

Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <http://orca.cf.ac.uk/117989/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Sevane, N., Martínez, R and Bruford, M. W. 2019. Genome-wide differential DNA methylation in tropically adapted Creole cattle and their Iberian ancestors. *Animal Genetics* 50 (1) , pp. 15-26. 10.1111/age.12731 file

Publishers page: <http://dx.doi.org/10.1111/age.12731> <<http://dx.doi.org/10.1111/age.12731>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



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

3 Natalia Sevane^{1*}, Rodrigo Martínez², Michael W. Bruford^{1,3}

4 ¹School of Biosciences, Cardiff University, Cardiff, UK

5 ²Corporación Colombiana De Investigación Agropecuaria (Corpoica), Centro de investigaciones
6 Tibaitatá, Bogotá, Colombia

7 ³Sustainable Places Research Institute, Cardiff University, Cardiff, UK

8

9 * Correspondence (NS): nsevane@ucm.com; Sevane-FernandezN@cardiff.ac.uk; Tel: +44 29
10 2087 5073

11

12 **Summary**

13 Enhancing climate resilience and sustainable production for animals in harsh environments are
14 important goals for the livestock industry given the predicted impacts of climate change. Rapid
15 adaptation to extreme climatic conditions has already been imposed on livestock species,
16 including those exported after Columbus' arrival in the Americas. We compared the
17 methylomes of two Creole cattle breeds living in tropical environments with their putative
18 Spanish ancestors to understand the epigenetic mechanisms underlying rapid adaptation of a
19 domestic species to a new and more physiologically challenging environment. Reduced
20 representation bisulfite sequencing (RRBS) was used to assess differences in methylation in
21 Creole and Spanish samples and revealed 334 differentially methylated regions (DMRs) using
22 high stringency parameters (p -value < 0.01, ≥ 4 CpGs within a distance of 200 bp, mean
23 methylation difference > 25%), annotated to 263 unique features. Gene ontology analysis
24 revealed candidates involved in tropical adaptation processes, including genes differentially
25 hyper- or hypomethylated above 80% in Creole samples displaying biological functions related
26 to immune response (*IRF6*, *PRGDR*, *FAM19A5*, *PRLYRP1*), nervous system (*GBX2*, *NKX2-8*,
27 *RPGR*), energy management (*BTD*), heat resistance (*CYB561*) and skin and coat attributes
28 (*LGR6*). Our results entail that major environmental changes imposed on Creole cattle has had
29 an impact on their methylomes measurable today, which affects genes implicated in important

30 pathways for adaptation. Although further work is needed, this first characterization of
31 methylation patterns driven by profound environmental change provides a valuable pointer for
32 the identification of biomarkers of resilience for improved cattle performance and welfare
33 under predicted climatic change models.

34

35 **Keywords** *Bos taurus*, Criollo, livestock, epigenomics, RRBS

36

37 **Running Head** Epigenomics of tropical adaptation in Creole cattle

38

39 **Introduction**

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

56 Events in human history have included episodes where rapid adaptation to extreme climatic
57 conditions have been imposed on a limited number of domestic animals. One example of such
58 an event is Columbus' arrival in the Americas. Livestock species were brought from the Iberian

59 Peninsula to the Americas on Columbus' second journey in 1493 (Rodero *et al.* 1992) and
60 spread throughout the continent, adapting to a wide range of alien environmental conditions
61 and giving rise to 'Creole' animal populations (Rouse 1997). The total number of Iberian cattle
62 brought to the Americas is estimated to have been less than 1,000 (Rodero *et al.* 1992). After
63 nearly 300 years of Creole cattle expansion, several other European breeds were introduced
64 and crossed with local populations (Willham 1982), as well as with Indian Zebu cattle, especially
65 in tropical areas (Santiago 1978). Creole cattle were subsequently displaced into marginal,
66 demanding environmental areas where they still occur. Examples include the Costeño con
67 Cuernos and San Martinero breeds of Colombia, which descend from Spanish cattle and also
68 have minor influences from Continental and/or Zebu breeds (Martínez *et al.* 2012; Ginja *et al.*
69 2013). These breeds are therefore the product of several centuries of adaptation to new, local
70 and challenging environments. The Costeño con Cuernos was developed in Caribbean Colombia
71 and tolerates high temperatures and humidity, being found from the swamp areas of Córdoba
72 and Magdalena to the dry savannah of Sucre and Bolívar (Pinzón 1984). The San Martinero was
73 developed in the Colombian Orinoquia region in the 17th century and is almost uniquely
74 adapted to tropical rainforests (Holdrige & Hunter 1961).

75 A central goal of evolutionary biology, and an increasingly relevant one to agriculture, is to
76 elucidate the genetic architecture of adaptation. The past decade has yielded an increasing
77 number of examples where regulatory changes have been shown to contribute to species-
78 specific adaptations and to reproductive isolation (Blekhman *et al.* 2008). There is mounting
79 evidence that heritable variation in relevant traits can be generated through a suite of
80 epigenetic mechanisms, even in the absence of genetic variation, which eventually might
81 promote permanent changes in DNA sequence (Varriale 2014; Fagny *et al.* 2015). Among
82 epigenetic mechanisms, DNA methylation via 5-methylcytosine is a key modification in
83 vertebrate genomes that imparts an additional layer of heritable regulatory information upon
84 DNA and is essential for viability in a myriad of biological processes (Lister & Ecker 2009).
85 Epigenomic studies in cattle include muscle and placental tissues analysed with non-base-
86 resolution methods (Su *et al.* 2014; Huang *et al.* 2014) and two recent studies using base-
87 resolution techniques, a low coverage whole genome bisulfite sequencing (WGBS) analysis of

88 bovine placenta (Schroeder *et al.* 2015) and reduced representation bisulfite sequencing (RRBS)
89 of ten bovine tissues, including blood (Zhou *et al.* 2016). However, these only described DNA
90 methylome landscapes, not exploring environmental influences on phenotypic variation. Thus,
91 despite increasing knowledge about the genes involved in bovine adaptation to tropical climate
92 (Gautier *et al.* 2009; Chan *et al.* 2010; Porto-Neto *et al.* 2014; Makina *et al.* 2015; Wang *et al.*
93 2016; Pitt *et al.* 2018), we lack understanding of relevant epigenetic function (see Varriale 2014
94 for a review).

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

107

108 **Material and methods**

109 *Samples and DNA extraction*

110 Five New and Old World cattle breeds were analysed in this study. The sample comprised
111 Colombian Creole cattle Costeño con Cuernos (n = 2) and San Martinero (n = 1) breeds and
112 Iberian cattle representing the main ancestors of these Creole populations including the
113 Asturiana de los Valles (n = 1), Lidia (n = 1) and Retinta (n = 1) breeds. Samples were collected
114 from adult males between 7 and 11 years old. Animals were reared in their native environment
115 under extensive conditions with access to characteristic local vegetation available, growing
116 under the particular climatic and dietary conditions that gave rise to the different breed

117 adaptations (Table 1). DNA was extracted from blood samples using the UltraClean BloodSpin
118 DNA Isolation Kit (MO BIO Laboratories, Inc) for the Creole samples and the QIAamp DNA Blood
119 Mini Kit (Qiagen) for the Spanish samples. The concentration and quality of genomic DNA was
120 evaluated using the Qubit dsDNA HS Assay Kit (Life Technologies).

121

122 *Reduced representation bisulfite sequencing (RRBS)*

123 Genomic DNA (0.5-1.0 µg) from each sample was restricted with *MspI* enzyme (New England
124 Biosciences), cleaned using DNA Clean and concentrator-25 columns (Zymo Research), and
125 eluted in 60 µl for library preparation. The sticky ends produced by *MspI* digestion were filled
126 with CG nucleotides and Illumina sequencing adapters. The TruSeq Nano DNA LT Library Prep
127 Kit (Illumina) was used for 3'adenylation and adapter ligation. The end-repaired samples were
128 purified using 2.5X AMPure XP Beads (Beckman Coulter) and eluted in 20 µl resuspension
129 buffer. After adapter ligation, samples were again purified using 1.0X AMPure XP Beads and
130 eluted in 40 µl resuspension buffer. Size-selection of DNA fragments (~175-225 bp) was
131 performed using a 2% Agarose gel (Invitrogen), and the selected fragments were purified twice
132 using 1.0X AMPure XP Beads and finally eluted in 22 µl resuspension buffer. Bisulfite conversion
133 of non-methylated cytosines was performed on 20 µl size-selected fragments using the EZ DNA
134 Methylation-Lightning Kit (Zymo Research). PCR (20 cycles) was performed to enrich the
135 sequencing library by using a TruSeq Nano DNA LT Library Prep Kit (Illumina). The Pfu Turbo Cx
136 Hotstart DNA polymerase (Agilent Technologies) and 10 mM dNTP mix (Life Technologies) were
137 used for PCR reactions. After enrichment, the library was purified twice using 1X AMPure XP
138 Beads (Beckman Coulter) and quantified using the Qubit dsDNA HS Assay Kit (Life
139 Technologies). The average library size was determined using an Agilent 2100 Bioanalyzer
140 (Agilent Technologies). The libraries were then pooled in equimolar ratios of 2 nM, and 6.0 pM
141 of the pool was clustered and sample tracked using the cBot (Illumina) and sequenced following
142 a 2 x 150 bp protocol for 300 cycles using the HiSeq 2500 system (Illumina).

143

144 *RRBS data analysis and genome-wide DNA methylation levels*

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

155 For CpG level comparison, percent methylation of individual CpGs was calculated using
156 MethylKit package in R (Akalin *et al.* 2012) and the coverage files from Bismark aligner. To
157 prevent PCR bias and increase the power of the statistical tests we discarded bases with high
158 (above 99.9th percentile of coverage in each sample) and low (below 10X coverage, CpG₁₀) read
159 coverage. Each sequenced and filtered CpG₁₀ site was assigned a percentage methylation score.
160 The CpG₁₀ bisulfite conversion rate was calculated as the number of thymines (non-methylated
161 cytosines) divided by coverage for each non-CpG cytosine as implemented in MethylKit.
162 Coverage and correlation plots were generated also by MethylKit. The pattern of methylation
163 around different components of the cattle genome, including gene bodies (defined as the
164 region from transcription start site -TSS- to transcription termination site -TTS-), TSSs, TTSSs, and
165 CpG islands (CpGI), was also investigated using the Seqmonk software (Babraham
166 bioinformatics, UK), from 20 kb upstream to 20 kb downstream. CpG₁₀ were annotated with the
167 closest/overlapping TSS (± 100 kb) (Miele & Dekker 2008; Sanyal *et al.* 2012) and CpGI using
168 identgenloc program from the DMAP package (Stockwell *et al.* 2014). Promoters were defined
169 as -0-2 kb of TSS, and CpGI shores and shelves as $\pm 0-2$ kb and $\pm 2-4$ kb flanking regions of CpGI,
170 respectively.

171

172 *Differentially methylated region (DMR) analysis*

173 DMRs were established among Creole and Spanish groups to account for breed similarities
174 related to their geographical location, i.e. the Iberian Peninsula (three breeds, three samples =
175 three biological replicates within the same group) and Colombia (two breeds, three samples =
176 three biological replicates within the same group). To compare spatially contiguous stretches of
177 methylated cytosines across the Creole and the Spanish genomes, DMRs were determined
178 using the R package dispersion shrinkage for sequencing data (DSS) (Feng *et al.* 2014), which
179 outperforms other methods when sample size per group is small owing to the adoption of Wald
180 test with shrinkage for determining differentially methylated cytosines (DMC) (Zhang *et al.*
181 2016). We identified DMRs using the coverage files from Bismark and the callDMR function with
182 a *p*-value threshold of 0.01, delta=0.1 and otherwise default parameters. To be considered
183 significant, a DMR was required to contain at least 3 CpG sites (default parameter, although the
184 smallest significant DMR included 4 CpGs) within a distance of 200 bp, and with an absolute
185 mean methylation difference greater than 25% when comparing Creole and Spanish samples
186 (Akalin *et al.* 2012). As CpG₁₀, DMRs were annotated with the closest/overlapping TSS (± 100 kb)
187 and CpGI using the identgenloc program from the DMAP package.

188

189 *Gene ontology (GO) analysis*

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

199

200 *Validation of RRBS data with HiSeq bisulfite sequencing PCR (HiSeq-BSP)*

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

214

215 **Results**

216 *Assessment of RRBS data and genome-wide DNA methylation levels*

217 Fragmentation with the restriction enzyme *MspI* of blood-extracted DNA from three Creole and
218 three Spanish samples resulted in high quality sequencing RRBS libraries enriched for high CG
219 regions. Illumina HiSeq 2500 sequencing generated between 15 and 38 million reads per
220 sample (accession number GSE101796) and a total of 136 million reads (Table S2). Quality
221 control analysis using Trim Galore and MethylKit indicated that the 150 bp sequences displayed
222 the expected nucleotide composition based on *MspI* digestion and bisulfite conversion (98%
223 average sodium bisulfite conversion efficiency). On average, 98.2% of reads passed the filtering
224 process (Table S2). The mean percentage of mapped reads was 85%, with 33-61% of reads
225 mapping to multiple locations of the genome and 29-43% mapping uniquely (Table S2).
226 Sequences that did not map, or did not map uniquely, were excluded from the analysis.
227 After alignment, we filtered the CpG dinucleotides based on a coverage of 10 or more reads
228 (CpG₁₀). The number of CpG₁₀ per sample ranged from 0.4 to 1.6 million, and the mean
229 coverage from 33 to 106 (Table S3). Of these sites, 20,234 were present in all six samples (Table

230 S4). We observed high positive correlations between all the samples analysed (mean Pearson's
231 correlation coefficient = 0.8), although clear variation was present between them (Figure S1).
232 The distribution of sequence read coverage of CpG₁₀ per sample is shown in Figure S2, and
233 highlights that despite the observation that the filtered CpG₁₀ displayed high mean coverage,
234 the libraries did not suffer from bias due to excessive amplification of a subset of fragments, as
235 reflected in the absence of peaks on the right-hand side of each histogram. The RRBS protocol
236 has been shown to enrich for CpGs and, as CpGs have been universally reported to be regions
237 of gene regulation via methylcytosine and are generally demethylated, the percent methylation
238 of CpGs in RRBS libraries is expected to be lower than the average methylation of the genome
239 (~80%). Accordingly, global CpG₁₀ methylation ranged from 51 to 57% across samples (Table
240 S3). The distribution of methylation at each CpG₁₀ site revealed a bimodal pattern, with heavy
241 methylation (>95%) of 39 to 53% CpG₁₀ and completely unmethylated bases (<5%) ranging
242 between 35% and 47% (Figure S3). The median methylation was high (84%) (Table S3),
243 reflecting the heavy hypermethylation of 48% of CpG₁₀ sites. However, hypomethylated CpG₁₀
244 sites were also evident, including 42% of the analysed CpG sites (Figure S3). The percentage of
245 CHG and CHH methylation was low in cattle blood, ranging from 0.9% to 1.6% (Table S3).
246 RRBS reads were detected in most chromosomal regions (chromosomes 1–29 and X) in each
247 group, although some gaps existed (Figure 1). This even read distribution indicated that cattle
248 blood methylomes can be detected by RRBS technology with good representation, thereby
249 ensuring accurate examination of variation in DNA methylation. The distribution of CpG₁₀
250 related to CpGIs (6,235, 31%) revealed that RRBS data is highly enriched in CpGI cores (85%),
251 while only a small amount is in CpGI shores (11%) and shelves (4%) (Table S4). The distribution
252 of CpG₁₀ in relation to genes (4,389, 22%) showed that almost 85% mapped to gene bodies
253 (92% located in introns and 8% in exons) and a much smaller percentage mapped to gene
254 promoters (15%), with the main amount located in intergenic regions (15,845, 78%) (Table S4).
255 DNA methylation levels sharply decreased in the 2 kb region upstream of TSSs and dropped to
256 the lowest point before TSSs (Figure 2A), corresponding with the distribution of gene
257 promoters, usually prone to transcription, whereas levels dramatically increased in the 3'
258 direction, peaking 5' to the TTS (Figure 2A), related with the methylation of gene bodies

259 contributing to chromatin structure stability and the regulation of gene expression (Bird 2002).
260 The level dropped slightly and plateaued after TTS (Figure 2A). As expected, the level of
261 methylation in CpGIs was lower than outside CpGIs (Figure 2B).

262

263 *Differential methylation between Creole and Spanish cattle samples*

264 Comparison between spatially contiguous stretches of DMCs from Creole and Spanish samples
265 revealed 334 DMRs (p -value < 0.01, ≥ 4 CpGs within a distance of 200 bp, mean methylation
266 difference > 25%, Table S5). Annotation of these DMRs showed that 275 sites (82%),
267 corresponding to 263 unique features, were overlapping a gene or within a distance of ± 100 kb
268 from the closest TSS. Approximately 37% of DMRs overlapped a gene, while $\sim 4\%$ were in
269 regions 2 kb upstream of TSS or promoters. Intragenic DMR were equally divided between
270 introns (52%) and exons (48%). Around 36% of DMRs were located in CpGIs, mainly in CpGI
271 cores (81%), whereas only 12% and 7% overlapped CpGIs shores and shelves, respectively.
272 Interestingly, a high proportion of DMRs (71%) displayed hypermethylation in Creole samples.
273 Table 2 shows the DMRs overlapping a gene or CpGI hyper- and hypomethylated above 80% in
274 Creole samples.

275

276 *Gene ontology (GO) analysis*

277 Among the 263 differentially methylated unique annotated features, functional data for 213
278 genes was obtained with PANTHER, including the GO classes molecular function (the primary
279 activities of gene products at the molecular level), biological process (sets of molecular events
280 or operations with a defined beginning and end) and cellular component (Figure S4). The
281 annotated DMRs were then analysed using DAVID and WebGeStalt tools. DAVID FAC analysis
282 produced 16 enriched functional clusters under high stringency conditions for 115 DAVID IDs
283 (Table S6). Among these enriched functional clusters, homeobox, epidermal growth factor (two
284 clusters) and immunoglobulin (two clusters) were identified. We analysed the distribution of
285 annotated DMRs along the cattle chromosomes, confirming one enriched genomic region in
286 chromosome 21 that comprised 12 genes related to the cellular component membrane (Table
287 S6). KEGG pathway analysis retrieved a total of 14 pathways (Table 3): ten from WebGeStalt,

288 including immune related processes such as leukocyte, T cell and lymphocyte differentiation
289 and activation, and circulatory system development or cell proliferation (Figure 3); and four
290 from DAVID tool –acute myeloid leukemia, insulin signalling pathway, Rap1 signaling pathway,
291 microRNAs in cancer (Figure S5).

292

293 *Validation of RRBS data with HiSeq bisulfite sequencing PCR (HiSeq-BSP)*

294 We used HiSeq-BSP to assess the methylation patterns of three gene annotated DMRs,
295 including regions displaying high and low differential methylation levels between Creole and
296 Spanish samples and implicated in immune (*SERPIN1*), cancer (*SHOX2*) and nervous system
297 (*NRXN2*) processes. The HiSeq-BSP methylation results were significant for the three amplified
298 regions ($p\text{-value} < 0.01$, ≥ 4 CpGs, mean methylation difference $\geq 10\%$) and concordant with the
299 methylation profiles obtained with the RRBS analysis (Table S7).

300

301 **Discussion**

302 Studies on adaptation are key to disentangling the evolutionary potential of organisms in
303 response to biotic and abiotic stress and other environmental challenges, which could
304 potentially be highly relevant in the context of global climate change. Tropical environments
305 are characterized by high temperature and humidity, episodes of feed and water scarcity and
306 virulent tropical diseases and parasite infections. Creole cattle demonstrate greater resistance
307 to such conditions, surviving, breeding and producing efficiently in the tropics (Hernández-
308 Cerón *et al.* 2004; Martínez *et al.* 2008). Two tropically adapted Creole breeds and their likely
309 Spanish ancestors were analysed to establish epigenomic differences among groups accounting
310 for breed similarities related to their geographical location. The Costeño con Cuernos and San
311 Martinero breeds have been developed under physiologically challenging tropical conditions.
312 The Iberian breed Retinta is distributed throughout central and southern Iberia, which is
313 characterized by a xeric climate. The Asturiana de los Valles breed reflects the northern Iberian
314 gene-pool and is exposed to a milder climate, mostly cold and damp. The Lidia breed (Spanish
315 fighting bull) has not been selected for productivity traits and thus may be the most
316 representative modern descendent of Iberian cattle herds back in the 15th century. We

317 detected 334 highly significant DMRs between the groups. The methylation profiles obtained
318 were consistent with previous studies (e.g. Zhou *et al.* 2016). High stringency parameters to
319 detect DMRs (p -value < 0.01, ≥ 4 CpGs within a distance of 200 pb, mean methylation
320 difference > 25%) when compared with other studies (e.g. Gao *et al.* 2014; Day *et al.* 2016;
321 Shankar *et al.* 2015; Baerwald *et al.* 2016), were taken as statistically significant to overcome
322 the relatively small number of biological replicates characteristic of many epigenomic
323 experiments (e.g. Miele *et al.* 2008; Zhou *et al.* 2016; Semik *et al.* 2017).

324 In concordance with previous studies on bovine adaptation to tropical climates, including both
325 taurine and indicine (Gautier *et al.* 2009; Chan *et al.* 2010; Porto-Neto *et al.* 2014; Makina *et al.*
326 2015; Wang *et al.* 2016; Pitt *et al.* 2018), we found a number of differentially methylated genes
327 between Creole and Spanish groups implicated in several biological processes key for survival in
328 harsh environments, such as immunity, nervous system processes, energy management, heat
329 resistance and skin and coat attributes (Table S5).

330 Tropical cattle carry lower burdens of ticks, have enhanced disease resistance and superior
331 innate immunity, which is reflected in a higher number of genes under selection related to the
332 immune system in studies on adaptation to tropical conditions (e.g. Amorim *et al.* 2015; Liu *et*
333 *al.* 2018; Pitt *et al.* 2018). Some of the genes hypermethylated above 80% in Creole samples are
334 implicated in immune processes (Table 2): i) *IRF6*, involved in inflammatory responses,
335 macrophage activation and dysregulation of metabolic and immunologic homeostasis (Li *et al.*
336 2017); ii) *PRGDR*, which plays an important role in the immune response found in allergic
337 diseases, apart from facilitating smooth muscle relaxation and vasodilatation, inhibiting platelet
338 aggregation and contributing to the regulation of pain perception and sleep (Pettipher *et al.*
339 2007); and iii) *FAM19A5*, a brain-specific chemokine or neurokinin that acts as regulator of
340 immune and nervous cells (Tom Tang *et al.* 2004). On the contrary, a DMR was found
341 hypomethylated above 80% in Creole samples in a CpGI core located inside an exon of the
342 *PRLYRP1* gene. The protein encoded by this gene has been reported to interact with microbes
343 to maintain intestinal homeostasis (Seabury *et al.* 2010) and has been associated with
344 resistance to *Mycobacterium avium ssp. paratuberculosis* (Pant *et al.* 2011), both in cattle. This
345 gene is also associated with several health, reproduction and body conformation traits in

346 Holstein cows (Cole *et al.* 2011). The ability to cope with parasitic and infectious diseases in the
347 adaptation to new environments also seems relevant at the multi-genic level given the high
348 enrichment of pathways such as regulation of leukocyte differentiation, T cell activation,
349 leukocyte cell-cell adhesion or lymphocyte activation (Figure 3, Table 3), and the presence of
350 two enriched functional clusters related to immunoglobulins (Table S6). These findings are also
351 in agreement with the work of Fagny *et al.* (2015) that describes the existence of epigenetic
352 variability on immune processes implicated in the adaptation to changes in habitat and lifestyle
353 in humans.

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

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

374 Genes involved in cardiovascular physiology can facilitate heat resistance in tropical climates.
375 We found a mean hypermethylation level of above 80% in Creole samples for the *PTGDR* gene,
376 which facilitates smooth muscle relaxation and vasodilatation (Pettipher *et al.* 2007), and the
377 *CYB561* gene, influencing cardiovascular responses to sympathetic activation (Fung *et al.* 2008),
378 as well as the high enrichment of the circulatory system development pathway (Table 2, Table
379 3, Figure 3). Skin and coat attributes are also important for adaptation to harsh conditions, with
380 a direct influence on the thermo-resistance to tropical conditions. The *IRF6* gene, also
381 implicated in immune homeostasis, promotes epithelial cell proliferation and differentiation
382 (Richardson *et al.* 2006) and was hypermethylated above 80% in the Creole group (Table 2). On
383 the contrary, *LGR6* gene, which establishes sebaceous glands and interfollicular epidermis
384 postnatally (Snippert *et al.* 2010), showed hypomethylation above 80% in Creole samples. In
385 addition, two enriched functional clusters related to epidermal growth factor were identified
386 with the DAVID FAC analysis (Table S6). Four microRNAs, regulation of which at the 3'
387 untranslated region plays important roles in the modulation of gene expression (Su *et al.* 2011),
388 were also differentially methylated between Creole and Spanish cattle groups.

389 Some cancers, especially in young animals, might be a by-product of novel adaptation and have
390 their origins in recent evolutionary changes in morphology and life-history, driving evolution of
391 many features of cellular behaviour and regulation (Leroi *et al.* 2003). Concordantly, rapid bouts
392 of evolution, such as artificial selection in domestic species, have been shown to make animals
393 prone to different cancers (Leroi *et al.* 2003). Epigenetic changes, especially DNA methylation,
394 alter signal-transduction pathways during the early stages of tumor development. Tumor cells,
395 opposed to normal cells, show local hypermethylation of some CpGI combined with global
396 genome demethylation (Bernstein *et al.* 2007). Taking into account that RRBS enriches for GC-
397 rich regions such as CpGIs (Laird 2010), the high proportion of hypermethylated DMRs in Creole
398 samples (71%), along with the high number of genes differentially hyper- or hypomethylated
399 above 80% in these samples and related to oncogenic processes (in particular, *NKX2-8*, *LATS2*,
400 *BRAT1*, *BLM*, *TP53/11* also known as *PIG11*, *TM4SF5*, *TRIM25*, *LGR6*) (Table 2), as well as the
401 high enrichment of several pathways implicated in cancer (acute myeloid leukemia, Rap1
402 signaling pathway, microRNAs in cancer) (Table 3, Figure S5), might reflect an on-going

403 adaptation process to tropical conditions on the descendants from the cattle brought from
404 Iberia to Colombia.

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

412 Although DNA methylation is universally associated with gene expression silencing (Bird 2002),
413 the complex gene and pathway connections coupled with the long-range interactions of
414 regulatory elements that cannot simply be predicted by genomic proximity (Miele & Dekker
415 2008; Sanyal et al. 2012), hinders the extrapolation of epigenomic and genomic factors, along
416 with environmental influences, to phenotypic transitions. Moreover, RRBS only covers a small
417 fraction of the genome and cellular heterogeneity is a major challenge when comparing DNA
418 methylation across samples. Blood samples consist of a mixture of immune cells in varying
419 proportions with unique methylation profiles that may have hindered the ability to detect
420 DMRs (Reinius *et al.* 2012). The presence of C/T SNPs at CpGs may be also a confounding
421 variable, especially when methylation levels are compared among individuals from genetically
422 nonhomogeneous populations (Daca-Roszak et al. 2015). Thus, these results should be
423 considered as a preliminary survey, highlighting the need for additional epigenomic studies on a
424 wider sample set under more standardized conditions, using more extensive techniques like
425 WGBS and strengthening the connection between epigenomic and phenotypic variability by
426 integrating also genomic and gene expression data sets.

427

428 In conclusion, we characterize differential methylation patterns between tropically adapted
429 bovine breeds and their main ancestors for the first time, and show that challenging climate
430 and environmental factors imposed on a reduced number of animals had an impact on their
431 methylome pattern still measurable today, affecting genes implicated in important signalling

432 pathways for adaptation and pointing towards the epigenetic fine-tuning on the regulation of
433 gene activity. The comparison between both sample groups identified DMRs annotated to
434 genes directly or indirectly involved in tropical adaptation processes, such as immunity, nervous
435 system processes, energy management, heat resistance and skin and coat attributes. The ability
436 of epigenetic changes to provide an initial rapid and flexible response to environmental
437 challenges, makes epigenetic studies a promising field to uncover alternative mechanisms
438 driven evolution of adaptive phenotypes, eventually generating permanent genetic changes.
439 The DMRs detected in this study, along with the tissue analysed, blood, that is easily accessible
440 and reflects the immune status of individuals, provide a valuable starting point for the
441 identification of epigenetic biomarkers of resilience for improved cattle performance and
442 welfare under predicted climatic change models.

443

444 **Data availability**

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

447

448 **Conflict of interest**

449 The authors declare that they have no conflict of interest.

450

451 **Acknowledgements**

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

460

461 **Refereces**

- 462 Akalin A., Kormaksson M., Li S., Garrett-Bakelmann F.E., Figueiroa M.E., Melnick A. & Mason CE. (2012)
463 methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles.
464 Genome Biology 13, R87.
- 465 Amorim C.E., Daub J.T., Salzano F.M., Foll M. & Excoffier L. (2015) Detection of convergent genome-wide
466 signals of adaptation to tropical forests in humans. PLoS One 10, e0121557.
- 467 Baerwald M.R., Meek M.H., Stephens M.R., Nagarajan R.P., Goodbla A.M., Tomalty K.M.H., Thorgaard
468 G.H., May B. & Nichols K.M. (2016) Migration-related phenotypic divergence is associated with
469 epigenetic modifications in rainbow trout. Molecular Ecology 25, 1785-1800.
- 470 Bernstein B.E., Meissner A. & Lander E.S. (2007) The mammalian epigenome. Cell 128, 669-681.
- 471 Bird A. (2002) DNA methylation patterns and epigenetic memory. Genes & Development 16, 6-21.
- 472 Blekhman R., Oshlack A., Chabot A.E., Smyth G.K. & Gilad Y. (2008) Gene regulation in primates evolves
473 under tissue-specific selection pressures. PLoS Genetics 4, e1000271.
- 474 Chan E.K., Nagaraj S.H. & Reverter A. (2010) The evolution of tropical adaptation: comparing taurine and
475 zebu cattle. Animal Genetics 41, 467-477.
- 476 Ciscar J.C., Feyen L., Soria A., et al. (2014) Climate Impacts in Europe. The JRC PESETA II Project. JRC
477 Scientific and Policy Reports; 2014. EUR 26586EN.
- 478 Cole J.B., Wiggans G.R., Ma L., Sonstegard T.S., Lawlor T.J. Jr., Crooker B.A., Van Tassell C.P., Yang J.,
479 Wang S., Matukumalli L.K. & Da Y. (2011) Genome-wide association analysis of thirty one
480 production, health, reproduction and body conformation traits in contemporary U.S. Holstein cows.
481 BMC Genomics 12, 408.
- 482 Daca-Rozsak P., Pfeifer A., Żebracka-Gala J., Rusinek D., Szybińska A., Jarzab B., Witt M. & Ziętkiewicz E.
483 (2015) Impact of SNPs on methylation readouts by Illumina Infinium HumanMethylation450
484 BeadChip Array: implications for comparative population studies. BMC Genomics 16, 1003.
- 485 Day S.E., Coletta R.L., Kim J.Y., Campbell L.E., Benjamin T.R., Roust L.R., De Filippis E.A., Dinu V., Shaibi
486 G.Q., Mandarino L.J. & Coletta D.K. (2016) Next-generation sequencing methylation profiling of
487 subjects with obesity identifies novel gene changes. Clinical Epigenetics 8,77.
- 488 Fagny M., Patin E., Maclsaac J.L., et al. (2015) The epigenomic landscape of African rainforest hunter-
489 gatherers and farmers. Nature Communications 6, 10047.
- 490 Feng H., Conneely K.N. & Wu H. (2014) A Bayesian hierarchical model to detect differentially methylated
491 loci from single nucleotide resolution sequencing data. Nucleic Acids Research 42, e69.

492 Fung M.M., Nguyen C., Mehtani P., Salem R.M., Perez B., Thomas B., Das M., Schork N.J., Mahata S.K.,
493 Ziegler M.G. & O'Connor D.T. (2008) Genetic variation within adrenergic pathways determines in
494 vivo effects of presynaptic stimulation in humans. *Circulation* 117, 517-525.

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

498 Gautier M., Flori L., Riebler A., Jaffrézic F., Laloé D., Gut I., Moazami-Goudarzi K. & Foulley J.L. (2009) A
499 whole genome Bayesian scan for adaptive genetic divergence in West African cattle. *BMC Genomics*
500 10, 550.

501 Ginja C., Gama L.T., Cortes O., Delgado J.V., Dunner S., García D., Landi V., Martín-Burriel I., Martínez-
502 Martínez A., Penedo M.C., Rodellar C., Zaragoza P., Cañon J. & BioBovis Consortium. (2013) Analysis
503 of conservation priorities of Iberoamerican cattle based on autosomal microsatellite markers.
504 *Genetics Selection Evolution* 30, 45-35.

505 Hahn G.L. (1999) Dynamic responses of cattle to thermal heat loads. *Journal of Animal Science* 77, 10-
506 20.

507 Hernández-Cerón J., Chase C.C. & Hansen P.J. (2004) Differences in sensitivity to heat-shock between
508 preimplantation embryos from heat-tolerant (Brahman and Romosinuano) and heat-sensitive
509 (Angus) breeds. *Journal of Dairy Science* 87, 53-58.

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

512 Huang D.W., Sherman B.T. & Lempicki R.A. (2009) Systematic and integrative analysis of large gene lists
513 using DAVID Bioinformatics Resources. *Nature Protocols* 4, 44-57.

514 Huang Y.Z., Sun J.J., Zhang L.Z., Li C.J., Womack J.E., Li Z.J., Lan X.Y., Lei C.Z., Zhang C.L., Zhao X. & Chen
515 H. (2014) Genome-wide DNA methylation profiles and their relationships with mRNA and the
516 microRNA transcriptome in bovine muscle tissue (*Bos taurine*). *Scientific Reports* 4, 6546.

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

520 Krueger F. & Andrews S.R. (2011) Bismark: a flexible aligner and methylation caller for Bisulfite-Seq
521 applications. *Bioinformatics* 27, 1571-1572.

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

524 Lean I.J. & Rabiee A.R. (2011) Effect of feeding biotin on milk production and hoof health in lactating
525 dairy cows: a quantitative assessment. *Journal of Dairy Science* 94, 1465-1476.

526 Leroi A.M., Koufopanou V. & Burt A. (2003) Cancer selection. *Nature Reviews Cancer* 3, 226-231.

527 Li C., Ying W., Huang Z., Brehm T., Morin A., Vella A.T. & Zhou B. (2017) IRF6 regulates alternative
528 activation by suppressing PPAR γ in male murine macrophages. *Endocrinology* 158, 2837-2847.

529 Lister R. & Ecker J.R. (2009) Finding the fifth base: Genome-wide sequencing of cytosine methylation.
530 *Genome Research* 19, 959-966.

531 Liu Y-H., Wang L. & Xu T. (2018) Whole-Genome Sequencing of African Dogs Provides Insights into
532 Adaptations against Tropical Parasites. *Molecular Biology and Evolution* 35, 287-298.

533 Lyraki R., Megaw R. & Hurd T. (2016) Disease mechanisms of X-linked retinitis pigmentosa due to RP2
534 and RPGR mutations. *Biochemical Society Transactions* 44, 1235-1244.

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

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

540 Martínez A.M., Gama L.T., Cañón J., et al. (2012) Genetic Footprints of Iberian Cattle in America 500
541 Years after the Arrival of Columbus. *PLoS One* 7, e49066.

542 Martínez R.A., García D., Gallego J.L., Onofre G., Pérez J. & Cañón J. (2008) Genetic variability in
543 Colombian Creole cattle populations estimated by pedigree information. *Journal of Animal Science*
544 86, 545-552.

545 Mi H., Poudel S., Muruganujan A., Casagrande J.T. (2016) Thomas PD. PANTHER version 10: expanded
546 protein families and functions, and analysis tools. *Nucleic Acids Research* 44, D336-342.

547 Miele A. & Dekker J. (2008) Long-range chromosomal interactions and gene regulation. *Molecular
548 Biosystems* 4, 1046-1057.

549 Pant S.D., Verschoor C.P., Schenkel F.S., You Q., Kelton D.F. & Karrow N.A. (2011) Bovine PGLYRP1
550 polymorphisms and their association with resistance to *Mycobacterium avium* ssp. *paratuberculosis*.
551 *Animal Genetics* 42, 354-360.

552 Pettipher R., Hansel T.T. & Armer R. (2007) Antagonism of the prostaglandin D2 receptors DP1 and
553 CRTH2 as an approach to treat allergic diseases. *Nature Reviews Drug Discovery* 6, 313-325.

554 Pinzón M.E. (1984) Historia de la ganadería bovina en Colombia. Suplemento ganadero. Banco
555 Ganadero. Santafé de Bogotá Colombia 4, (1):208.

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

559 Porto-Neto L.R., Reverter A., Prayaga K.C., Chan E.K., Johnston D.J., Hawken R.J., Fordyce G., Garcia J.F.,
560 Sonstegard T.S., Bolormaa S., Goddard M.E., Burrow H.M., Henshall J.M., Lehnert S.A. & Barendse
561 W. (2014) The genetic architecture of climatic adaptation of tropical cattle. *PLoS One* 9, e113284.

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

565 Richardson R.J., Dixon J., Malhotra S., Hardman M.J., Knowles L., Boot-Handford R.P., Shore P.,
566 Whitmarsh A. & Dixon M.J. (2006) *Irf6* is a key determinant of the keratinocyte proliferation-
567 differentiation switch. *Nature Genetics* 38, 1329-1334.

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

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

572 Safra N., Bassuk A.G., Ferguson P.J., Aguilar M., Coulson R.L., Thomas N., Hitchens P.L., Dickinson P.J.,
573 Vernau K.M., Wolf Z.T. & Bannasch D.L. (2013) Genome-wide association mapping in dogs enables
574 identification of the homeobox gene, *NKX2-8*, as a genetic component of neural tube defects in
575 humans. *PLoS Genetics* 9, e1003646.

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

577 Sanyal A., Lajoie B.R., Jain G. & Dekker J. (2012) The long-range interaction landscape of gene promoters.
578 *Nature* 489, 109-413.

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

582 Seabury C.M., Seabury P.M., Decker J.E., Schnabel R.D., Taylor J.F. & Womack J.E. (2010) Diversity and
583 evolution of 11 innate immune genes in *Bos taurus taurus* and *Bos taurus indicus* cattle. *Proceedings*
584 *of the National Academy of Sciences USA* 107, 151-156.

585 Semik E., Ząbek T., Gurgul A., Fornal A., Szmatoła T., Pawlina K., Wnuk M., Klukowska-Rötzler J., Koch C.,
586 Mählmann K. & Bugno-Poniewierska M. (2017) Comparative analysis of DNA methylation patterns of
587 equine sarcoid and healthy skin samples. *Veterinary and Comparative Oncology* 16, 37-46.

588 Shankar K., Kang P., Zhong Y., Borengasser S.J., Wingfield C., Saben J., Gomez-Acevedo H. & Thakali K.M.
589 (2015) Transcriptomic and epigenomic landscapes during cell fusion in BeWo trophoblast cells.
590 *Placenta* 36, 1342-1351.

591 Snippert H.J., Haegebarth A., Kasper M., Jaks V., van Es J.H., Barker N., van de Wetering M., van den
592 Born M., Begthel H., Vries R.G., Stange D.E., Toftgård R. & Clevers H. (2010) Lgr6 marks stem cells in
593 the hair follicle that generate all cell lineages of the skin. *Science* 327, 1385-1389.

594 Stockwell P.A., Chatterjee A., Rodger E.J. & Morison I.M. (2014) DMAP: differential methylation analysis
595 package for RRBS and WGBS data. *Bioinformatics* 30, 1814-022.

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

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

601 Tom Tang Y., Emtage P., Funk W.D., Hu T., Arterburn M., Park E.E. & Rupp F. (2004) Tafa: a novel
602 secreted family with conserved cysteine residues and restricted expression in the brain. *Genomics*
603 83, 727-734.

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

606 Wang J., Duncan D., Shi Z. & Zhang B. (2013) WEB-based GENE SeT Analysis Toolkit (WebGestalt):
607 update 2013. *Nucleic Acids Research* 41, W77-W83.

608 Wang M.D., Dzama K., Rees D.J. & Muchadeyi F.C. (2016) TROPICALLY ADAPTED CATTLE OF AFRICA:
609 PERSPECTIVES ON POTENTIAL ROLE OF COPY NUMBER VARIATIONS. *Animal Genetics* 47, 154-164.

610 Weyrich A., Lenz D., Jeschek M., Chung T.H., Rübensam K., Göritz F., Jewgenow K. & Fickel J. (2016)
611 Paternal intergenerational epigenetic response to heat exposure in male Wild guinea pigs.
612 *Molecular Ecology* 25, 1729-1740.

613 Wilde D. (2006) Influence of macro and micro minerals in the peri-parturient period on fertility in dairy
614 cattle. *Animal Reproduction Science* 96, 240-249.

615 Willham R.L. (1982) Genetic improvement of beef cattle in the United States: cattle, people and their
616 interaction. *Journal of Animal Science* 54, 659-666.

617 Wolf B. (2012) Biotinidase deficiency: if you have to have an inherited metabolic disease, this is the one
618 to have. *Genetics in Medicine* 14, 565-575.

619 Zhang Y., Baheti S. & Sun Z. (2016) Statistical method evaluation for differentially methylated CpGs in
620 base resolution next-generation DNA sequencing data. *Briefings in Bioinformatics* pii, bbw133.
621 Zhou Y., Xu L., Bickhart D.M., Abdel Hay E.H., Schroeder S.G., Connor E.E., Alexander L.J., Sonstegard T.S.,
622 Van Tassell C.P., Chen H., Liu G.E. (2016) Reduced representation bisulphite sequencing of ten
623 bovine somatic tissues reveals DNA methylation patterns and their impacts on gene expression.
624 *BMC Genomics* 17, 779.
625

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

Breed	Location	MASL ¹	MAT ² (°C)	MARH ³ (%)	MAR ⁴ (mm)
Costeño con Cuernos	Department of Córdoba (Sinú river valley, Colombia)	300	30	80	2,500
San Martinero	Department of Meta (Colombia)	700	21	70	1,800
Asturiana de los Valles	Mieres (Asturias, Spain)	380	11	80	1,000
Lidia (Casta Navarra)	Igúzquiza (Navarra, Spain)	450	12	67	600
Retinta	Tierra de Barros (Badajoz, Spain)	400	17	66	450

627 ¹ Metres above sea level
628 ² Mean annual temperature
629 ³ Mean annual relative humidity
630 ⁴ Mean annual rainfall
631

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

Symbol	Gene name	Gene overlap	CpGI relation	Meth diff ¹
<i>Hypermethylated in Creole samples</i>				
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
<i>Hypomethylated in Creole samples</i>				
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

634 ¹Methylation differences averaged from all CpG sites within the defined region. Negative differential methylation values indicate
 635 hypermethylation in Creole samples; positive differential methylation values indicate hypomethylation in Creole samples.
 636

637

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

GO term	Description	Gene count	Enrichment score	Genes
GO:1902105	Regulation of leukocyte differentiation	7	5,717	NRARP,LIF,PGLYRP1,PRKCZ,RUNX1,NKAP,CD83
GO:0050863	Regulation of T cell activation	7	5,449	NRARP,CD5,PRKCZ,MAD1L1,LMO1,NKAP,CD83
GO:1903037	Regulation of leukocyte cell-cell adhesion	7	5,128	NRARP,CD5,PRKCZ,MAD1L1,LMO1,NKAP,CD83
GO:0051249	Regulation of lymphocyte activation	8	4,689	NRARP,CD5,PGLYRP1,PRKCZ,MAD1L1,LMO1,NKAP,CD83
GO:0022407	Regulation of cell-cell adhesion	8	4,634	NRARP,CD5,PRKCZ,ALOX12,MAD1L1,LMO1,NKAP,CD83
GO:0050865	Regulation of cell activation	10	4,837	NRARP,CD5,PGLYRP1,PRKCZ,ALOX12,PDGFA,MAD1L1,LMO1,NKAP,CD83
GO:0072359	Circulatory system development	14	3,013	NRARP,LIF,ALOX12,PDGFA,GBX2,MAP2K2,SMAD6,CYP1B1,EOMES,BAK1,SHOX2,FLRT2,ADM2,BCOR
GO:2000026	Regulation of multicellular organismal development	19	2,334	NRARP,LIF,PGLYRP1,PRKCZ,ALOX12,PDGFA,MAP2K2,CYP1B1,EOMES,SHOX2,RUNX1,RFX4,BHLHE23,FLRT2,NKAP,CD83,ADM2,PHOX2B,BCOR
GO:0008283	Cell proliferation	21	2,341	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	1,915	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	6,448	MAP2K2, KIT, RUNX1
bta04910	Insulin signalling pathway	4	3,540	PRKCZ, SOCS2, MAP2K2, PRKAR1B
bta04015	Rap1 signalling pathway	5	2,812	PRKCZ, MAPK12, PDGFA, MAP2K2, KIT
bta05206	MicroRNAs in cancer	7	3,397	DNMT3A, CYP24A1, BAK1, CYP1B1, PDGFA, MAP2K2, PAK4

640

641

642

643 **Figures**

644 **Figure 1.** Chromosomal distribution of reads in the Creole (A) and Spanish (B) grouped samples.

645 The distribution of reads is shown in a gradient from blue (low) to red (high).

646 **Figure 2.** DNA methylation levels in relation to gene bodies (A) and CpG islands (CpGI) (B). Gene

647 bodies were defined as the region from the transcription start site (TSS) to transcription
648 termination site (TTS).

649 **Figure 3.** KEGG pathway overrepresentation enrichment analysis (ORA) of differentially
650 methylated genes between Creole and Spanish cattle samples performed with WebGeStalt.

651

652 **Supporting information**

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

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

656 **Table S3.** Number, coverage and methylation distribution of CpG₁₀ in Creole (SM1, CCC1, CCC2)
657 and Spanish (LD1, RAV, RET) cattle RRBS methylomes.

658 **Table S4.** CpG₁₀ present in all six samples; annotated with the closest/overlapping transcription
659 start sites (TSS) (± 100 kb), including promoters (-2 kb), introns and exons; and with the
660 closest/overlapping CpG island (CpGI), including CpGI shores ($\pm 0-2$ kb) and shelves ($\pm 2-4$ kb).

661 **Figure S1.** Scatter plot and correlation of CpG₁₀ methylation between Creole and Spanish cattle
662 samples. Numbers above the diagonal denote pair-wise Pearson's correlation scores. The
663 histograms on the diagonal are the methylation distribution of CpG₁₀ sites for each sample.

664 Below the diagonal, the scatter plots of percentage methylation values for each pair in Creole
665 (SM1, CCC1, CCC2) and Spanish (LD1, RAV, RET) RRBS libraries are shown.

666 **Figure S2.** CpG₁₀ site coverage histogram of (A) Creole (SM1, CCC1, CCC2) and (B) Spanish (LD1,
667 RAV, RET) cattle RRBS libraries.

668 **Figure S3.** CpG₁₀ methylation distribution in (A) Creole (SM1, CCC1, CCC2) and (B) Spanish (LD1,
669 RAV, RET) cattle RRBS libraries.

670 **Table S5.** Differentially methylated regions (DMRs) showing a p-value < 0.01 , ≥ 4 CpGs within a
671 distance of 200 bp and mean methylation difference $> 25\%$ between Creole and Spanish sample

672 groups; annotation with the closest/overlapping transcription start sites (TSS) (± 100 kb),
673 including promoters (-2kb), introns and exons; annotation with the closest/overlapping CpG
674 island (CpGI), including CpGI shores ($\pm 0-2$ kb) and shelves ($\pm 2-4$ kb). Positive differential
675 methylation values indicate hypomethylation in Creole samples; negative differential
676 methylation values indicate hypermethylation in Creole samples.

677 **Figure S4.** Functional annotation of genes differentially methylated between Creole and
678 Spanish cattle samples using PANTHER. A) Molecular function; B) Biological process; C) Cellular
679 component.

680 **Table S6.** Detailed functional annotation of the differentially methylated regions (DMRs)
681 showing a *p-value* < 0.01 , ≥ 4 CpGs within a distance of 200 bp and a mean methylation
682 difference $> 25\%$ between Creole and Spanish cattle samples using DAVID Functional
683 Annotation Cluster (FAC) analysis under high stringency ease scores. Distribution of enriched
684 genomic regions along the cattle chromosomes obtained with DAVID.

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

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

688