ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12731

Genome-wide differential DNA methylation in tropically adapted Creole cattle and their Iberian ancestors

N. Sevane*, R. Martínez† and M. W. Bruford*‡

*School of Biosciences, Cardiff University, Cathays Park, Cardiff, CF10 3AX, UK. †Corporación Colombiana De Investigación Agropecuaria (Corpoica), Centro de Investigaciones Tibaitatá, km 14 via Bogotá, 250047 Mosquera, Colombia. ‡Sustainable Places Research Institute, Cardiff University, Cardiff, CF10 3BA, UK.

Summary

Enhancing climate resilience and sustainable production for animals in harsh environments are important goals for the livestock industry given the predicted impacts of climate change. Rapid adaptation to extreme climatic conditions has already been imposed on livestock species, including those exported after Columbus's arrival in the Americas. We compared the methylomes of two Creole cattle breeds living in tropical environments with their putative Spanish ancestors to understand the epigenetic mechanisms underlying rapid adaptation of a domestic species to a new and more physiologically challenging environment. Reduced representation bisulfite sequencing was used to assess differences in methylation in Creole and Spanish samples and revealed 334 differentially methylated regions using high stringency parameters (*P*-value <0.01, ≥4 CpGs within a distance of 200 bp, mean methylation difference >25%) annotated to 263 unique features. Gene ontology analysis revealed candidate genes involved in tropical adaptation processes, including genes differentially hyper- or hypomethylated above 80% in Creole samples displaying biological functions related to immune response (*IRF6, PTGDR, FAM19A5, PGLYRP1*), nervous system (*GBX2, NKX2-8, RPGR*), energy management (*BTD*), heat resistance (*CYB561*) and skin and coat attributes (*LGR6*). Our results entail that major environmental changes imposed on Creole cattle have had an impact on their methylomes measurable today, which affects genes implicated in important pathways for adaptation. Although further work is needed, this first characterization of methylation patterns driven by profound environmental change provides a valuable pointer for the identification of biomarkers of resilience for improved cattle performance and welfare under predicted climatic change models.

Keywords *Bos taurus*, Criollo, epigenomics, livestock, reduced-representation bisulfite sequencing

Introduction

Assessment of climate change impacts predicts a progressive upward trend in average temperatures over the coming century with climatic fluctuations that may lead to a simplification of vegetation, a decrease in forage production and quality, and changes in organismal life cycles (Ciscar *et al.* 2014). For animal health, the distribution and extent of parasitic and infectious disease may increase, as natural control via low winter temperatures will be reduced. This increased disease risk and the adverse effects of extreme humidity on health may also affect temperate regions, where rainfall is predicted to increase. Moreover, stress generated by adaptation to changing conditions coupled with temperature increase may compromise immune responses to pathogens and external challenges and may lead to reduction in food intake, growth, milk yield and reproductive efficiency (Hahn 1999), jeopardizing animal welfare. Thus, improving climate resilience and enhancing sustainable production are important goals for the livestock industry. However, classical breeding programs may not provide efficient medium- to long-term strategies equipped to counter the expected pace of climate change. Relying on short-term responses coupled with the ability to convey heritable phenotypic plasticity to future generations (Weyrich *et al.* 2016) could provide a better alternative for facing this imminent challenge.

Events in human history have included episodes when rapid adaptation to extreme climatic conditions has been imposed on a limited number of domestic animals. One

Address for correspondence

N. Sevane, School of Biosciences, Cardiff University, Cathays Park, Cardiff, CF10 3AX, UK.
E-mails: nsevane@ucm.es; sevane-fernandezn@cardiff.ac.uk

Accepted for publication 28 August 2018

example of such an event is Columbus's arrival in the Americas. Livestock species were brought from the Iberian Peninsula to the Americas on Columbus's second journey in 1493 (Rodero *et al.* 1992) and spread throughout the continent, adapting to a wide range of alien environmental conditions and giving rise to 'Creole' animal populations (Rouse 1997). The total number of Iberian cattle brought to the Americas is estimated to have been less than 1000 (Rodero *et al.* 1992). After nearly 300 years of Creole cattle expansion, several other European breeds were introduced and crossed with local populations (Willham 1982) as well as with Indian Zebu cattle, especially in tropical areas (Santiago 1978). Creole cattle were subsequently displaced into marginal, demanding environmental areas where they still exist. Examples include the Costeño con Cuernos and San Martinero breeds of Colombia, which descend from Spanish cattle and also have minor influences from continental and/or zebu breeds (Martínez *et al.* 2012; Ginja *et al.* 2013). These breeds are therefore the product of several centuries of adaptation to new, local and challenging environments. The Costeño con Cuernos breed, which tolerates high temperatures and humidity, was developed in Caribbean Colombia and is found from the swamp areas of Córdoba and Magdalena to the dry savannah of Sucre and Bolívar (Pinzón 1984). The San Martinero was developed in the Colombian Orinoquia region in the 17th century and is almost uniquely adapted to tropical rainforests (Holdrige & Hunter 1961).

A central goal of evolutionary biology, and an increasingly relevant one to agriculture, is to elucidate the genetic architecture of adaptation. The past decade has yielded an increasing number of examples for which regulatory changes have been shown to contribute to species-specific adaptations and to reproductive isolation (Blekhman *et al.* 2008). There is mounting evidence that heritable variation in relevant traits can be generated through a suite of epigenetic mechanisms, even in the absence of genetic variation, which eventually might promote permanent changes in DNA sequence (Varriale 2014; Fagny *et al.* 2015). Among epigenetic mechanisms, DNA methylation via 5-methylcytosine is a key modification in vertebrate genomes that imparts an additional layer of heritable regulatory information upon DNA and is essential for viability in a myriad of biological processes (Lister & Ecker 2009).

Epigenomic studies in cattle include muscle and placental tissues analysed with non-base-resolution methods (Huang *et al.* 2014; Su *et al.* 2014) and two recent studies using base-resolution techniques, a low-coverage whole-genome bisulfite sequencing analysis of bovine placenta (Schroeder *et al.* 2015) and reduced-representation bisulfite sequencing (RRBS) of 10 bovine tissues, including blood (Zhou *et al.* 2016). However, in these studies only DNA methylome landscapes were described and environmental influences on phenotypic variation were not explored. Thus, despite increasing knowledge about the genes involved in bovine adaptation to tropical climate (Gautier *et al.* 2009; Chan *et al.* 2010; Porto-Neto *et al.* 2014; Makina *et al.* 2015; Wang *et al.* 2016; Pitt *et al.* 2018), we lack understanding of relevant epigenetic function (see Varriale 2014 for a review).

In this study therefore, we aimed to address the role of epigenetic regulation on tropical adaptation in cattle by comparing the methylomes of modern tropical Creole bovine breeds with modern day samples from breeds, including their putative Iberian ancestors. Although the number of samples analysed was relatively small, as in many similar studies (e.g. Korkmaz & Kerr 2017; Semik *et al.* 2017), we included five different breeds to establish epigenomic differentiation among groups, accounting for breed similarities related to their geographical location [i.e. the Iberian Peninsula (three breeds) and Colombia (two breeds)], and used high stringency parameters to detect significant differentially methylated regions (DMRs). We generated a genome-wide map of DNA methylation at single-nucleotide resolution in cattle that provides, apart from the inherent advance in knowledge about the bovine epigenome, insights into the biology and evolution of a species under profound climate change and a base for future climate-related research in cattle.

Materials and methods

Samples and DNA extraction

Five New and Old World cattle breeds were analysed in this study. The sample comprised the Colombian Creole cattle Costeño con Cuernos ($n = 2$) and San Martinero ($n = 1$) breeds and Iberian cattle representing the main ancestors of these Creole populations including the Asturiana de los Valles ($n = 1$), Lidia ($n = 1$) and Retinta ($n = 1$) breeds. Samples were collected from adult males between 7 and 11 years old. Animals were reared in their native environment under extensive conditions with access to characteristic available local vegetation, growing under the particular climatic and dietary conditions that gave rise to the different breed adaptations (Table 1). DNA was extracted from blood samples using the UltraClean Blood-Spin DNA Isolation Kit (MO BIO Laboratories, Inc.) for the Creole samples and the QIAamp DNA Blood Mini Kit (Qiagen) for the Spanish samples. The concentration and quality of genomic DNA were evaluated using the Qubit dsDNA HS Assay Kit (Life Technologies).

Reduced representation bisulfite sequencing

Genomic DNA (0.5–1.0 μg) from each sample was restricted with the MspI enzyme (New England Biosciences), cleaned using DNA Clean and concentrator-25 columns (Zymo Research) and eluted in 60 μl for library preparation. The

Table 1 Geographic and climatic conditions of Creole and Spanish breeds.

Breed	Location	MASL	MAT (°C)	MARH (%)	MAR (mm)
Costeño con Cuernos	Department of Córdoba (Sinú River Valley, Colombia)	300	30	80	2500
San Martinero	Department of Meta (Colombia)	700	21	70	1800
Asturiana de los Valles	Mieres (Asturias, Spain)	380	11	80	1000
Lidia (Casta Navarra)	Igúzquiza (Navarra, Spain)	450	12	67	600
Retinta	Tierra de Barros (Badajoz, Spain)	400	17	66	450

MASL, metres above sea level; MAT, mean annual temperature; MARH, mean annual relative humidity; MAR, mean annual rainfall.

sticky ends produced by MspI digestion were filled with CG nucleotides and Illumina sequencing adapters. The TruSeq Nano DNA LT Library Prep Kit (Illumina) was used for 3′ adenylation and adapter ligation. The end-repaired samples were purified using 2.5× AMpure XP Beads (Beckman Coulter) and eluted in 20 μl of resuspension buffer. After adapter ligation, samples were again purified using 1.0× AMpure XP Beads and eluted in 40 μl of resuspension buffer. Size selection of DNA fragments (~175–225 bp) was performed using a 2% agarose gel (Invitrogen), and the selected fragments were purified twice using 1.0× AMpure XP Beads and finally eluted in 22 μl of resuspension buffer. Bisulfite conversion of non-methylated cytosines was performed on 20 μl of size-selected fragments using the EZ DNA Methylation-Lightning Kit (Zymo Research). PCR (20 cycles) was performed to enrich the sequencing library using a TruSeq Nano DNA LT Library Prep Kit (Illumina). The Pfu Turbo Cx Hotstart DNA polymerase (Agilent Technologies) and 10 mM dNTP mix (Life Technologies) were used for PCR reactions. After enrichment, the library was purified twice using 1× AMpure XP Beads (Beckman Coulter) and quantified using the Qubit dsDNA HS Assay Kit (Life Technologies). The average library size was determined using an Agilent 2100 Bioanalyzer (Agilent Technologies). The libraries were then pooled in equimolar ratios of 2 nM, and 6.0 pM of the pool was clustered and sample tracked using the cBot (Illumina) and sequenced following a 2 × 150-bp protocol for 300 cycles using the HiSeq 2500 system (Illumina).

RRBS data analysis and genome-wide DNA methylation levels

Quality assessment and control were performed using TRIM GALORE software (Babraham Bioinformatics, UK). For adapter trimming, the minimum required adapter overlap was 1 bp. To remove potential methylation-biased bases from the MspI digestion end-repair reaction, RRBS reads were trimmed an additional 2 bp when adapter contamination was detected and by 2 bp at the start when the read started with CAA or CGA. Trimming was performed on all reads using a minimum Phred quality score of 20. Sequences were mapped with single end mapping to bovine genome assembly UMD3.1.1 using BISMARK (Krueger & Andrews 2011). Following optimisation, a seed length of 20 bp was chosen and only one mismatch was allowed. The minimum alignment score function was set at L,0,−0.6. Only the reads that were aligned to a unique region in the genome were used for further analysis.

For CpG level comparison, percentage of methylation of individual CpGs was calculated using the METHYLKIT package in R (Akalin *et al.* 2012) and the coverage files from the BISMARK aligner. To prevent PCR bias and increase the power of the statistical tests, we discarded bases with high (above 99.9th percentile of coverage in each sample) and low (below 10× coverage, CpG_{10}) read coverage. Each sequenced and filtered CpG_{10} site was assigned a percentage of methylation score. The CpG_{10} bisulfite conversion rate was calculated as the number of thymines (non-methylated cytosines) divided by coverage for each non-CpG cytosine, as implemented in METHYLKIT. Coverage and correlation plots were also generated by METHYLKIT. The pattern of methylation around different components of the cattle genome, including gene bodies [defined as the region from transcription start site (TSS) to transcription termination site (TTS)], TSSs, TTSs and CpG islands (CpGI), was also investigated using SEQMONK software (Babraham Bioinformatics) from 20 kb upstream to 20 kb downstream. CpG_{10} were annotated with the closest/overlapping TSS (±100 kb) (Miele & Dekker 2008; Sanyal *et al.* 2012) and CpGI using the IDENTGENLOC program from the DMAP package (Stockwell *et al.* 2014). Promoters were defined as −0 to −2 kb of TSS, and CpGI shores and shelves as ±0–2 and ±2–4 kb flanking regions of CpGI respectively.

Differentially methylated region analysis

Differentially methylated regions were established among Creole and Spanish groups to account for breed similarities related to their geographical location, that is, the Iberian Peninsula (three breeds, three samples = three biological replicates within the same group) and Colombia (two breeds, three samples = three biological replicates within the same group). To compare spatially contiguous stretches of methylated cytosines across the Creole and Spanish genomes, DMRs were determined using the R package DSS (dispersion shrinkage for sequencing data) (Feng *et al.* 2014), which outperforms other methods when sample size per group is small owing to the adoption of the Wald test with shrinkage for determining differentially methylated cytosines (Zhang *et al.* 2016). We identified DMRs using the coverage files from BISMARK and the CALLDMR

function with a *P*-value threshold of 0.01, delta equal to 0.1 and otherwise default parameters. To be considered significant, a DMR was required to contain at least three CpG sites (default parameter, although the smallest significant DMR included four CpGs) within a distance of 200 bp and with an absolute mean methylation difference greater than 25% when comparing Creole and Spanish samples (Akalin *et al.* 2012). As CpG_{10}, DMRs were annotated with the closest/overlapping TSS (±100 kb) and CpGI using the IDENTGENLOC program from the DMAP package.

Gene ontology analysis

Annotated DMRs were subjected to Gene Ontology (GO) enrichment using the PANTHER v.10 web resource (Mi *et al.* 2016). This GO classification system was used to assign a putative function to each gene by way of biological process, molecular function and cellular components. The Database for Annotation, Visualization and Integrated Discovery (DAVID) v6.8 (Huang *et al.* 2009) was used to determine processes of major biological significance through the FUNCTIONAL ANNOTATION CLUSTER (FAC) tool based on the GO annotation function. High stringency ease score parameters were selected to obtain confident enrichment scores. KEGG pathway analyses were performed using both DAVID and the WEBGESTALT overrepresentation enrichment analysis (ORA) (Wang *et al.* 2013) to map clusters of genes involved in common pathways and processes.

Validation of RRBS data with HiSeq bisulfite sequencing PCR

We performed validation of RRBS data with HiSeq bisulfite sequencing PCR (HiSeq-BSP) for three DMRs annotated to immune, cancer and nervous system genes, displaying both hyper- and hypomethylation patterns (Table S1). The initial concentration of genomic DNA was measured using the Qubit dsDNA HS Assay Kit (Life Technologies). The samples were then diluted accordingly to achieve the recommended DNA input of 500 ng at a concentration of 25 ng/μl for bisulfite treatment. The samples were bisulfite-treated using the EZ DNA Methylation-Lightning Kit (Zymo Research). The treated DNA was PCR-amplified using specific primers for BSP designed and validated by Zymo Research (Table S1). The amplified products for these three assays were pooled together for each sample, and sequencing libraries were made by using the TruSeq Nano DNA LT Library Prep Kit (Illumina). Following the library preparation, the final concentration of the library was measured using the Qubit dsDNA HS Assay Kit (Life Technologies). The libraries were diluted to 12 pM and were sequenced by using the 600 Cycles v3 Reagent Kit (Illumina) on the MiSeq (Illumina) on a 150-base paired-end run. Sequence reads were trimmed, aligned and analysed as described above.

Results

Assessment of RRBS data and genome-wide DNA methylation levels

Fragmentation with the restriction enzyme MspI of blood-extracted DNA from three Creole and three Spanish samples resulted in high quality sequencing RRBS libraries enriched for high CG regions. Illumina HiSeq 2500 sequencing generated between 15 and 38 million reads per sample (accession no. GSE101796) and a total of 136 million reads (Table S2). Quality control analysis using TRIM GALORE and METHYLKIT indicated that the 150-bp sequences displayed the expected nucleotide composition based on MspI digestion and bisulfite conversion (98% average sodium bisulfite conversion efficiency). On average, 98.2% of reads passed the filtering process (Table S2). The mean percentage of mapped reads was 85%, with 33–61% of reads mapping to multiple locations of the genome and 29–43% mapping uniquely (Table S2). Sequences that did not map, or did not map uniquely, were excluded from the analysis.

After alignment, we filtered the CpG dinucleotides based on a coverage of 10 or more reads (CpG_{10}). The number of CpG_{10} per sample ranged from 0.4 to 1.6 million, and the mean coverage ranged from 33 to 106 (Table S3). Of these sites, 20 234 were present in all six samples (Table S4). We observed high positive correlations between all the samples analysed (mean Pearson's correlation coefficient = 0.8), although clear variation was present between them (Fig. S1). The distribution of sequence read coverage of CpG_{10} per sample is shown in Fig. S2 and highlights that, despite the observation that the filtered CpG_{10} displayed high mean coverage, the libraries did not suffer from bias due to excessive amplification of a subset of fragments, as reflected in the absence of peaks on the right-hand side of each histogram. The RRBS protocol has been shown to enrich for CpGIs and, as CpGIs have been universally reported to be regions of gene regulation via methylcytosine and are generally demethylated, the percentage of methylation of CpGs in RRBS libraries is expected to be lower than the average methylation of the genome (~80%). Accordingly, global CpG_{10} methylation ranged from 51% to 57% across samples (Table S3). The distribution of methylation at each CpG_{10} site revealed a bimodal pattern, with heavy methylation (>95%) of 39–53% CpG_{10} and completely unmethylated bases (<5%) ranging between 35% and 47% (Fig. S3). The median methylation was high (84%) (Table S3), reflecting the heavy hypermethylation of 48% of CpG_{10} sites. However, hypomethylated CpG_{10} sites were also evident, including 42% of the analysed CpG sites (Fig. S3). The percentage of CHG and CHH methylation was low in cattle blood, ranging from 0.9% to 1.6% (Table S3).

Although some gaps existed, RRBS reads were detected in most chromosomal regions (chromosomes 1–29 and X)

in each group (Fig. 1). This even-read distribution indicates that cattle blood methylomes can be detected by RRBS technology with good representation, thereby ensuring accurate examination of variation in DNA methylation. The distribution of CpG_{10} related to CpGIs (6235; 31%) revealed that RRBS data are highly enriched in CpGI cores (85%), whereas only a small amount is in CpGI shores (11%) and shelves (4%) (Table S4). The distribution of CpG_{10} in relation to genes (4389; 22%) showed that almost 85% mapped to gene bodies (92% located in introns and 8% in exons) and a much smaller percentage mapped to gene promoters (15%), with the main amount located in intergenic regions (15 845; 78%) (Table S4).

DNA methylation levels sharply decreased in the 2-kb region upstream of TSSs and dropped to the lowest point before TSSs (Fig. 2a), corresponding with the distribution of gene promoters, usually prone to transcription, whereas levels dramatically increased in the 3′ direction, peaking 5′ to the TTS (Fig. 2a), related to the methylation of gene bodies contributing to chromatin structure stability and the regulation of gene expression (Bird 2002). The level dropped slightly and plateaued after TTS (Fig. 2a). As expected, the level of methylation in CpGIs was lower than outside CpGIs (Fig. 2b).

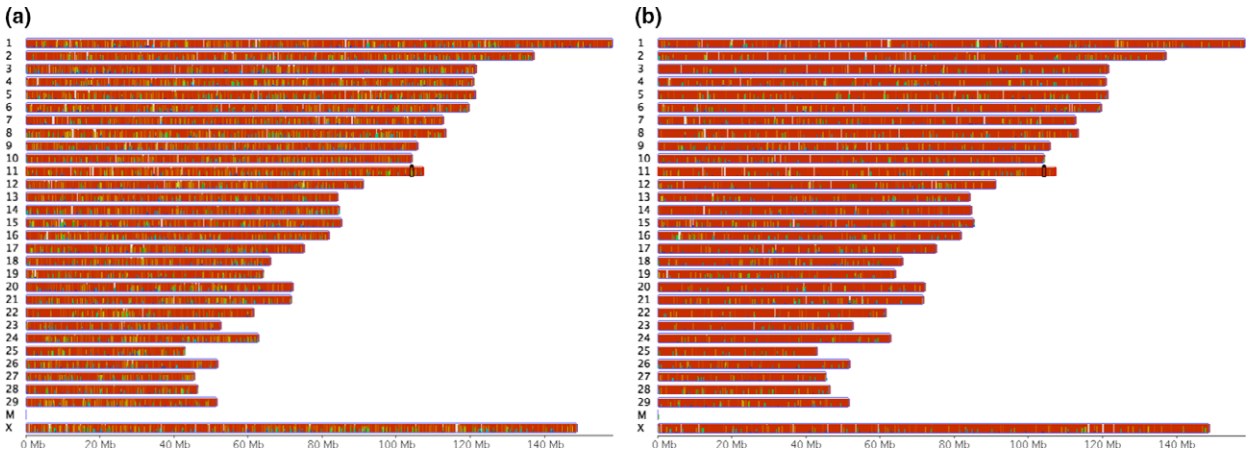

Figure 1 Chromosomal distribution of reads in the (a) Creole and (b) Spanish grouped samples. The distribution of reads is shown in a gradient from blue (low) to red (high).

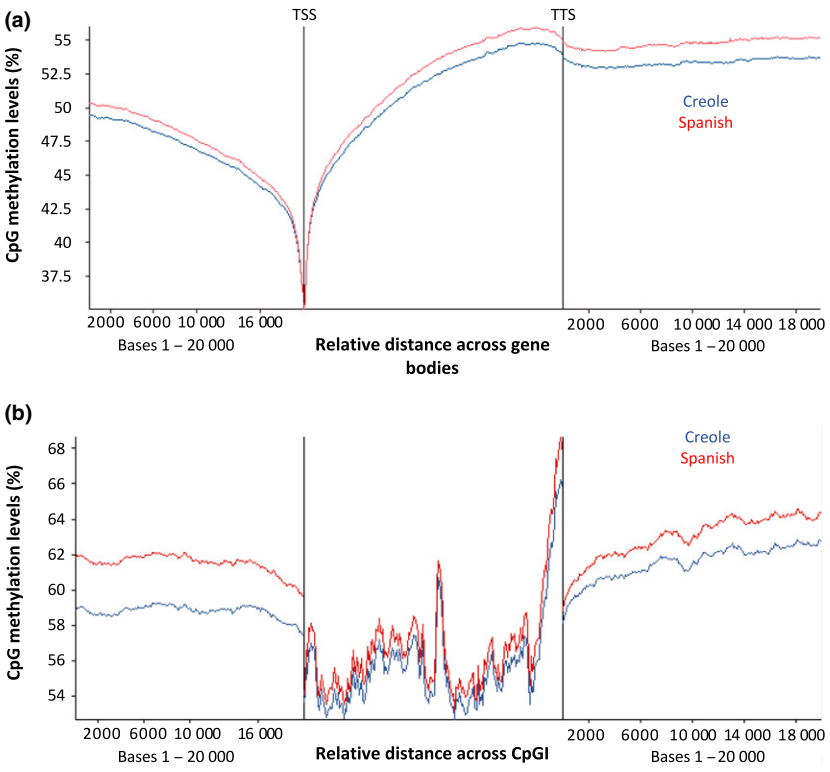

Figure 2 DNA methylation levels in relation to (a) gene bodies and (b) CpG islands. Gene bodies were defined as the region from the transcription start site to transcription termination site.

Table 2 Differentially methylated regions (DMRs) overlapping a gene or CpGI and showing hyper- and hypomethylated levels above 80% in Creole samples.

Symbol	Gene name[1]	Gene overlap	CpGI relation	Meth diff[2]
Hypermethylated in Creole samples				
GBX2	*Gastrulation brain homeobox 2* (E1BJ47)	–	CpGI core	−0.84
LATS2	*Large tumor suppressor kinase 2*	On intron	–	−0.83
BRAT1	*BRCA1 associated ATM activator 1*	–	CpGI shelf	−0.83
BTD	*Biotinidase* (F1MJM4)	Intron–exon boundary	–	−0.82
BLM	*Bloom syndrome RecQ like helicase*	On intron	–	−0.82
NKX2-8	*NK2 homeobox 8* (E1BAC5)	On exon	CpGI core	−0.81
IRF6	*Interferon regulatory factor 6*	On intron	CpGI core	−0.81
PTGDR	*Prostaglandin D2 receptor* (PD2R)	On exon	CpGI core	−0.81
TP53I11	*Tumor protein p53 inducible protein 11* (PIG11)	On intron	–	−0.81
TM4SF5	*Transmembrane 4 L six family member 5* (T4S5)	Intron–exon boundary	–	−0.81
ATP13A3	*ATPase 13A3* (E1BG26)	–	CpGI core	−0.81
TRIM25	*Tripartite motif containing 25 (A6QLA8)*	On intron	CpGI shelf	−0.80
CYB561	*Cytochrome b-561* (CY561)	–	CpGI core	−0.80
FAM19A5	*Family with sequence similarity 19 member A5, C-C motif chemokine like* (F19A5)	On exon	–	−0.80
RPGR	*Retinitis pigmentosa GTPase regulator*	–	CpGI core	−0.80
SNX13	*Sorting nexin 13*	On intron	–	−0.80
Hypomethylated in Creole samples				
TNRC18	*Trinucleotide repeat containing 18*	On intron	–	0.80
PAPLN	*Papilin, proteoglycan like sulfated glycoprotein*	Exon–intron boundary	–	0.80
LGR6	*Leucine rich repeat containing G protein-coupled receptor 6* (LOC100336662)	On intron	–	0.81
PGLYRP1	*Peptidoglycan recognition protein 1* (PGRP1)	On exon	CpGI core	0.84

[1]UniProtKB/TrEMBL record displayed within parentheses.
[2]Methylation differences averaged from all CpG sites within the defined region. Negative differential methylation values indicate hypermethylation in Creole samples; positive differential methylation values indicate hypomethylation in Creole samples.

Differential methylation between Creole and Spanish cattle samples

Comparison of spatially contiguous stretches of differentially methylated cytosines between Creole and Spanish samples revealed 334 DMRs (*P*-value <0.01, ≥4 CpGs within a distance of 200 bp, mean methylation difference >25%; Table S5). Annotation of these DMRs showed that 275 sites (82%), corresponding to 263 unique features, were overlapping a gene or within a distance of ±100 kb from the closest TSS. Approximately 37% of DMRs overlapped a gene, whereas ~4% were in regions 2 kb upstream of a TSS or promoters. Intragenic DMR were equally divided between introns (52%) and exons (48%). Around 36% of DMRs were located in CpGIs, mainly in CpGI cores (81%), whereas only 12% and 7% overlapped CpGIs shores and shelves respectively. Interestingly, a high proportion of DMRs (71%) displayed hypermethylation in Creole samples. The DMRs overlapping a gene or CpGI hyper- and hypomethylated above 80% in Creole samples are shown in Table 2.

Gene ontology analysis

Among the 263 differentially methylated unique annotated features, functional data for 213 genes were obtained with PANTHER, including the GO classes Molecular Function (the primary activities of gene products at the molecular level), Biological Process (sets of molecular events or operations with a defined beginning and end) and Cellular Component (Fig. S4). The annotated DMRs were then analysed using DAVID and WEBGESTALT tools. DAVID FAC analysis produced 16 enriched functional clusters under high stringency conditions for 115 DAVID IDs (Table S6). Among these enriched functional clusters, homeobox, epidermal growth factor (two clusters) and immunoglobulin (two clusters) were identified. We analysed the distribution of annotated DMRs along the cattle chromosomes, confirming one enriched genomic region on chromosome 21 that comprised 12 genes related to the cellular component membrane (Table S6). KEGG pathway analysis retrieved a total of 14 pathways (Table 3): 10 from WEBGESTALT, including immune-related processes such as leukocyte, T cell and lymphocyte differentiation and activation, and circulatory system development or cell proliferation (Fig. 3) and four from the DAVID tool, including acute myeloid leukemia, insulin signalling pathway, Rap1 signaling pathway and microRNAs in cancer (Fig. S5).

Validation of RRBS data with HiSeq-BSP

We used HiSeq-BSP to assess the methylation patterns of three gene-annotated DMRs, including regions displaying high and low differential methylation levels between Creole and Spanish samples and implicated in immune

Table 3 KEGG pathway enrichment analysis of differentially methylated genes between Creole and Spanish cattle samples using WEBGESTALT and DAVID tools.

Gene ontology (GO) term	Description	Gene count	Enrichment score	Genes
GO:1902105	Regulation of leukocyte differentiation	7	5717	*NRARP, LIF, PGLYRP1, PRKCZ, RUNX1, NKAP, CD83*
GO:0050863	Regulation of T cell activation	7	5449	*NRARP, CD5, PRKCZ, MAD1L1, LMO1, NKAP, CD83*
GO:1903037	Regulation of leukocyte cell-cell adhesion	7	5128	*NRARP, CD5, PRKCZ, MAD1L1, LMO1, NKAP, CD83*
GO:0051249	Regulation of lymphocyte activation	8	4689	*NRARP, CD5, PGLYRP1, PRKCZ, MAD1L1, LMO1, NKAP, CD83*
GO:0022407	Regulation of cell-cell adhesion	8	4634	*NRARP, CD5, PRKCZ, ALOX12, MAD1L1, LMO1, NKAP, CD83*
GO:0050865	Regulation of cell activation	10	4837	*NRARP, CD5, PGLYRP1, PRKCZ, ALOX12, PDGFA, MAD1L1, LMO1, NKAP, CD83*
GO:0072359	Circulatory system development	14	3013	*NRARP, LIF, ALOX12, PDGFA, GBX2, MAP2K2, SMAD6, CYP1B1, EOMES, BAK1, SHOX2, FLRT2, ADM2, BCOR*
GO:2000026	Regulation of multicellular organismal development	19	2334	*NRARP, LIF, PGLYRP1, PRKCZ, ALOX12, PDGFA, MAP2K2, CYP1B1, EOMES, SHOX2, RUNX1, RFX4, BHLHE23, FLRT2, NKAP, CD83, ADM2, PHOX2B, BCOR*
GO:0008283	Cell proliferation	21	2341	*NRARP, LIF, NPR3, ALOX12, PDGFA, GBX2, MAP2K2, SMAD6, CYP1B1, P3H2, BAK1, MAD1L1, DAGLA, SHOX2, LMO1, LTBP3, NKX2-8, MAB21L2, NKAP, IRF6, PHOX2B*
GO:0048513	Animal organ development	28	1915	*NRARP, LIF, PGLYRP1, PRKCZ, ALOX12, PDGFA, GBX2, MAP2K2, SMAD6, CYP1B1, EOMES, BAK1, MAD1L1, SHOX2, IFITM5, HOXB1, RUNX1, RFX4, NKX2-8, BHLHE23, MAB21L2, FLRT2, NKAP, TMEM14C, IRF6, CD83, PHOX2B, BCOR*
bta05221	Acute myeloid leukemia	3	6448	*MAP2K2, KIT, RUNX1*
bta04910	Insulin signalling pathway	4	3540	*PRKCZ, SOCS2, MAP2K2, PRKAR1B*
bta04015	Rap1 signalling pathway	5	2812	*PRKCZ, MAPK12, PDGFA, MAP2K2, KIT*
bta05206	MicroRNAs in cancer	7	3397	*DNMT3A, CYP24A1, BAK1, CYP1B1, PDGFA, MAP2K2, PAK4*

(*SERPINB1*), cancer (*SHOX2*) and nervous system (*NRXN2*) processes. The HiSeq-BSP methylation results were significant for the three amplified regions (*P*-value <0.01, ≥4 CpGs, mean methylation difference ≥10%) and concordant with the methylation profiles obtained with the RRBS analysis (Table S7).

Discussion

Studies on adaptation are key to disentangling the evolutionary potential of organisms in response to biotic and abiotic stress and other environmental challenges, which could potentially be highly relevant in the context of global climate change. Tropical environments are characterized by high temperature and humidity, episodes of feed and water scarcity and virulent tropical diseases and parasitic infections. Creole cattle demonstrate greater resistance to such conditions, surviving, breeding and producing efficiently in the tropics (Hernández-Cerón *et al.* 2004; Martínez *et al.* 2008). Two tropically adapted Creole breeds and their likely Spanish ancestors were analysed to establish epigenomic differences among groups accounting for breed similarities related to their geographical location. The Costeño con Cuernos and San Martinero breeds have been developed under physiologically challenging tropical conditions. The Iberian breed Retinta is distributed throughout central and southern Iberia, which is characterized by a xeric climate. The Asturiana de los Valles breed reflects the northern Iberian gene pool and is exposed to a milder climate, mostly cold and damp. The Lidia breed (Spanish fighting bull) has not been selected for productivity traits and thus may be the most representative modern descendent of Iberian cattle herds present back in the 15th century. We detected 334 highly significant DMRs between the groups. The methylation profiles obtained were consistent with previous studies (e.g. Zhou *et al.* 2016). High stringency parameters to detect DMRs (*P*-value <0.01, ≥4 CpGs within a distance of 200 pb, mean methylation difference >25%) when compared with other studies (e.g. Gao *et al.* 2014; Shankar *et al.* 2015; Baerwald *et al.* 2016; Day *et al.* 2016) were taken as statistically significant to overcome the relatively small number of biological replicates characteristic of many epigenomic experiments (e.g. Miele & Dekker 2008; Zhou *et al.* 2016; Semik *et al.* 2017).

In concordance with previous studies on bovine adaptation to tropical climates, including both taurine and indicine (Gautier *et al.* 2009; Chan *et al.* 2010; Porto-Neto *et al.*

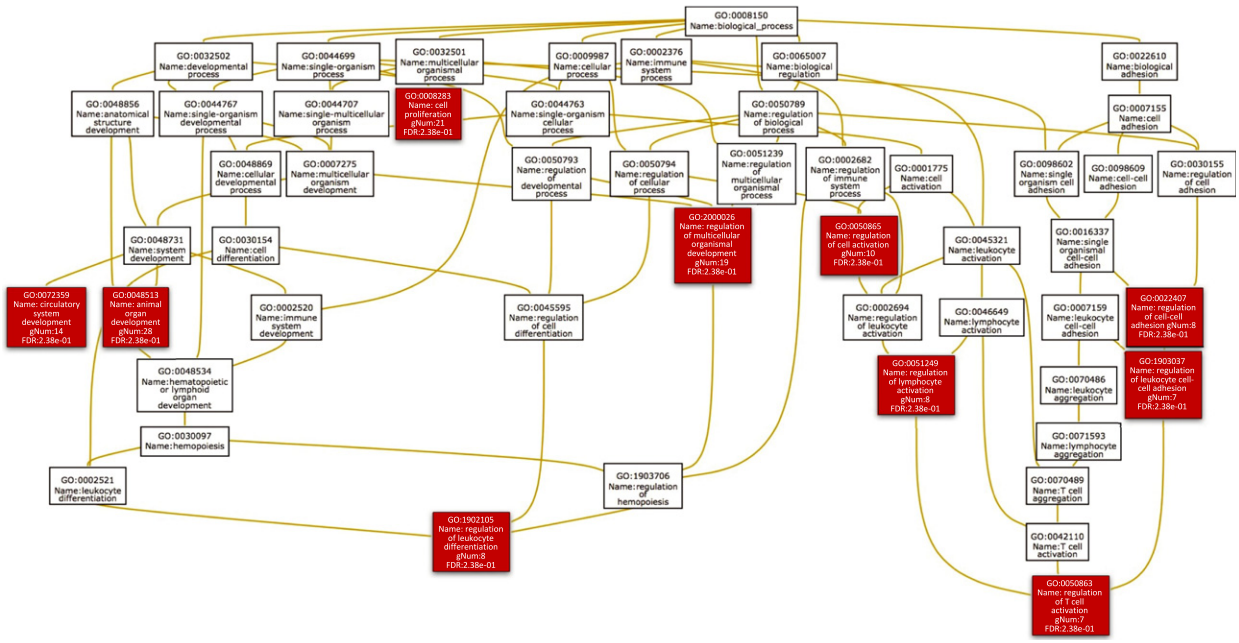

Figure 3 KEGG pathway overrepresentation enrichment analysis of differentially methylated genes between Creole and Spanish cattle samples, performed with WEBGESTALT.

2014; Makina *et al.* 2015; Wang *et al.* 2016; Pitt *et al.* 2018), we found a number of differentially methylated genes between Creole and Spanish groups implicated in several biological processes, such as immunity, nervous system processes, energy management, heat resistance and skin and coat attributes, which are key for survival in harsh environments (Table S5).

Tropical cattle carry lower burdens of ticks, and they have enhanced disease resistance and superior innate immunity, which is reflected in a higher number of genes under selection related to the immune system in studies on adaptation to tropical conditions (e.g. Amorim *et al.* 2015; Liu *et al.* 2018; Pitt *et al.* 2018). Some of the genes hypermethylated above 80% in Creole samples are implicated in immune processes (Table 2): (i) *IRF6*, which is involved in inflammatory responses, macrophage activation and dysregulation of metabolic and immunologic homeostasis (Li *et al.* 2017); (ii) *PTGDR*, which plays an important role in the immune response found in allergic diseases, apart from facilitating smooth muscle relaxation and vasodilatation, inhibiting platelet aggregation and contributing to the regulation of pain perception and sleep (Pettipher *et al.* 2007); and (iii) *FAM19A5*, a brain-specific chemokine or neurokine that acts as regulator of immune and nervous cells (Tom Tang *et al.* 2004). On the contrary, a DMR was found hypomethylated above 80% in Creole samples in a CpGI core located within an exon of the *PGLYRP1* gene. The protein encoded by this gene has been reported to interact with microbes to maintain intestinal homeostasis (Seabury *et al.* 2010) and has been associated with resistance to *Mycobacterium avium* ssp. paratuberculosis (Pant *et al.* 2011), both in cattle. This gene is also associated with several health, reproduction and body conformation traits in Holstein cows (Cole *et al.* 2011). The ability to cope with parasitic and infectious diseases in the adaptation to new environments also seems relevant at the multi-genic level given the high enrichment of pathways such as regulation of leukocyte differentiation, T cell activation, leukocyte cell–cell adhesion or lymphocyte activation (Fig. 3, Table 3) and the presence of two enriched functional clusters related to immunoglobulins (Table S6). These findings are also in agreement with the work of Fagny *et al.* (2015), who describe the existence of epigenetic variability on immune processes implicated in the adaptation to changes in habitat and lifestyle in humans.

Nervous system processes, including changes in behaviour, circadian clock, olfactory and eye function or chemosensory perception, are key for animals adapting to new light, food, reproduction or predatory conditions. Genes with roles in nervous system processes hypermethylated above 80% in Creole samples include (Table 2): (i) *GBX2*, a modulator of thalamus cells development (Mallika *et al.* 2015); (ii) *NKX2-8*, a regional homeobox gene with functions in neuronal development (Safra *et al.* 2013) as well as in tumour suppression; and (iii) *FAM19A5* (see above). Another gene showing the same methylation pattern and implicated in eye function is *RPGR*. The protein encoded by this gene localizes to the outer of rod photoreceptors and is crucial for their viability; its deficiency causes X-linked retinitis pigmentosa (Lyraki *et al.* 2016).

The efficient management of energy storage and mobilization, during wet and dry seasons respectively, provides a

greater ability to tolerate poor feed in harsh environments (Amorim *et al.* 2015). Methylation differences in the insulin signalling pathway (Table 3) or in genes such as *DAGLA* (*diacylglycerol lipase alpha*), *FADS2* (*fatty acid desaturase 2*) or *LMF1* (*lipase maturation factor 1*) (Table S5) may determine variations in energy metabolism. A gene hypermethylated above 80% in Creole samples is *BTD*, which encodes a protein that catalyses the recycling of biotin from biocytin (Wolf 2012) (Table 2). Biotin is a member of the B Vitamin group and is an essential nutrient in the formation of keratin as well as for gluconeogenesis, lipogenesis and protein synthesis. Biotin treatments have been reported to have beneficial effects on milk production, hoof health and reproduction traits (Wilde 2006; Lean & Rabiee 2011).

Genes involved in cardiovascular physiology can facilitate heat resistance in tropical climates. We found a mean hypermethylation level above 80% in Creole samples for the *PTGDR* gene, which facilitates smooth muscle relaxation and vasodilatation (Pettipher *et al.* 2007), and the *CYB561* gene, influencing cardiovascular responses to sympathetic activation (Fung *et al.* 2008) as well as the high enrichment of the circulatory system development pathway (Tables 2 & 3, Fig. 3). Skin and coat attributes are also important for adaptation to harsh conditions, with a direct influence on thermo-resistance to tropical conditions. The *IRF6* gene, also implicated in immune homeostasis, promotes epithelial cell proliferation and differentiation (Richardson *et al.* 2006) and was hypermethylated above 80% in the Creole group (Table 2). On the contrary, the *LGR6* gene, which establishes sebaceous glands and interfollicular epidermis postnatally (Snippert *et al.* 2010), showed hypomethylation above 80% in Creole samples. In addition, two enriched functional clusters related to epidermal growth factor were identified with the DAVID FAC analysis (Table S6). Four microRNAs, which may play important roles in the modulation of gene expression (Su *et al.* 2011), were also differentially methylated between Creole and Spanish cattle groups.

Some cancers, especially in young animals, might be a by-product of novel adaptation and have their origins in recent evolutionary changes in morphology and life history (Leroi *et al.* 2003). Concordantly, rapid bouts of evolution, such as artificial selection in domestic species, have been shown to make animals prone to different cancers (Leroi *et al.* 2003). Epigenetic changes, especially DNA methylation, alter signal-transduction pathways during the early stages of tumour development. Tumour cells, as opposed to normal cells, show local hypermethylation of some CpGI combined with global genome demethylation (Bernstein *et al.* 2007). Taking into account that RRBS enriches for GC-rich regions such as CpGIs (Laird 2010), the high proportion of hypermethylated DMRs in Creole samples (71%), along with the high number of genes differentially hyper- or hypomethylated above 80% in these samples and related to oncogenic processes (in particular, *NKX2-8; LATS2; BRAT1; BLM; TP53I11,* also known as *PIG11*; *TM4SF5; TRIM25;* and *LGR6*) (Table 2), as well as the high enrichment of several pathways implicated in cancer (acute myeloid leukemia, Rap1 signaling pathway and microRNAs in cancer) (Table 3, Fig. S5), might reflect an on-going adaptation process to tropical conditions in descendants from the cattle brought from Iberia to Colombia.

Three regions were chosen to verify RRBS methylation levels including both hyper- and hypo-methylated DMRs and genes related to the main biological processes immunity (*SERPINB1*), cancer (*SHOX2*) and nervous system (*NRXN2*). Levels of methylation were higher overall in the RRBS dataset than were obtained by HiSeq-BSP, but the observed direction and tendency of changes were consistent for all the regions under analysis (Table S7). The difference in magnitude between RRBS and HiSeq-BSP results may have been due to the lower bisulfite conversion rate obtained for the HiSeq-BSP protocol (89%) compared with RRBS (98%) and/or PCR bias.

Although DNA methylation is universally associated with gene expression silencing (Bird 2002), the complex gene and pathway connections, coupled with the long-range interactions of regulatory elements that cannot simply be predicted by genomic proximity (Miele & Dekker 2008; Sanyal *et al.* 2012), hinders the extrapolation of epigenomic and genomic factors, along with environmental influences, to phenotypic transitions. Moreover, RRBS covers only a small fraction of the genome, and cellular heterogeneity is a major challenge when comparing DNA methylation across samples. Blood samples consist of a mixture of immune cells in varying proportions with unique methylation profiles that may have hindered the ability to detect DMRs (Reinius *et al.* 2012). The presence of C/T SNPs within CpGs may be also a confounding variable, especially when methylation levels are compared among individuals from genetically differentiated populations (Daca-Roszak *et al.* 2015). Thus, these results should be considered a preliminary survey, highlighting the need for additional epigenomic studies on a wider sample set under more standardized conditions, using more extensive techniques, such as whole-genome bisulfite sequencing, and strengthening the connection between epigenomic and phenotypic variability by also integrating genomic and gene expression datasets.

Conclusion

In conclusion, we characterized differential methylation patterns between tropically adapted bovine breeds and their main ancestors for the first time and showed that challenging climate and environmental factors imposed on a reduced number of animals had an impact on their methylome pattern still measurable today, affecting genes implicated in important signalling pathways for adaptation and pointing towards epigenetic fine-tuning of the regulation of gene activity. The comparison between the sample

groups identified DMRs annotated to genes directly or indirectly involved in tropical adaptation processes, such as immunity, nervous system processes, energy management, heat resistance, and skin and coat attributes. The ability of epigenetic changes to provide an initial rapid and flexible response to environmental challenges makes epigenetic studies a promising field for uncovering alternative mechanism-driven evolution of adaptive phenotypes, eventually generating permanent genetic changes. The DMRs detected in this study, along with the tissue analysed—blood—which is easily accessible and reflects the immune status of individuals, provide a valuable starting point for the identification of epigenetic biomarkers of resilience for improved cattle performance and welfare under predicted climatic change models.

Acknowledgements

We thank Prof. J. Cañón and Prof. S. Dunner from Universidad Complutense de Madrid (UCM) for their contribution of the Spanish samples included in the study. We also thank M. Barbato and P. Orozco-terWengel from Cardiff University and M. Montes from UCM Genetic Service for their collaboration in data analysis. NS is a recipient of a Marie Skłodowska-Curie Individual Fellowship, and this project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No DLV-655100. MWB is funded by BBSRC through the FACCE-JPI ERA-NET Climate Smart Agriculture project CLIMGEN (BB/M019276/1).

Conflict of interest

The authors declare that they have no conflict of interest.

Data availability

The datasets supporting the results of this article were deposited in the Gene Expression Omnibus (GEO) with accession GSE101796.

References

Akalin A., Kormaksson M., Li S., Garrett-Bakelmann F.E., Figeuroa M.E., Melnick A. & Mason C.E. (2012) METHYLKIT: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biology* **13**, R87.

Amorim C.E., Daub J.T., Salzano F.M., Foll M. & Excoffier L. (2015) Detection of convergent genome-wide signals of adaptation to tropical forests in humans. *PLoS ONE* **10**, e0121557.

Baerwald M.R., Meek M.H., Stephens M.R., Nagarajan R.P., Goodbla A.M., Tomalty K.M.H., Thorgaard G.H., May B. & Nichols K.M. (2016) Migration-related phenotypic divergence is associated with epigenetic modifications in rainbow trout. *Molecular Ecology* **25**, 1785–800.

Bernstein B.E., Meissner A. & Lander E.S. (2007) The mammalian epigenome. *Cell* **128**, 669–81.

Bird A. (2002) DNA methylation patterns and epigenetic memory. *Genes and Development* **16**, 6–21.

Blekhman R., Oshlack A., Chabot A.E., Smyth G.K. & Gilad Y. (2008) Gene regulation in primates evolves under tissue-specific selection pressures. *PLoS Genetics* **4**, e1000271.

Chan E.K., Nagaraj S.H. & Reverter A. (2010) The evolution of tropical adaptation: comparing taurine and zebu cattle. *Animal Genetics* **41**, 467–77.

Ciscar J.C., Feyen L., Soria A. *et al.* (2014) *Climate Impacts in Europe: the JRC PESETA II Project* (JRC Scientific and Policy Reports EUR 26586EN). Publications Office of the European Union, Luxembourg.

Cole J.B., Wiggans G.R., Ma L. *et al.* (2011) Genome-wide association analysis of thirty-one production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows. *BMC Genomics* **12**, 408.

Daca-Roszak P., Pfeifer A., Żebracka-Gala J., Rusinek D., Szybińska A., Jarząb B., Witt M. & Ziętkiewicz E. (2015) Impact of SNPs on methylation readouts by Illumina Infinium HumanMethylation450 BeadChip Array: implications for comparative population studies. *BMC Genomics* **16**, 1003.

Day S.E., Coletta R.L., Kim J.Y. *et al.* (2016) Next-generation sequencing methylation profiling of subjects with obesity identifies novel gene changes. *Clinical Epigenetics* **8**, 77.

Fagny M., Patin E., MacIsaac J.L. *et al.* (2015) The epigenomic landscape of African rainforest hunter-gatherers and farmers. *Nature Communications* **6**, 10047.

Feng H., Conneely K.N. & Wu H. (2014) A Bayesian hierarchical model to detect differentially methylated loci from single nucleotide resolution sequencing data. *Nucleic Acids Research* **42**, e69.

Fung M.M., Nguyen C., Mehtani P. *et al.* (2008) Genetic variation within adrenergic pathways determines *in vivo* effects of presynaptic stimulation in humans. *Circulation* **117**, 517–25.

Gao F., Zhang J., Jiang P., Gong D., Wang J.W., Xia Y., Østergaard M.V., Wang J. & Sangild P.T. (2014) Marked methylation changes in intestinal genes during the perinatal period of preterm neonates. *BMC Genomics* **15**, 716.

Gautier M., Flori L., Riebler A., Jaffrézic F., Laloé D., Gut I., Moazami-Goudarzi K. & Foulley J.L. (2009) A whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC Genomics* **10**, 550.

Ginja C., Gama L.T., Cortes O. *et al.* (2013) Analysis of conservation priorities of Iberoamerican cattle based on autosomal microsatellite markers. *Genetics Selection Evolution* **45**, 35.

Hahn G.L. (1999) Dynamic responses of cattle to thermal heat loads. *Journal of Animal Science* **77**, 10–20.

Hernández-Cerón J., Chase C.C. & Hansen P.J. (2004) Differences in sensitivity to heat-shock between preimplantation embryos from heat-tolerant (Brahman and Romosinuano) and heat-sensitive (Angus) breeds. *Journal of Dairy Science* **87**, 53–8.

Holdrige L.R. & Hunter J.R. (1961) Clave de las Asociaciones Climáticas del mundo y guía para el uso de la tierra en los trópicos. *Suplemento de la Revista Académica Colombiana de Ciencias* **11**, 14.

Huang D.W., Sherman B.T. & Lempicki R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature Protocols* **4**, 44–57.

Huang Y.Z., Sun J.J., Zhang L.Z. *et al.* (2014) Genome-wide DNA methylation profiles and their relationships with mRNA and the microRNA transcriptome in bovine muscle tissue (*Bos taurine*). *Scientific Reports* **4**, 6546.

Korkmaz F.T. & Kerr D.E. (2017) Genome-wide methylation analysis reveals differentially methylated loci that are associated with an age-dependent increase in bovine fibroblast response to LPS. *BMC Genomics* **18**, 405.

Krueger F. & Andrews S.R. (2011) BISMARK: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–2.

Laird P.W. (2010) Principles and challenges of genomewide DNA methylation analysis. *Nature Reviews Genetics* **11**, 191–203.

Lean I.J. & Rabiee A.R. (2011) Effect of feeding biotin on milk production and hoof health in lactating dairy cows: a quantitative assessment. *Journal of Dairy Science* **94**, 1465–76.

Leroi A.M., Koufopanou V. & Burt A. (2003) Cancer selection. *Nature Reviews Cancer* **3**, 226–31.

Li C., Ying W., Huang Z., Brehm T., Morin A., Vella A.T. & Zhou B. (2017) IRF6 regulates alternative activation by suppressing PPARγ in male murine macrophages. *Endocrinology* **158**, 2837–47.

Lister R. & Ecker J.R. (2009) Finding the fifth base: genome-wide sequencing of cytosine methylation. *Genome Research* **19**, 959–66.

Liu Y.-H., Wang L. & Xu T. (2018) Whole-genome sequencing of African dogs provides insights into adaptations against tropical parasites. *Molecular Biology and Evolution* **35**, 287–98.

Lyraki R., Megaw R. & Hurd T. (2016) Disease mechanisms of X-linked retinitis pigmentosa due to *RP2* and *RPGR* mutations. *Biochemical Society Transactions* **44**, 1235–44.

Makina S.O., Muchadeyi F.C., van Marle-Köster E., Taylor J.F., Makgahlela M.L. & Maiwashe A. (2015) Genome-wide scan for selection signatures in six cattle breeds in South Africa. *Genetics Selection Evolution* **47**, 92.

Mallika C., Guo Q. & Li J.Y. (2015) Gbx2 is essential for maintaining thalamic neuron identity and repressing habenular characters in the developing thalamus. *Developmental Biology* **407**, 26–39.

Martínez R.A., García D., Gallego J.L., Onofre G., Pérez J. & Cañón J. (2008) Genetic variability in Colombian Creole cattle populations estimated by pedigree information. *Journal of Animal Science* **86**, 545–52.

Martínez A.M., Gama L.T., Cañón J. *et al.* (2012) Genetic footprints of Iberian cattle in America 500 years after the arrival of Columbus. *PLoS ONE* **7**, e49066.

Mi H., Poudel S., Muruganujan A., Casagrande J.T. & Thomas P.D. (2016) PANTHER version 10: expanded protein families and functions, and analysis tools. *Nucleic Acids Research* **44**, D336–42.

Miele A. & Dekker J. (2008) Long-range chromosomal interactions and gene regulation. *Molecular Biosystems* **4**, 1046–57.

Pant S.D., Verschoor C.P., Schenkel F.S., You Q., Kelton D.F. & Karrow N.A. (2011) Bovine *PGLYRP1* polymorphisms and their association with resistance to *Mycobacterium avium* ssp. paratuberculosis. *Animal Genetics* **42**, 354–60.

Pettipher R., Hansel T.T. & Armer R. (2007) Antagonism of the prostaglandin D2 receptors DP1 and CRTH2 as an approach to treat allergic diseases. *Nature Reviews Drug Discovery* **6**, 313–25.

Pinzón M.E. (1984) *Historia de la Ganadería Bovina en Colombia. Suplemento Ganadero. Banco Ganadero*, Bogotá, Colombia.

Pitt D., Bruford M.W., Barbato M., Orozco-terWengel P., Martínez R. & Sevane N. (2018) Demography and rapid local adaptation shape Creole cattle genome diversity in the tropics. *Evolutionary Applications* https://doi.org/10.1111/eva.12641

Porto-Neto L.R., Reverter A., Prayaga K.C. *et al.* (2014) The genetic architecture of climatic adaptation of tropical cattle. *PLoS ONE* **9**, e113284.

Reinius L.E., Acevedo N., Joerink M., Pershagen G., Dahlén S.E., Greco D., Söderhäll C., Scheynius A. & Kere J. (2012) Differential DNA methylation in purified human blood cells: implications for cell lineage and studies on disease susceptibility. *PLoS ONE* **7**, e41361.

Richardson R.J., Dixon J., Malhotra S., Hardman M.J., Knowles L., Boot-Handford R.P., Shore P., Whitmarsh A. & Dixon M.J. (2006) Irf6 is a key determinant of the keratinocyte proliferation-differentiation switch. *Nature Genetics* **38**, 1329–34.

Rodero E., Rodero A. & Delgado J.V. (1992) Primitive Andalusian livestock and their implications in the discovery of America. *Archivos de Zootecnia* **41**, 383–400.

Rouse J.E. (1997) *The Criollo: Spanish Cattle in the Americas*. University of Oklahoma Press, Norman, OK.

Safra N., Bassuk A.G., Ferguson P.J. *et al.* (2013) Genome-wide association mapping in dogs enables identification of the homeobox gene, *NKX2-8*, as a genetic component of neural tube defects in humans. *PLoS Genetics* **9**, e1003646.

Santiago A.A. (1978) Evolution of Zebu cattle in Brazil. *The Zebu Journal* **1**, 6.

Sanyal A., Lajoie B.R., Jain G. & Dekker J. (2012) The long-range interaction landscape of gene promoters. *Nature* **489**, 109–13.

Schroeder D.I., Jayashankar K., Douglas K.C., Thirkill T.L., York D. & Dickinson P.J. (2015) Early developmental and evolutionary origins of gene body DNA methylation patterns in mammalian placentas. *PLoS Genetics* **11**, e1005442.

Seabury C.M., Seabury P.M., Decker J.E., Schnabel R.D., Taylor J.F. & Womack J.E. (2010) Diversity and evolution of 11 innate immune genes in *Bos taurus taurus* and *Bos taurus indicus* cattle. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 151–6.

Semik E., Ząbek T., Gurgul A. *et al.* (2017) Comparative analysis of DNA methylation patterns of equine sarcoid and healthy skin samples. *Veterinary and Comparative Oncology* **16**, 37–46.

Shankar K., Kang P., Zhong Y., Borengasser S.J., Wingfield C., Saben J., Gomez-Acevedo H. & Thakali K.M. (2015) Transcriptomic and epigenomic landscapes during cell fusion in BeWo trophoblast cells. *Placenta* **36**, 1342–51.

Snippert H.J., Haegebarth A., Kasper M. *et al.* (2010) Lgr6 marks stem cells in the hair follicle that generate all cell lineages of the skin. *Science* **327**, 1385–9.

Stockwell P.A., Chatterjee A., Rodger E.J. & Morison I.M. (2014) DMAP: differential methylation analysis package for RRBS and WGBS data. *Bioinformatics* **30**, 1814–22.

Su Z., Xia J. & Zhao Z. (2011) Functional complementation between transcriptional methylation regulation and post-transcriptional microRNA regulation in the human genome. *BMC Genomics* **12**, S15.

Su J., Wang Y., Xing X., Liu J. & Zhang Y. (2014) Genome-wide analysis of DNA methylation in bovine placentas. *BMC Genomics* **15**, 12.

Tom Tang Y., Emtage P., Funk W.D., Hu T., Arterburn M., Park E.E. & Rupp F. (2004) TAFA: a novel secreted family with conserved cysteine residues and restricted expression in the brain. *Genomics* **83**, 727–34.

Varriale A. (2014) DNA methylation, epigenetics, and evolution in vertebrates: facts and challenges. *International Journal of Evolutionary Biology*, **2014**, 475981.

Wang J., Duncan D., Shi Z. & Zhang B. (2013) WEB-based GEne SeT AnaLysis Toolkit (WEBGESTALT): update 2013. *Nucleic Acids Research* **41**, W77–83.

Wang M.D., Dzama K., Rees D.J. & Muchadeyi F.C. (2016) Tropically adapted cattle of Africa: perspectives on potential role of copy number variations. *Animal Genetics* **47**, 154–64.

Weyrich A., Lenz D., Jeschek M., Chung T.H., Rübensam K., Göritz F., Jewgenow K. & Fickel J. (2016) Paternal intergenerational epigenetic response to heat exposure in male Wild guinea pigs. *Molecular Ecology* **25**, 1729–40.

Wilde D. (2006) Influence of macro and micro minerals in the periparturient period on fertility in dairy cattle. *Animal Reproduction Science* **96**, 240–9.

Willham R.L. (1982) Genetic improvement of beef cattle in the United States: cattle, people and their interaction. *Journal of Animal Science* **54**, 659–66.

Wolf B. (2012) Biotinidase deficiency: if you have to have an inherited metabolic disease, this is the one to have. *Genetics in Medicine* **14**, 565–75.

Zhang Y., Baheti S. & Sun Z. (2016) Statistical method evaluation for differentially methylated CpGs in base resolution next-generation DNA sequencing data. *Briefings in Bioinformatics* **19**, 374–86.

Zhou Y., Xu L., Bickhart D.M. *et al.* (2016) Reduced representation bisulphite sequencing of ten bovine somatic tissues reveals DNA methylation patterns and their impacts on gene expression. *BMC Genomics* **17**, 779.

Supporting information

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

Figure S1 Scatter plot and correlation of CpG_{10} methylation between Creole and Spanish cattle samples.

Figure S2 CpG_{10} site coverage histogram of (a) Creole (SM1, CCC1, CCC2) and (b) Spanish (LD1, RAV, RET) cattle RRBS libraries.

Figure S3 CpG_{10} methylation distribution in (a) Creole (SM1, CCC1, CCC2) and (b) Spanish (LD1, RAV, RET) cattle RRBS libraries.

Figure S4 Functional annotation of genes differentially methylated between Creole and Spanish cattle samples using PANTHER.

Figure S5 KEGG signalling pathways of differentially methylated genes between Creole and Spanish cattle samples obtained with the DAVID tool.

Table S1 Bisulfite PCR primer sequences used for HiSeq-BSP validation of RRBS data on Creole and Spanish cattle samples.

Table S2 Creole (SM1, CCC1, CCC2) and Spanish (LD1, RAV, RET) cattle RRBS data summary.

Table S3 Number, coverage and methylation distribution of CpG_{10} in Creole (SM1, CCC1, CCC2) and Spanish (LD1, RAV, RET) cattle RRBS methylomes.

Table S4 CpG_{10} present in all six samples: annotated with the closest/overlapping transcription start sites (TSS) ($\pm$100 kb), including promoters ($-$0 to 2 kb), introns and exons; and with the closest/overlapping CpG island (CpGI), including CpGI shores ($\pm$0–2 kb) and shelves ($\pm$2–4 kb).

Table S5 Differentially methylated regions (DMRs) showing a P-value <0.01, $\geq$4 CpGs within a distance of 200 bp and mean methylation difference >25% between Creole and Spanish sample groups; annotation with the closest/overlapping transcription start sites (TSS) ($\pm$100 kb), including promoters ($-$0 to 2 kb), introns and exons; annotation with the closest/overlapping CpG island (CpGI), including CpGI shores ($\pm$0–2 kb) and shelves ($\pm$2–4 kb).

Table S6 Detailed functional annotation of the differentially methylated regions (DMRs) showing a P-value <0.01, $\geq$4 CpGs within a distance of 200 bp and a mean methylation difference >25% between Creole and Spanish cattle samples using DAVID Functional Annotation Cluster (FAC) analysis under high stringency ease scores.

Table S7 Comparison between HiSeq-BSP and RRBS DMR results.