1 **The bovine alveolar macrophage DNA methylome is resilient to infection with**

2 ***Mycobacterium bovis***

3 Running Title: Impact of mycobacterial infection on DNA methylation

4 Alan Mark O'Doherty [1,*,¶], Kevin Rue-Albrecht [2,¶], David Andrew Magee [1], Simone Ahting [3],

5 Rachelle Elizabeth Irwin [4], Thomas Johnathan Hall [1], John Arthur Browne [1], Nicolas Claude

6 Nalpas [5], Colum Patrick Walsh [4], Stephen Vincent Gordon [6,7], Marcin Wlodzimierz

7 Wojewodzic [8] and David Evan MacHugh [1,7]

8 [1] Animal Genomics Laboratory, UCD School of Agriculture and Food Science, University

9 College Dublin, Belfield, Dublin, D04 V1W8, Ireland.

10 [2] Kennedy Institute of Rheumatology, Nuffield Department of Orthopaedics, Rheumatology

11 and Musculoskeletal Sciences, University of Oxford, Headington, Oxford OX3 7FY, UK.

12 [3] Institute of Molecular Medicine, Trinity College Dublin, Dublin, D08 W9RT, Ireland.

13 [4] Genomic Medicine Research Group, Biomedical Sciences Research Institute, Centre for

14 Molecular Biosciences, University of Ulster, Coleraine, BT52 1SA UK.

15 [5] Quantitative Proteomics and Proteome Centre Tübingen, Interfaculty Institute for Cell

16 Biology, University of Tübingen, 72076 Tübingen, Germany.

17 [6] UCD School of Veterinary Medicine, University College Dublin, Belfield, Dublin, D04

18 V1W8, Ireland.

19 [7] UCD Conway Institute of Biomolecular and Biomedical Research, University College

20 Dublin, Belfield, Dublin, D04 V1W8, Ireland.

21 [8] University of Birmingham, School of Biosciences, Edgbaston, Birmingham, B15 2TT, UK.

22   ¶ These authors contributed equally to this work.

23   * Corresponding author: alan.odoherty@ucd.ie

32

33

34

35

36

37

38

39

40   **Abstract**

41   DNA methylation is pivotal in orchestrating gene expression patterns in various mammalian

42   biological processes. Perturbation of the bovine alveolar macrophage (bAM) transcriptome, due

43   to *Mycobacterium bovis* (*M. bovis*) infection, has been well documented; however, the impact of

44   this intracellular pathogen on the bAM epigenome has not been determined. Here, whole genome

45   bisulfite sequencing (WGBS) was used to assess the effect of *M. bovis* infection on the bAM

46   DNA methylome. The methylomes of bAM infected with *M. bovis* were compared to those of

47   non-infected bAM 24 hours post-infection (hpi). No differences in DNA methylation (CpG or

48   non-CpG) were observed. Analysis of DNA methylation at proximal promoter regions uncovered

49   >250 genes harbouring intermediately methylated (IM) promoters (average methylation of 33–

50   66%). Gene ontology analysis, focusing on genes with low, intermediate or highly methylated

51   promoters, revealed that genes with IM promoters were enriched for immune-related GO

52   categories; this enrichment was not observed for genes in the high or low methylation groups.

53   Targeted analysis of genes in the IM category confirmed the WGBS observation. This study is

54   the first in cattle examining genome-wide DNA methylation at single nucleotide resolution in an

55   important bovine cellular host-pathogen interaction model, providing evidence for IM promoter

56   methylation in bAM.

57

Introduction

Infection with *Mycobacterium bovis*, the causative agent of bovine tuberculosis (BTB), accounts annually for more than $3 billion of losses to global agriculture through lost productivity and disease control costs [1]. There is also evidence suggesting that the burden of *M. bovis* as the cause of zoonotic tuberculosis in humans may be underestimated [2], which highlights the need for a more detailed understanding of the impact of *M. bovis* in both cattle and humans. Unravelling host cellular processes that are perturbed or manipulated by intracellular pathogens is an important area of research in infection biology, particularly for disease control and the development of next-generation diagnostics and prognostics. In this regard, host cell epigenetic modifications induced, either as a component of the response to *M. bovis* infection, or as an immunoevasion strategy by the pathogen itself, remain to be fully elucidated [3].

Modifications to the genome, such as DNA methylation and histone tail modifications, in combination with RNA-mediated regulatory mechanisms are fundamental in modulating tissue-specific gene expression [4-6]. Epigenetic gene regulation represents an important framework for understanding how environmental stimuli are disseminated to the transcriptome and preserved through subsequent somatic cell divisions [5]. DNA methylation (5-methylcytosine), the most widely studied genome modification, is involved in a variety of cellular processes including genomic imprinting, X-chromosome inactivation, chromosome stability and gene transcription [7] and has been proposed to be influenced by external stimuli across a wide range of biological contexts [8-12]. Therefore, we hypothesised that changes to DNA methylation may be involved in the bovine host response to infection with *M. bovis*; this mechanism has previously been proposed for human tuberculosis caused by infection with *Mycobacterium tuberculosis* [13].

80   Host epigenomic plasticity to *M. tuberculosis* has been reported previously[14,15]. Sharma

81   and colleagues showed that non-CpG loci in the host genome were hypermethylated following

82   reduced representation bisulfite sequencing (RRBS) analysis of THP-1 macrophages (a human

83   monocytic cell line) infected with *M tuberculosis* [14]. In addition, Zheng *et al*. [15] demonstrated

84   that interleukin gene promoter sequences, and their receptors, were associated with

85   hypermethylation following analysis of THP-1 cells infected with clinical strains of *M.*

86   *tuberculosis*, using the human inflammatory response methyl-profiler DNA methylation PCR

87   array. Furthermore, it has been demonstrated that DNA methylation is associated with hypoxic

88   survival of *M. tuberculosis* [16]. Most recently, Doherty *et al*. reported that there were in excess of

89   750 differentially methylated regions between *M. bovis*-infected and healthy cattle in a study

90   using RRBS to examine CD4$^+$ T lymphocytes isolated from circulating blood samples [17].

91   A range of studies have highlighted the impact of infecting microorganisms on host DNA

92   methylation patterns. For example, distinct DNA methylation changes have been observed in

93   macrophages infected with the intracellular protozoan *Leishmania donovani*, the causative agent

94   of visceral leishmaniasis [18]. In addition, global DNA methylation changes have been detected in

95   human neutrophils infected with *Anaplasma phagocytophilum*, which causes granulocytic

96   anaplasmosis [19]. Finally, it has been proposed that, during chronic *Helicobacter pylori* infection

97   in humans, functional *H. pylori* DNA methyltransferases enter host epithelial cells and methylate

98   their recognition sequences in chromosomal DNA, potentially contributing to the pathogenesis of

99   gastric adenocarcinoma or lymphoma of the mucosa-associated lymphoid tissue [20].

100   Our group has previously revealed the impact of *M. bovis* infection on the mammalian

101   alveolar macrophage gene expression, demonstrating that the bAM transcriptome is substantially

102   reprogrammed as a consequence of both host-driven defence responses and mycobacterial-

103   induced perturbation and manipulation of cellular processes [21-24]. However, the effect of *M. bovis*

104   on the bovine host epigenome, specifically the DNA methylome of bAM, remains unexplored.

105   Recent work has shown that intracellular microbial infection can lead to alterations of the host

106   DNA methylome; therefore, for the present study we used WGBS to test the hypothesis that

107   bAM DNA methylation patterns are altered during the earliest stage of *M. bovis* infection in

108   cattle.

109   **Materials and Methods**

110   **Ethics statement**

111   All animal procedures were performed according to the provisions of the Cruelty to

112   Animals Act of 1876 and EU Directive 2010/63/EU. Ethical approval was obtained from the

113   University College Dublin Animal Ethics Committee (protocol number AREC-13-14-Gordon).

114   **Isolation and infection of bovine alveolar macrophages**

115   Isolation and purification of bAM from cattle was performed as previously described by

116   our group [21,23] and is summarized in Fig. 1. Briefly, total lung cells were harvested by pulmonary

117   lung lavage with Hank's Balanced Salt Solution (Invitrogen, Life Technologies) following the

118   removal of lungs from eight unrelated Holstein-Friesen male calves. Total lung cells were

119   washed and cultured for 24 h at 37 °C in R10$^+$ media (RPMI 1640 medium supplemented with

120   antibiotics [Invitrogen]). After incubation, cells were prepared for infection by dissociation and

121   seeding at 5×10^5 viable cells/well, for each biological replicate. The purity of the seeded

122   macrophages was confirmed by flow cytometry using anti-CD14 antibody. bAM were infected

123   with *M.bovis* strain AF2122/97 at a multiplicity of infection (MOI) of 10 bacilli per alveolar

124   macrophage as described in detail previously [21,23]. These previous studies used comparative

125   RNA-seq-based transcriptomics and targeted quantitative assays (RT-qPCR and multiplex

126   ELISA) of several NF-κB-inducible pro- and anti-inflammatory cytokines and chemokines,

127   including CCL-4, IL-1β, IL-6, IL-10 and IL-12, to verify that at 24 hours post-infection (hpi) *M.*

128   *bovis*-treated bAM cells were infected and had internalised bacilli [21,23].

7

129   **Isolation of DNA and library preparation**

130   DNA was extracted from *M. bovis*-infected bAM 24 hpi ($n = 8$) and from control bAM

131   ($n = 8$) at the same time point using the DNeasy kit (Qiagen) according to the manufacturer's

132   recommendations. DNA was quantified using the Qubit dsDNA HS Assay Kit (ThermoFisher

133   Scientific). Libraries were prepared for WGBS using the post–bisulfite conversion library

134   preparation method for methylation analysis (EpiGnome™ Methyl-Seq Kit, Epicentre, Illumina)

135   according to the manufacturer's instructions. Genomic DNA (50 ng) isolated from *M. bovis*-

136   infected or non-infected bAM (24 hpi) was bisulfite-modified (EZ Methylation-Direct Kit,

137   Zymo) according to the manufacturer's guidelines. DNA synthesis was performed by mixing

138   bisulfite-converted DNA with 2 µl DNA synthesis primer, incubating at 95 °C for 5 min, cooling

139   on ice, followed by addition of 4 µl EpiGnome DNA Synthesis Premix 0.5 µl, 100 mM DTT and

140   0.5 µl EpiGnome Polymerase. Reactions were incubated at 25 °C for 5 min followed by 42 °C

141   for 30 min, then cooled to 37 °C for 2 min before addition of 1 µl Exonuclease I to each reaction.

142   Following this, reaction mixtures were incubated at 37 °C for 10 min, 95 °C for 3 min and then

143   held at 25 °C. DNA was di-tagged by adding 7.5 µl EpiGnome TT Premix and 0.5 µl DNA

144   polymerase to each reaction and incubating at 25 °C for 30 min, 95 °C for 3 min and cooling to

145   4 °C. Tagged DNA was purified using the using the AMPure XP (1.6× beads, 40 µl) system. A

146   PCR step was performed to generate the second strand of DNA, complete the addition of the

147   Illumina adaptor sequences and incorporate an index sequence. 22.5 µl of di-tagged DNA was

148   mixed with 25 µl FailSafe PCR PreMix E, 1 µl EpiGnome Forward PCR Primer, 1 µl EpiGnome

149   Index PCR Primer and 0.5 µl FailSafe PCR Enzyme (1.25 U) and subjected to an initial

150   denaturation of ds DNA at 95 °C for 1 min followed by 10 cycles of 95 °C for 30 sec, 55 °C for

151   30 sec and 68 °C for 3 min. Following PCR, the reactions were incubated at 68 °C for 7 min.

152 EpiGnome libraries were purified using the AMPure (1× beads, 50 µl) system to remove primer

153 dimers. Libraries were quantified by Qubit using the Qubit dsDNA HS Assay Kit (ThermoFisher

154 Scientific) and library quality was assessed on an Agilent BioAnalyzer using the High sensitivity

155 DNA assay kit (Agilent Technologies).

156 **Pyrosequencing**

157 Genomic DNA was extracted from *M. bovis*-infected and control bAM (isolated from a

158 parallel set of four animals to those used for WGBS) and quantified with the High-Sensitivity

159 DNA Assay Kit (Agilent Technologies). DNA (200 ng) was bisulfite-modified using the EZ

160 Methylation-Direct Kit (Zymo) and eluted in 50 µl elution buffer. Bisulfite PCR reactions were

161 performed in 25 µl consisting of 0.2 µm each primer, 2 mM $MgCl_2$, 1× PCR buffer (minus

162 magnesium), 0.2 mM dNTPs, Platinum *Taq* DNA polymerase (Invitrogen), and 3 µl bisulfite-

163 modified DNA. Primer sequences are detailed in Table 1. PCR cycling conditions were as

164 follows: 95 °C for 5 min followed by 40 cycles of 30 sec each at 95 °C; either 55 °C (*TNF,*

165 *NFKB2* and *IL12A*) 56 °C (*DTX4*, *C1QB* and *NOS2*) or 58 °C (*TLR2*) for 30 sec; 72 °C for

166 30 sec, and a final elongation step of 5 min at 72 °C. PCR products were verified by

167 electrophoresis on a 2% w/v agarose gel before pyrosequencing (Pyromark Q24, Qiagen).

168 Pyrosequencing assays were designed in-house and carried out as previously described [25,26].

169 Only pyrosequencing reactions that passed Pyromark Q24 internal controls for bisulfite

170 modification were included in the analysis. Two-tailed paired sample *t*-tests were used to assess

171 statistically significant DNA methylation between control and *M. bovis*-infected samples.

172

173

9

174 **Bisulfite PCR, cloning, sequencing and combined bisulfite restriction analysis**

175 Bisulfite-converted DNA from control and infected bAM was amplified in 25 µl

176 reactions containing 0.2 µM primers, 1× buffer, 0.2 mM dNTPs, 2.5 U Platinum *Taq* DNA

177 polymerase and 3 mM MgCl$_2$. Primer sequences are detailed in Table 1. PCR cycling conditions

178 were as follows: 95 °C for 3 min followed by 35 cycles of 30 sec each at 95 °C; 58 °C, 72 °C for

179 30 sec, and a final elongation step of 5 min at 72 °C. PCR products were purified using the

180 Wizard clean up kit (Promega) and cloned into the pJET1.2/blunt vector (Fermentas). Insertion

181 of PCR products was verified by digestion with *Bgl*II and positive clones were sequenced using

182 conventional Sanger sequencing (Eurofins Genomics). Combined bisulfite restriction analysis

183 (COBRA) was carried out using *Taqα*I, and/or *Aci*I as outlined in [27]. Sequence analysis and

184 alignment was performed using DNAStar EditSeq, MegAlign (www.dnastar.com) and BiQ Meth

185 Analyzer (http://biq-analyzer.bioinf.mpi-inf.mpg.de). During this analysis, sequences with low

186 C-T conversion rate ($< 95\%$) and with a high number of sequencing errors (sequence identity

187 with genomic sequence less than 80%) were excluded from the alignment. Identical clones were

188 also excluded from the analysis.

189 **Illumina sequencing and initial quality control**

190 Pooled libraries were sequenced at the Michigan State University Research Technology

191 Support Facility. Paired-end reads (2×150 bp) were obtained by Illumina sequencing of each

192 pooled library on four lanes of a HiSeq 2500 sequencer, in rapid run mode. After pooling data

193 from all lanes, bisulfite-treated (BS) libraries yielded 45.3–67.4 million read pairs per sample;

194 comparably, non-bisulfite (NON-BS) libraries yielded a total of 40–50 million read pairs per

195 sample across all lanes (Supplementary Table 1).

196   Quality control of raw read pairs using *FastQC*
197   (www.bioinformatics.babraham.ac.uk/projects/fastqc) revealed similar QC metrics for both
198   infected and control samples (Supplementary Table 2). Although samples from animals 1, 4, 5,
199   and 6 raised warnings of *Over-Represented Sequences*, this warning was systematically triggered
200   by N-polymers in the second mate, a technical issue resolved by quality trimming. As expected,
201   BS libraries raised significantly more *Per Base Sequence Content* and *Per Base GC Content* than
202   NON-BS libraries, due to the nature of the bisulfite treatment (Supplementary Table 2).

203   **Adapter and quality trimming**

204   Stringent adapter trimming (overlap ≥ 1 bp at the 3' end of each read), and quality
205   trimming (Phred ≤ 20 from the 3' end of each read) using *Trim Galore!* [version 0.4.1]
206   (www.bioinformatics.babraham.ac.uk/projects/trim_galore) left 97.5–98.4% of raw read pairs in
207   bisulfite-treated samples, and 93.1–94.3% of raw read pairs for NON-BS samples. However, this
208   trend was reversed at the nucleotide level, with 77.5–86.9% of sequenced bases left in BS
209   libraries, against 84.8–87.3% in NON-BS libraries. Notably, BS libraries generally showed
210   higher levels of adapter contamination (54.2–77.3% of raw reads) relative to NON-BS libraries
211   (48.6–58.2% of raw reads), based on the stringent detection rule described above. Second read
212   mates displayed a larger proportion of low-quality sequenced bases trimmed (7.6–13.5% of raw
213   sequenced bases) relative to first mates (1.9–4.5% of raw sequenced bases), in both BS and
214   NON-BS libraries (Supplementary Table 1).

215   Notably, quality control of trimmed libraries revealed a significant improvement of *Over-*
216   *Represented Sequences*, and full resolution of *Adapter Content* warnings (data not shown). As a
217   result of stringent adapter trimming (even a single trailing A at the 3' end was trimmed;
218   following Trim Galore! default settings), all samples raised warnings of *Per Base Sequence*

11

219   *Content* caused by the severe under-representation of A nucleotides at the 3' end of reads, a

220   known artefact of the stringent trimming process with no notable repercussion on the subsequent

221   alignments and methylation calls.

222   **Alignment of bisulfite-treated libraries**

223   BS libraries were aligned using *Bismark* [version 0.15.2] [28] and the *Bowtie2* aligner

224   [version 2.2.6] [29] in strand-specific (directional) mode to computationally generate bisulfite-

225   converted copies of the top and bottom strands of the *Bos taurus* UMD3.1 genome assembly [30].

226   Alignment efficiency (*i.e.*, read pairs aligned to a unique locus) reached 59.7–68.3%, for a total

227   of 28.8–42.7 million read pairs aligned uniquely per sample. Aligned reads were found evenly

228   distributed between the top and bottom strands of the BS-converted genome (Supplementary

229   Table 3). Bismark methylation calls revealed methylation levels in the range of 69.2–73.8% in

230   CpG context, for a total of 113–156 million methylation calls per sample. In contrast, non-CpG

231   context displayed markedly low methylation levels (0.7–1.7%), with orders of magnitude larger

232   counts of methylation calls, owing to their broader definition of methylation context (392–543

233   million calls in CHG context; 1.1–1.5 billion calls in CHH context: H corresponds to A, T or C).

234   **Deduplication of aligned bisulfite-treated libraries**

235   Paired-end alignments where both mates aligned to the same position in the genome were

236   removed from the Bismark alignment output using the *deduplicate_bismark* script to mitigate the

237   impact of duplicate DNA fragments sequenced. This procedure discarded 6.1–18.8% aligned

238   read pairs, leaving 25.5–38.7 million aligned read pairs for subsequent methylation calls

239   (Supplementary Table 4).

240

Methylation calls

241

242   The *bismark_methylation_extractor* script was used in a two-pronged approach. First,

243   methylation calls extracted from the full sequence of aligned read pairs were used to evaluate *M*-

244   bias across the aligned mates. *M*-bias plots show the methylation proportion across each possible

245   position in the read, and reveal anomalies at any position of the sequenced reads, often found

246   toward ends of the sequenced reads. After analysis of the *M*-bias plots generated in the first pass,

247   the second call to the *bismark_methylation_extractor* script was set to ignore the first seven

248   bases at the 5' end of both read mates. Collated reports of the Bismark pipeline leading to the

249   final methylation calls are available as HTML files in Supplementary File 1.

Statistical analyses

250

251   Methylation calls in CpG context were imported from the individual Bismark CpG

252   reports, combined and processed in a *BSseq* container of the *bsseq* Bioconductor package

253   [version 1.9.2; www.bioconductor.org/packages/bsseq] [31]. Genomic coordinates of CpG islands

254   on all sequences (*i.e.*, unmasked sequence CpG island track) in the UMD3.1/bosTau6 assembly

255   were obtained from the UCSC Table Browser [32].

Non CpG methylation

256

257   Non-CpG methylation (CHH and CHG) present in the WGBS reads were analysed using

258   the methylKit R package [33]. De-duplicated bam files produced during alignment with Bismark

259   were sorted and saved as sam formatted files. Individual CHH or CHG were imported separately.

260   These files were imported to the methylKit object using strict criteria: at least $10\times$ coverage per

261   feature and the feature must be present across all samples. This resulted in a total of 685,311 and

262   284,641 features for CHH and CHG respectively, with a median coverage of $40\times$ for CHH and

13

263 41× for CHG. Median coverage between all samples was used to calculate a scaling factor to

264 normalize the coverage across samples. Differential methylation was determined for individual

265 features with an overdispersion parameter included (shrinkMN) and Benjamini-Hochberg (B-H)

266 FDR adjustment for multiple comparisons [34]. All downstream analyses were carried out

267 separately for CHH and CHG methylation. Average methylation between groups was tested

268 using a paired *t*-test.

Expression dynamics of genes associated with chromatin configuration and DNA

270 **methylation**

271   For the bAM samples used for the WGBS analysis at 24 hpi described here, differentially

272 expressed genes were previously identified in the same *M. bovis*-infected bAM, relative to the

273 non-infected bAM at 2, 6, 24 and 48 hpi using RNA-seq [23]. A comprehensive list (EPI-list) of

274 151 genes previously identified as being involved with histone modifications or DNA

275 methylation was generated from the literature (Supplementary File 2). RNA-seq transcriptomics

276 data was mined, at each time point, using the EPI-list. Ultimately, 86 genes were identified from

277 the RNA-seq data (Supplementary File 2). These 86 genes were denoted as genes of interest

278 (GOI) and their expression was determined at 2, 6, 24 and 48 hpi using the previously published

279 lists of differentially expressed genes ($P < 0.05$, B-H FDR-adjusted).

Gene ontology enrichment analyses

281   Gene ontology enrichment analyses were performed using the Bioconductor *GOseq*

282 software package [35] and the annotation package *org.Bt.eg.db*

283 (https://bioconductor.org/packages/org.Bt.eg.db). Notably, the probability weighting function

284 (PWF) supplied to *GOseq* was calculated without length bias for the analysis of promoters, as

14

285   those were defined in this study to a constant width of 2 kb (1.5 kb upstream and 500 bp

286   downstream of TSS).

287   **Results**

288   **WGBS summary statistics**

289   In summary, 16 individually barcoded WGBS libraries, prepared using bAM DNA

290   extracted from eight *M. bovis*-infected and eight non-infected samples, were sequenced on an

291   Illumina HiSeq 2500 sequencer in rapid run mode. This generated 45.3–67.4 million read pairs

292   per sample and ~32× sequencing depth per condition (*M. bovis*-infected and non-infected bAM).

293   These data satisfy previously defined criteria for WGBS, with respect to the number of

294   independent   biological   replicates   and   the   sequencing   depth

295   (www.roadmapepigenomics.org/protocols) [36]. An assessment of bisulfite conversion rates was

296   performed using non-CpG methylation according to Clark *et al.* [37]. Based on this approach,

297   conversion efficiencies were $> 99\%$ using CHH methylation values and $> 98\%$ using CHG

298   methylation values.

299   When all CpG dimers in the reference bovine genome (UMD3.1) were considered—

300   approximately 55.1 million stranded loci, including potentially unmappable CpG dimers—the

301   25.5–38.7 million aligned read pairs used for methylation calls led to an average 1.1–1.4

302   methylation call per individual strand-specific CpG dimer in each individual sample. As a result

303   of collapsing methylation calls as unstranded CpG loci, the even distribution of aligned read

304   pairs on both strands of the reference genome (Supplementary Table 3) doubled coverage to

305   2.12–2.78× per unstranded CpG dimer (~27.5 million unstranded loci). Unstranded methylation

306   calls were used from this point onwards. While the mean coverage of CpG dimers covered in at

307   least one sample was similar (2.13–2.81×; ~27.3 million loci), CpG dimers covered in all

308   samples was larger for each sample (4.6–5.9×; 4.7 million loci). Notably, the average coverage

309   of all CpG dimers in known CpG islands (CGIs)—including CpG dimers with null coverage—

310   was similar to those latter values (3.9–5.4×; 2.9 million loci), suggesting a consistent coverage of

311   CGIs across all samples.

312   **Genome-wide scan for differentially methylated regions 24 hpi**

313   An unbiased genome-wide scan was performed to identify potential differentially

314   methylated regions (DMRs), including only CpG loci with at least two methylation calls for at

315   least six of the eight biological replicates in each sample group, thereby ensuring at least 12×

316   coverage for any CpG dimer in both sample groups. As a comparison, the analysis was repeated

317   after randomising samples from both infection groups to produce a distribution of t-statistics

318   under the null hypothesis. The *bsseq* package was used to calculate t-statistics in a paired design

319   for both original and randomised sets of sample (Fig. 2A and 2B).

320   Potential DMRs were identified as genomic regions including at least three loci with

321   absolute t-statistics greater than 4.6 and a mean difference in methylation level (across samples

322   and loci) greater than 10% between the two groups. This analysis did not reveal significant

323   differences in methylation between infected and non-infected bAM; regions identified in the

324   original data were comparable to randomised data in number of regions identified and their

325   properties (*e.g.*, width, number of methylation loci, sum of t-statistics, proportion of regions

326   showing increased and decreased methylation level) (Supplementary Table 5).

327   **Distribution of methylation across different genomic regions**

328   Following the genome-wide scan, DNA methylation was determined at the following

329   defined functional genomic elements: gene bodies, intergenic sequences and proximal promoters.

330   Genomic elements were defined as described previously by Peat and colleagues [38] and the

17

331   number of regions compared in each category are outlined in Table 2. As expected for a

332   differentiated/somatic cell type, the majority of CpGs within intragenic sequences, gene bodies

333   and CpG-deficient promoters were widely methylated [39,40]. CpG-rich promoters containing a

334   CpG island (CGI) or overlapping a CGI (Promoter CGI and CGI promoter) were mostly

335   hypomethylated (Fig. 3). Interestingly, CGIs remote from annotated gene promoters (non-

336   promoter CGIs) showed variable methylation—most were hypermethylated (>75% methylated,

337   18,581 CGIs) with 9,103 non-promoter CGIs hypomethylated (<25%).

338   **Pyrosequencing validation of WGBS at key immune function genes**

339   To confirm the WGBS observation that DNA methylation was not different between

340   control and infected bAM, 24 hpi, a small panel of key immune genes, *TNF*, *IL12A*, *TLR2*,

341   *NFKB2*, *C1QB*, *NOS2* and *DTX4* were selected for targeted analysis by pyrosequencing.

342   Transcription of these genes has previously been shown to be upregulated in bAM 24 hpi with

343   *M. bovis* [21,23]; the specific loci that were analysed by pyrosequencing are detailed in Fig. 4. Four

344   of the loci are hypomethylated (*TNF*, *IL12A*, *TLR2* and *NFKB2*), one is intermediately

345   methylated (*NOS2*) and two are highly methylated (*C1QB*). Using DNA isolated from a parallel

346   set of control ($n = 4$) and infected ($n = 4$) bAM samples, average methylation levels at the

347   proximal promoter regions of *NFKB2*, *TLR2*, *IL12A* and *TNF*, and the gene bodies of *C1QB*,

348   *NOS2* and *DTX4*, were determined. Statistical analysis using a paired *t*-test did not reveal

349   significant differences ($P \geq 0.05$) in mean methylation levels between the examined loci of

350   infected and non-infected bAM samples (Fig. 5), supporting the WGBS observation that *M.*

351   *bovis* does not have an effect on the CpG methylation in bAM.

352

Promoter methylation level and gene ontology in bovine alveolar macrophages

353

354   Leveraging the absence of significant DMRs between infected and non-infected bAM

355 samples, methylation calls were pooled across all sixteen samples (eight *M. bovis*-infected and

356 eight controls) to analyse methylation levels in bAM gene promoters, with maximal coverage. In

357 this analysis, promoters were defined as regions spanning 1.5 kb upstream and 500 bp

358 downstream of each transcription start site (TSS), with a minimum of 10 CpGs each associated

359 with at least five methylation calls were included (26.8 million loci). Of the 24,616 genes

360 annotated in the bovine genome (Ensembl BioMart March 2016 archive), 22,964 were retained

361 for this analysis, on the basis that their promoter contained at least ten loci, each locus having at

362 least five methylation calls. For those genes, mean promoter methylation was estimated and

363 summarised, alongside average gene body methylation, in Fig. 2C.

364   Notably, the vast majority of gene promoters were found at either extreme of the

365 methylation range. Indeed, 18,438 promoters (80.3%) display methylation levels greater than

366 75% or lower than 25% (8,145 $\geq$ 75% methylated; 10,293 $\leq$ 25% methylated). However, 2,580

367 promoters (9.7%) displayed an average intermediate methylation level (IM, 33–66%). Strikingly,

368 gene ontology (GO) analysis of the genes associated with IM promoters (33–66%) using the

369 *GOseq* package [35] revealed a marked enrichment for immune-related GO categories including

370 "defense response" ($P < 10^{-08}$), "defense response to bacterium" ($P < 10^{-07}$), "response to

371 bacterium" ($P < 10^{-07}$), "chemokine-mediated signaling pathway" ($P < 10^{-06}$) and "chemokine

372 activity" ($P < 10^{-04}$), among others (Supplementary Table 6). In contrast, no significant

373 enrichment for immune-related GO categories was found for promoters with methylation levels

374 0–1% (759 promoters), 0–10% (5,605), 10–20% (3,255), or 90–99% (1,997) (Supplementary

375   Table 6). Instead, the latter only suggested enrichment for generic GO categories (*e.g.*,

376   "intracellular organelle", "transcription regulatory region DNA binding").

377   A hallmark of some imprinted genes is that they contain a 5' differentially methylated

378   region that is IM (resulting from parent-of-origin specific methylation patterns); therefore, the

379   IM   promoter   list   was   interrogated   for   known   bovine   imprinted   genes

380   (www.geneimprint.com/site/genes-by-species.Bos+taurus). This analysis confirmed IM at the

381   promoters of the following imprinted genes; *PLAGL1*, *SNRPN*, *MEST*, *PEG10*, *GNAS* and *NNAT*

382   (Supplementary Fig. 1).

Targeted analysis of intermediately methylated (IM) gene promoters

384   To confirm the presence of IM at immune gene promoters, COBRA and clonal analysis

385   of bisulfite PCR products was performed. Firstly, proximal promoter alignment plots for the IM

386   group were visually screened to remove promoters that were included due to averaging of

387   sequences with high and low methylation (example of this in Fig. 6). This analysis was restricted

388   to IM promoters containing a minimum of 30 CpGs (1,034 loci), to ensure sufficient CpG

389   coverage during COBRA and clonal bisulfite sequencing analysis. Of the 1,034 IM promoters,

390   267 promoters remained in the IM group and 60/267 (22.5%) of them had a promoter CGI

391   (Supplementary File 3). GO analysis of these 267 IM gene promoters with $\geq$ 30 CpGs revealed

392   enrichment for NADH dehydrogenase-associated activity (Supplementary Table 6). Two

393   immune-related gene promoters with the highest CpG content, *C1QB* and *IL2RA*, were selected

394   for further analysis (Fig. 7). Clonal analysis revealed that, although there are clearly

395   hypermethylated and hypomethylated *C1QB* and *IL2RA* alleles, the prominent allelic

396   methylation pattern is mosaic (Figs. 8 and 9); suggesting that the IM promoters analysed are

20

397   methylated in an allele-independent as opposed to an allele-specific pattern, an observation that

398   has been previously reported [41]. To further confirm our WGBS and clonal bisulfite sequencing

399   results we carried out COBRA on *C1QB* and *IL2RA* IM regions, using *M. bovis*-infected and

400   non-infected bAM (Fig. 8 and 9). Results from COBRA support our observation that the *C1QB*

401   and *IL2RA* proximal promoters were IM. Additionally, bovine sperm, kidney, liver and heart

402   samples were assessed using COBRA to determine whether IM might be tissue-specific.

403   COBRA indicated that *IL2RA* was almost completely methylated in sperm and predominantly

404   methylated in the kidney, liver and heart; suggesting a potential tissue-specific IM in bAM (Fig.

405   9). This possible tissue-specific IM pattern was not observed at the *C1QB* locus (Fig. 8).

406   **Non-CpG methylation analysis**

407   Overall, we found a low level of methylation in the context of CHH: mean values of

408   0.98% and 0.96% were estimated for control and *M. bovis*-infected bAM, respectively (Fig. 10).

409   CHH methylation was not different between control and infected-bAM (*t*-statistic 1.32,

410   df = 12.7, $P > 0.05$). Similarly, for CHG methylation we found an overall low mean methylation

411   of 1.53% and 1.49% for control and *M. bovis*-infected bAM, respectively (Fig. 10). There was no

412   difference for this mean methylation in CHH context between groups (*t*-statistic 1.26, df = 13.5,

413     $P > 0.05$). Neither clustering on all data nor top 5,000 most variable features revealed any

414   patterns in these data sets for CHH and CHG methylation. This was also concordant with

415   differential methylation tests showing no loci as significantly differentially methylated between

416   groups by a methylation difference greater than 1% and q-value = 0.01.

417

418

419 **Relationship between *M. bovis* infection and expression of chromatin and DNA modifiers**

420   Based on our WGBS results, bAM DNA methylation is not affected by infection with

421 *M. bovis* at 24 hpi; therefore, we next determined whether *M. bovis* infection has an effect on

422 chromatin. To do this, transcription analysis of chromatin and DNA modifying enzymes was

423 carried out using our previously published RNA-seq data from *M. bovis*-infected bAM [23] and a

424 similar approach to that detailed by Nestorov and colleagues [42]. A list of 151 genes (EPI-list) that

425 encode chromatin and DNA modifying enzymes was assembled from the literature

426 (Supplementary File 2). To identify chromatin and DNA modifying-associated genes that were

427 detected by RNA-seq, differentially expressed genes ($P < 0.05$, B-H FDR-adjusted) at each time

428 point were compiled and searched using the list of 151 known genes. This identified a list of 86

429 genes of interest (GOI). The number of GOIs was determined at each time point and the results

430 were as follows: 2 hpi 0/86, 6 hpi 8/86 (3 upregulated, 5 downregulated), 24 hpi 37/86 (16

431 upregulated, 21 downregulated) and 48 hpi 48/86 (19 upregulated, 29 downregulated)

432 (Supplementary File 2). *HDAC5*, *KDM2B*, *EZH1*, *PRDM2*, *SETMAR*, *SMYD4* and *USP12* were

433 differentially expressed at all time points post-infection (excluding 2 hpi).

434 **Discussion**

435   Here we present genome-wide DNA methylation profiles of bAM infected with *M. bovis*

436 versus non-infected controls at 24 hpi. We show that CpG methylation in bAM is not altered in

437 response to *M. bovis* at 24 hpi. Since previous studies suggest that DNA methylation changes are

438 established relatively late in the silencing pathway and are preceded by alterations to histone

439 modifications and chromatin packing [43], our results may reflect the early post-infection time

440 point examined in this study. Examination of the WGBS data, focusing on the relationship

441 between DNA methylation and proximal promoters, revealed an enrichment of gene promoters

442 that were intermediately methylated.

443   Global methylation patterns were first analysed using an unbiased genome-wide scan to

444 identify differentially methylated loci between *M. bovis*-infected and control bAM. Following

445 this, we examined the impact of infection on the bAM methylome in greater detail by assessing

446 DNA methylation at specific genomic features. Given the relationship between promoter

447 methylation, gene body methylation and transcription [44], these genomic features comprised the

448 main focus of these analyses. Promoters and CGIs were separated into the following categories

449 as previously described [38]: CGI promoters (*i.e.*, gene promoters overlapping a CpG island), non-

450 CGI promoters, promoter CGIs and non-promoter CGIs. None of these approaches revealed any

451 differentially methylated loci between *M. bovis*-infected and non-infected control bAM at 24 hpi.

452 Therefore, it is unlikely that the substantial transcriptomic perturbation observed in bAM during

453 the first 24 h of *M. bovis* infection [23] is due to reconfiguration of CpG methylation patterns. On

454 the other hand, the results presented here indicate that cell signalling and transcription factor-

455 driven gene regulatory transduction cascades lead to the rapid transcriptional activation of

456 immune- and other genes. This observation is supported by previous work showing that DNA

23

457   methylation changes in THP-1 macrophages infected with *M. tuberculosis* do not occur at CpGs

458   [14]. In their study, Sharma and colleagues demonstrated that methylation was perturbed at non-

459   CpGs. Similarly, Lyu *et al.* recently demonstrated that infection of human THP-1 macrophages

460   with virulent and avirulent *M. tuberculosis* is not associated with host DNA methylation changes

461   [45]. Unlike THP-1 macrophages infected with *M. tuberculosis*, we show that non-CpG

462   methylation is not altered in bAM infected with *M. bovis*. The lack of differences in non-CpG

463   methylation may be explained, in part, by the differences in cell types used—comparing a

464   macrophage-like human cell line (THP-1) to a primary, differentiated bovine macrophage.

465   Mycobacteria have recently been reported to modulate the host immune response through

466   chromatin modifications [46,47]. Given the absence of differential CpG methylation between non-

467   infected and *M. bovis*-infected bAM at 24 hpi, and the differential expression of genes encoding

468   chromatin modifiers observed in the current study, it is reasonable to hypothesise that chromatin

469   reconfiguration may have a role in regulating host gene expression in response to infection with

470   *M. bovis*.

471   To comprehensively annotate gene promoter methylation in bAM we quantified average

472   DNA methylation at proximal promoter regions spanning the TSS (1,500 bp upstream and

473   500 bp downstream). As expected, the majority of promoters containing or overlapping a CGI

474   were hypomethylated and those promoters not associated with CGIs were, generally, highly

475   methylated [48]. However, a large number of promoters (2,580) exhibited mean methylation levels

476   ranging between 33–66% (intermediately methylated; IM). Interestingly, in addition to this, gene

477   ontology analysis of the genes proximal to these promoters indicated a marked enrichment for

478   immune-function related categories. Further analysis of the IM promoter group revealed that

479   most promoters were included due to averaging of methylated and unmethylated CpGs within

24

480   the 2 kb promoter regions. After removing these promoters, 267 promoters remained that

481   exhibited IM. Validation experiments, using clonal analysis and COBRA, confirmed

482   intermediate DNA methylation at the proximal promoter of two of these non-imprinted IM

483   genes, *C1QB* and *IL2RA*. Six of the 267 promoters were proximal to known bovine imprinted

484   genes, displaying predominant intermediate methylation of 5' CGIs; as expected for imprinted

485   genes in an adult somatic cell type [49]. The remaining promoters are IM non-imprinted genes.

486   Previous work by Weber and co-workers demonstrated that, in somatic cells, the concentration

487   of CpGs within a gene promoter is related to the level of DNA methylation; promoters with a

488   high frequency of CpGs (HCP) tend to be unmethylated and promoters with a lower CpG content

489   (LCP) tend to be methylated [48]. Sixty of the 267 IM promoters identified in this study contained

490   high frequencies of CpGs (CGIs) normally associated with unmethylated HCPs, suggesting that

491   promoter IM in bAM is occurring irrespective of CpG density. It has been suggested that

492   intermediate DNA methylation is a conserved signature of genome regulation associated with

493   intermediately active rather than suppressed gene expression [41]. It is possible that these

494   intermediately methylated promoters are a hallmark of bAM and functionally associated with

495   this particular cell type. However, Elliot and colleagues demonstrated that different tissues and

496   cell types are intermediately methylated equally [41]; therefore, the function of intermediate

497   methylation at these genomic loci remains to be fully elucidated.

498   **Conclusion**

499   This is the first comprehensive analysis of the mammalian alveolar macrophage DNA

500   methylome in response to infection with a mycobacterial pathogen. Although the epigenome of

501   host bAM was not perturbed by a 24 h exposure to the pathogenic bacterium, *M. bovis*, this work

502   provides the first annotation of genome-wide DNA methylation patterns in the bovine genome

25

503   and is directly aligned with the goal of the Functional Annotation of Animal Genomes (FAANG)

504   project to '*produce comprehensive maps of functional elements in the genomes of domesticated*

505   *animal species*' [50]. Furthermore, this work also provides evidence for differential methylation at

506   the proximal promoter regions of more than 200 non-imprinted genes.

507

508

509   **Acknowledgments**

510   We would like to thank Felix Krueger of the Babraham Institute for advice on WGBS data

511   processing and quality control and for providing permission to use the Bismark Reports and

512   Babraham Bioinformatics logos, located in the supplementary information. In addition, we

513   would like to thank Han Haige for assistance with translation of a Chinese journal article.

514   **Author Contributions** Conceived and designed the experiments: AMOD, DAM, SVG and

515   DEM. Prepared the samples: DAM and JB. Performed the experiments: AMOD, DAM, SA and

516   REI. Provided RNA-seq data: NCN. Analysed the data: KRA, MWW, AMOD, SA and TJH.

517   Prepared the manuscript: AMOD, KRA and DEM. All authors read and approved the

518   manuscript.

519   **Disclosure of interest** The authors report no conflict of interest.

520   **Data availability** All WGBS data is available from NCBI GEO [51] (Accession Number

521   GSE110412).

522   **Figure Legends**

523   **Figure 1. Schematic representation of sample preparation.** Bovine alveolar macrophages

524   (bAM) were isolated, post-mortem, from the lungs of age-matched male Holstein-Friesian calves

525   by lavage. Purity of the cells was confirmed using flow cytometry with anti-CD14. Isolated cells

526   were washed and seeded for 24 h prior to infection. Infected bAM were exposed to *M. bovis* at a

527   multiplicity of infection ratio of 10:1 for 2 h. After 2 h the media was replaced in control and

528   infected samples and cells were harvested after 24 h for analysis of DNA methylation. This

529   figure was prepared by A.M.O'D. using the Biomedical PPT toolkit suite (www.motifolio.com).

27

530 **Figure 2. (A)** Distribution of paired *t*-statistics between *M. bovis*-infected and control non-

531 infected bAM samples based on smoothed WGBS data. **(B)** Distribution of paired *t*-statistics

532 between randomised samples based on smoothed WGBS data. **(C)** Distribution of average

533 methylation level (%) in promoters and gene bodies across all samples. Only CpG loci with

534 coverage greater or equal to 10 were considered. Only genes where gene body and promoter both

535 had 10 or more sufficiently covered CpG loci were considered.

536 **Figure 3. Distribution of DNA methylation in different genomic contexts in non-infected**

537 **and *M. bovis*-infected bovine alveolar macrophages (24 hpi).** Analysis of WGBS data from *M.*

538 *bovis*-infected and non-infected bovine alveolar macrophages (bAM) revealed that genomic

539 methylation, in the context of CpGs, was not altered at any of the sequence features outlined

540 (intergenic regions, gene bodies, or promoters with or without CpG islands (CGIs) in the host

541 following infection. Blue and red violins represent non-infected and *M. bovis*-infected bAM,

542 respectively.

543 **Figure 4. Schematic representation of WGBS data at loci related to immune function.**

544 WGBS proximal promoter plots. Representative plots showing the average methylation spanning

545 a 10 kb region at the 5' end of the *TNF*, *IL12A*, *TLR2*, *NFKB2*, *DTX4*, *C1QB* and *NOS2* genes.

546 The red and blue lines represent average methylation levels for infected and control samples,

547 respectively.

548 **Figure 5. Pyrosequencing validation of whole-genome bisulfite sequencing (WGBS) results.**

549 Locations of the pyrosequencing assays are denoted by 'PCR' in Fig. 4. Methylation was not

550 different at any of the loci tested (paired *t*-test $P \geq 0.05$) between *M. bovis*-infected and non-

551 infected control bAM. The number of CpG dinucleotides analysed at each loci were *TNF* (9

28

552   CpGs), *IL12A* (11 CpGs), *TLR2* (9 CpGs), *NFKB2* (8 CpGs), *C1QB* (4 CpGs), *DTX4* (3 CpGs)

553   and *NOS2* (1 CpG).

554   **Figure 6. Analysis of promoters with highly methylated and unmethylated sequence.** 1,034

555   proximal promoters, with a minimum of 30 CpGs, shown to be intermediately methylated (IM)

556   in the WGBS analysis were visually inspected to remove false positives. 767 were eliminated

557   from the IM group due to averaging of highly methylated and unmethylated CpGs in the

558   proximal promoter region (green dashed box). Four examples are presented here: the *GUCY2D*,

559   *SYN1*, *RRP1B* and *FOXRED2* genes. Red and blue lines represent average methylation levels for

560   infected and control samples, respectively.

561   **Figure 7. Gene promoters for combined bisulfite restriction analysis (COBRA) and clonal**

562   **analysis.** WGBS aligments at the *C1QB* and *IL2RA* gene loci. Each panel represents a 10 kb

563   region at the 5' end of the gene. Green dashed boxes illustrate the *C1QB* and *IL2RA* proximal

564   promoter regions (the TSS minus 1.5 kb, plus 500 bp) identified as intermediately methylated

565   (IM) during WGBS data analysis (average methylation 33–66%). PCR: region analysed using

566   bisulfite PCR, cloning and Sanger sequencing; CGIs: CpG islands; Gene: transcribed region;

567   Exons: shows the location of the first exon; Red line: *M. bovis*-infected bAM; Blue line: non-

568   infected control bAM.

569   **Figure 8. Confirmation of an intermediately methylated promoter region at the *C1QB* gene**

570   **locus. (A)** Clonal analysis of seven CpG dinucleotides in a 269 bp fragment of the bovine *C1QB*

571   5' promoter region, *a–d* represent sequencing of four biological replicates. Closed and open

572   circles denote methylated and unmethylated CpGs, respectively. **(B)** Aggregated representation

573   of methylation status at CpGs 1-7 in the *C1QB* proximal promoter region; (a)-(d) represent

29

574   animals A-D, numbers between boxes indicate genomic distance between CpGs while numbers

575   above boxes indicate the position of the CpG within the analysed region; BLUE = methylated,

576   BLACK = unmethylated, GREY = not present; **(C)** Schematic representation of the analysed

577   C1QB region and the recognition sites of *Aci*I and *Taqα*I as obtained by NEBcutter V2.0; length

578   is displayed in bp; **(D) – (F)** COBRA results of Uninfected **(D)**, Infected **(E)** and tissue samples

579   **(F)** digested with AciI, TaqαI or undigested (Ctrl).

580   **Figure 9**. **Confirmation of an intermediately methylated promoter region at the *IL2RA***

581   **gene locus. (A)** Clonal analysis of 10 CpG dinucleotides in a 378 bp fragment of the bovine

582   *IL2RA* 5' promoter region, *a–d* represent sequencing of four biological replicates. Closed and

583   open circles denote methylated and unmethylated CpGs, respectively. **(B)** Aggregated

584   representation of methylation status at CpGs 1-10 in the *IL2RA* proximal promoter region; *a-d*

585   represent animals A-D, numbers between boxes indicate genomic distance between CpGs while

586   numbers above boxes indicate the position of the CpG within the analysed region; BLUE =

587   methylated, BLACK = unmethylated, GREY = not present; **(C)** Schematic representation of the

588   analysed *IL2RA* region and the recognition sites of *Taqα*I as obtained by NEBcutter V2.0; length

589   is displayed in bp; **(D) – (F)** COBRA results of Uninfected **(D)**, Infected **(E)** and tissue samples

590   **(F)** digested with *Taqα*I or undigested (Ctrl).

591   **Figure 10. Non-CpG methylation levels differ but not significantly in bovine alveolar**

592   **macrophages (bAM) infected with *Mycobacterium bovis*.** Mean methylation at the non-CpG

593   contexts CHG (left) and CHH (right) for control and *M. bovis*-infected bAMs, respectively:

594   differences were not significant by *t*-statistic.

595

596 **Table 1** Primers use for targeted DNA methylation analysis.

Gene Name	Ensembl ID	Forward primer 5'-3'	Reverse Primer 5'-3'	Sequencing Primer 5'-3'
Tumour necrosis factor alpha	*TNF* ENSBTAG00000025471	AGTAATTGGTTTAGAGA AGTTTATTTAGAA	CTTCCTTAATAAAAAAACC CATAAACTCAT **BIOTIN***	GGTTTAGAGAAGTTTAT TTAGAAT
Interleukin 12A	*IL12A* ENSBTAG00000015150	TAATTAGAGAGTTAGGTT G GTTATTTATTG **BIOTIN***	ATAAAAATATAACCCCT AATTTAACCTCC	CAACCACCACCCTCA
Toll-like receptor 2	*TLR2* ENSBTAG00000008008	GGGGATGTTAGAGGATTT TAATTTTTGAT **BIOTIN***	ACCCCAACCCCCTC CTCC	CTAAACCACAAAAATTAC
Complement C1q B Chain (pyroseq)	*C1QB* ENSBTAG00000011196	GGGGGTTTTGGGTAA TGG	AACTAAACTAATCTCC TTTAAAACTCAC	GGAGATATTAGAGTAAA GGTT
Nitric oxide synthase 2	*NOS2* ENSBTAG00000006894	GGGGTTTGGTGTAG TTATTGT	CTACCTAATTCTAACCAC TAACCTCTACT **BIOTIN***	TGTGAAGGAGGAAGG
Deltex E3 ubiquitin ligase 4	*DTX4* ENSBTAG00000004046	GAAGTTTTAGAGTTAG GGTGGATATTAGTT	TCCCAATCCTCAACATCC TCTCAT **BIOTIN***	GTTAGGGTGGATATTA GTTT
Nuclear Factor Kappa-B, Subunit 2	*NFKB2* ENSBTAG00000006017	TTTGGTGGTGGGAG AGGT **BIOTIN***	CCTCCTCCCACCCTT ACC	ACCACCCAAAAATCTAA
Complement C1q B Chain	*C1QB* ENSBTAG00000011196	AGAATTTGAATTAGGGTT TTTGAT T	AAACACTTTCAAATCCC ATTTCTA	n/a
Interleukin 2 receptor subunit alpha	*IL2RA* ENSBTAG00000020892	TTAGGGTATTATGGTGAG AGAATTAAG	AAAAAAACAAAAAAATT CCCACTAC	n/a

597 * Denotes which primer in the pyrosequencing assay that was HPLC purified and 5'Biotinylated

Table 2 Number of tiles/regions in each category.

Region	Control	*M. bovis*
Intergenic	347,561	347,561
Gene Body	166,466	166,466
CGI Promoters	12,047	12,047
Non-CGI Promoters	13,413	13,413
Promoter CGIs	11,222	11,222
Non-Promoter CGIs	30,284	30,284

References

1   Garnier, T. *et al.* The complete genome sequence of *Mycobacterium bovis*. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 7877-7882, doi:10.1073/pnas.1130426100 (2003).

2   Olea-Popelka, F. *et al.* Zoonotic tuberculosis in human beings caused by *Mycobacterium bovis*-a call for action. *Lancet Infect. Dis.* **17**, e21-e25, doi:10.1016/S1473-3099(16)30139-6 (2017).

3   Kathirvel, M. & Mahadevan, S. The role of epigenetics in tuberculosis infection. *Epigenomics* **8**, 537-549, doi:10.2217/epi.16.1 (2016).

4   Holoch, D. & Moazed, D. RNA-mediated epigenetic regulation of gene expression. *Nat. Rev. Genet.* **16**, 71-84, doi:10.1038/nrg3863 (2015).

5   Jaenisch, R. & Bird, A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat. Genet.* **33 Suppl**, 245-254, doi:10.1038/ng1089 (2003).

6   O'Doherty, A. M. & McGettigan, P. A. Epigenetic processes in the male germline. *Reprod. Fertil. Dev.* **27**, 725-738, doi:10.1071/rd14167 (2015).

7   Yong, W. S., Hsu, F. M. & Chen, P. Y. Profiling genome-wide DNA methylation. *Epigenetics Chromatin* **9**, 26, doi:10.1186/s13072-016-0075-3 (2016).

8   Gonzalez, R. M., Ricardi, M. M. & Iusem, N. D. Epigenetic marks in an adaptive water stress-responsive gene in tomato roots under normal and drought conditions. *Epigenetics* **8**, 864-872, doi:10.4161/epi.25524 (2013).

9   Kucharski, R., Maleszka, J., Foret, S. & Maleszka, R. Nutritional control of reproductive status in honeybees via DNA methylation. *Science* **319**, 1827-1830, doi:10.1126/science.1153069 (2008).

10   Navarro-Martin, L. *et al.* DNA methylation of the gonadal aromatase (*cyp19a*) promoter is involved in temperature-dependent sex ratio shifts in the European sea bass. *PLoS Genet.* **7**, e1002447, doi:10.1371/journal.pgen.1002447 (2011).

11   O'Doherty, A. M. *et al.* Negative energy balance affects imprint stability in oocytes recovered from postpartum dairy cows. *Genomics* **104**, 177-185, doi:10.1016/j.ygeno.2014.07.006 (2014).

12   Waterland, R. A. *et al.* Season of conception in rural Gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.* **6**, e1001252, doi:10.1371/journal.pgen.1001252 (2010).

13   Esterhuyse, M. M., Linhart, H. G. & Kaufmann, S. H. Can the battle against tuberculosis gain from epigenetic research? *Trends Microbiol.* **20**, 220-226, doi:10.1016/j.tim.2012.03.002 (2012).

14   Sharma, G. *et al.* Genome-wide non-CpG methylation of the host genome during *M. tuberculosis* infection. *Sci. Rep.* **6**, 25006, doi:10.1038/srep25006 (2016).

15   Zheng, L. *et al.* Unraveling methylation changes of host macrophages in *Mycobacterium tuberculosis* infection. *Tuberculosis (Edinb)* **98**, 139-148, doi:10.1016/j.tube.2016.03.003 (2016).

16   Shell, S. S. *et al.* DNA methylation impacts gene expression and ensures hypoxic survival of *Mycobacterium tuberculosis*. *PLoS Pathog.* **9**, e1003419, doi:10.1371/journal.ppat.1003419 (2013).

17   Doherty, R. *et al.* The CD4(+) T cell methylome contributes to a distinct CD4(+) T cell transcriptional signature in *Mycobacterium bovis*-infected cattle. *Sci. Rep.* **6**, 31014, doi:10.1038/srep31014 (2016).

18   Marr, A. K. *et al. Leishmania donovani* infection causes distinct epigenetic DNA methylation changes in host macrophages. *PLoS Pathog.* **10**, e1004419, doi:10.1371/journal.ppat.1004419 (2014).

19   Sinclair, S. H., Yegnasubramanian, S. & Dumler, J. S. Global DNA methylation changes and differential gene expression in *Anaplasma phagocytophilum*-infected human neutrophils. *Clin. Epigenetics* **7**, 77, doi:10.1186/s13148-015-0105-1 (2015).

20   Sitaraman, R. *Helicobacter pylori* DNA methyltransferases and the epigenetic field effect in cancerization. *Front. Microbiol.* **5**, 115, doi:10.3389/fmicb.2014.00115 (2014).

21   Magee, D. A. *et al.* Innate cytokine profiling of bovine alveolar macrophages reveals commonalities and divergence in the response to *Mycobacterium bovis* and *Mycobacterium tuberculosis* infection. *Tuberculosis (Edinb)* **94**, 441-450, doi:10.1016/j.tube.2014.04.004 (2014).

22   Malone, K. M. *et al.* Comparative 'omics analyses differentiate *Mycobacterium tuberculosis* and *Mycobacterium bovis* and reveal distinct macrophage responses to infection with the human and bovine tubercle bacilli. *Microb. Genom.*, doi:10.1099/mgen.0.000163 (2018).

23   Nalpas, N. C. *et al.* RNA sequencing provides exquisite insight into the manipulation of the alveolar macrophage by tubercle bacilli. *Sci. Rep.* **5**, 13629, doi:10.1038/srep13629 (2015).

24   Vegh, P. *et al.* MicroRNA profiling of the bovine alveolar macrophage response to *Mycobacterium bovis* infection suggests pathogen survival is enhanced by microRNA regulation of endocytosis and lysosome trafficking. *Tuberculosis (Edinb)* **95**, 60-67 (2015).

25   O'Doherty, A. M. *et al.* DNA methylation dynamics at imprinted genes during bovine pre-implantation embryo development. *BMC Dev. Biol.* **15**, 13, doi:10.1186/s12861-015-0060-2 (2015).

26   O'Doherty, A. M. *et al.* DNA methylation plays an important role in promoter choice and protein production at the mouse *Dnmt3L* locus. *Dev. Biol.* **356**, 411-420, doi:10.1016/j.ydbio.2011.05.665 (2011).

27   O'Doherty, A. M., O'Shea, L. C. & Fair, T. Bovine DNA methylation imprints are established in an oocyte size-specific manner, which are coordinated with the expression of the DNMT3 family proteins. *Biol. Reprod.* **86**, 67, doi:10.1095/biolreprod.111.094946 (2012).

28   Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571-1572, doi:10.1093/bioinformatics/btr167 (2011).

29   Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357-359, doi:10.1038/nmeth.1923 (2012).

30   Zimin, A. V. *et al.* A whole-genome assembly of the domestic cow, Bos taurus. *Genome Biol.* **10**, R42, doi:10.1186/gb-2009-10-4-r42 (2009).

31   Hansen, K. D., Langmead, B. & Irizarry, R. A. BSmooth: from whole genome bisulfite sequencing reads to differentially methylated regions. *Genome Biol.* **13**, R83, doi:10.1186/gb-2012-13-10-r83 (2012).

32   Karolchik, D. *et al.* The UCSC Table Browser data retrieval tool. *Nucleic Acids Res.* **32**, D493-496, doi:10.1093/nar/gkh103 (2004).

33   Akalin, A. *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87, doi:10.1186/gb-2012-13-10-r87 (2012).

34   Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B Method.* **57**, 289-300 (1995).

35   Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14, doi:10.1186/gb-2010-11-2-r14 (2010).

36   Ziller, M. J., Hansen, K. D., Meissner, A. & Aryee, M. J. Coverage recommendations for methylation analysis by whole-genome bisulfite sequencing. *Nat. Methods* **12**, 230-232, doi:10.1038/nmeth.3152 (2015).

37   Clark, S. J. *et al.* Genome-wide base-resolution mapping of DNA methylation in single cells using single-cell bisulfite sequencing (scBS-seq). *Nat. Protoc.* **12**, 534-547, doi:10.1038/nprot.2016.187 (2017).

38   Peat, J. R. *et al.* Genome-wide bisulfite sequencing in zygotes identifies demethylation targets and maps the contribution of TET3 oxidation. *Cell Rep.* **9**, 1990-2000, doi:10.1016/j.celrep.2014.11.034 (2014).

39   Chen, Z. X. & Riggs, A. D. DNA methylation and demethylation in mammals. *J. Biol. Chem.* **286**, 18347-18353, doi:10.1074/jbc.R110.205286 (2011).

40   Deaton, A. M. *et al.* Cell type-specific DNA methylation at intragenic CpG islands in the immune system. *Genome Res.* **21**, 1074-1086, doi:10.1101/gr.118703.110 (2011).

41   Elliott, G. *et al.* Intermediate DNA methylation is a conserved signature of genome regulation. *Nat. Commun.* **6**, 6363, doi:10.1038/ncomms7363 (2015).

42   Nestorov, P., Hotz, H. R., Liu, Z. & Peters, A. H. Dynamic expression of chromatin modifiers during developmental transitions in mouse preimplantation embryos. *Sci. Rep.* **5**, 14347, doi:10.1038/srep14347 (2015).

43   Smith, Z. D. & Meissner, A. DNA methylation: roles in mammalian development. *Nat. Rev. Genet.* **14**, 204-220, doi:10.1038/nrg3354 (2013).

44   Jones, P. A. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484-492, doi:10.1038/nrg3230 (2012).

45   Lyu, L. N., Jia, H. Y., Li, Z. H., Liu, Z. Q. & Zhang, Z. D. [Changes and differences of DNA methylation in human macrophages infected with virulent and avirulent *Mycobacterium tuberculosis*]. *Zhonghua Jie He He Hu Xi Za Zhi* **40**, 509-514, doi:10.3760/cma.j.issn.1001-0939.2017.07.006 (2017).

46   Sharma, G., Upadhyay, S., Srilalitha, M., Nandicoori, V. K. & Khosla, S. The interaction of mycobacterial protein Rv2966c with host chromatin is mediated through non-CpG methylation and histone H3/H4 binding. *Nucleic Acids Res.* **43**, 3922-3937, doi:10.1093/nar/gkv261 (2015).

47   Yaseen, I., Kaur, P., Nandicoori, V. K. & Khosla, S. Mycobacteria modulate host epigenetic machinery by Rv1988 methylation of a non-tail arginine of histone H3. *Nat. Commun.* **6**, 8922, doi:10.1038/ncomms9922 (2015).

48   Weber, M. *et al.* Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat. Genet.* **39**, 457-466, doi:10.1038/ng1990 (2007).

49   Pervjakova, N. *et al.* Imprinted genes and imprinting control regions show predominant intermediate methylation in adult somatic tissues. *Epigenomics* **8**, 789-799, doi:10.2217/epi.16.8 (2016).

50   Andersson, L. *et al.* Coordinated international action to accelerate genome-to-phenome with FAANG, the Functional Annotation of Animal Genomes project. *Genome Biol.* **16**, 57, doi:10.1186/s13059-015-0622-4 (2015).

51   Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207-210 (2002).

Lavage

Trachea

Bronchi

Right lung
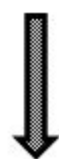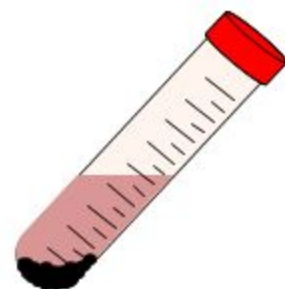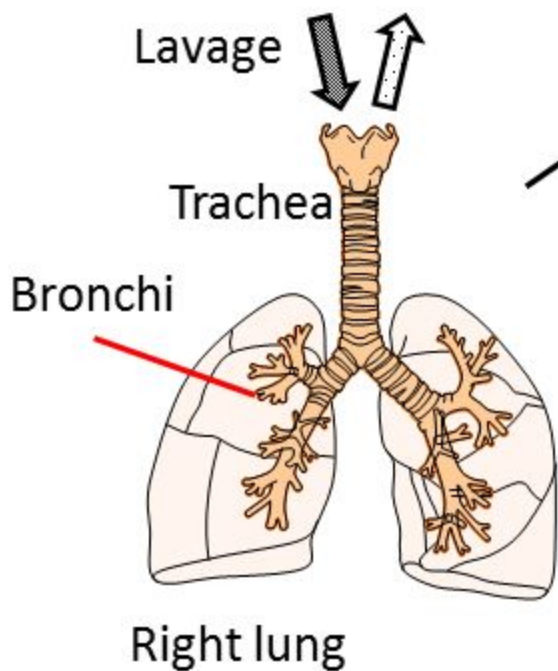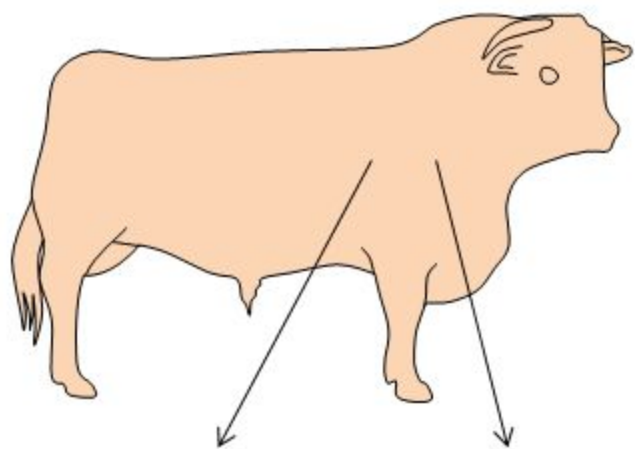
Wash and seed
Alveolar macrophages

Control   *M. bovis*

Whole Genome
Bisulfite Sequencing

Pyrosequencing

A

Distribution of *t*-statistics

N = 20225053   Bandwidth = 0.02148

B

Randomised samples

N = 20225053   Bandwidth = 0.02059

C

23: 27,536,175 − 27,538,174 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
TNF
Exons

1: 108,291,484 − 108,293,483 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
IL12A
Exons

17: 3,962,533 − 3,964,532 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
TLR2
Exons

26: 22,889,878 − 22,891,877 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
NFKB2
Exons
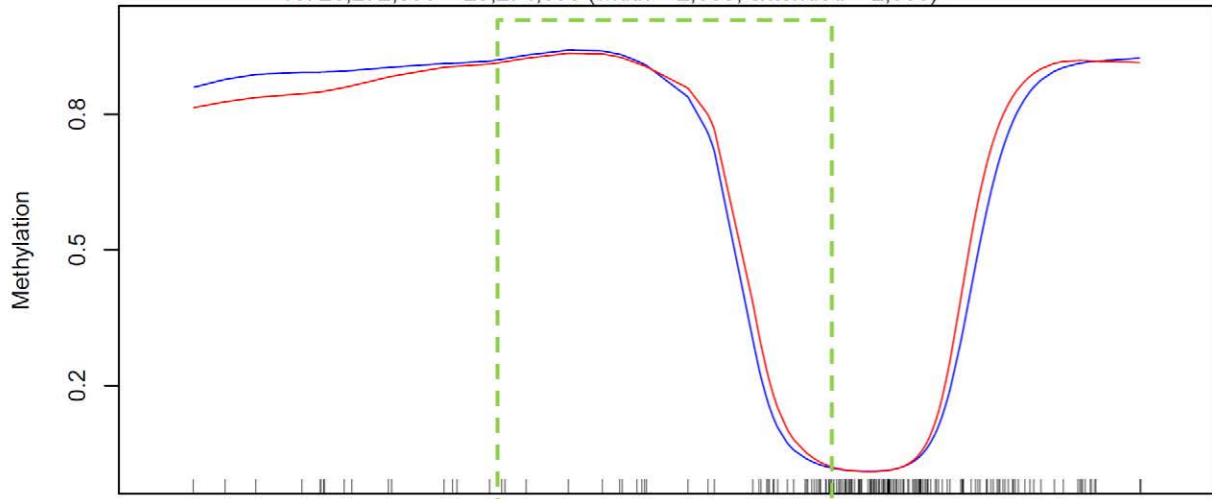
15: 83,629,277 − 83,631,276 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
DTX4
Exons

2: 130,775,412 − 130,777,411 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
C1QB
Exons

19: 19,956,050 − 19,958,049 (width = 2,000, extended = 4,000)

PCR
CGIs
Promoter
NOS2
Exons

2: 130,775,412 − 130,777,411 (width = 2,000, extended = 4,000)

13: 17,539,461 − 17,541,460 (width = 2,000, extended = 4,000)

A

(a) (b) (c) (d)

B

(a) (b) (c) (d)

C

1 — 378

TaqI   TaqI

D Control **E** Infected

Ctrl *Taqαl*   Ctrl *Taqαl*

600 bp
300 bp
200 bp
100 bp

F Sperm   Kidney   Liver   Heart

Ctrl *Taqαl*   Ctrl *Taqαl*   Ctrl *Taqαl*   Ctrl *Taqαl*

600 bp
300 bp
200 bp
100 bp