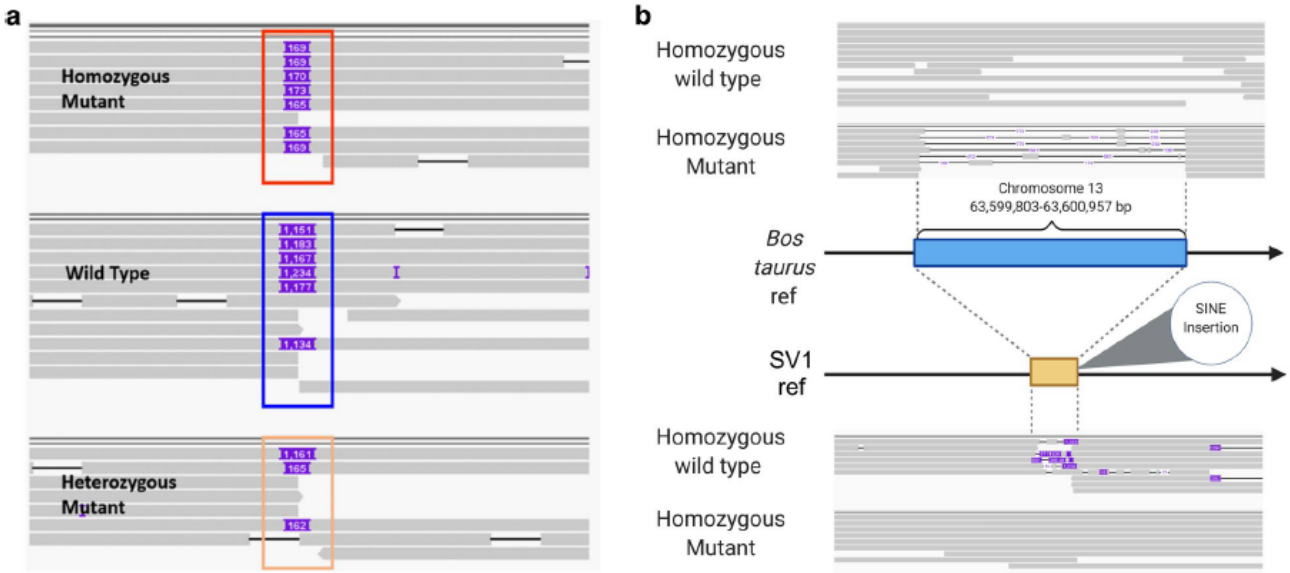
Using coalescent analysis, we estimated that the carrier and non-carrier haplotypes of this structural variant derived from a common ancestor that lived around 864 generations back, which translates into 3,420 years ago assuming a generation interval of five years in cattle. Taking this estimate as a lower boundary for the age of the mutation, *ASIP*-SV1 is most likely an ancient derived allele that segregated in the *B. indicus* lineage. Therefore, *ASIP*-SV1 is expected to be present in other *B. indicus* breeds or even *B. taurus* populations that were historically introgressed with *B. indicus*.

2.5.6 Analysis of nanopore sequence data of Brahman heifers refines the *ASIP*-SV1 structural variant

To confirm the *ASIP*-SV1 structural variant, we inspected the candidate region in ~8x ONT data of 13 Brahman heifers from Australia (Fig. 6). Alignment of these long reads against the bovine genome assembly ARS-UCD1.2 revealed at least four carriers of the mutant haplotype. Due to the low genome coverage, we

were unable to confidently determine the homozygous wild type, heterozygous and homozygous mutant genotypes across all the samples, but the list of the most likely genotypes are in Table S2 [Additional file 4 Table S2]. Nevertheless, these data allowed us to confirm the existence of the *ASIP*-SV1 structural variant in Australian Brahman, and to verify that it segregates independently from rs519457228 (r^2 = 0.04). By aligning the ONT reads against an augmented mutant genome in which the 1155-bp deletion was incorporated but not the insertion, we were able to infer that the inserted sequence is unlikely to be longer than 180 bp, although its exact sequence length remains unknown because of the high error rate related to single-molecule sequencing technologies.

Figure 6 - Oxford Nanopore Technologies long read alignments spanning *ASIP*-SV1 in Brahman cattle.



(**a**) To improve visualization the sequences were first aligned to a modified ARS-UCD1.2 assembly where the CHR13:63,599,803-63,600,957 segment was deleted. Wild type sequences are displayed as more than 1.1-kb insertions, whereas mutant sequences are displayed as insertions ranging from 162 to 173 bp (variation likely due to ONT sequencing errors). The only ONT read that mapped to the deletion position in the wild type animal was verified to be an alignment artifact (data not shown). (**b**) The long reads were further mapped to a modified ARS-UCD1.2 sequence that contained both the deletion and insertion of the reverse complement of CHR13:63,283,374-63,283,523.

2.6 Discussion

In this paper, we report the identification of positional candidate variants that affect color patterns in Nellore (*Bos indicus*) cattle. Our study revealed a single major signal on chromosome 13 in the vicinity of the *ASIP* gene that is associated with DHC particularly on the head, neck, hump, and knees of male animals. Refinement of the signal with whole-genome short-read sequence data showed that the causal variant is likely to have a regulatory effect and is located in the 5'-region of *ASIP*, rather than affecting coding regions. This fits with the expectation that regulatory effects rather than changes to the protein are the cause of the observed quantitative pattern since alterations in the protein sequence encoded by autosomal genes would likely affect the whole coat qualitatively and would less likely behave in a sex-specific manner.

Among the possible causal variants, a complex structural rearrangement (*ASIP*-SV1) consisting of a 1155-bp deletion combined with an insertion of more than 150 bp including a SINE element seemed to be the most plausible candidate due to its size and location. The ancestral sequence was associated with lighter hair-coat, whereas the derived sequence was strongly correlated with darker hair. Furthermore, we showed that the *ASIP*-SV1 mutant haplotype segregates also in Australian Brahman cattle, and we were able to refine this *ASIP*-SV1 structural arrangement using ONT data generated for 13 Brahman heifers. In our study, we also found another important candidate variant, i.e. an in-frame deletion located in the last exon of the *ASIP* gene. Although this variant was not included in the segment with strong associations for DHC, it was poorly genotyped in our study due to PCR bias during library preparation and it could represent a sequencing or alignment artifact rather than a real variant. Since we were unable to discard it as a causal variant, it was retained as a functional candidate mutation. Studies in other species have already observed frameshift deletions that lead to uniformly eumelanistic (black) phenotypes [37-40]. However, a regulatory variant such as the *ASIP*-SV1 structural variant represents a more plausible candidate than a coding variant to explain the quantitative differences between light and dark Nellore bulls, as well as the specificity of the variation to certain parts of the body as opposed to the whole body.

The agouti signaling protein (ASIP) plays a crucial role in decreasing eumelanin and increasing pheomelanin production by blocking the melanocortin 1 receptor (MC1R) [41, 42]. ASIP controls the ratio of pheomelanin (yellow/red) to eumelanin (black) pigments and can act locally as an extracellular color modifier that influences the distribution of these pigments on the body [41, 42]. The primary ligand of MC1R is the α-melanocyte-stimulating hormone (α-MSH), which promotes eumelanin synthesis. ASIP acts as a competitive inhibitor of MC1R and promotes pheomelanin synthesis [40, 41], which means that decreased expression or loss of function of the *ASIP* gene would result in a black coat phenotype.

The bovine *ASIP* gene is composed of three coding exons (2, 3 and 4) and six additional 5'-UTR exons (1A, 2A, 3A, 1B, 1C and 2C) that, through permutation, can result in six transcripts (1A, 1B, 1C, 2C and 1C2C) that use different start sites [30-32]. The location of the *ASIP*-SV1 structural variant suggests that it may affect the expression of some, but not all, of these transcripts. An example of *ASIP* structural variants that impact the expression and diversity of the transcripts is found in Chen et al. [43], who report a ~3.1-kb element duplicated in reverse orientation that is located ~15 kb upstream of *ASIP* and causes a light-bellied phenotype in mice. Large structural variations involving the entire *ASIP* gene have also been shown to cause the white coat color in Merino sheep and Saanen goats [44-46]. Structural variations that affect only the 5'-regulatory region of *ASIP* are also responsible for three characteristic mutant coat patterns in goats [46], namely Swiss markings (A^{sm}), badgerface (A^{b}) and peacock (A^{pc}), and for color dilution in quails [47].

Our findings indicate that the *ASIP*-SV1 structural variant was likely formed over 3,000 years ago through a retrotransposition event of a Bov-tA SINE, which replaced a ~1-kb region of one of the cryptic introns of *ASIP*. Such retrotranspositions that affect host gene expression and transcript diversity, associated with variation in hair phenotypes, have been reported in other animal species. For example, Demars et al. [48] found that fleece variation in sheep was explained by the insertion of a *EIF2S2* antisense RNA into the 3′ UTR of *IRF2BP2*, leading to abnormal *IRF2BP2* transcription.

A recent study explored the genetics of white coat in buffaloes using whole-genome and RNA sequencing data [49]. The authors combined GWAS,

biological experiments and population genomics to show that a ~2-kb LINE-1 insertion between exons 1C and 2 was responsible for an increased expression of *ASIP* and consequently for a white coat color. This insertion seems to act as a strong proximal promoter that increases the transcription of *ASIP* and affects melanocyte maturation. In contrast, our findings point to a ~1-kb deletion followed by a small SINE-1 insertion between the 1B and 1C non-coding exons of *ASIP* as the cause of dark hair on specific parts of the body in Nellore bulls. The two mutations are similar in the sense that they involve mobile DNA elements affecting alternative transcripts of *ASIP* that recruit non-coding exons. However, while the mutation reported here most likely increases eumelanin production, the mutation described in Liang et al. [49] seems to decrease it. In addition, the variant reported here shortens the regulatory sequence of the *ASIP* gene, whereas the variant in Liang et al. [49] lengthens it. These two mutations have reciprocal effects and provide further support that the candidate variant found here most likely affects one of the critical promoters of the *ASIP* gene.

A recent genetic divergence analysis conducted by Senczuk et al [10] compared 15 white/grey Central and Eastern European *B. taurus* breeds with four non-grey northwestern cattle populations and identified three loci that displayed substantial divergence between the two groups, namely CHR2:6,510,630-7,010,630, CHR14:22,531,305-25,722,332 and CHR26:22,789,524-23,289,524. However, in our study, none of these loci were associated with DHC in Nellore cattle, which could be explained by one of the following alternative hypotheses: (a) the trait investigated differed between the studies, i.e. DHC on specific parts of the body analyzed in our work and white/grey hair as the predominant color of the animals' body in the other study; (b) our study contained false negatives due to limitations in statistical power; (c) differences in the study design and data analysis could account for the different results, with one study performing GWAS on phenotypes measured in one breed and the other performing F_{ST} analysis without phenotypes in multiple breeds; (d) some of the loci reported by Senczuk et al. [10] are truly divergent between Central/Eastern and northwestern European breeds, but unrelated to coat color; or (e) different mutations affect the same phenotype in different breeds and sub-species.

Although we could not test the above-mentioned hypotheses directly, there is evidence in the Nellore data alone to support hypothesis (d) for the

CHR14:22,531,305-25,722,332 region. Briefly, this chromosome 14 region has been shown to overlap with a large haplotype of northwestern *B. taurus* origin spanning the *pleomorphic adenoma gene 1* (*PLAG1*), which is associated with body size [14]. This haplotype is rare in Central and Eastern cattle, but segregates at high frequency (~18%) in Nellore cattle as a result of historical *B. taurus* introgression. Therefore, if the chromosome 14 region was strongly associated with white/grey hair, in our study we should have been able to identify an association signal for DHC on chromosome 14, or at least to detect non-grey hair-coat color in carrier bulls. Since neither of these possibilities was realized here, this chromosome 14 region is unlikely to contain variants that affect white/grey hair in the studied breeds.

Holland [12] analyzed the segregation of white/grey coat color in a Nellore-Angus crossbred population. Significant associations were found on chromosome 6 within an interval containing the *KIT* gene, which was previously implicated in coat color. As for the results by Senczuk et al. [10], we did not observe associations on chromosome 6 and this is most likely related to the above hypothesis (a), given that the traits under investigation differed between the two studies. Another possible explanation is that a variant in the *KIT* gene causes white/grey coat in Nellore cattle and the *ASIP* haplotype identified here further promotes dark hair on the head, neck, hump, and knees in males. Separate epigenetic mechanisms may underlie the red-to-black transition in pigmentation observed from birth to yearling and age-related darkening in males (e.g. methylation patterns, expression of miRNA and histone acetylation), and may also explain the sexual dimorphism in coat color (e.g. sex-specific *ASIP* regulatory elements, such as promoters and enhancers).

2.7 Conclusions

We found a single statistically significant GWAS signal for darkness of hair coat in Nellore cattle, which mapped to the *ASIP* gene. A structural variant (*ASIP*-SV1) located upstream of *ASIP* was strongly associated with darker hair on the head, neck, hump and knee regions of males, which suggests that this variant is involved with decreased expression of *ASIP* and consequently a higher production of eumelanin. Although other candidate variants, including a 6-bp in-frame deletion in the last exon of *ASIP*, were not found in the region with the strongest association,

they could not be excluded as causal variants. Overall, our study provides strong evidence that functional variants within or near the *ASIP* gene account for variation in regional darkness of hair coat in Nellore and Brahman cattle.

2.8 Declarations

2.8.1 Availability of data and materials

The data used in this study were obtained under license and thus are not publicly available. However, the data sets are available for academic use from the corresponding author upon reasonable request.

2.8.2 Funding

This research was funded by Sao Paulo Research Foundation (FAPESP, process 2010/52030-2 and 2016/05787-7), National Council for Scientific and Technological Development (CNPq, process 560922/2010-8 and 483590/2010-0) and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) – Finance Code 001. The Oxford Nanopore long reads sequencing was funded by Meat and Livestock Australia (Scholarship project number: B.STU.2001, NG project number: P.PSH.0833, Ageing project number: L.GEN.1808).

2.8.3 Competing interests

By the time this study was completed, AF was employed by Personal-PEC; CP was employed by CRV-Lagoa; TaS was employed by Recombinetics Inc; AU, RT and MM were shareholders of AgroPartners Consulting; and YU and JG were members of the scientific board of AgroPartners Consulting. Mention of trade names or commercial products in this publication is solely for information and does not imply recommendation or endorsement by USDA. USDA is an equal opportunity provider and employer.

2.8.4 Authors' contributions

YU and JG conceived and designed the study; RP, ThS, LZ, RC and MC scored coat phenotypes; AF revised all phenotype scores and calculated the final average score for each animal; CP collected samples of white Nellore bulls for sequencing; MM and BT coordinated sequencing of white Nellore bulls; JG and TaS coordinated genotyping and sequencing of the remaining Nellore animals; BT, AU and YU performed and discussed the association analyses; YU performed haplotype

and coalescent analyses; BT, MM, RT and YU performed and discussed the analysis of Illumina reads; YU, BT, TL, DaB, DeB, TS and FL contributed to the dissection of the *ASIP*-SV variant from alignment data. HL, LN, ER and BH generated and performed analyses of Oxford Nanopore reads. HL, YU and TL performed simulations and comparative genomics analyses. BT and YU wrote the manuscript. All authors read and approved the final version of the manuscript.

2.8.5 Acknowledgements

We would like to thank Laiza Helena de Souza Lung for her helpful contributions in selecting white Nellore bulls for sequencing. The authors are also thankful to the editor and two anonymous reviewers for their suggestions to improve the manuscript.

2.9 References

1. ABIEC - Associação Brasileira das Indústrias Exportadoras de Carnes. Estatísticas. http://abiec.com.br/. Accessed 18 Nov 2020.

2. Ajmone-Marsan P, Garcia JF, Lenstra JA, The Globaldiv Consortium. On the origin of cattle: how aurochs became cattle and colonized the world. Evol Anthropol. 2010;19:148–57.

3. Karthickeyan SMK, Kumarasamy P, Sivaselvam SN, Saravanan R, Thangaraju P. Analysis of microsatellite markers in Ongole breed of cattle. Indian J Biotechnol. 2008;7:113-6.

4. Fadare AO, Peters SO, Yakubu A, Sonibare AO, Adeleke MA, Ozoje MO, et al. Physiological and haematological indices suggest superior heat tolerance of white-coloured West African Dwarf sheep in the hot humid tropics. Trop Anim Health Prod 2013;45:157-65.

5. Leite JHGM, Silva RG, da Silva WST, da Silva WE, Paiva RDM, Sousa JER, et al. Locally adapted Brazilian ewes with different coat colors maintain homeothermy during the year in an equatorial semiarid environment. Int J Biometeorol.

2018;62:1635–44.

6. Da Silva RG, La Scala Junior N, Tonhati H. Radiative properties of the skin and haircoat of cattle and other animals. Trans ASAE. 2003;46:913-8.

7. Fonseca VO, Souza CF, Azevedo NA, Oliveira LZ, Monteiro GA, Cavalcanti LFL, et al. Reproductive parameters of pasture-raised Nelore bulls (*Bos taurus indicus*) in different age. Arq Bras Med Vet Zootec. 2019;71:385-92.

8. Decker JE, McKay SD, Rolf MM, Kim JW, Molina Alcala A, Sonstegard TS, et al. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. PLoS Genet. 2014;10:e1004254.

9. Barbato M, Hailer F, Upadhyay M, Del Corvo M, Colli L, Negrini R, et al. Adaptive introgression from indicine cattle into white cattle breeds from Central Italy. Sci Rep. 2020;101:1279.

10. Senczuk G, Guerra L, Mastrangelo S, Campobasso C, Zoubeyda K, Imane M, et al. Fifteen shades of grey: Combined analysis of genome-wide SNP data in steppe and Mediterranean grey cattle sheds new light on the molecular basis of coat color. Genes. 2020;11:932.

11. Rosen BD, Bickhart DM, Schnabel RD, Koren S, Elsik CG, Tseng E, et al. De novo assembly of the cattle reference genome with single-molecule sequencing. GigaScience. 2020;9:giaa021.

12. Holland PW. Identification of the regions of the bovine genome associated with grey coat color in a Nellore-Angus cross population. Master Thesis, Texas A&M University. 2015.

13. Carvalheiro R, Boison SA, Neves HHR, Sargolzaei M, Schenkel FS, Utsunomiya YT, et al. Accuracy of genotype imputation in Nelore cattle. Genet Sel Evol. 2014;46:69.

14. Utsunomiya YT, Milanesi M, Utsunomiya ATH, Torrecilha RBP, Kim ES, Costa MS, et al. A *PLAG1* mutation contributed to stature recovery in modern cattle. Sci Rep. 2017;7:17140.

15. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007;81:559–75.

16. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience. 2015;4:7.

17. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet. 2011;88:76–82.

18. Yang J, Zaitlen NA, Goddard ME, Visscher PM, Price AL. Advantages and pitfalls in the application of mixed-model association methods. Nat Genet. 2014;46:100–6.

19. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2020. https://www.R-project.org/

20. Falconer DS, Mackay TFC. Introduction to quantitative genetics. 4th ed. Harlow: Longman Group Limited; 1996. p.464.

21. Vitezica ZG, Varona L, Legarra A. On the additive and dominant variance and covariance of individuals within the genomic selection scope. Genetics. 2013;195:1223-30.

22. Loh P-R, Danecek P, Palamara PF, Fuchsberger C, Reshef YA, Finucane HK et al. Reference-based phasing using the Haplotype Reference Consortium panel. Nat Genet. 2016;48:1443–8.

23. Utsunomiya YT, Milanesi M, Utsunomiya ATH, Ajmone-Marsan P, Garcia JF. GHap: an R package for genome-wide haplotyping. Bioinformatics. 2016;32:2861-2.

24. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;251:1754–60.

25. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.

26. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, et al. The Ensembl variant effect predictor. Genome Biol. 2016;17:122.

27. Wang X, Zheng Z, Yudong C, Chen T, Li C, Fu W, et al. CNVCaller: highly efficient and widely applicable software for detecting copy number variations in large populations. GigaScience. 2017;6:1-12.

28. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14:178–92.

29. Heng L. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics. 2018;34:3094–100.

30. Girardot M, Martin J, Guibert S, Leveziel H, Julien R, Oulmouden A. Widespread expression of the bovine *Agouti* gene results from at least three alternative promoters. Pigment Cell Res. 2005;18:34–41.

31. Girardot M, Guibert S, Laforet MP, Gallard Y, Larroque H, Oulmouden A. The insertion of a full-length *Bos taurus* LINE element is responsible for a transcriptional deregulation of the Normande Agouti gene. Pigment Cell Res. 2006;19:346-355.

32. Albrecht E, Komolka K, Kuzinski J, Maak S. Agouti revisited: transcript quantification of the ASIP gene in bovine tissues related to protein expression and

localization. PLoS One. 2012;7:e35282.

33. Liu Y, Fang X, Zhao Z, Li J, Albrecht E, Schering L, et al. Polymorphisms of the ASIP gene and the haplotype are associated with fat deposition traits and fatty acid composition in Chinese Simmental steers. Arch Anim Breed. 2019;62:135-42.

34. Albers PK, McVean G. Dating genomic variants and shared ancestry in population-scale sequencing data. PLoS Biol. 2020; 18:e3000586.

35. Zavarez LB, Utsunomiya YT, Carmo AS, Neves HHR, Carvalheiro R, Ferencakovic M, et al. Assessment of autozygosity in Nellore cows (*Bos indicus*) through high-density SNP genotypes. Front Genet. 2015;6:5.

36. Lan JH, Yin Y, Reed EF, Moua K, Thomas K, Zhang Q. Impact of three Illumina library construction methods on GC bias and HLA genotype calling. Hum Immunol. 2015;76:166-75.

37. Rieder S, Taourit S, Mariat D, Langlois B, Guérin G. Mutations in the agouti (*ASIP*), the extension (*MC1R*), and the brown (*TYRP1*) loci and their association to coat color phenotypes in horses (*Equus caballus*). Mamm Genome. 2001;12:450-5.

38. Hiragaki T, Inoue-Murayama M, Miwa M, Fujiwara A, Mizutani M, Minvielle F, et al. Recessive black is allelic to the yellow plumage locus in Japanese Quail and associated with a frameshift deletion in the *ASIP* gene. Genetics. 2008;178:771-5.

39. Lai W. Hu M, Zhu W, Yu F, Bai C, Zhang J, et al. A 4-bp deletion in the *ASIP* gene is associated with the recessive black coat colour in domestic guinea pigs (*Cavia porcellus*). Anim Genet. 2019;50:190-1.

40. Letko A, Ammann B, Jagannathan V, Henkel J, Leuthard F, Schelling C, et al. A deletion spanning the promoter and first exon of the hair cycle-specific *ASIP* transcript isoform in black and tan rabbits. Anim Genet. 2019;51:137-40.

41. Barsh G, Gunn T, He L, Schlossman S, Duke-Cohan J. Biochemical and genetic studies of pigment-type switching. Pigment Cell Res. 2000;13:48–53.

42. Cieslak M, Reissmann M, Hofreiter M, Ludwig A. Colours of domestication. Biol Rev Camb Philos Soc. 2011;86:885-99.

43. Chen Y, Duhl DMJ, Barsh G. Opposite orientations of an inverted duplication and allelic variation at the mouse *agouti* locus. Genetics. 1996;144:265-77.

44. Norris BJ, Whan VA. A gene duplication affecting expression of the ovine ASIP gene is responsible for white and black sheep. Genome Res. 2008;18:1282-93.

45. Fontanesi L, Beretti F, Riggio V, Gómez González E, Dall'Olio S, Davoli R, et al. Copy number variation and missense mutations of the agouti signaling protein (*ASIP*) gene in goat breeds with different coat colors. Cytogenet Genome Res. 2009;126:333-47.

46. Henkel J, Saif R, Jagannathan V, Schmocker C, Zeindler F, Bangerter E, et al. Selection signatures in goats reveal copy number variants underlying breed-defining coat color phenotypes. PLoS Genetics. 2019;15:e1008536.

47. Robic A, Morisson M, Leroux S, Gourichon D, Vignal A, Thebault N, et al. Two new structural mutations in the 5′ region of the *ASIP* gene cause diluted feather color phenotypes in Japanese quail. Genet Sel Evol. 2019;51:12.

48. Demars J, Cano M, Drouilhet L, Plisson-Petit F, Bardou P, Fabre S, et al. Genome-wide identification of the mutation underlying fleece variation and discriminating ancestral hairy species from modern woolly sheep. Mol Biol Evol. 2017;34:1722-9.

49. Liang D, Zhao P, Si J, Fang L, Pairo-Castineira E, Hu X, et al. Genomic analysis revealed a convergent evolution of LINE-1 in coat color: A case study in water buffaloes (*Bubalus bubalis*). Mol Biol Evol. 2021;38:1122-36.

3 CAPÍTULO 2 – THE *ASIP*-SV1 VARIANT ASSOCIATED WITH HAIR COAT DARKNESS SEGREGATES IN MULTIPLE *Bos indicus* CATTLE BREEDS

Beatriz Batista Trigo[1,2], Natalia Fonseca Alves[1,2], Marco Milanesi[3], José Fernando Garcia[1,2,3,4], Abdulfatai Tijjani [5], Yuri Tani Utsunomiya[1,2,3]*

[1]Department of Production and Animal Health, School of Veterinary Medicine of Araçatuba, São Paulo State University (Unesp), 16050-680. R. Clóvis Pestana 793 - Dona Amelia, Araçatuba/SP Brazil;
[2]International Atomic Energy Agency (IAEA) Collaborating Centre on Animal Genomics and Bioinformatics, 16050-680. R. Clóvis Pestana 793 - Dona Amelia, Araçatuba/SP, Brazil;
[3]AgroPartners Consulting. R. Floriano Peixoto, 120 - Sala 43A - Centro, Araçatuba/SP, 16010-220
[4]Department of Preventive Veterinary Medicine and Animal Reproduction, School of Agricultural and Veterinarian Sciences, São Paulo State University (Unesp), 14884-900 Via de Acesso Prof. Paulo Donato Castellane s/n - Jaboticabal/SP, Brazil.
[5]International Livestock Research Institute (ILRI), Nairobi, Kenya
*Corresponding author: ytutsunomiya@gmail.com

3.1 Resumo

As raças zebuínas (*Bos indicus*) estão bem adaptadas e permanecem bastante produtivas em ambientes que apresentam altas cargas de doenças, alimentação pouco nutritiva e altas temperaturas. Os padrões na coloração da pele e da pelagem têm sido associados à tolerância ao calor e hipotetizados por contribuírem para a adaptação dos zebuínos nas regiões tropicais e sub-tropicais. Uma variação estrutural no gene proteína de sinalização aguti (*ASIP*), nomeado *ASIP*-SV1, foi encontrada e fortemente correlacionada com o escurecimento da pelagem em regiões específicas do corpo dos gados Nelore e Brahman, sendo então predita por ser de origem *B. indicus.* Analisamos de maneira visual os sequenciamentos de genoma completo de zebuínos e não-zebuínos para elucidar a segregação da variante *ASIP*-SV1 em gados. Das sequências de 216 animais analisados, 62 zebuínos e 5 não-zebuínos tinham a *ASIP*-SV1 em seu genoma. As raças de *B.*

taurus portadoras da *ASIP*-SV1 em nossa análise eram previamente conhecidas por carrearem introgressão zebuína e também serem bem adaptadas em regiões tropicais e sub-tropicais.

Palavras-chave: Coloração de pelagem. Zebuíno. Taurino. Agouti. Introgressão. Adaptação climática.

3.2 Abstract

Zebu breeds (*Bos indicus*) are well adapted to and remain fairly productive in environments presenting a high disease burden, lowly nutritious food and high temperatures. Skin and coat color patterns have been associated with heat tolerance and hypothesized to contribute to the adaptation of zebu animals to the tropical and subtropical regions of the world. A structural variation (SV) in the agouti signaling protein gene (*ASIP*), namely *ASIP*-SV1, has been found to strongly correlate with the darkness of coat in specific regions of the body of Nellore and Brahman bulls, and thus predicted to be of *B. indicus* origin. Here we visually analyzed WGS of indicine and non-indicine cattle to elucidate *ASIP*-SV1 segregation in cattle. Of 216 animals' sequences analyzed, 62 zebu and 5 non-zebu animals had *ASIP*-SV1 in their genome. *B. taurus* breeds carrying the SV in our analysis were previously known to carry indicine introgression and to be also well adapted to tropical and subtropical regions.

Keywords: Coat color. Zebu. Taurine. Agouti. Introgression. Climate adaptation.

3.3 Main text [2]

A complex structural rearrangement was found in Nellore cattle and Australian Brahman cattle in the *ASIP* locus on Chromosome 13 and was named *ASIP*-SV1 (Trigo *et al.* 2021). This structural variation (SV) is a 1,155-bp deletion combined with more than 150-bp of a SINE element insertion. The derived sequence obtained by *ASIP*-SV1 was strongly correlated with darker hair coat, while the ancestral sequence was associated with lighter hair. Furthermore, coalescent analysis indicated that the *ASIP*-SV1 structural variant was formed over 3,400 years ago in a retrotransposition event with a Bov-tA SINE replacing one of the cryptic introns of *ASIP* in ~1kb region. The hypothesis around this variant is that *ASIP*-SV1

[2] Artigo de acordo com as normas da revista *Animal Genetics* (ANEXO B)

is an ancient derived allele currently segregating in *B. indicus* breeds and in *B. taurus* populations that carry *B. indicus* introgression (Trigo *et al.* 2021). Utsunomiya *et al.* (2022), searching for *B. taurus* haplotypes that segregate in Nellore cattle, confirmed that *ASIP*-SV1 is of *B. indicus* origin, reinforcing the hypothesis that other *B. indicus* breeds might also present the variant. Here, we aimed at elucidating the segregation of *ASIP*-SV1 in *B. indicus* and *B. taurus* breeds.

In this study, we used whole-genome sequence (WGS) data of 216 animals, from which 137 were *B. indicus* cattle from Brazil, Africa and Pakistan, 77 were *B. taurus* and 2 were *B. grunniens* animals. Zebu sequences from Brazil was obtained from previous study (Trigo *et al.* 2021), other zebu sequences were obtained by a partnership (not available publicly) and non-zebu sequences were obtained from the public sequence database of NCBI - Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra/). The *ASIP*-SV1 region (CHR13:63,599,803- 63,600,957) was extracted from zebu sequences using SAMtools v1.8 (Li *et al.* 2009) and the *ASIP*-SV1 region from the other animals was extracted in NCBI as BAM output. Sequences aligned to the UMDv3.1 assembly had their coordinates lifted-over to the ARS-UCDv1.2 assembly, and the *ASIP*-SV1 region sequence was visually analyzed using the Integrative Genomics Viewer (IGV) software (Thorvaldsdóttir *et al.* 2013) to look for the structural variant. Coat color phenotypes of animals used in this study were unknown.

Using IGV software we visually analyzed *ASIP*-SV1 looking for an abrupt decrease in genome coverage indicating the deletion of ~1.1kb of bp, and in addition atypical reads and soft-clipped bases in the deletion region indicating the insertion. It is also possible to infer if the animal is heterozygous or homozygous for the ASIP-SV1 looking if there are both variant and regular reads in sequence analyzed.

Of 137 zebu sequences analyzed, it was possible to identify the *ASIP*-SV1 variant in 62 animals (45.26%). From 79 non-zebu sequences analyzed, only 5 (6.3%) presented the *ASIP*-SV1 variant. Considering both zebu and non-zebu sequences, 31% of the animals presented the ASIP-SV1 variant. Of all 68 animals with *ASIP*-SV1 deletion variation 98.52% have the *ASIP*-SV1 insertion variation observed in IGV software (Table 1). When we analyzed the number of SV copies, 52 animals (78%) seemed to be heterozygous, while 15 (22%) were homozygous for *ASIP*-SV1 (Table S1).

Table 1 - Species, breeds and the number of analyzed animals carrying *ASIP*-SV1 deletion and insertion

Species	Breed	# of animals analyzed	# of animals with Deletion	# of animals with Insertion
Bos indicus	Afar	14	5	5
	Begait	9	2	2
	Ethiopia-Boran	10	4	4
	Kenya-Boran	16	6	6
	Butana	10	5	5
	Fogera	12	3	3
	Kenana	10	2	2
	Nellore	27	19	19
	Ogaden	9	4	4
	Achai	2	2	2
	Bhagnari	3	0	0
	Cholistani	2	2	2
	Dajal	1	0	0
	Dhanni	2	1	1
	Gabrali	2	1	1
	HisarH	2	1	1
	Lohani	1	1	1
	Red Sindhi	1	1	1
	Sahiwal	2	2	2
	Tharparkar	2	2	2
Bos grunniens	Yak	2	0	0
Bos taurus	Angus	10	0	0
	Belgian Blue	5	0	0
	Charolais	10	0	0
	Hereford	10	0	0
	Holstein	10	0	0
	Jersey	7	0	0
	Limousin	10	0	0
	Romagnola	5	4	3
	Simmental	10	1	1
Total	216	68	67	

Coordinates of *ASIP*-SV1 from all animals were similar and consistent with our previous study (Trigo *et al.* 2021), even with the adjusted coordinates between old and new assemblies. Furthermore, the deletion and the insertion appeared together in our analysis, reinforcing that the SINE insertion and the deletion occurred in a single mutation event.

The indicine breeds analyzed (except for Bhagnari and Dajal) had at least one animal presenting *ASIP*-SV1, while from the ten (10) non-zebu breeds analyzed only two (2) showed the *ASIP*-SV1 in genome sequence.

These Bhagnari and Dajal results can be explained by the fact that we had a limited number of genome sequences available in this analysis (3 and 1 respectively) and coat color phenotype of them are unknown to expect *ASIP*-SV1. It is important to highlight that the quality of sequence coverage can also impact in IGV visualization of ASIP-SV1 deletion and insertion.

B. taurus breeds presenting *ASIP*-SV1 in IGV visualization were Romagnola (4 animals from 5 analyzed) and Simmental (1 animal from 10 analyzed). Romagnola is known to have been introgressed with *B. indicus* haplotypes in the past (McTavish *et al.* 2013; Barbato *et al.* 2020). Barbato *et al.* (2020) described indicine introgression to Romagnola, and chromosome 13 had one of the highest levels of zebu component. Another study also described SVs shared between zebu and Italian cattle, including the Romagnola breed (Upadhyay *et al*. 2021). The *ASIP* locus was not emphasized in their study but our analysis showed that *B. indicus* *ASIP*-SV1 region was introgressed in the Romagnola genome. Of note, the coat color pattern in this breed showed darkness of hair coat around the head, neck, shoulders and knees, similar to the coat color pattern seen in Nellore and Brahman cattle.

In previous research Trigo *et al.* (2021) used coalescent analysis and estimated that *ASIP*-SV1 was derived over 3,000 years ago (YA). In addition, studies suggested that some migration routes could cause indicine introgression in some European taurine breeds ~3,000 YA (Barbato et al. 2020; Pellecchia et al. 2007), the same period estimated for *ASIP*-SV1 mutation. This information can be useful to suggest that *ASIP*-SV1, maybe together with other genes, should play some role on cattle climate adaptation. In addition, as cited before the *ASIP*-SV1 causing a coat color pattern that in a black skin was described to contribute in heat tolerance and adaptation to tropical and subtropical regions (Da Silva *et al.* 2003).

Most animals analyzed here had only one copy of *ASIP*-SV1, and heterozygous cattle were associated with the middle-range darkness of coat shadow (Trigo *et al.* 2021). Some animals have their intensity of darkness changed depending on seasons. This change in the darkness intensity is not clearly solved but could be explained by one of the following hypotheses: (a) occasioned by the variation in sunlight intensity and duration across seasons, where the darkness intensity needs to be adapted to reflect or absorb light wavelengths; (b) controlled by hormones related to male cattle: (b.1) in mating season making this animal more attractive to females or (b.2) related to fertility.

The sexual dimorphism and the range of darkness intensity should be tested in the future to understand how it works, but our findings provide strong evidence that *ASIP*-SV1 is an indicine SV that is prevalent in tropical cattle. We also found the *ASIP*-SV1 in a breed with *B. indicus* introgression suggesting that this variant was also introgressed and selected for in some *B. taurus* breeds.

3.4 Acknowledgements

The researchers BBT and NFA were funded by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior—Brasil (CAPES)—Finance Code 001 and Sao Paulo Research Foundation (FAPESP, process 2020/07051-3) respectively.

3.5 Availability of data

The *B. indicus* data used in this study were obtained under license and thus are not publicly available, non-zebu data are available at NCBI - Sequence Read Archive (SRA) (https://www.ncbi.nlm.nih.gov/sra/).

3.6 References

ABCZ – Associação Brasileira de Criadores de Zebuínos. Raças Zebuínas (2021). Available at: https://www.abcz.org.br/a-abcz/racas-zebuinas [accessed October 17, 2021]

ABIEC – Associação Brasileira das Indústrias Exportadoras de Carnes (2021). Beef report 2021. http://abiec.com.br/publicacoes/beef-report-2021/ [accessed October 11, 2021]

Barbato M., Hailer F., Upadhyay M., Del Corvo M., Colli L., Negrini R. *et al.* (2020) Adaptive introgression from indicine cattle into white cattle breeds from Central Italy. *Scientific Report* **10**, 1279.

Da Silva R.G., La Scala J.N. & Tonhati H. (2003) Radiative properties of the skin and haircoat of cattle and other animals. *Transactions of the* ASAE **46**, 913–8.

Fadare A.O., Peters S.O., Yakubu A., Sonibare A.O., Adeleke M.A., Ozoje M.O. *et al.* (2013) Physiological and haematological indices suggest superior heat tolerance of white-coloured West African Dwarf sheep in the hot humid tropics. Tropical Animal Health and Production **45**, 157–65.

Leite J.H.G.M., Silva R.G., da Silva W.S.T., da Silva W.E., Paiva R.D.M., Sousa J.E.R. *et al.* (2018) Locally adapted Brazilian ewes with different coat colors maintain homeothermy during the year in an equatorial semiarid environment. International *Society of* Biometeorology **62**, 1635–44.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N. *et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–9.

McTavish E.J., Decker J.E., Schnabel R.D., Taylor J.F. & Hillis D.M. (2013) New World cattle show ancestry from multiple independent domestication events. *Proceedings of the National Academy of Sciences of the United States of America* **110**(15), E1398-E1406.

Pellecchia M., Negrini R., Colli L., Patrini M., Milanesi E., Achilli A., *et al*. (2007) The mystery of Etruscan origins: novel clues from Bos taurus mitochondrial DNA. *Proceedings. Biological sciences* **274**(1614), 1175-1179.

Thorvaldsdóttir H., Robinson J.T. & Mesirov J.P. (2013) Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Briefings in Bioinformatics* **14**, 178–92.

Trigo B.B., Utsunomiya A.T.H., Fortunato A.A.A.D., Milanesi M., Torrecilha R.B., Lamb H. *et al.* (2021) Variants at the ASIP locus contribute to coat color darkening in Nellore cattle. *Genetics Selection Evolution* **53**, 40.

Upadhyay M., Derks M.F.L., Andersson G., Medugorac I., Groenen M.A.M., Crooijmans R.P.M.A. (2021) Introgression contributes to distribution of structural variations in cattle. *Genomics* **113**(5), 3092-3102.

Utsunomiya Y.T., Fortunato A.A.A.D., Milanesi M., Trigo B.B., Alves N.F., Sonstegard T.S. *et al.* (2022) *Bos taurus* haplotypes segregating in Nellore (*Bos indicus*) cattle. *Animal Genetics* **53**, e13164.

APÊNDICE A– REFERÊNCIAS DA INTRODUÇÃO GERAL

ABCZ. Associação Brasileira de Criadores de Zebuínos. Raças Zebuínas. Disponível em:<https://www.abcz.org.br/a-abcz/racas-zebuinas>. Acesso em: 12 jan. 2022.

ABIEC. Associação Brasileira das Indústrias Exportadoras de Carnes. Estatísticas. Disponível em:<http://abiec.com.br/>. Acesso em: 12 jan. 2022.

AJMONE-MARSAN, P.; GARCIA, J. F.; LENSTRA, J. A. On the origin of cattle: how aurochs became cattle and colonized the world. **Evol. Anthropol.,** v. 19, p. 148-157, 2010.

ANDRADE, A. B. F.; SILVA, R. G.; COSTA, A. J.; ROCHA, U. F.; LANDIM, V. J. C. Genetic and environmental aspects of the resistance of zebu cattle to the tick *Boophilus microplus*. In: 6th World Congress on Genetics Applied to Livestock Production, Armildade, 1998. n.27, p.339-342.

ARORA, N.; SIDDIQUI, E. M.; MEHAN, S. Involvement of adenylate cyclase/cAMP/CREB and SOX9/MITF in melanogenesis to prevent vitiligo. **Mol. Cell. Biochem.,** v. 476, p. 1401–1409, 2021.

BARBATO, M. et al. Adaptive introgression from indicine cattle into white cattle breeds from Central Italy. **Sci. Rep.,** v. 10, n. 1, p. 1279, 2020.

BARSH, G.; GUNN, T.; HE, L.; SCHLOSSMAN, S.; DUKE-COHAN, J. Biochemical and genetic studies of pigment-type switching. **Pigment. Cell. Res.,** v. 13, p. 48–53, 2000.

CARDOSO, C. P.; SILVA, B. F.; GONÇALVES, D. S.; TAGLIARI, N. J.; SAITO, M. E.; AMARANTE, A. F. T. Resistência contra ectoparasitas em bovinos da raça Crioula Lageana e meio-sangue Angus avaliada em condições naturais. **Pesqui. Vet. Bras.,** v. 34, n. 2, p. 141-146, 2014.

CAVALCANTI, L. C. G. **Avaliação genética de padres de pelagem em ovinos da raça crioula.** 2015, 64 p. Dissertação (Mestrado em Ciências Animais –Faculdade de Agronomía e Veterinária, Universidade de Brasília, Brasília, 2015.

CHEN, W.; YANG, C. C.; LIAO, C. Y.; HUNG, C. L.; TSAI, S. J.; CHEN, K. F.; SHEU, H. M.; ZOUBOULIS, C. C: Expression of sex-determining genes in human sebaceous glands and their possible roles in pathogenesis of acne. **J. Eur. Acad. Dermatol. Venereol.,** v. 20, p. 846-852, 2006.

CIESLAK, M.; REISSMANN, M.; HOFREITER, M.; LUDWIG, A. Colours of domestication. **Biol. Rev. Camb. Philos. Soc.,** v. 86, n. 4, p. 885-899, 2011.

DA SILVA, R. G.; LA SCALA JUNIOR, N.; TONHATI, H. Radiative properties of the skin and haircoat of cattle and other animals. **Trans. ASAE.,** v. 46, p. 913-918, 2003.

DECKER, J. E. et al. Worldwide patterns of ancestry, divergence, and admixture in domesticated cattle. **PLoS Genet.,** v. 10, p. e1004254, Mar, 2014.

D'ORAZIO, J.; JARRETT, S.; AMARO-ORTIZ, A.; SCOTT, T. UV Radiation and the Skin. **Int. J. Mol. Sci.,** v. 14, n. 6, p. 12222-12248, 2013.

ERDOĞAN, B. Anatomy and Physiology of Hair. In: KUTLUBAY, Z.; SERDAROGLU, S. **Hair and Scalp Disorders**. Londres: IntechOpen, 2017. DOI: 10.5772/67269. Disponível em:<https://www.intechopen.com/chapters/53880>. Acesso em: 15 jan. 2022.

FADARE, A. O. et al. Physiological and haematological indices suggest superior heat tolerance of white-coloured West African Dwarf sheep in the hot humid tropics. **Trop. Anim. Health. Prod.,** v. 45, p. 157-165, 2013.

FONSECA, V. O. et al. Reproductive parameters of pasture-raised Nelore bulls (*Bos taurus indicus*) in different age. **Arq. Bras. Med. Vet. Zootec.,** v.;71, p. 385-392, 2019.

KARTHICKEYAN, S. M. K.; KUMARASAMY, P.; SIVASELVAM, S. N.; SARAVANAN, R.; THANGARAJU, P. Analysis of microsatellite markers in Ongole breed of cattle. **Indian J. Biotechnol.,** v. 7, p. 113-116, 2008.

LEITE, J. H. G. M. et al. Locally adapted Brazilian ewes with different coat colors maintain homeothermy during the year in an equatorial semiarid environment. **Int. J. Biometeorol.,** v. 62, p. 1635–1644, 2018.

LETKO, A. et al. A deletion spanning the promoter and first exon of the hair cycle-specific *ASIP* transcript isoform in black and tan rabbits. **Anim. Genet.,** v. 51, p. 137-40, 2019.

LIM, J.; NG, K. J.; CLAVEL, C. Chapter four – Dermal papilla regulation of hair growth and pigmentation. In: PEREZ-MORENO, M. **Advances in stem cells and their niches**. Volume 3. Elsevier, 2019. p. 115-138.

NGUYEN, N. T.; FISHER, D. E. MITF and UV responses in skin: From pigmentation to addiction. **Pigment. Cell. Melanoma. Res.,** v. 32, n. 2, p. 224–236, 2019.

PASSERON, T.; VALENCIA, J. C.; BERTOLOTTO, C.; HOASHI, T.; LE PAPE, E.; TAKAHASHI, K.; BALLOTTI, R.; HEARING, V. J. SOX9 is a key player in ultraviolet B-induced melanocyte differentiation and pigmentation. **Proc. Natl. Acad. Sci. USA.,** v. 104, n. 35, p. 13984-9, 2007.

ROCHUS, C. M.; WESTBERG SUNESSON, K.; JONAS, E.; MIKKO, S.; JOHANSSON, A. M. Mutations in ASIP and MC1R: dominant black and recessive black alleles segregate in native Swedish sheep populations. **Anim. Genet.,** v.50, n. 6, p. 712-717, 2019.

SCIENSKI, K. **Characterizing a region of bovine chromosome 6 associated with three cattle coat color phenotypes.** 2018. 78 p. Master's thesis (Master of Science) - Texas A&M University, Texas, 2018.

SHARMA, S.; WAGH, S.; GOVINDARAJAN, R. Melanosomal proteins--role in melanin polymerization. **Pigment. Cell. Res.,** v. 15, n. 2, p. 127-133, 2002.

STENN, K. S.; PAUS, R. Controls of hair follicle cycling. **Physiol. Rev.,** v. 81, v. 1, p. 449-494, 2001.

SUZUKI, H. Evolutionary and phylogeographic views on Mc1r and Asip variation in mammals. ***Genes. Genet. Syst*.,** v. 88, n. 3, p. 155-164, 2013.

Trigo, B. B. et al. Variants at the ASIP locus contribute to coat color darkening in Nellore cattle. **Genet. Sel. Evol.,** v. 53, n. 1, p. 40, 2021.

ZOUBOULIS, C. C.; CHEN, W. C.; THORNTON, M. J.; QIN, K.; ROSENFIELD, R. Sexual hormones in human skin. **Horm. Metab. Res.,** v. 39, n. 2, p. 85-95, 2007.

APÊNDICE B – MATERIAL SUPLEMENTAR REFERENTE AO CAPÍTULO 1

Figure S1 - Histogram and summary statistics of average visual scores for darkness of hair coat in 432 Nellore bulls

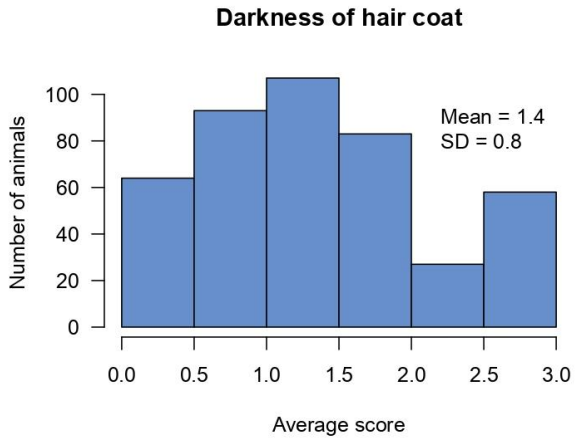


Figure S2 - Genome-wide association analysis for DHC in Nellore cattle including the BovineHD1300018322 SNP as a fixed effect.

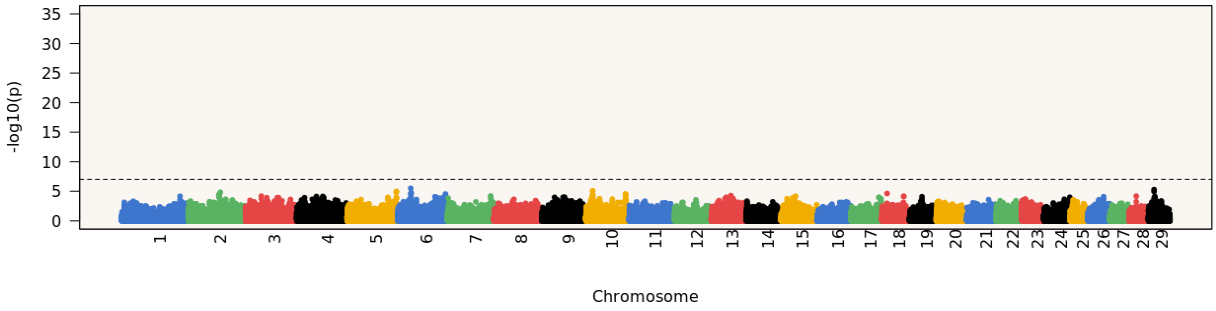


Table S1 - List of positional candidate variants spanning the *ASIP* locus.

Format: XLSX (112 kb). According to archive size and numbers of lines and columns it is only available at: https://gsejournal.biomedcentral.com/articles/10.1186/s12711-021-00633-2#Sec23.

Table S2 - Inferred genotypes for *ASIP*-SV1 and rs519457228 in 13 Australian Brahman heifers sequenced with Oxford Nanopore Technologies.

Sample	Likely ASIP-CNVR1 genotype[a]	Number of Reads Spanning ASIP-CNVR1	Likely rs519457228 genotype[b]	Number of Reads Spanning rs519457228
BR001	**heterozygous**	9	**heterozygous**	4
BR002	homozygous ref	12	homozygous ref	10
BR003	**homozygous alt**	4	homozygous ref	5
BR004	homozygous ref	5	homozygous ref	6
BR005	homozygous ref	2	homozygous ref	2
BR006	homozygous ref	5	homozygous ref	5
BR007	homozygous ref	4	homozygous ref	1
BR008	homozygous ref	4	homozygous ref	1
BR009	**homozygous alt**	8	homozygous ref	4
BR010	**heterozygous**	3	homozygous ref	1
BR011	homozygous ref	1	homozygous ref	4
BR012	unclear	4	**heterozygous**	5
BR013	homozygous ref	4	homozygous ref	3

[a]Alternative allele is the deletion of the 13:63,599,803-63,600,957 sequence

[b]Alternative allele is the deletion of the 13:63,667,797-63,667,802 sequence

Carriers are indicated in bold

Supplementary methods - Comparative genomics analysis to infer ancestral and derived *ASIP* alleles.

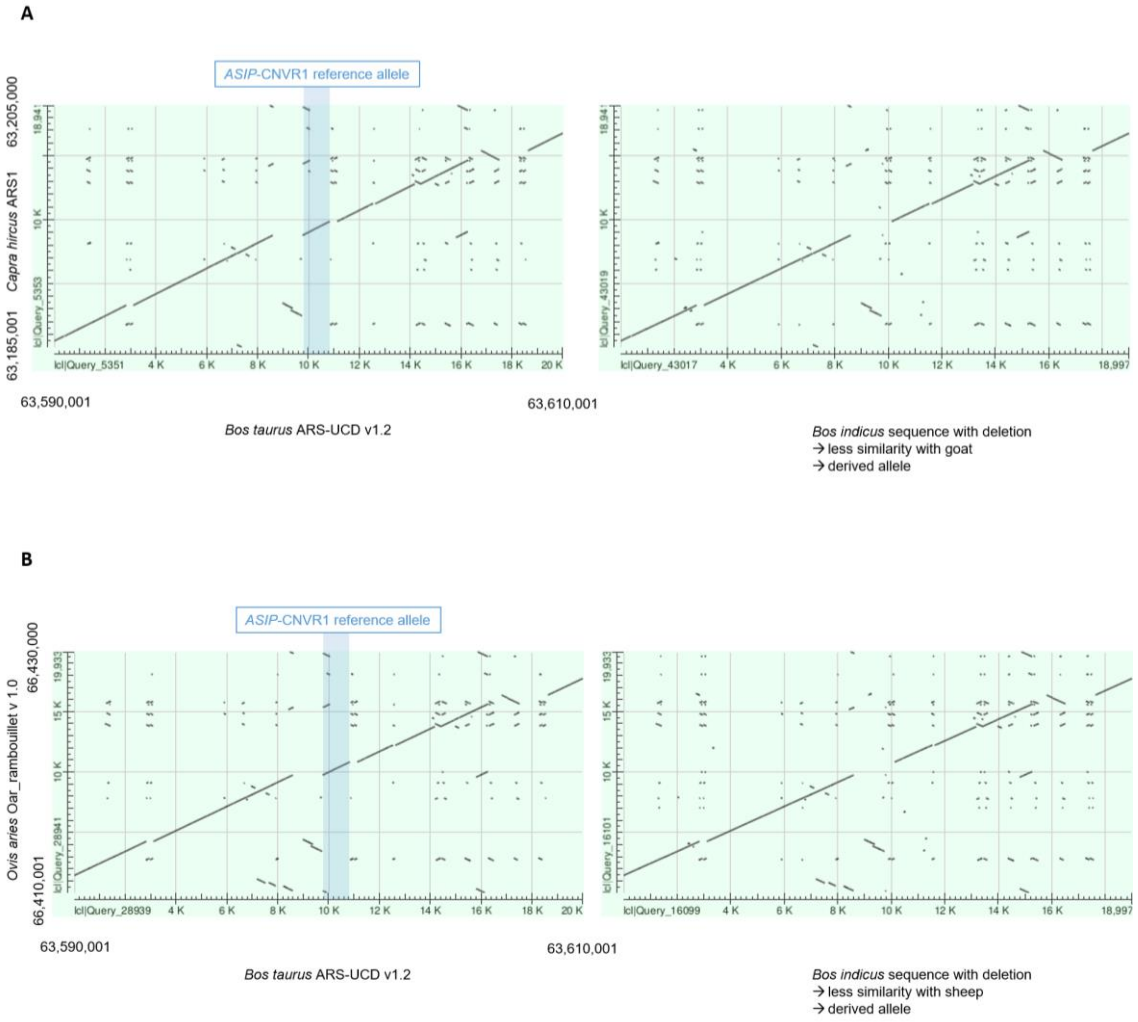
In order to determine the ancestral and derived alleles at the *ASIP*-SV1 locus, we performed three separate analyses. In the first one, the bovine reference sequence (*Bos taurus* ARS-UCDv1.2) was aligned against the yak (*Bos grunniens* BosGru_v2.0), sheep (*Ovis aries* Oar_rambouillet_v1.0) and goat (*Capra hircus* ARS1) orthologous sequences using Clustal Omega v1.2.4 (available at https://www.ebi.ac.uk/Tools/msa/clustalo/). In the second one, FASTA files containing the whole bovine chromosome 13 sequence were aligned against the sheep and goat reference genome assemblies using blast2seq (available at https://blast.ncbi.nlm.nih.gov/Blast.cgi). One of the files had the bovine reference sequence unchanged, whereas the other one had the deletion at the *ASIP*-SV1 locus. Finally, simulated and real Oxford Nanopore Technologies (ONT) long reads containing the reference and alternative bovine *ASIP*-SV1 alleles were aligned against the yak genome using the alignment pipeline described in the article.

Results

The Clustal Omega alignments are displayed at the end of this file. They show that the bovine reference sequence at the *ASIP*-SV1 is also found in yak, sheep and goat (excluding small breakpoints of few bp), indicating that the 1,155 bp deletion associated with darker hair in Nellore cattle is a derived allele, whereas the reference bovine sequence is the ancestral state. **Figure SM1** shows the alignment of the whole bovine chromosome 13 – both in the reference and alternative forms – against the sheep and goat reference genomes, which further establishes the bovine reference as wild type and the deletion as mutant. Lastly, real and simulated ONT

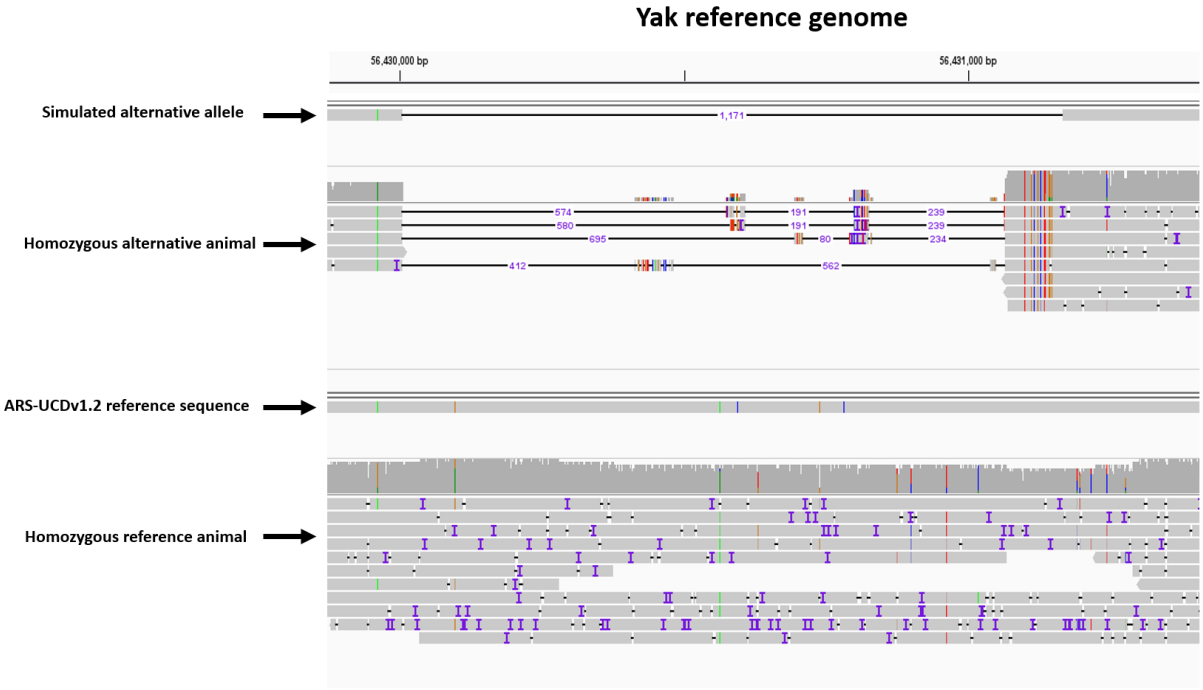
long reads were mapped against the yak genome (**Figure SM2**) to reinforce that the deletion associated with darker hair is the derived allele.

Figure SM1 - Dot plots of bovine chromosome 13 alignments against the sheep



(**A**) and goat (**B**) reference genomes. Both the reference (left side plots) and alternative (right side plots) bovine alleles were used. The reference allele is marked in blue. The plots show that the reference bovine allele is also found in sheep and goat, indicating that the deletion is the mutant sequence, and the reference allele is the wild type sequence.

Figure SM2 - Oxford Nanopore Technologies long reads mapped against the yak reference genome.



Reads containing the bovine reference allele map through the yak sequence, whereas reads carrying the alternative allele display a ~1 kbp deletion at the *ASIP*-SV1 location. This indicates that the deletion is a derived sequence, whereas the bovine reference allele is ancestral.

```
CLUSTAL O(1.2.4) multiple sequence alignment

#OBS = The last ~200 bp were clipped due to breakpoints in sheep and goat

bos_taurus/11353                ----TTGGAAGGACTGATGCTGAAGCTTCAATACTTTGGCCACCTGATGCAAACAGCTGA  56
bos_grunniens/11353             ----TTGGAAAGACTGATGCTGAAGCTTCAATACTTTGGCCACCTGATGCAAACAGCTGA  56
capra_hircus/11353              GGTAGGACTGATGCTGAAGCTGAAGCTTCAATACTTTGGCTACCTGATGCAAACAGCTGA  60
ovis_aries_rambouillet/11353    GGTAGGACTGATGCTGAAGCTGAAGCTTCAATACTTTGGCTACCTGATGCAAACAGCTGA  60
                                     **** ********************* ******************

bos_taurus/11353                CACATTGGAAAAGGCCCTGATGGTGGGAAAGACTGAAGGCAGGAGGAGAAGGGGATGACA  116
bos_grunniens/11353             CACATTGGAAAAGGCCCTGATGGTGGGAAAGACTGAAAGCAGGAGGAGAAGGGGATGACA  116
capra_hircus/11353              CACATTGGAAAAGACCCTAACGCTGGGAAAGATTGAAGGCAGGAGGAGCAGGGGATGACA  120
ovis_aries_rambouillet/11353    CACATTGGAAAAGACCCTAATGCTAGGAAAGATTGAAGGCAGGAGGAGCAGGGGATGACG  120
                                ************ **** * * * ******* **** ********** **********

bos_taurus/11353                GAGGATGAGATGGTTGGATAACATCACCATTTCAATGGACATGAGTTTGAGCAAACTCTG  176
bos_grunniens/11353             GAGGATGAGATGGTTGGATAACATCACCATTTCAATGGACATGAGTTTGAGCAAACTCTG  176
capra_hircus/11353              GAGGATGAGATGGTTGGATGGCAACACCATTTCAATGGACATGAGCTTGAGCAAACTCTG  180
ovis_aries_rambouillet/11353    GAGGATGAGATGGTTGGATGGCATCACCATTTCAATGGACATGAGCTTGAGCAAACTCTG  180
                                ******************  ** ********************* **************

bos_taurus/11353                GGAGACAGTGAAGGACAGGGAAGCCCGGTGTGTTGCAGACCATGGGGTCACAAAGAGTCG  236
bos_grunniens/11353             GGAGACAGTGAAGGACAGGGAAGCCCGGTGTGTTGCAGACCATGGGGTCACAAAGAGTTG  236
capra_hircus/11353              GGAGTCAGTGAAGGACAGGGAAACCCAGTGTGTTGCAGACTATGGGATCACAAAGAACTG  240
ovis_aries_rambouillet/11353    GGAGGCAGTGAAGGACAGGGAAGCCCAGTGTGTTGCAGACTATGGGATCACAAAGAATTG  240
                                **** ***************** *** ************* ***** *********   *

bos_taurus/11353                GACATGACTTAGCAACTAAACTCACACAGTGTCCAGAGGAACTTGTGAAAAAGAAGCTTT  296
bos_grunniens/11353             GACATGACTTAGCAACTAAACTCACACAGTGTCCAGAGGAACTTGTGAAAAAGAAGCTTT  296
capra_hircus/11353              GACATGACTTAGCAACTAAACTCACACAGTGTACAGAGAAACTTGTGAAAAAGAAGCTTT  300
ovis_aries_rambouillet/11353    GACATGACTTAGCAACTAAACTCACACAGTGTACAGAGAAACTTGTGAAAAAGAAGCTTT  300
                                ******************************** ***** ********************

bos_taurus/11353                AAGGACCTAGATTCTGAGTCAT-AACTTGTAGAACTGAAGTGG------------GACCA  343
bos_grunniens/11353             AAGGACCTAGATTCTGAGTCAT-AACTTGTAGAACTGAAGTGG------------GACCA  343
capra_hircus/11353              AAGGACCTAGATTGTGAAGTCCTAACTTGCAGAACTGAAGTGGGACCGGGGGAATCTTTA  360
ovis_aries_rambouillet/11353    GAGGACCTAGATTGTGAAGTCCTAACTTGCAGAACTGAAGTGGGACCGGGGGAATCTTTA  360
                                 ************ ***      ****** *************               *

bos_taurus/11353                GTGTTTCTTAAGGTGATGCTAATAAAAAAGATGCTCAGGAGACCCTTGGAGAAACATTTA  403
bos_grunniens/11353             GTGTTTCTTAAGGTGATGCTAATAAAAAAGATGCTCAGGAGACCCTTGGAGAAACATTTA  403
capra_hircus/11353              CCGTTTCTTAAGGTGATGCTAATAAAGAAGATGCTCAGGAGACCCTTGGAGAAACATTTA  420
ovis_aries_rambouillet/11353    CCATTTCTTAAGGTGATGCTAATAAAGAAGATGCTCAGGAGACCCTTGGAGAAACATTTA  420
                                   ************************ ***********************************
```

```
bos_taurus/11353                GAGAAAGAAACTGCCTGCTCAGGGAAGACTGGACTGATATCTGCATGTATGCCTTCTCCA  463
bos_grunniens/11353             GAGAAAGAAACTGCCTGCTCAGGGAAGACTGGACTGATATCTGCATGTATGCCTTCTCCA  463
capra_hircus/11353              GAGAAAGAAACTACCTGCTCAGGGAAGACTGGACTGGTATCTGCATGTATGCCTTCTCCA  480
ovis_aries_rambouillet/11353    GAGGAAGAAACTACCTGCTCAGGGAAGACTGGACTGGTATCTGCATGTATGCCTTCTCCA  480
                                *** ******** ********************** **********************

bos_taurus/11353                GATCCACGTGACTTTACCACACCTGGCCTCAGGACAGCACGAGCATTCTATCCCATCAAA  523
bos_grunniens/11353             GATCCACGTGACTTTACCACACCTGGCCTCAGGACAGCACGAGCATTCTATCCCATCAAA  523
capra_hircus/11353              GATCCACATGACTTCACCACATCTGGCCTCAGGACAGCATGAGCATTCTATCCCATCAAA  540
ovis_aries_rambouillet/11353    GATCCACATGACTTCACCACATCTGGCCTCAGGACAGCACGAGCATTCTATCCCGTCAAA  540
                                ******* ****** ****** ***************** ************* *****

bos_taurus/11353                ATTATTCTGCTCCACTGGCTGTGAAGGAGAAGGAGAGATGCATGTTCCAGGACATTGGCT  583
bos_grunniens/11353             ATTATTCTGCTCCACTGGCTGTGAAGGAGAAGGAGAGATGCCTGTTCCAGGACATTGGCT  583
capra_hircus/11353              GTTATTCTGCTCCACTGGCTGTGAAGGAGAAGGAGATACGCATGTTCCAGGACATTGGCT  600
ovis_aries_rambouillet/11353    GTTATTCTGCTCCACTGGCTGTGAAG---AAGGAGATACGCATGTTCCAGGACATTGGCT  597
                                 ************************   ******* * ** *****************

bos_taurus/11353                CAGACAGGCCTACTGAGTGACTCTAAAGACTTGTAGACACCAAAAGGGTTTTTAAGAATA  643
bos_grunniens/11353             CAGACAGGCCTATTGAGTGACTCTAAAGACTTGTAGACACCAAAAGGGTTTTTAAGAATA  643
capra_hircus/11353              CAGACAGGCCTACTGAGTGACTCTAAAGACTTGCAGACATCAAAAGGGTTTTGAAGAATA  660
ovis_aries_rambouillet/11353    TAGACAGGCCTACTGAGTGACTCTAAAGACTTGCAGACATCAAAAGGGTTTTGAAGAATA  657
                                 ********** ******************* ***** ************ *******

bos_taurus/11353                ATCATCATAATATAGTAATGAATCAAAAACAAGCAGACATGTGCTTAAATGAGCCG-GAG  702
bos_grunniens/11353             ATCATCATAATATAGTAATGAATCAAAAACAAGCAGACATGTGCTTAAATGAGCCTGGAG  703
capra_hircus/11353              TTCATCATAACATAGTAATGAATCAAAAACAGGCAGACATATGCTTAAATGAGCCTGGAG  720
ovis_aries_rambouillet/11353    TTCATCATAACATAGTAATGAATCAAAAACAGGCAGACATATGCTTAAATGAGCCTGGAG  717
                                 ********* ******************* ******** **************  ***

bos_taurus/11353                GGACTGGAGTTAAAAGAAGGAAGGGGCTGTTCAGATTCGGGGGTCTAGCAGAATGGACTT  762
bos_grunniens/11353             GGACTGGAGTTAAAAGAAGGAAGGGGCTGTTCAGATTCAGGGGTCTAGCAGAATGGACTT  763
capra_hircus/11353              GGACTGGAGTTAAAAGAAGGAAGGGGCTGTTCAGATTCAGGGGTCTAGCAGAATGGACTT  780
ovis_aries_rambouillet/11353    GGACTGGAGTTAAAAGAAGGAAGGGGCTGTTCAGACTCAGGGGTCTAGCAGAATGGACTT  777
                                ********************************** ** ********************

bos_taurus/11353                GGTTGAGAAAATCCAGCAAGACAACTGAAGAGGACTCTTGGGATTACCCCACCAGAACTT  822
bos_grunniens/11353             GGTTGAGAAAATCCAGCAAGATAACTGAAGAGGACTCTTGGGATTACCCCACCAGAACTT  823
capra_hircus/11353              GCTTGAGAAAATCCAGCAAGACAACTGAAGAGGACTCTTGGGATTACCCCACCAGAACTT  840
ovis_aries_rambouillet/11353    GCTTGAGAAAATCCAGCAAGACAACTGAAGAGGACTCTTGGGATTACCCCACCAGAACTT  837
                                * ****************** ************************************

bos_taurus/11353                CTGAGGCAAGTCTAATGAAATGACTTACAAGTGTTAACTGTAGAGTAGCCTCTAAGGCGC  882
bos_grunniens/11353             CTGAGGCAAGTCTAATGAAATGACTTACAAGTGTTAACTGTAGAGTAGCCTCTAAGGCGC  883
capra_hircus/11353              CTGAGGCAAGTCTAATGAAATGACTTACAAGTGTTAACTGTAGAGTAGCCGCTAAGGCGC  900
ovis_aries_rambouillet/11353    CTGAGGCAAGTCTAATGAAATGACTTACAATTGTTAACTGTAGAGTAGCCGCTAAGGCGC  897
                                ***************************** ****************** *********

bos_taurus/11353                CCCCATTGGCCTAATTCTGGTTGACATTGCACAACCAGGGTGAGCTGACA    932
bos_grunniens/11353             CCCCATTGGCCTAATTCTGGTTGACATTGCACAACCAGGGTGAGCTGACA    933
capra_hircus/11353              CCACATTGGCCTAATTCTGGTTGACATTGCACAACCAGGGTGAGCTGACA    950
ovis_aries_rambouillet/11353    CCACATTGGCCTAATTCTGGTTGACATTGCACAACCAGGGTGAGCTGACA    947
                                ** ********************************************
```

APÊNDICE C – MATERIAL SUPLEMENTAR REFERENTE AO CAPÍTULO 2

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continua)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
1	*Bos indicus*	Afar	AFRETHF000000000002	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
2	*Bos indicus*	Afar	AFRETHF000000000003	Africa	ARS-UCD1.2	No		No
3	*Bos indicus*	Afar	AFRETHF000000000004	Africa	ARS-UCD1.2	No		No
4	*Bos indicus*	Afar	AFRETHF000000000005	Africa	ARS-UCD1.2	No		No
5	*Bos indicus*	Afar	AFRETHF000000000006	Africa	ARS-UCD1.2	No		No
6	*Bos indicus*	Afar	AFRETHF000000000007	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
7	*Bos indicus*	Afar	AFRETHF000000000008	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
8	*Bos indicus*	Afar	AFRETHM0000000000011	Africa	ARS-UCD1.2	No		No
9	*Bos indicus*	Afar	AFRETHM0000000000012	Africa	ARS-UCD1.2	No		No
10	*Bos indicus*	Afar	AFRETHX0000000000013	Africa	ARS-UCD1.2	No		No
11	*Bos indicus*	Afar	AFRETHX0000000000014	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
12	*Bos indicus*	Afar	AFRETHX0000000000015	Africa	ARS-UCD1.2	No		No
13	*Bos indicus*	Afar	AFRETHX0000000000016	Africa	ARS-UCD1.2	Yes	Homozygous	Yes
14	*Bos indicus*	Afar	AFRETHX0000000000017	Africa	ARS-UCD1.2	No		No
15	*Bos indicus*	Begait	BEGETHF000000000010	Africa	ARS-UCD1.2	No		No
16	*Bos indicus*	Begait	BEGETHF000000000011	Africa	ARS-UCD1.2	No		No
17	*Bos indicus*	Begait	BEGETHF000000000013	Africa	ARS-UCD1.2	No		No
18	*Bos indicus*	Begait	BEGETHF000000000014	Africa	ARS-UCD1.2	No		No
19	*Bos indicus*	Begait	BEGETHF000000000016	Africa	ARS-UCD1.2	No		No
20	*Bos indicus*	Begait	BEGETHF000000000018	Africa	ARS-UCD1.2	No		No
21	*Bos indicus*	Begait	BEGETHF000000000019	Africa	ARS-UCD1.2	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
22	*Bos indicus*	Begait	BEGETHF000000000025	Africa	ARS-UCD1.2	Yes	Homozygous	Yes
23	*Bos indicus*	Begait	BEGETHF000000000027	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
24	*Bos indicus*	Ethiopia-Boran	BORETHF000000000001	Africa	ARS-UCD1.2	No		No
25	*Bos indicus*	Ethiopia-Boran	BORETHF000000000002	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
26	*Bos indicus*	Ethiopia-Boran	BORETHF000000000003	Africa	ARS-UCD1.2	No		No
27	*Bos indicus*	Ethiopia-Boran	BORETHF000000000004	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
28	*Bos indicus*	Ethiopia-Boran	BORETHF000000000005	Africa	ARS-UCD1.2	No		No
29	*Bos indicus*	Ethiopia-Boran	BORETHF000000000007	Africa	ARS-UCD1.2	No		No
30	*Bos indicus*	Ethiopia-Boran	BORETHF000000000010	Africa	ARS-UCD1.2	No		No
31	*Bos indicus*	Ethiopia-Boran	BORETHM000000000012	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
32	*Bos indicus*	Ethiopia-Boran	BORETHM000000000013	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
33	*Bos indicus*	Ethiopia-Boran	BORETHM000000000015	Africa	ARS-UCD1.2	No		No
34	*Bos indicus*	Kenya-Boran	BORKENX000000000130	Africa	ARS-UCD1.2	No		No
35	*Bos indicus*	Kenya-Boran	BORKENX000000000168	Africa	ARS-UCD1.2	No		No
36	*Bos indicus*	Kenya-Boran	BORKENX000000000261	Africa	ARS-UCD1.2	No		No
37	*Bos indicus*	Kenya-Boran	BORKENX000000000375	Africa	ARS-UCD1.2	No		No
38	*Bos indicus*	Kenya-Boran	BORKENX000000000392	Africa	ARS-UCD1.2	No		No
39	*Bos indicus*	Kenya-Boran	BORKENX000000000439	Africa	ARS-UCD1.2	No		No
40	*Bos indicus*	Kenya-Boran	BORKENX000000000467	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
41	*Bos indicus*	Kenya-Boran	BORKENX000000000563	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
42	*Bos indicus*	Kenya-Boran	BORKENX000000000587	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
43	*Bos indicus*	Kenya-Boran	BORKENX000000000672	Africa	ARS-UCD1.2	No		No
44	*Bos indicus*	Kenya-Boran	BORKENX000000002684	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
45	*Bos indicus*	Kenya-Boran	BORKENX000000002754	Africa	ARS-UCD1.2	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
46	*Bos indicus*	Kenya-Boran	BORKENX000000002794	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
47	*Bos indicus*	Kenya-Boran	BORKENX000000003265	Africa	ARS-UCD1.2	No		No
48	*Bos indicus*	Kenya-Boran	BORKENX000000003321	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
49	*Bos indicus*	Kenya-Boran	BORKENX000000003324	Africa	ARS-UCD1.2	No		No
50	*Bos indicus*	Butana	BUTSUDX000000000011	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
51	*Bos indicus*	Butana	BUTSUDX000000000012	Africa	ARS-UCD1.2	No		No
52	*Bos indicus*	Butana	BUTSUDX000000000013	Africa	ARS-UCD1.2	No		No
53	*Bos indicus*	Butana	BUTSUDX000000000014	Africa	ARS-UCD1.2	No		No
54	*Bos indicus*	Butana	BUTSUDX000000000015	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
55	*Bos indicus*	Butana	BUTSUDX000000000016	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
56	*Bos indicus*	Butana	BUTSUDX000000000017	Africa	ARS-UCD1.2	No		No
57	*Bos indicus*	Butana	BUTSUDX000000000018	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
58	*Bos indicus*	Butana	BUTSUDX000000000019	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
59	*Bos indicus*	Butana	BUTSUDX000000000020	Africa	ARS-UCD1.2	No		No
60	*Bos indicus*	Fogera	FOGETHF000000000002	Africa	ARS-UCD1.2	No		No
61	*Bos indicus*	Fogera	FOGETHF000000000003	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
62	*Bos indicus*	Fogera	FOGETHF000000000005	Africa	ARS-UCD1.2	No		No
63	*Bos indicus*	Fogera	FOGETHF000000000006	Africa	ARS-UCD1.2	No		No
64	*Bos indicus*	Fogera	FOGETHF000000000007	Africa	ARS-UCD1.2	No		No
65	*Bos indicus*	Fogera	FOGETHF000000000008	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
66	*Bos indicus*	Fogera	FOGETHF000000000009	Africa	ARS-UCD1.2	No		No
67	*Bos indicus*	Fogera	FOGETHF000000000010	Africa	ARS-UCD1.2	No		No
68	*Bos indicus*	Fogera	FOGETHM000000000011	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
69	*Bos indicus*	Fogera	FOGETHX000000000052	Africa	ARS-UCD1.2	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
70	*Bos indicus*	Fogera	FOGETHX000000000054	Africa	ARS-UCD1.2	No		No
71	*Bos indicus*	Fogera	FOGETHX000000000058	Africa	ARS-UCD1.2	No		No
72	*Bos indicus*	Kenana	KENSUDX000000000003	Africa	ARS-UCD1.2	No		No
73	*Bos indicus*	Kenana	KENSUDX000000000007	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
74	*Bos indicus*	Kenana	KENSUDX000000000010	Africa	ARS-UCD1.2	Yes	Heterozygous	Yes
75	*Bos indicus*	Kenana	KENSUDX000000000011	Africa	ARS-UCD1.2	No		No
76	*Bos indicus*	Kenana	KENSUDX000000000012	Africa	ARS-UCD1.2	No		No
77	*Bos indicus*	Kenana	KENSUDX000000000013	Africa	ARS-UCD1.2	No		No
78	*Bos indicus*	Kenana	KENSUDX000000000014	Africa	ARS-UCD1.2	No		No
79	*Bos indicus*	Kenana	KENSUDX000000000016	Africa	ARS-UCD1.2	No		No
80	*Bos indicus*	Kenana	KENSUDX000000000018	Africa	ARS-UCD1.2	No		No
81	*Bos indicus*	Kenana	KENSUDX000000000020	Africa	ARS-UCD1.2	No		No
82	*Bos indicus*	Nellore	NEL1	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
83	*Bos indicus*	Nellore	NEL2	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
84	*Bos indicus*	Nellore	NEL3	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
85	*Bos indicus*	Nellore	NEL4	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
86	*Bos indicus*	Nellore	NEL5	Brazil	ARS-UCD1.2	No		No
87	*Bos indicus*	Nellore	NEL6	Brazil	ARS-UCD1.2	No		No
88	*Bos indicus*	Nellore	NEL7	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
89	*Bos indicus*	Nellore	NEL8	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
90	*Bos indicus*	Nellore	NEL9	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
91	*Bos indicus*	Nellore	NEL10	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
92	*Bos indicus*	Nellore	NEL11	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
93	*Bos indicus*	Nellore	NEL12	Brazil	ARS-UCD1.2	No		Yes

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
94	*Bos indicus*	Nellore	NEL13	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
95	*Bos indicus*	Nellore	NEL14	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
96	*Bos indicus*	Nellore	NEL15	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
97	*Bos indicus*	Nellore	NEL16	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
98	*Bos indicus*	Nellore	NEL17	Brazil	ARS-UCD1.2	No		No
99	*Bos indicus*	Nellore	NEL18	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
100	*Bos indicus*	Nellore	NEL19	Brazil	ARS-UCD1.2	No		No
101	*Bos indicus*	Nellore	NEL20	Brazil	ARS-UCD1.2	No		No
102	*Bos indicus*	Nellore	NEL21	Brazil	ARS-UCD1.2	No		No
103	*Bos indicus*	Nellore	NEL22	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
104	*Bos indicus*	Nellore	NEL23	Brazil	ARS-UCD1.2	Yes	Homozygous	Yes
105	*Bos indicus*	Nellore	NEL24	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
106	*Bos indicus*	Nellore	NEL8131	Brazil	ARS-UCD1.2	Yes	Heterozygous	Yes
107	*Bos indicus*	Nellore	NEL8243	Brazil	ARS-UCD1.2	No		No
108	*Bos indicus*	Nellore	NEL8254	Brazil	ARS-UCD1.2	No		No
109	*Bos indicus*	Ogaden	OGDETHX000000000001	Pakistan	ARS-UCD1.2	No		No
110	*Bos indicus*	Ogaden	OGDETHX000000000002	Pakistan	ARS-UCD1.2	No		No
111	*Bos indicus*	Ogaden	OGDETHX000000000004	Pakistan	ARS-UCD1.2	No		No
112	*Bos indicus*	Ogaden	OGDETHX000000000005	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
113	*Bos indicus*	Ogaden	OGDETHX000000000006	Pakistan	ARS-UCD1.2	No		No
114	*Bos indicus*	Ogaden	OGDETHX000000000007	Pakistan	ARS-UCD1.2	Yes	Homozygous	Yes
115	*Bos indicus*	Ogaden	OGDETHX000000000008	Pakistan	ARS-UCD1.2	No		No
116	*Bos indicus*	Ogaden	OGDETHX000000000009	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
117	*Bos indicus*	Ogaden	OGDETHX000000000010	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
118	*Bos indicus*	Achai	PKTACH0007	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
119	*Bos indicus*	Achai	PKTACH0064	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
120	*Bos indicus*	Bhagnari	PKTBGN0018	Pakistan	ARS-UCD1.2	No		No
121	*Bos indicus*	Bhagnari	PKTBGN0020	Pakistan	ARS-UCD1.2	No		No
122	*Bos indicus*	Bhagnari	PKTBGN0023	Pakistan	ARS-UCD1.2	No		No
123	*Bos indicus*	Cholistani	PKTCHO3702	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
124	*Bos indicus*	Cholistani	PKTCHO6314	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
125	*Bos indicus*	Dajal	PKTDAJ0001	Pakistan	ARS-UCD1.2	No		No
126	*Bos indicus*	Dhanni	PKTDHA0287	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
127	*Bos indicus*	Dhanni	PKTDHA0363	Pakistan	ARS-UCD1.2	No		No
128	*Bos indicus*	Gabrali	PKTGBR0003	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
129	*Bos indicus*	Gabrali	PKTGBR0027	Pakistan	ARS-UCD1.2	No		No
130	*Bos indicus*	HisarH	PKTHSH0044	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
131	*Bos indicus*	HisarH	PKTHSH0046	Pakistan	ARS-UCD1.2	No		No
132	*Bos indicus*	Lohani	PKTLHN0018	Pakistan	ARS-UCD1.2	Yes	Homozygous	Yes
133	*Bos indicus*	Red Sindhi	PKTRSD0303	Pakistan	ARS-UCD1.2	Yes	Homozygous	Yes
134	*Bos indicus*	Sahiwal	PKTSHW0001	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
135	*Bos indicus*	Sahiwal	PKTSHW0002	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
136	*Bos indicus*	Tharparkar	PKTTHP0138	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
137	*Bos indicus*	Tharparkar	PKTTHP0158	Pakistan	ARS-UCD1.2	Yes	Heterozygous	Yes
138	*Bos grunniens*	Yak	SRR1948183	Public database	UMD3.0	No		No
139	*Bos grunniens*	Yak	SRR1948186	Public database	UMD3.1	No		No
140	*Bos taurus*	Angus	SRR7946664	Public database	UMD3.1	No		No
141	*Bos taurus*	Angus	SRR7946665	Public database	UMD3.1	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
142	*Bos taurus*	Angus	SRR7946705	Public database	UMD3.1	No		No
143	*Bos taurus*	Angus	SRR7946702	Public database	UMD3.1	No		No
144	*Bos taurus*	Angus	SRR7946663	Public database	UMD3.1	No		No
145	*Bos taurus*	Angus	SRR7946708	Public database	UMD3.1	No		No
146	*Bos taurus*	Angus	SRR7946658	Public database	UMD3.1	No		No
147	*Bos taurus*	Angus	SRR1262648	Public database	UMD3.1	No		No
148	*Bos taurus*	Angus	SRR1262624	Public database	UMD3.1	No		No
149	*Bos taurus*	Angus	SRR7946657	Public database	UMD3.1	No		No
150	*Bos taurus*	Belgian Blue	SRR7946640	Public database	UMD3.1	No		No
151	*Bos taurus*	Belgian Blue	SRR7946638	Public database	UMD3.1	No		No
152	*Bos taurus*	Belgian Blue	SRR7946641	Public database	UMD3.1	No		No
153	*Bos taurus*	Belgian Blue	SRR7946707	Public database	UMD3.1	No		No
154	*Bos taurus*	Belgian Blue	SRR7946706	Public database	UMD3.1	No		No
155	*Bos taurus*	Charolais	SRR7946743	Public database	UMD3.1	No		No
156	*Bos taurus*	Charolais	SRR7946749	Public database	UMD3.1	No		No
157	*Bos taurus*	Charolais	SRR7946746	Public database	UMD3.1	No		No
158	*Bos taurus*	Charolais	SRR7946725	Public database	UMD3.1	No		No
159	*Bos taurus*	Charolais	SRR7946728	Public database	UMD3.1	No		No
160	*Bos taurus*	Charolais	SRR7946745	Public database	UMD3.1	No		No
161	*Bos taurus*	Charolais	SRR7946729	Public database	UMD3.1	No		No
162	*Bos taurus*	Charolais	SRR7946667	Public database	UMD3.1	No		No
163	*Bos taurus*	Charolais	SRR7946760	Public database	UMD3.1	No		No
164	*Bos taurus*	Charolais	SRR7946643	Public database	UMD3.1	No		No
165	*Bos taurus*	Hereford	SRR7946722	Public database	UMD3.1	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
166	*Bos taurus*	Hereford	SRR7946732	Public database	UMD3.1	No		No
167	*Bos taurus*	Hereford	SRR7946723	Public database	UMD3.1	No		No
168	*Bos taurus*	Hereford	SRR7946773	Public database	UMD3.1	No		No
169	*Bos taurus*	Hereford	SRR7946737	Public database	UMD3.1	No		No
170	*Bos taurus*	Hereford	SRR7946740	Public database	UMD3.1	No		No
171	*Bos taurus*	Hereford	SRR7946734	Public database	UMD3.1	No		No
172	*Bos taurus*	Hereford	SRR7946733	Public database	UMD3.1	No		No
173	*Bos taurus*	Hereford	SRR1972977	Public database	UMD3.1	No		No
174	*Bos taurus*	Hereford	SRR7946741	Public database	UMD3.1	No		No
175	*Bos taurus*	Holstein	SRR7994021	Public database	UMD3.1	No		No
176	*Bos taurus*	Holstein	SRR7994017	Public database	UMD3.1	No		No
177	*Bos taurus*	Holstein	SRR7994018	Public database	UMD3.1	No		No
178	*Bos taurus*	Holstein	SRR7994027	Public database	UMD3.1	No		No
179	*Bos taurus*	Holstein	SRR7994003	Public database	UMD3.1	No		No
180	*Bos taurus*	Holstein	SRR7994016	Public database	UMD3.1	No		No
181	*Bos taurus*	Holstein	SRR7994015	Public database	UMD3.1	No		No
182	*Bos taurus*	Holstein	SRR7994014	Public database	UMD3.1	No		No
183	*Bos taurus*	Holstein	SRR7994013	Public database	UMD3.1	No		No
184	*Bos taurus*	Holstein	SRR7994024	Public database	UMD3.1	No		No
185	*Bos taurus*	Jersey	SRR1262797	Public database	UMD3.1	No		No
186	*Bos taurus*	Jersey	SRR1262802	Public database	UMD3.1	No		No
187	*Bos taurus*	Jersey	SRR1262799	Public database	UMD3.1	No		No
188	*Bos taurus*	Jersey	SRR1262803	Public database	UMD3.1	No		No
189	*Bos taurus*	Jersey	SRR1262793	Public database	UMD3.1	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Continuação)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
190	*Bos taurus*	Jersey	SRR1262801	Public database	UMD3.1	No		No
191	*Bos taurus*	Jersey	SRR1262798	Public database	UMD3.1	No		No
192	*Bos taurus*	Limousin	SRR7946696	Public database	UMD3.1	No		No
193	*Bos taurus*	Limousin	SRR7946721	Public database	UMD3.1	No		No
194	*Bos taurus*	Limousin	SRR7946694	Public database	UMD3.1	No		No
195	*Bos taurus*	Limousin	SRR7946690	Public database	UMD3.1	No		No
196	*Bos taurus*	Limousin	SRR7946752	Public database	UMD3.1	No		No
197	*Bos taurus*	Limousin	SRR7946700	Public database	UMD3.1	No		No
198	*Bos taurus*	Limousin	SRR7946715	Public database	UMD3.1	No		No
199	*Bos taurus*	Limousin	SRR7946701	Public database	UMD3.1	No		No
200	*Bos taurus*	Limousin	SRR7946698	Public database	UMD3.1	No		No
201	*Bos taurus*	Limousin	SRR7946693	Public database	UMD3.1	No		No
202	*Bos taurus*	Romagnola	ERR466547	Public database	UMD3.1	Yes	Heterozygous	NA
203	*Bos taurus*	Romagnola	ERR1747046	Public database	UMD3.1	No		No
204	*Bos taurus*	Romagnola	ERR1766318	Public database	UMD3.1	Yes	Homozygous	Yes
205	*Bos taurus*	Romagnola	ERR1766319	Public database	UMD3.1	Yes	Heterozygous	Yes
206	*Bos taurus*	Romagnola	ERR1766320	Public database	UMD3.1	Yes	Heterozygous	Yes
207	*Bos taurus*	Simmental	SRR7946648	Public database	UMD3.1	No		No
208	*Bos taurus*	Simmental	SRR7946770	Public database	UMD3.1	No		No
209	*Bos taurus*	Simmental	SRR7971552	Public database	UMD3.1	No		No
210	*Bos taurus*	Simmental	SRR7946676	Public database	UMD3.1	No		No
211	*Bos taurus*	Simmental	SRR7946651	Public database	UMD3.1	Yes	Heterozygous	Yes
212	*Bos taurus*	Simmental	SRR7976539	Public database	UMD3.1	No		No
213	*Bos taurus*	Simmental	SRR7946653	Public database	UMD3.1	No		No

Table S1 - Information about species, breed, register, origin, reference assembly and ASIP-SV1 deletion and insertion analyzed.

(Conclusão)

Number	Species	Breed	Register	Origin	Reference Assembly	Deletion		Isertion
						Presence	Heterozygous/ Homozygous	
214	*Bos taurus*	Simmental	SRR7946647	Public database	UMD3.1	No		No
215	*Bos taurus*	Simmental	SRR7946650	Public database	UMD3.1	No		No
216	*Bos taurus*	Simmental	SRR1262805	Public database	UMD3.1	No		No

ANEXO A – NORMAS DE PUBLICAÇÃO DA REVISTA *GENOMICS, SELECTION AND EVOLUTION*

Research article

Criteria

Please use this template when preparing your Research article in a Word document. If preparing your manuscript in LaTeX, please use the template provided under the specific instructions for LaTeX further down the page.

You can suggest potential peer reviewers for your manuscript in your cover letter. If you wish to do so, please provide institutional email addresses where possible, or information which will help the Editor to verify the identity of the reviewer (for example an ORCID or Scopus ID).

Genetics Selection Evolution strongly encourages that all datasets on which the conclusions of the paper rely should be available to readers. We encourage authors to ensure that their datasets are either deposited in publicly available repositories (where available and appropriate) or presented in the main manuscript or additional supporting files whenever possible. Please see Springer Nature's information on recommended repositories. Where a widely established research community expectation for data archiving in public repositories exists, submission to a community-endorsed, public repository is mandatory. A list of data where deposition is required, with the appropriate repositories, can be found on the Editorial Policies Page.

Presubmission enquiries

If you wish to make a presubmission enquiry about the suitability of your manuscript, please email the editors who will respond to your enquiry as soon as possible.

Preparing your manuscript

The information below details the section headings that you should include in your manuscript and what information should be within each section.

Please note that your manuscript must include a 'Declarations' section including all of the subheadings (please see below for more information).

Title page

The title page should:

- present a title that includes, if appropriate, the study design
- list the full names and institutional addresses for all authors
    - if a collaboration group should be listed as an author, please list the Group name as an author. If you would like the names of the individual members of the Group to be searchable through their individual PubMed records, please include this information in the "Acknowledgements" section in accordance with the instructions below
- indicate the corresponding author

Abstract

The Abstract should not exceed 350 words. Please minimize the use of abbreviations and do not cite references in the abstract. The abstract must include the following separate sections:

- **Background:** the context and purpose of the study
- **Results:** the main findings
- **Conclusions:** a brief summary and potential implications

Background

The Background section should explain the background to the study, its aims, a summary of the existing literature and why this study was necessary.

Methods

The methods section should include:

- the aim, design and setting of the study
- the characteristics of participants or description of materials
- a clear description of all processes, interventions and comparisons. Generic names should generally be used. When proprietary brands are used in research, include the brand names in parentheses
- the type of statistical analysis used, including a power calculation if appropriate

Results

This should include the findings of the study including, if appropriate, results of statistical analysis which must be included either in the text or as tables and figures.

Discussion

For research articles this section should discuss the implications of the findings in context of existing research and highlight limitations of the study. For study protocols and methodology manuscripts this section should include a discussion of any practical or operational issues involved in performing the study and any issues not covered in other sections.

Conclusions

This should state clearly the main conclusions and provide an explanation of the importance and relevance of the study to the field.

List of abbreviations

If abbreviations are used in the text they should be defined in the text at first use, and a list of abbreviations can be provided.

Declarations

All manuscripts must contain the following sections under the heading 'Declarations':

- Ethics approval and consent to participate
- Consent for publication
- Availability of data and materials
- Competing interests
- Funding
- Authors' contributions
- Acknowledgements
- Authors' information (optional)

Please see below for details on the information to be included in these sections.

If any of the sections are not relevant to your manuscript, please include the heading and write 'Not applicable' for that section.

Ethics approval and consent to participate

Manuscripts reporting studies involving human participants, human data or human tissue must:

- include a statement on ethics approval and consent (even where the need for approval was waived)
- include the name of the ethics committee that approved the study and the committee's reference number if appropriate

Studies involving animals must include a statement on ethics approval and for experimental studies involving client-owned animals, authors must also include a statement on informed consent from the client or owner.

See our editorial policies for more information.

If your manuscript does not report on or involve the use of any animal or human data or tissue, please state "Not applicable" in this section.

Consent for publication

If your manuscript contains any individual person's data in any form (including any individual details, images or videos), consent for publication must be obtained from that person, or in the case of children, their parent or legal guardian. All presentations of case reports must have consent for publication.

You can use your institutional consent form or our consent form if you prefer. You should not send the form to us on submission, but we may request to see a copy at any stage (including after publication).

See our editorial policies for more information on consent for publication.

If your manuscript does not contain data from any individual person, please state "Not applicable" in this section.

Availability of data and materials

All manuscripts must include an 'Availability of data and materials' statement. Data availability statements should include information on where data supporting the results reported in the article can be found including, where applicable, hyperlinks to publicly archived datasets analysed or generated during the study. By data we mean the minimal dataset that would be necessary to interpret, replicate and build upon the findings reported in the article. We recognise it is not always possible to share research data publicly, for instance when individual privacy could be compromised, and in such instances data availability should still be stated in the manuscript along with any conditions for access.

Data availability statements can take one of the following forms (or a combination of more than one if required for multiple datasets):

- The datasets generated and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS]
- The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.
- All data generated or analysed during this study are included in this published article [and its supplementary information files].
- The datasets generated and/or analysed during the current study are not publicly available due [REASON WHY DATA ARE NOT PUBLIC] but are available from the corresponding author on reasonable request.
- Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.
- The data that support the findings of this study are available from [third party name] but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of [third party name].

- Not applicable. If your manuscript does not contain any data, please state 'Not applicable' in this section.

More examples of template data availability statements, which include examples of openly available and restricted access datasets, are available here.

BioMed Central also requires that authors cite any publicly available data on which the conclusions of the paper rely in the manuscript. Data citations should include a persistent identifier (such as a DOI) and should ideally be included in the reference list. Citations of datasets, when they appear in the reference list, should include the minimum information recommended by DataCite and follow journal style. Dataset identifiers including DOIs should be expressed as full URLs. For example:

Hao Z, AghaKouchak A, Nakhjiri N, Farahmand A. Global integrated drought monitoring and prediction system (GIDMaPS) data sets. figshare. 2014. http://dx.doi.org/10.6084/m9.figshare.853801

With the corresponding text in the Availability of data and materials statement:

The datasets generated during and/or analysed during the current study are available in the [NAME] repository, [PERSISTENT WEB LINK TO DATASETS].[Reference number]

If you wish to co-submit a data note describing your data to be published in *BMC Research Notes*, you can do so by visiting our submission portal. Data notes support open data and help authors to comply with funder policies on data sharing. Co-published data notes will be linked to the research article the data support (example).

Competing interests

All financial and non-financial competing interests must be declared in this section.

See our editorial policies for a full explanation of competing interests. If you are unsure whether you or any of your co-authors have a competing interest please contact the editorial office.

Please use the authors initials to refer to each authors' competing interests in this section.

If you do not have any competing interests, please state "The authors declare that they have no competing interests" in this section.

Funding

All sources of funding for the research reported should be declared. The role of the funding body in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript should be declared.

Authors' contributions

The individual contributions of authors to the manuscript should be specified in this section. Guidance and criteria for authorship can be found in our editorial policies.

Please use initials to refer to each author's contribution in this section, for example: "FC analyzed and interpreted the patient data regarding the hematological disease and the transplant. RH performed the histological examination of the kidney, and was a major contributor in writing the manuscript. All authors read and approved the final manuscript."

Acknowledgements

Please acknowledge anyone who contributed towards the article who does not meet the criteria for authorship including anyone who provided professional writing services or materials.

Authors should obtain permission to acknowledge from all those mentioned in the Acknowledgements section.

See our editorial policies for a full explanation of acknowledgements and authorship criteria.

If you do not have anyone to acknowledge, please write "Not applicable" in this section.

Group authorship (for manuscripts involving a collaboration group): if you would like the names of the individual members of a collaboration Group to be searchable through their individual PubMed records, please ensure that the title of the collaboration Group is included on the title page and in the submission system and also include collaborating author names as the last paragraph of the "Acknowledgements" section. Please add authors in the format First Name, Middle initial(s) (optional), Last Name. You can add institution or country information for each author if you wish, but this should be consistent across all authors.

Please note that individual names may not be present in the PubMed record at the time a published article is initially included in PubMed as it takes PubMed additional time to code this information.

Authors' information

This section is optional.

You may choose to use this section to include any relevant information about the author(s) that may aid the reader's interpretation of the article, and understand the standpoint of the author(s). This may include details about the authors' qualifications, current positions they hold at institutions or societies, or any other relevant background information. Please refer to authors using their initials. Note this section should not be used to describe any competing interests.

Footnotes

Footnotes can be used to give additional information, which may include the citation of a reference included in the reference list. They should not consist solely of a reference citation, and they should never include the bibliographic details of a reference. They should also not contain any figures or tables.

Footnotes to the text are numbered consecutively; those to tables should be indicated by superscript lower-case letters (or asterisks for significance values and other statistical data). Footnotes to the title or the authors of the article are not given reference symbols.

Always use footnotes instead of endnotes.

References

Examples of the Vancouver reference style are shown below.

See our editorial policies for author guidance on good citation practice

Web links and URLs: All web links and URLs, including links to the authors' own websites, should be given a reference number and included in the reference list rather than within the text of the manuscript. They should be provided in full, including both the title of the site and the URL, as well as the date the site was accessed, in the following format: The Mouse Tumor Biology Database. http://tumor.informatics.jax.org/mtbwi/index.do. Accessed 20 May 2013. If an author or group of authors can clearly be associated with a web link, such as for weblogs, then they should be included in the reference.

Example reference style:

Article within a journal

Smith JJ. The world of science. Am J Sci. 1999;36:234-5.

Article within a journal (no page numbers)

Rohrmann S, Overvad K, Bueno-de-Mesquita HB, Jakobsen MU, Egeberg R, Tjønneland A, et al. Meat consumption and mortality - results from the European Prospective Investigation into Cancer and Nutrition. BMC Medicine. 2013;11:63.

Article within a journal by DOI

Slifka MK, Whitton JL. Clinical implications of dysregulated cytokine production. Dig J Mol Med. 2000; doi:10.1007/s801090000086.

Article within a journal supplement

Frumin AM, Nussbaum J, Esposito M. Functional asplenia: demonstration of splenic activity by bone marrow scan. Blood 1979;59 Suppl 1:26-32.

Book chapter, or an article within a book

Wyllie AH, Kerr JFR, Currie AR. Cell death: the significance of apoptosis. In: Bourne GH, Danielli JF, Jeon KW, editors. International review of cytology. London: Academic; 1980. p. 251-306.

OnlineFirst chapter in a series (without a volume designation but with a DOI)

Saito Y, Hyuga H. Rate equation approaches to amplification of enantiomeric excess and chiral symmetry breaking. Top Curr Chem. 2007. doi:10.1007/128_2006_108.

Complete book, authored

Blenkinsopp A, Paxton P. Symptoms in the pharmacy: a guide to the management of common illness. 3rd ed. Oxford: Blackwell Science; 1998.

Online document

Doe J. Title of subordinate document. In: The dictionary of substances and their effects. Royal Society of Chemistry. 1999. http://www.rsc.org/dose/title of subordinate document. Accessed 15 Jan 1999.

Online database

Healthwise Knowledgebase. US Pharmacopeia, Rockville. 1998. http://www.healthwise.org. Accessed 21 Sept 1998.

Supplementary material/private homepage

Doe J. Title of supplementary material. 2000. http://www.privatehomepage.com. Accessed 22 Feb 2000.

University site

Doe, J: Title of preprint. http://www.uni-heidelberg.de/mydata.html (1999). Accessed 25 Dec 1999.

FTP site

Doe, J: Trivial HTTP, RFC2169. ftp://ftp.isi.edu/in-notes/rfc2169.txt (1999). Accessed 12 Nov 1999.

Organization site

ISSN International Centre: The ISSN register. http://www.issn.org (2006). Accessed 20 Feb 2007.

Dataset with persistent identifier

Zheng L-Y, Guo X-S, He B, Sun L-J, Peng Y, Dong S-S, et al. Genome data from sweet and grain sorghum (Sorghum bicolor). GigaScience Database. 2011. http://dx.doi.org/10.5524/100012.

Figures, tables and additional files

See General formatting guidelines for information on how to format figures, tables and additional files.

General formatting information

Manuscripts must be written in concise English. For help on scientific writing, or preparing your manuscript in English, please see Springer's Author Academy.

Quick points:

- Use double line spacing
- Include line and page numbering
- Use SI units: Please ensure that all special characters used are embedded in the text, otherwise they will be lost during conversion to PDF
- Do not use page breaks in your manuscript

File formats

The following word processor file formats are acceptable for the main manuscript document:

- Microsoft word (DOC, DOCX)
- Rich text format (RTF)
- TeX/LaTeX (use BioMed Central's TeX template)

Please note: editable files are required for processing in production. If your manuscript contains any non-editable files (such as PDFs) you will be required to re-submit an editable file if your manuscript is accepted.

Note that figures must be submitted as separate image files, not as part of the submitted manuscript file.

Additional information for TeX/LaTeX users

Please use a recommended TeX template and BibTeX stylefile if you use TeX format. When submitting TeX submissions, please submit your TeX file as the main manuscript file and your bib/bbl file as a dependent file. Please also convert your TeX file into a PDF and submit this PDF as an additional file with the name 'Reference PDF'. This PDF will be used by our production team as a reference point to check the layout of the article as the author intended. Please also note that all figures must be coded at the end of the TeX file and not inline.

The referencing style used by this journal is Vancouver, so we recommend using the vancouver.bst template, available in the BioMedCentral_article file below.

All relevant editable source files must be uploaded during the submission process. Failing to submit these source files will cause unnecessary delays in the production process.

TeX templates:

- BioMedCentral_article (ZIP format) - preferred template
- Springer article svjour3 (ZIP format) - preferred template
- birkjour (Birkhäuser, ZIP format)

- article (part of the standard TeX distribution)
- amsart (part of the standard TeX distribution)

Style and language

Manuscripts submitted to most journals do not undergo copyediting for style and language. Please check individual journal 'About' pages to confirm whether accepted manuscripts will undergo copyediting for style and language.

You can use a professional language editing service of your choice if you want to. Such services include:

- Edanz Language Editing. BioMed Central authors can obtain a 10% discount to the fee charged by Edanz if they choose to use this service.
- Nature Publishing Group Language Editing. Authors can use this coupon code to claim a 10% discount: LE_BM15

Contact the service providers directly to make arrangements for editing, and for pricing and payment details. Use of an editing service is neither a requirement nor a guarantee of acceptance for publication.

Data and materials

For all journals, BioMed Central strongly encourages all datasets on which the conclusions of the manuscript rely to be either deposited in publicly available repositories (where available and appropriate) or presented in the main paper or additional supporting files, in machine-readable format (such as spread sheets rather than PDFs) whenever possible. Please see the list of recommended repositories in our editorial policies.

For some journals, deposition of the data on which the conclusions of the manuscript rely is an absolute requirement. Please check the Instructions for Authors for the relevant journal and article type for journal specific policies.

For all manuscripts, information about data availability should be detailed in an 'Availability of data and materials' section. For more information on the content of this section, please see the Declarations section of the relevant journal's Instruction for Authors. For more information on BioMed Central's policies on data availability, please see our editorial policies.

Formatting the 'Availability of data and materials' section of your manuscript

The following format for the 'Availability of data and materials section of your manuscript should be used:

"The dataset(s) supporting the conclusions of this article is(are) available in the [repository name] repository, [unique persistent identifier and hyperlink to dataset(s) in http:// format]."

The following format is required when data are included as additional files:

"The dataset(s) supporting the conclusions of this article is(are) included within the article (and its additional file(s))."

For databases, this section should state the web/ftp address at which the database is available and any restrictions to its use by non-academics.

For software, this section should include:

- Project name: e.g. My bioinformatics project
- Project home page: e.g. http://sourceforge.net/projects/mged
- Archived version: DOI or unique identifier of archived software or code in repository (e.g. enodo)
- Operating system(s): e.g. Platform independent

- Programming language: e.g. Java
- Other requirements: e.g. Java 1.3.1 or higher, Tomcat 4.0 or higher
- License: e.g. GNU GPL, FreeBSD etc.
- Any restrictions to use by non-academics: e.g. licence needed

Information on available repositories for other types of scientific data, including clinical data, can be found in our editorial policies.

References

See our editorial policies for author guidance on good citation practice.

Genetics Selection Evolution requires references to be formatted in Vancouver referencing style . If you are submitting your manuscript as a Tex/LaTex file, please select the relevant template. For more information on how to format your references, please see BioMed Central's submission guidelines.

What should be cited?

Only articles, clinical trial registration records and abstracts that have been published or are in press, or are available through public e-print/preprint servers, may be cited.

Unpublished abstracts, unpublished data and personal communications should not be included in the reference list, but may be included in the text and referred to as "unpublished observations" or "personal communications" giving the names of the involved researchers. Obtaining permission to quote personal communications and unpublished data from the cited colleagues is the responsibility of the author. Journal abbreviations follow Index Medicus/MEDLINE.

Any in press articles cited within the references and necessary for the reviewers' assessment of the manuscript should be made available if requested by the editorial office.

Web links and URLs: All web links and URLs, including links to the authors' own websites, should be given a reference number and included in the reference list rather than within the text of the manuscript. They should be provided in full, including both the title of the site and the URL, as well as the date the site was accessed, in the following format: The Mouse Tumor Biology Database. http://tumor.informatics.jax.org/mtbwi/index.do. Accessed 20 May 2013. If an author or group of authors can clearly be associated with a web link, such as for weblogs, then they should be included in the reference.

Authors may wish to make use of reference management software to ensure that reference lists are correctly formatted. An example of such software is Papers, which is part of Springer Science+Business Media.

ANEXO B – NORMAS DE PUBLICAÇÃO DA REVISTA *ANIMAL GENETICS*

ANIMAL GENETICS
Immunogenetics, Molecular Genetics and Functional Genomics

Author Guidelines

Scope

Animal Genetics reports frontline research on molecular genetics and functional genomics of economically important and domesticated animals. Publications include variability at genome, gene, expression, metabolome and protein levels; mapping of traits and quantitative trait loci (QTL); associations between genes or gene expression and traits; genetic diversity and epidemiology; epigenetics and epigenomics; and genome-wide characterization of gene expression. We welcome, amongst others, manuscripts that report novel findings relating to:

- Evidence for novel causative genes and their allelic variants
- Molecular background of traits revealed by genomic selection
- Comprehensive analysis of genetic and epigenetic variation associated with complex traits
- Effect of causative mutations
- Geographic distribution of adaptive variation
- Breed relationships and population history derived from mitochondrial DNA, Y-chromosomal, genome-wide SNP panels or whole-genome sequences
- Development of bioinformatic tools, including comprehensive and user-friendly software packages

Reports of gene sequences that are not accompanied by novel genetic findings will not be considered. Studies of association of gene variants with traits should preferably be based on haplotypes. Associations of a single silent genetic variant with a single trait within one population without support of genome-wide association studies (GWAS) and reporting a significant but only marginal effect will normally not be considered for publication. We also do not encourage separate publications based on the same dataset. Genetic diversity studies should have a wider scope than a few breeds on a regional scale (e.g. originating from a single country). If based on microsatellites, typically at least 20 markers are required, but we do not accept microsatellite-based studies on genetic diversity of species for which SNP bead-arrays are available. Gene expression studies should have a genetic component (e.g. the effect of a mutation; evidencing new candidate genes) and/or be at the genome-wide level; we do not publish molecular studies of development. Comparison of transcriptomes should be based on at least three samples (biological replicates) per category being compared (breed, tissue, condition, treatment, etc.); these should be sequenced separately and the individual variation of the expression profiles should be shown.

General

Manuscripts should be submitted through the *Animal Genetics* – ScholarOne Manuscripts (S1M) electronic editorial office. Note that prior to submission, all relevant datasets, including those required for reproduction of the data, should be submitted to public repositories and that submission codes should be cited in the section entitled Data Availability. If this is seriously problematic, contact the Editor-in-Chief.

Free format submission

Animal Genetics now offers Free format submission for a simplified and streamlined submission process.

Before you submit, you will need:

- Your manuscript: this should be an editable file including text, figures, and tables, or separate files – whichever you prefer. All required sections should be contained in your manuscript, including abstract, introduction, methods, results, and discussion. Figures and tables should have legends. Figures should be uploaded in the highest resolution possible. References may be submitted in any style or format, as long as it is consistent throughout the manuscript. Supporting information should be submitted in separate files. If the manuscript, figures or tables are difficult for you to read, they will also be difficult for the editors and reviewers, and the editorial office will send it back to you for revision. Your manuscript may also be sent back to you for revision if the quality of English language is poor.
- An ORCID ID, freely available at https://orcid.org. *(Why is this important? Your article, if accepted and published, will be attached to your ORCID profile. Institutions and funders are increasingly requiring authors to have ORCID IDs.)*
- The title page of the manuscript, including:
    - A title and running title
    - Your co-author details, including affiliation and email address. *(Why is this important? We need to keep all co-authors informed of the outcome of the peer review process.)*
    - Statements relating to our ethics and integrity policies, which may include any of the following *(Why are these important? We need to uphold rigorous ethical standards for the research we consider for publication):*

- data availability statement
- funding statement
- conflict of interest disclosure
- ethics approval statement
- patient consent statement
- permission to reproduce material from other sources
- clinical trial registration
- a running title

To submit, login at **https://mc.manuscriptcentral.com/angen** and create a new submission. Follow the submission steps as required and submit the manuscript.

The first time accessing the site, the corresponding author will need to create an account (top left corner). The author must provide a full postal address, email address, and telephone and fax numbers. *Animal Genetics* – S1M will then create a user name and password, which should be retained for future access to the site. Once the author is logged into the system, the Main Menu will be displayed. Clicking on the Author Centre will bring up instructions for uploading the manuscript and associated files.

Contact data for coauthors must be entered during manuscript submission. It is the responsibility of the corresponding author to inform the coauthors about this. Data for up to 10 authors can be entered during submission, but additional authors listed in the manuscript will also be shown in the paper as published. We encourage obtaining and using ORCID codes as unique personal identifiers (**https://orcid.org/**).

In the cover letter, authors should summarize shortly the message of their study and describe its novelty. In addition, the authors may inform the Editor of conflicts of interest or problems with data availability or any other relevant information. If a previous version of the manuscript was rejected while encouraging resubmission, the cover letter should specify what has been changed relative to the previous version. The journal will not accept for publication manuscripts that describe experiments showing lack of concern for present ethical and welfare standards. The Editor's decision in this regard is final. Authors must comply with ethical guidelines as described in **http://publicationethics.org/resources**. The Editor reserves the right to report gross misconduct to the institutions of the authors.

Manuscripts that are not prepared in accordance with the instructions as detailed below will be returned to the authors. Manuscripts must not be under consideration for publication elsewhere. Revision of the

manuscript must be received by the Editor within three months of his first decision; otherwise the revision will be considered a new submission.Manuscripts must be grammatically and linguistically correct (see below for advice for non-native English speakers) and also conform to the standardized scientific notations and correct style.

- Spelling should conform to the *Concise Oxford English Dictionary* (UK spelling) or *Merriam-Webster's Collegiate Dictionary* (US spelling), and either US or UK spelling must be used consistently throughout.
- Units of measurements, symbols and abbreviations should follow those in *Units, Symbols and Abbreviations* (6^{th}, 2008, Royal Society of Medicine, London) specifying the use of SI (Système International) units.
- Use the symbol × for multiplication or crossing, not the letter x; 37 °C (with space); 95% (without space).
- Number items in the text using (i), (ii), (iii), (iv), etc.
- 1900s, 2000s, etc. (no apostrophe before the 's'); dataset (one word); data are (not: data is); cattle are (not: cattle is);
- DNA sequences should be closed-up (CGCAGAACT, not CGC AGA ACT). See below for nucleotide abbreviations.
- Use *italics* for *et al.*, *P* (value of significance), variable symbols (H_o for observed heterozygosity, F_{ST}, *N* or *n* (number), etc.; note that a suffix is NOT in italics. Do not use italics for e.g. and i.e. (both not followed by a comma) and do not italicize parentheses surrounding italic text ('(*taurus*)', not '*(B. taurus)*'). Also italicize text in another language, such as the Linnaean species names in Latin.
- Use small capitals for m (molar), but not for M (mole); use small capitals for program names (e.g. Admixture), but not for AMOVA or R (computing language).
- All gene names (lower case) and gene symbols (capitals) should be italicized throughout the text, tables and figures. Locus symbols used in *Animal Genetics* publications must be confirmed at the NCBI or Ensembl websites (https://www.ncbi.nlm.nih.gov/gene/ or ensemble.org).
- Use for DNA and protein sequence variants (SNPs, indels, etc.) the Mutnomen nomenclature (http://varnomen.hgvs.org/.)
- 'Only' usually precedes the verb; 'in order to' can often by replaced by 'to'; 'very' can almost always be omitted. 'Since': do not use instead of 'because'; use 'while' only for happening at the same time, not instead of 'although' or 'whereas'.
- Use 'that' (not preceded by a comma) for a defining clause and ', which' for a clause that adds information.
- Define non-standard abbreviations at their first use. RT-PCR = reverse-transcriptase PCR, not real-time PCR. RNAseq or RNA-seq, not RNA-Seq.
- Abbreviations that do not need a definition or reference

1. Common scientific abbreviations

- AD = Anno Domini
- aDNA = ancient DNA
- AFLP = amplified fragment length polymorphism
- AMOVA = analysis of molecular variance
- BAC = bacterial artificial chromosome
- BC, BCE = Before Christian era
- blast = Basic Local Alignment Search Tool (also blastn, blastx). As software names these terms are in small caps.
- BLUP = Best Linear Unbiased Prediction
- BSA = bovine serum albumin
- bp = basepairs (always singular, not bps)
- cM = centiMorgans

- cpm = counts per minute
- CNV = copy number variant
- CRISPR = Clustered Regularly Interspaced Short Palindromic Repeats
- cDNA = copy DNA or complementary DNA
- dd- = dideoxy-
- DNA = deoxyribonucleic acid
- DNase = deoxyribonuclease
- dNTP = deoxynucleotide triphosphate
- ds = double-stranded
- ELISA = enzyme-linked immunosorbent assay
- EST = expressed sequence tag
- FAANG = functional annotation of animal genomes.
- FISH = fluorescence *in situ* hybridization
- FRET = Förster resonance energy transfer
- Gb = gigabases
- GBS = genotyping by sequencing
- GFP = green fluorescent protein
- GO = gene ontology
- GWAS = Genome-wide association study
- h = hour(s)
- HD-SNP = high density SNP (bead array)
- HPLC = high-performance liquid chromatography
- HWE = Hardy-Weinberg equilibrium
- IBD = identity-by-descent
- IBS = identity-by-state
- kb = kilobases
- kDa = kilodalton
- KEGG = Kyoto Encyclopedia of Genes and Genomes
- KO = knock-out (mutant)
- l (lower case) = liter/litre(s)
- LINE = long interspersed element
- LD = linkage disequilibrium
- lncRNA = long non-coding RNA
- LTR = long terminal repeat
- LOD = logarithm of odds
- MAF = minor allele frequency
- MAS = marker-assisted selection
- Mb = megabases
- MDS = multidimensional scaling
- miRNA = microRNA
- MCMC = Markov chain- Monte Carlo (sampling of probability distribution)
- MHC = major histocompatibility complex
- MRCA = most recent common ancestor
- MSY = male-specific part of Y-chromosome
- mtDNA = mitochondrial DNA
- mya =million years ago
- NJ = neighbor-joining
- NRY = non-recombining part (of Y-chromosome)
- nt = nucleotide
- OMIA = on-line Mendelian inheritance in animals
- OMIM = on-line Mendelian inheritance in men

- ORF = open reading frame
- PAGE = polyacrylamide gel electrophoresis
- PAR = pseudo-autosomal region
- PCA = principal component analysis
- PCR = polymerase chain reaction
- qPCR = quantitative PCR
- QTL = quantitative trait loci (or locus); singular or plural (not QTLs)
- RACE = rapid amplification of cDNA ends
- REML = restricted maximum likelihood
- RFLP = restriction fragment length polymorphism
- RNA = ribonucleic acid
- RNase = ribonuclease
- RNA-seq = transcriptome sequencing
- ROH(s) = run(s) of homozygosity
- RT-PCR = reverse transcriptase-PCR
- s = second(s)
- SD = standard deviation
- SDS = sodium dodecyl sulfate
- SSCP = single-strand conformation polymorphism
- SINE = short interspersed element
- SNP(s) = single nucleotide polymorphism(s)
- ss = single-stranded
- SSR = simple sequence repeat
- STS = sequence site
- Tn = transposon
- u = unit
- UTR = untranslated region
- UV = ultraviolet
- VCF = variant call format
- VNTR = variable number of tandem repeats
- WGS = whole-genome sequence
- wt = wild-type
- YBP = years before present
- Zoo-FISH = FISH with heterologous chromosome-specific probes

2. Common chemical abbreviations (EDTA, Tris, etc.), abbreviations for fluorescent compounds (FAM, TET, SYBR, etc.)
3. Abbreviations for restriction enzymes and other enzymes used in experiments (e.g. EcoRI, MspI; Taq polymerase). Do NOT italicize.
4. Common buffers or media (TE, SSC, TAE, TBE, LB, FCS).
5. Nucleotides sequence abbreviations

- A = adenosine
- B = C, G or T
- C = cytidine
- D = A, G or T
- G = guanosine
- H = A, C or T
- I = inosine
- K = G or T
- M = A or C
- N = any nucleotide

- R = A or G (purine)
- S = C or G
- T = thymidine
- U = uridine
- V = A, C or G
- W = A or T
- Y = C or T (pyrimidine)

6. Chromosomal abbreviations followed by a number (e.g. BTA1 = cattle chromosome 1)

- BTA = *Bos taurus* (cattle) chromosome
- CFA = *Canis familiaris* (dog) chromosome
- ECA = *Equus caballus* (horse) chromosome
- GGA = *Gallus gallus* (chicken) chromosome
- HSA = *Homo sapiens* (human) chromosome
- MMU = *Mus musculus* (mouse) chromosome
- OAR = *Ovis aries* (sheep) chromosome
- SSC = *Sus scrofa* (pig) chromosome

7. Institutions/organizations

- EMBL = European Molecular Biology Laboratory
- GenBank = NIH genetic sequence database
- INRA = Institut National de la Recherche Agronomique
- MARC = Meat Animal Research Center
- NCBI = National Center for Biotechnology Information
- USDA = United States Department of Agriculture

Article Preparation Support

Wiley Editing Services offers expert help with English Language Editing, as well as translation, manuscript formatting, figure illustration, figure formatting, and graphical abstract design – so you can submit your manuscript with confidence.

Also, check out our resources for Preparing Your Article for general guidance about writing and preparing your manuscript.

Types of manuscripts

Research Articles should consist of the following sections: title page, Summary, Keywords, Introduction, Materials and methods, Results, Discussion, Acknowledgements, and References, followed by tables and figures with full captions/legends and then legends for Supporting information (with full supporting information provided in a separate document, as described below). Results and Discussion may be combined into one section. Research Articles should normally not exceed 5000 words in length, including all text, references, figures, legends, etc.

Short Communications should consist of the following sections: title page, Summary, Keywords, main text, Acknowledgements, Availability of data, and References, followed by tables and figures with full captions/legends and then legends for supporting information (with full supporting information provided in a separate document, as described below). Short Communications manuscripts must not exceed 1500 words in total and two figures and/or tables. GWAS without follow-up studies of candidate genes and not offering novel insights are usually published as Short Communications.

Mini-reviews are invited on issues of interest to *Animal Genetics* readers. Mini-review manuscripts should include title page, Summary, Keywords, main text with a limited number of optional sections,

Acknowledgments and References, followed by tables and figures with full captions/legends and then captions/legends for supporting information (with full supporting information provided in a separate document, as described below) and would usually not exceed 5000 words in total.

Brief Reports report technical information such as panels of polymorphisms with allele frequencies, clear associations between non-causative polymorphisms and a trait if this offers no further insights, new causative mutations in known causative genes, and software development. Brief Reports should not exceed 500 words in total. Subheadings are adapted to the specific information. References are abbreviated and numbered sequentially in the text and are followed by tables and figures with full captions/legends and then captions/legends for Supporting information (with full supporting information provided in a separate document, as described below).

Brief Reports may also report the availability via a link of curated and annotated datasets (e.g., compilations of high-density SNP genotypes from a wide panel of animals, mitochondrial DNA sequences from different sources, sequence variants of specific genes or group of genes from one or more species with corresponding phenotypes) that would constitute useful reference data for future animal genetic studies.

Please click here if you wish to view an example of a **Research Article**, **Short communication** or **Brief Report**.

Arrangement of the manuscript

Animal Genetics now offers **Free format submission** for a simplified and streamlined submission process.

For the initial submission, the manuscript can be a single file including text, figures, and tables, or separate files – whichever you prefer. All required sections, including abstract and for Research Articles, introduction, methods, results, and discussion, should be contained in your manuscript. Figures and tables should have legends. References may be submitted in any style or format, as long as it is consistent throughout the manuscript.

For revised submissions, the manuscript should be submitted as a single file, prepared in MSWord, containing the sections and material listed above for the type of manuscript submitted (e.g. Research Article, Short Communication, etc.), double-spaced at 12-point size. Highest quality figures (at least 300 dpi) may also be submitted as separate files. Supporting information should be submitted as described below.

- Leave the right-hand margin unjustified
- Turn the hyphenation option off
- Use tabs, not spaces, to separate data in tables or use the table option in Word
- Number the lines of text

Title page

The title page should include a short but informative title, the initials and surnames of each author followed by his or her department, institution, city, postal code and country. The fax and telephone numbers and email address of the corresponding author should also be included. Any change of address should be given in footnotes. A running title should also be included. Acknowledgements of funding and institutions and manuscript reference number should not appear on the title page.

Summary

Research Articles, Short Communications and Mini-reviews should include a summary of not more than 250 words. The summary should summarize the main findings in a single paragraph and be

comprehensible on its own. References should not be used in the summary, and abbreviations should be used sparingly and be explained.

Keywords (3 to 10)

In order to expand word search capabilities, try to avoid terms already in the title.

Introduction

Studies on Mendelian traits should cite the corresponding OMIA code (https://omia.org/home).

Acknowledgements

Acknowledgements should follow the main text but precede the Availability of data section. Within this section, personal acknowledgements should precede those of institutions and agencies.

Availability of data

Availability of data should be specified in a separate section preceding the References. Data that are integral to the article must be made available in such a way as to enable readers to replicate, verify and build upon the conclusions published in the article. Data availability also contributes to visibility and impact. Any restriction on the availability of this data must be disclosed at the time of submission.

Data for which public repositories are widely used (see en.wikipedia.org/wiki/List_of_biological_databases; genome.cshlp.org/site/misc/ifora_weblinks.xhtml) should be deposited in such a repository prior to publication. The appropriate linking details and identifier(s), such as active accession codes, should be obtained before submission.

Nucleotide data (genomes, genes, coding and non-coding RNA) must be submitted and deposited in the DDBJ, EMBL or GenBank databases. Submission to any one of the three collaborating databases is sufficient to ensure data entry in all. Contact information for nucleotide sequence:

DDBJ: ddbj@ddbj.nig.ac.jp; www.ddbj.nig.ac.jp

EMBL: datasubs@ebi.ac.uk; www.ebi.ac.uk

GenBank: info@ncbi.nlm.nih.gov web; www.ncbi.nlm.nih.gov

For special types of submission (e.g. genomes, bulk submission, etc.), additional submission protocols are available from the above websites.

Protein sequences that have been determined by direct amino acid sequencing must be submitted to SWISS-PROT at EMBL-EBI (www.ebi.ac.uk, datasubs@ebi.ac.uk). Protein sequences that are obtained by translation of nucleic acid sequences should not be submitted but are assigned SWISS-PROT accession numbers on incorporation into TrEMBL. Experimental results revealing function, subcellular location, subunit etc. should also be submitted to SWISS-PROT (update@ebi.ac.uk).

SNPs, length variation (microsatellites, indels, CNVs) and rearrangements should be submitted to the European Variation Archive (EMBL-EBI), which accepts files in VCF format.

We recommend https://osf.io/ or www.figshare.com for GWAS genotypic and phenotypic datasets. QTL mapping data for cattle, sheep, pigs, horses and rainbow trout should be submitted to the AnimalQTLdb (http://www.animalgenome.org/cgi-bin/QTLdb/index) in MIQAS format.

References

References should be supplied in a consistent format and will be converted to the following format:

References in the text follow the Harvard system in style [i.e. name followed by date, Brown & Smith (1966); if there are three or more authors, Brown *et al.* 1966]. References at the end of the text are listed in alphabetical order of (first) authors, giving both the full title of the article and full name of the journal, and should appear as follows:

Giger U., Reilly M.P., Asakura T., Baldwin C.J. & Harvey J.W. (1986) Autosomal recessive phosphofructokinase deficiency in English springer spaniel dogs. *Animal Genetics* **17**, 15–23, doi 10.1111/j.1365-2052.1986.tb03184.x.

McKeam T.J. (1980) Method for growing hybridomas in rats or mice. In: *Monoclonal Antibodies* (ed. by R.H. Kennet, T.J. McKearn & K.B. Bechtol), pp. 403–4. Plenum Press, New York.

Personal communications (e.g. K. MacPherson, personal communication) must be authorized in writing by those involved. Unpublished data should be cited in the text as (unpublished data). References to manuscripts submitted, but not yet accepted, should be cited in the text as (M. Bruce & A. Laing, in preparation) and should not be included in the list of references.

We recommend the use of a tool such as EndNote, Reference Manager, Mendeley or Zotero for reference management and formatting.

For Brief Reports, text citations should appear as superscript numbers (in numerical order), and an abbreviated format should be used for references at the end of the text [i.e. a maximum of two authors, no article title and with abbreviated journal titles (e.g. 1 Giger U. *et al.* (1986) *Anim Genet* **17**, 15–23)].

Tables

Tables in the main text should contain only data that are directly relevant for the main conclusions of the article or facilitate the understanding of the text. Other supporting data (e.g. sampling details, primer sequences) can be listed as supporting information (see below). Values listed in tables should not contain insignificant digits after the decimal point (e.g., 123.4567 ± 0.8910 = 123.5 ± 0.9). Use only horizontal lines and check that the layout is adequate with the independent variable in the first column. In the caption, define non-standard abbreviations even if already defined in the text.

Figures

Figures shoud be referred to in the text by Fig. 1, Figs, 1 and 2, Fig. S1. Figures with more than one part should be labelled with lower case letters [e.g. (a), (b), (c)]. Lettering on graphs and other illustrations should be horizontal whenever possible and be of adequate size. Figures should be largely self-explanatory with symbols and colours preferably explained within the figure. Figures showing sequence data will not generally be published unless key features of the sequence need to be highlighted. Please save vector graphics (e.g. line artwork) in Encapsulated Postscript (eps), PowerPoint (pptx) or pdf format and bitmap files (e.g. half-tones) in PDF, JPG or Tagged Image File Format (TIFF) with a resolution of at least 300 dpi. Please consult the instructions for electronic artwork formats and resolutions (digital illustration standards).

In the caption, define non-standard abbreviations even if already defined in the text.

Animal Genetics reproduces colour figures free of charge.

Supporting information

Supporting information is- published as web material at the *Animal Genetics* website and may include, amongst others, primer sequences, specification of samples, methodological aspects, allele frequencies, summary statistics, GWAS-implicated candidate genes, sequence alignments, additional analyses, movie files, audio clips, 3D structures or other multimedia files. Supporting information files should be cited within the text in sequential order (as Table S1, Table S2, etc.; Fig. S1, Fig. S2, etc.; Appendix S1; Appendix S2, etc.), and a short descriptive legend for each supplementary item should be included at the end of the document after the tables and figures. Authors are responsible for the preparation of supporting information, which for text, figures and normal tables should be supplied in a single, separate MSWord document;) with full captions/legends ABOVE each image. Large data files may be supplied as MSExcel documents. Supplementary files will be published as supplied by the authors, and a proof is not made

available prior to publication; thus, authors should provide any Supporting information in the desired final format.

Availability of Materials

It is understood that by publishing an article in *Animal Genetics* the authors agree to make available upon reasonable request to other scientists engaged in academic (non-profit) research any of the nucleic acids, antibodies, cells, etc. that are not available from commercial suppliers and are required to substantiate the scientific conclusions of the article.

Proofs

The corresponding author will receive an email containing a link to a PDF of the proof. Thus, the corresponding author should make sure that the email address on file with Manuscript Central is current and working. Acrobat Reader will be required to read and edit the PDF. This software can be downloaded (free of charge) from http://www.adobe.com/products/acrobat/readstep2.html.

Early View

Animal Genetics is covered by Wiley's Early View service. Early View articles are complete full-text articles published online in advance of their publication. Articles are therefore available as soon as they are ready, rather than after allocation to an issue. Early View articles are complete and final. They have been fully reviewed, revised and edited for publication, and the authors' final corrections have been incorporated. Because they are in final form, no changes can be made after online publication. Early View articles are fully citable and carry an online publication date and DOI for citation. More information about DOIs can be found at: http://www.doi.org/faq.html.

Offprints

Free access to the final PDF offprint of your article will be available via Author Services. Please therefore sign up for Author Services if you would like to access your PDF offprint and enjoy the many other benefits the service offers.

Author Services

Author Services enables authors to track their article – once it has been accepted – through the production process to publication online. Authors can check the status of their article online and choose to receive automated emails at key stages of production. The corresponding author will receive an email with a unique link that enables him/her to register and have the article automatically added to the system. Please ensure that a complete and working email address is provided when submitting the manuscript. Visit this page for more details on online production tracking and for a wealth of resources including FAQs and tips on article preparation, submission and more.

Article Promotion Support

Wiley Editing Services offers professional video, design, and writing services to create shareable video abstracts, infographics, conference posters, lay summaries, and research news stories for your research – so you can help your research get the attention it deserves.

We recommend that authors promote the impact of their paper by using the Article Share option and uploading the email addresses of up to 10 colleagues, who will receive free access to the final version of the paper. As soon as authors have received the email notifying them that their article has been received and the production process is now underway, they should visit https://authorservices.wiley.com/author-resources/Journal-Authors/Promotion/sharing-your-research.html.

Author material archive policy

Please note that unless specifically requested, Wiley Blackwell will dispose of all electronic material submitted two months after publication.

Copyright

If your manuscript is accepted, the author identified as the formal corresponding author for the manuscript will receive an email prompting them to login into Author Services; where via the Wiley Author Licensing Service (WALS) they will be able to complete the license agreement on behalf of all authors on the manuscript.

For authors signing the copyright transfer agreement:

(a) If the open access option is not selected, the corresponding author will be presented with the copyright transfer agreement (CTA) to sign. The terms and conditions of the CTA can be previewed in the samples associated with the Copyright FAQs at CTA Terms and Conditions.

(b) If the open access option is selected, the corresponding author will have a choice of the following Creative Commons License Open Access Agreements (OAA):

- Creative Commons Attribution License OAA
- Creative Commons Attribution NonCommercial License OAA
- Creative Commons Attribution NonCommercial-NoDerivatives License OAA

To preview the terms and conditions of these open access agreements, please visit the Copyright FAQs hosted on Wiley Author Services. For more information regarding Creative Commons License, please visit Copyright & License hosted at Wiley Open Access.

For authors funded by The Wellcome Trust and members of the Research Councils UK (RCUK) or the Austrian Science Fund (FWF):

If you choose open access, you will be given the opportunity to publish your article under a CC-BY license, supporting you in complying with your Funder requirements. For more information on this policy and the Journal's compliant self-archiving policy, please visit: http://www.wiley.com/go/funderstatement and view this video.

Authors who did not select open access when they originally accessed the copyright form via Author Services but who subsequently wish to make **their article open access should see the section open access. Similarly, authors who wish to switch to the open access** selection after their article is published online as Early View should see the section open access.

Note to Contributors on Deposit of Accepted Version

(a) Funder arrangements

Certain funders, including the NIH, members of the Research Councils UK (RCUK) and Wellcome Trust require deposit of the Accepted Version in a repository after an embargo period. Details of funding arrangements are set out at the following website: http://www.wiley.com/go/funderstatement. Please contact the Production Editor if you have additional funding requirements.

(b) Institutions

Wiley has arrangements with certain academic institutions to permit the deposit of the Accepted Version in the institutional repository after an embargo period. Details of such arrangements are set out at the

following website: http://www.wiley.com/go/funderstatement.

Open access

With open access the authors, the author's funding agency, or the author's institution pays a fee to ensure that the article is made available to non-subscribers upon publication via Wiley Online Library, as well as deposited in the funding agency's preferred archive. For the full list of terms and conditions, see http://www.wileyopenaccess.com/view/terms.html.

Upon acceptance, authors wishing to send their article to open access will be required to complete the online payment form at: https://authorservices.wiley.com/bauthor/onlineopen_order.asp.

Wiley's Author Name Change Policy

In cases where authors wish to change their name following publication, Wiley will update and republish the paper and redeliver the updated metadata to indexing services. Our editorial and production teams will use discretion in recognizing that name changes may be of a sensitive and private nature for various reasons including (but not limited to) alignment with gender identity, or as a result of marriage, divorce, or religious conversion. Accordingly, to protect the author's privacy, we will not publish a correction notice to the paper, and we will not notify co-authors of the change. Authors should contact the journal's Editorial Office with their name change request.

Last updated on 19 February 2021