


RESEARCH ARTICLE

Open Access



Epigenetic analysis of high and low motile sperm populations reveals methylation variation in satellite regions within the pericentromeric position and in genes functionally related to sperm DNA organization and maintenance in *Bos taurus*

Emanuele Capra¹, Barbara Lazzari¹, Federica Turri¹, Paola Cremonesi¹, Antônia Moemia Rodrigues Portela², Paolo Ajmone-Marsan^{3,4}, Alessandra Stella¹ and Flavia Pizzi^{1*} 

Abstract

Background: Sperm epigenetics is an emerging area of study supported by observations reporting that abnormal sperm DNA methylation patterns are associated with infertility. Here, we explore cytosine-guanine dinucleotides (CpGs) methylation in high (HM) and low motile (LM) *Bos taurus* sperm populations separated by Percoll gradient. HM and LM methylation patterns were investigated by bisulfite sequencing.

Results: Comparison between HM and LM sperm populations revealed that methylation variation affects genes involved in chromatin organization. CpG Islands (CGIs), were highly remodelled. A high proportion of CGIs was found to be methylated at low/intermediate level (20–60%) and associated to the repetitive element BTSAT4 satellite. The low/intermediate level of methylation in BTSAT4 was stably maintained in pericentric regions of chromosomes. BTSAT4 was hypomethylated in HM sperm populations.

Conclusions: The characterization of the epigenome in HM and LM *Bos taurus* sperm populations provides a first step towards the understanding of the effect of methylation on sperm fertility. Methylation variation observed in HM and LM populations in genes associated to DNA structure remodelling as well as in a repetitive element in pericentric regions suggests that maintenance of chromosome structure through epigenetic regulation is probably crucial for correct sperm functionality.

Keywords: Sperm, Motility, Methylation, Epigenetic, Satellite

Background

Male infertility is a complex disorder affecting humans as well as other animals. Infertility is partially explained by physiological and biochemical factors, such as low sperm counts and poor sperm quality. The genetic basis of male infertility accounts for about 15% of infertile cases [1, 2]. The etiology of this disorder remains unclear both in

human and other species. For example, bulls considered of high-merit based on different sperm traits such as spermatozoa motility and morphology, are sometimes unable to produce successful full-term pregnancies [3, 4]. Different molecular parameters related to sperm nuclear and mitochondrial DNA, plasma membrane and lipid composition affect the ability of spermatozoa to fertilize oocytes and contribute to normal embryo development [5–7]. Therefore, much remains to be understood and novel molecular approaches may help to unravel the molecular basis of infertility.

* Correspondence: pizzi@ibba.cnr.it

¹Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, via Einstein, 26900 Lodi, Italy

Full list of author information is available at the end of the article



Among the known epigenetic processes in mammalian cells, DNA methylation has been identified as an important regulatory mechanism of genome function in normal embryonic development, X-chromosome inactivation and genomic imprinting [8, 9]. DNA methylation of the 5-carbon position in cytosine residues was reported to be predominantly present in cytosine-guanine dinucleotides (CpG) and especially in GC rich regions called CpG islands (CGIs) [10]. CGIs methylation in different genomic features impacts gene expression, i.e., promoter demethylation is associated with gene expression, while methylation in gene bodies influences splicing [11]. Methylation is also observed in Repetitive Elements (RE) of adult cells playing a role in the maintenance of chromosome structure and genome integrity [12].

Sperm epigenetic marks are unique, thus the factors that determine the patterns of DNA methylation differ between male germ cells and somatic cells. Although RE are highly methylated in both germ and somatic cells, elements from several subfamilies show different levels of methylation in the two cell types [13]. Centromeric regions in spermatogonia are known to be less methylated compared to somatic tissues [14]. This methylation pattern is supposed to play a role in germ-cell chromatin organization, rather than in the control of gene expression [15]. Most of the epigenetic signatures in germ cells are erased after conception from the morula stage to the blastocyst stage in the inner cell mass (ICM); successively, a sharp increase in the level of methylation in the embryo is observed following implantation [16, 17]. However, a proper regulation of epigenetic processes during spermatogenesis is necessary to ensure embryonic development in addition to sperm function. The level of DNA methylation of round spermatid was reported to be different from that of mature spermatozoa. Round spermatid rather than mature spermatozoa microinsemination was also observed to profoundly influence epigenetic marks in the embryo, thus affecting embryonic development and male fertility [18].

Aberrant locus specific or global methylation has been associated to abnormal semen parameters, as well as to male infertility. Oligospermic patients were reported to present a hypomethylation or unmethylation pattern at the H19 gene encoding insulin-like growth factor 2 (IGF2) imprinting control region 1 (ICR1); furthermore, hypermethylation at the Mesoderm-specific transcript (MEST) imprinted locus as well as a reduced sperm quality, as compared with normozoospermic men [19]. Broad DNA hypermethylation across many loci, also including the Satellite 2 repetitive element, was associated to poor sperm concentration and motility and to morphology alterations in abnormal human sperm [20]. The level of DNA methylation in human sperm, determined through an ELISA-like method, was correlated to conventional sperm parameters,

e.g. concentration and motility, as well as sperm chromatin and DNA integrity, but not to sperm viability and morphology [21]. DNA methylation in human spermatozoa was higher in low quality spermatozoa [22]. Pyrosequencing analysis of long interspersed elements (LINE) in human sperm, after bisulfite conversion, estimated an overall global methylation of about 75%, increasing with age. At the same time, targeted bisulfite sequencing of different selected genes showed a lower methylation level with a strong trend toward age-associated hypomethylation in some genomic regions [23]. Targeted bisulfite sequencing also revealed different levels of methylation in the promoter regions between high and low motile human sperm [24].

In farm animals, several studies showed altered sperm methylation to be associated with male infertility. DNA methylation pattern were found to be different between spermatozoa from high-fertile and sub-fertile buffalo bulls [25]. Recently, assessment of the epigenetic signature of bull spermatozoa using a human DNA methylation microarray [26] and Methyl-Binding Domain (MBD) Sequencing [27] revealed differentially methylated CpG sites and regions associated to bull fertility rate.

In the present study, the 5-methyl cytosine variations in CpGs were evaluated in high and low motility bull sperm populations following methyl enrichment and bisulfite sequencing approach. The objective is to produce a genome-wide methylation profile in the two populations, and to identify differential epigenetic signatures between high (HM) and low motile (LM) sperm.

Results

Isolation of spermatozoa and evaluation of sperm characteristics

Sperm cells were successfully fractionated into HM and LM populations; a significant ($P < 0.05$) improvement of several sperm quality parameters was observed in HM population in comparison to semen at thawing considering the following parameters: straight-line velocity VSL, curvilinear velocity VCL, average path velocity VAP and amplitude of lateral head displacement ALH variables [VSL ($\mu\text{m/s}$): 46.08 ± 4.11 , 61.24 ± 2.91 ; VCL ($\mu\text{m/s}$): 76.35 ± 6.02 , 110.37 ± 4.25 ; VAP ($\mu\text{m/s}$): 55.38 ± 4.27 , 74.02 ± 3.01 ; ALH (μm) 2.53 ± 0.14 , 3.72 ± 0.10 ; respectively in semen at thawing and HM population (Additional file 1).

Sequencing statistic and CpG methylation distribution

Since methylation levels in sperm are expected to be generally high [13], the Methyl-binding domain (MDB) approach was used to select hypermethylated regions [28]. Bisulfite sequencing was then applied to the methylation-enriched genomic fraction to investigate CpG methylation level at single base resolution in the highly methylated regions. To exclude the presence of technical biases caused

by unbalanced sequencing between groups of samples due to MBD enrichment, we evaluated cytosine coverage consistency between the HM and LM groups; no technical bias related to the combined approach (MDB and bisulfite sequencing) was observed (Additional file 2).

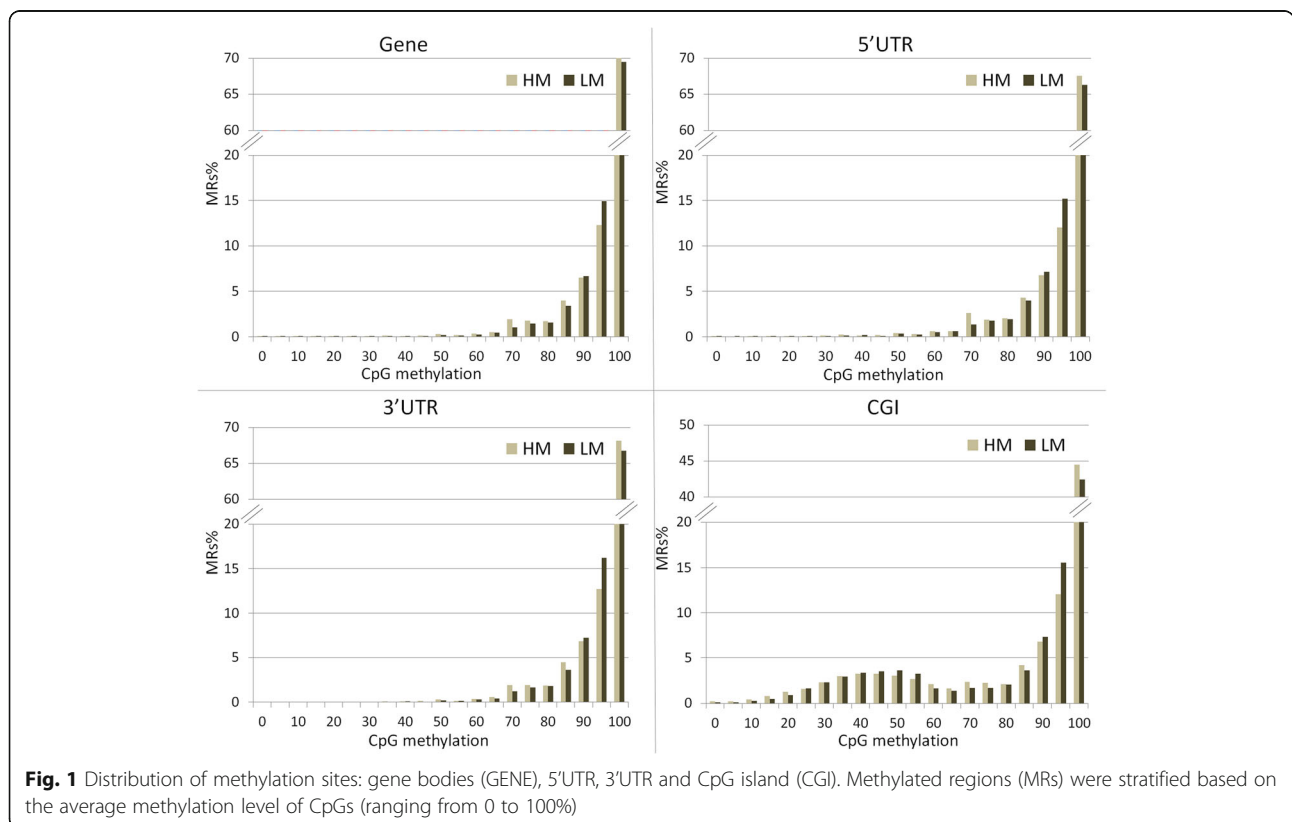
The average number of reads per sample was 28.1 M (range: 13.2 M–37.5 M). Mapping efficiency was high for all samples (83.1–90.6%). After calculating cytosine methylation conversion, a high percentage (93.7%) of the cytosines in the CpG enriched regions was methylated in both sperm populations (see Additional file 3 for statistics). After applying a threshold of at least 5X coverage per cytosine, a total of 26.6 M methylated regions (MR) (100 bp tiles with sliding window size of 100 bp) were identified spanning across the whole bovine genome. Among these, 1,086,748 methylated regions (MRs), shared between at least 3 out of 4 for both HM and LM sperm populations, were selected for DNA cytosine methylation profile comparison.

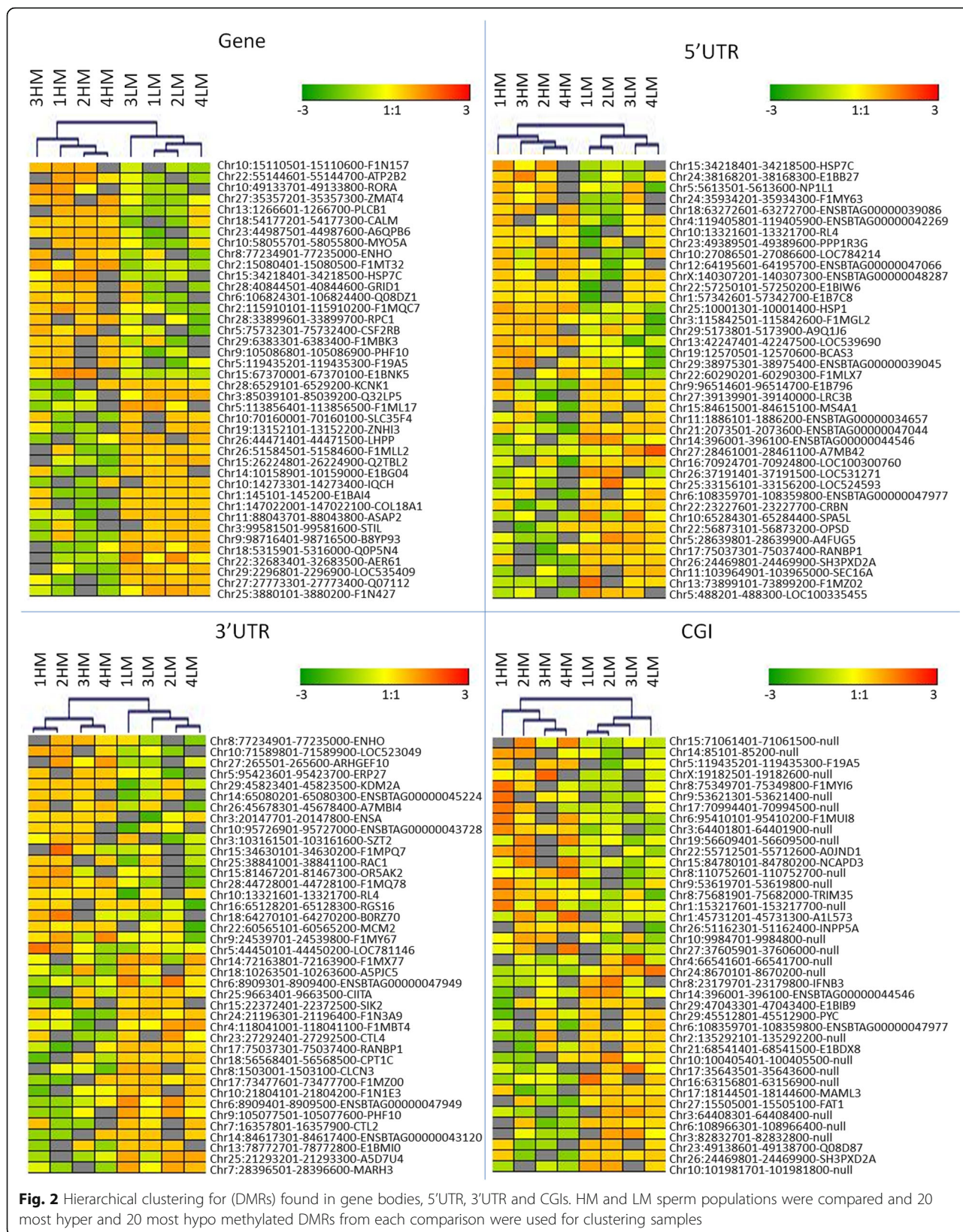
Among these, a total of 423,673 MRs mapped in 14,071 out of 23,970 annotated genes. Furthermore, 12,744 MRs mapped upstream (–2Kb) and 19,475 MRs downstream (+2Kb) of gene regions. A total of 9397 MRs were located within the 23,431 annotated CpG islands (CGIs) (Additional files 4, 5, 6 and 7). Gene bodies, 5' and 3' UTRs were prevalently hyper-methylated in both

sperm populations. Intriguingly, probes overlapping CGIs showed a peculiar distribution, with a relevant proportion of cytosines having an intermediate level of methylation (between 30 and 60%) (Fig. 1).

Differentially methylated regions between HM and LM sperm populations

A genome-wide analysis of genes and regulatory elements revealed that a small percentage of CpGs shows a significant variation in the methylation level (differentially methylated regions (DMRs)/MRs percentage) between HM and LM sperm populations in gene bodies (1.45%), 5' untranslated regions (5'UTRs) (3.12%) and 3'UTRs (2.72%). Considering CGIs, a higher proportion of the methylome (9.77%) was remodelled in HM vs LM sperm populations. Hierarchical analysis of the 20 most hyper and hypo methylated DMRs found in CGIs, in gene bodies, 5'UTR and 3'UTR well discriminated HM from LM samples (Fig. 2). Annotation of 6131 DMRs that overlapped gene bodies resulted in 3278 differentially methylated genes (DMGs) (Additional file 8). In addition, 398, 538 and 918 DMRs located near 5'UTR, 3'UTRs and CGIs, were close (\pm 2Kb) to 355, 484 and 297 DMGs, respectively (Additional files 9, 10 and 11). Gene ontology (GO) analysis was performed on genes found to be differentially methylated in 5'UTR, 3'UTRs





and CGIs, and on a selection of 423 genes differentially methylated in gene bodies (468 DMRs with false discovery rate (FDR) < 10exp-10). Variation in CpG methylation in different gene features and CGIs affect GO terms related to DNA replication, repair, and DNA and telomere organization and maintenance. In addition, GO terms related to hindbrain function, epithelia and endothelia migration metabolic processes were also observed to differ between HM and LM sperm populations. Unexpectedly, a large part of significant gene ontology terms are related to 3'UTRs, whereas only few terms are affected by CpG variation in 5'UTR (Table 1).

Methylation distribution in CpG islands

To further explore bovine sperm CpG methylation in CGIs, the global level of cytosine methylation was calculated in each CGI. Out of 23,431 CGIs annotated in the bovine genome, 3869 were detected (in at least 3 out of 4 samples for HM and LM) in our dataset. Based on CpG methylation level in CGIs (Fig. 1), profiles were grouped in two classes (20–60% and 80–100%), and distribution of CGIs length was calculated in each class in HM and LM sperm populations (Fig. 3). Differences in length distribution can be observed between the two classes, both in short (4–10 Kb, Fig. 3a) and long (10–240 Kb,

Table 1 GO terms identified for the differentially methylated genes (DMGs) found to differ between high motile (HM) and low motile (LM) sperm populations in gene bodies (GENE), 5' untranslated regions (5'UTRs), 3' untranslated regions (3'UTRs) and CpG islands (CGIs)

GO-ID	GO-Term	Associated Genes Found	P-Value*
GENE			
GO:0000723	Telomere maintenance	[ERCC4, LRIG1, PRKDC, TEP1, WRN]	1.88E-02
GO:0032200	Telomere organization	[ERCC4, LRIG1, PRKDC, TEP1, WRN]	1.92E-02
GO:0019722	Calcium, mediated multicellular organism signaling	[ASPH, HDAC4, ITPR1, KSR2, P2RX3, PLCE1]	2.13E-02
GO:0033555	Response to stress	[CACNA1B, GRIN2B, P2RX3]	2.46E-02
GO:0043954	cellular component maintenance	[ABL2, MTMR2, MTSS1]	2.81E-02
GO:0021575	Hindbrain morphogenesis	[ABL2, ATP2B2, DLC1, LDB2]	3.48E-02
GO:0031623	Receptor internalization	[CAV3, MTMR2, PICALM]	4.76E-02
5'UTR			
GO:0010634	Positive regulation of epithelial cell migration	[BCAR1, BCAS3, ENPP2, WNT7A]	3.70E-03
GO:0010595	Positive regulation of endothelial cell migration	[BCAR1, BCAS3, WNT7A]	5.16E-03
3'UTR			
GO:0035162	Embryonic haemopoiesis	[GATA3, KAT6A, KMT2A, PBX1, STK4]	1.31E-04
GO:0006516	Glycoprotein catabolic process	[FBXO6, GPC1, NEU4]	4.45E-03
GO:0046470	Phosphatidylcholine metabolic process	[LIPC, LPCAT3, PLA2G2E, SLC44A2, SLC44A4]	1.11E-02
GO:0051569	Regulation of histone H3-K4 methylation	[AUTS2, GATA3, KMT2A]	2.37E-02
GO:1901616	Organic hydroxy compound catabolic process	[IMPA1, LIPC, NUDT3]	2.83E-02
GO:0042439	Ethanolamine, containing compound metabolic process	[LIPC, LPCAT3, PLA2G2E, SLC44A2, SLC44A4]	3.05E-02
GO:0006303	Double, strand break repair via nonhomologous end joining	[KDM2A, PRPF19, WHSC1]	3.85E-02
GO:0071353	Cellular response to interleukin-4	[GATA3, MCM2, MCM7]	4.24E-02
GO:0032508	DNA duplex unwinding	[FBXO18, MCM2, MCM7, MRPL36]	4.29E-02
GO:0032392	DNA geometric change	[FBXO18, MCM2, MCM7, MRPL36]	4.31E-02
GO:0000726	non-recombinational repair	[KDM2A, PRPF19, WHSC1]	4.97E-02
CGI			
GO:0035637	Multicellular organismal signaling	[CACNA1C, DMRT3, DPP6, KCNQ1, NFASC, P2RX3]	8.53E-04
GO:0032288	Myelin assembly	[GPC1, NFASC, TENM4]	1.30E-03
GO:0006942	regulation of striated muscle contraction	[CACNA1C, KCNQ1, PDE5A, TNNT3]	3.22E-03
GO:0019226	transmission of nerve impulse	[DMRT3, DPP6, NFASC, P2RX3]	3.62E-03
GO:0032200	Telomere organization	[ERCC4, LRIG1, TERT, WRN]	3.80E-03
GO:0000723	Telomere maintenance	[ERCC4, LRIG1, TERT, WRN]	5.45E-03

Indicated are gene ontology IDs (GO-ID), gene ontology terms (GO-term), associated genes found and corrected p-values as determined by ClueGO (<http://apps.cytoscape.org/apps/cluego>)

* Term P-Value Corrected with Bonferroni step down

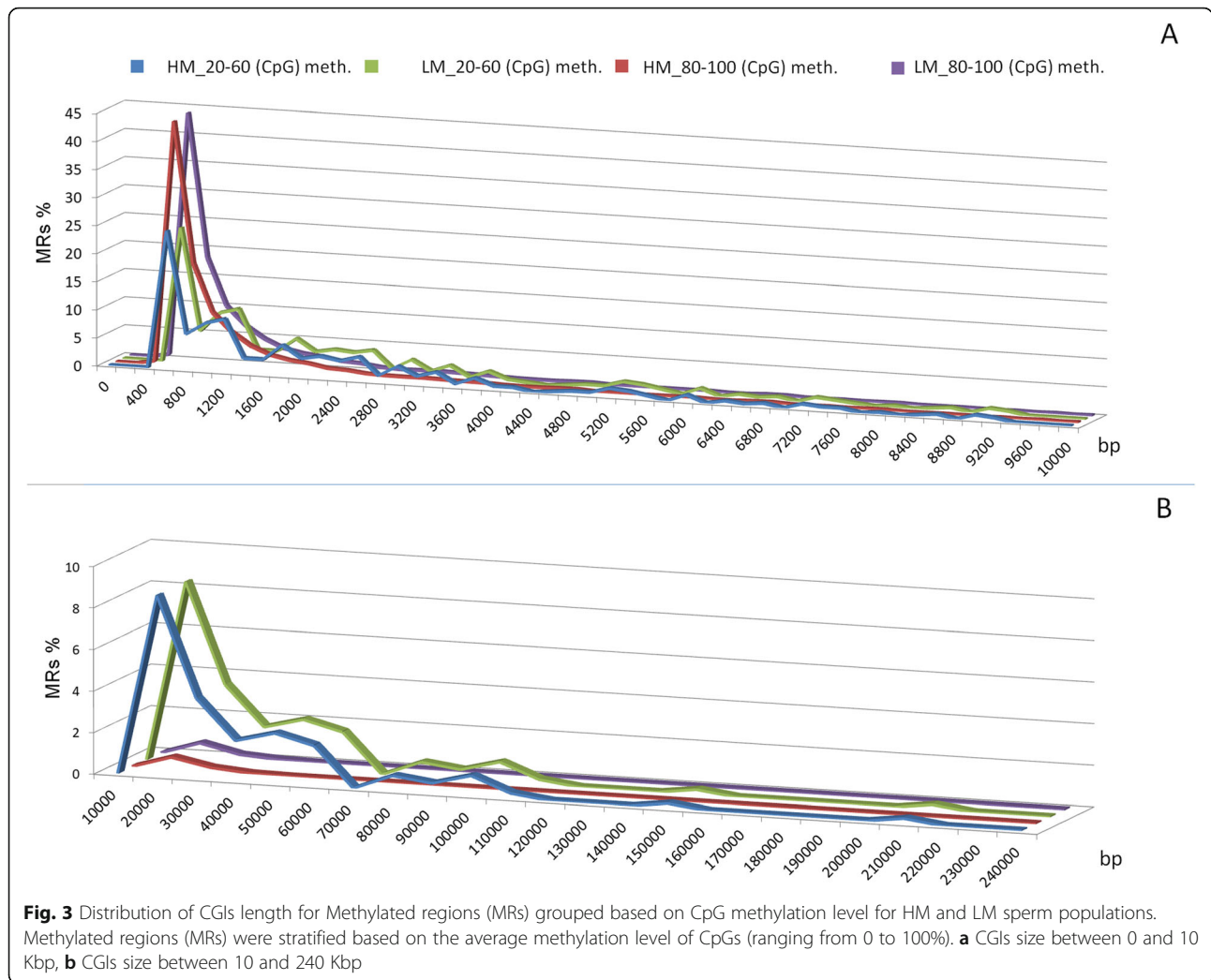


Fig. 3b) CGIs. Notably, 20–60% methylated CGIs exhibit peaks in length distribution that are not observable in highly methylated ones. This behaviour supports the hypothesis that genomic elements (i.e. repetitive elements motifs) present in single or multiple copies could be the target for methylation at low/intermediate levels. In particular, the 1400-bp peak observable in Fig. 3a is compatible with the expected length of Bovidae alpha satellites centromere repeats [29, 30].

Methylation distribution in repetitive DNA elements

Analysis of CGIs size distribution in low/intermediate methylated regions suggests that the atypical methylation profile observed is likely associated to repetitive DNA elements.

To further test this hypothesis, low/intermediate and highly methylated sequences were used as a query to the Database of repetitive rDNA elements Repbase. Database interrogation returned the bovine satellite BTSAT4 for about 75% of intermediate methylated sequences, whereas

the percentage of BTSAT4 in hypermethylated regions was close to zero. Other repetitive elements such as BTSAT3 and OSSAT2 satellites and BTLTR1 and ERV2–1-LTR transposable elements were also methylated prevalently at intermediate level (20–60%) (Table 2). Out of 2434 BTSAT4 elements annotated in the bovine genome 720 were detected (in at least 3 out of 4 samples for HM and LM) in our dataset.

Analysis of CpG methylation outlined an overall low level of BTSAT4 methylation in the HM sperm population. Considering 159 DMRs in the BTSAT4 regions, 122 were more methylated in LM sperm populations (Additional file 12) (Fig. 4).

To further support the hypothesis of a targeted methylation pattern of repetitive DNA elements, the genomic distribution of CGIs and BTSAT4 elements was observed. The overall CGIs covered by MBD sequencing represent about 1% of the total cow genome, while sequenced BTSAT4 regions represent 0.2%. The intersection of genomic absolute positions of CGIs and BTSAT4 repeats

Table 2 Frequency of occurrence for Repetitive Elements (REs) overlapping CGIs with different methylation levels (20–60% methylation and 80–100% methylation) in high motile (HM) and low motile (LM) sperm populations. Frequency of occurrence for REs is also reported for *Bos taurus* genome

HM				LM				GENOME	
CGIs methyl. 20–60%		CGIs methyl. 80–100%		CGIs methyl. 20–60%		CGIs methyl. 80–100%		Methyl. ref. genome	
RE type	%	RE type	%	RE type	%	RE type	%	RE type	%
BTSAT4	75.3	GC-rich	24	BTSAT4	75.9	GC-rich	23.5	Bov-tA2	8.6
SSU-rRNA	4.5	Bov-tA2	3.8	SSU-rRNA	4.6	Bov-tA2	4.1	ART2A	8.3
BTSAT2	3.8	MIRb	3.4	BTSAT2	3.9	MIRb	3.3	BovB	6.7
GC-rich	3	ART2A	2.6	GC-rich	3.3	C-rich	2.7	BOV-A2	5.1
OSSAT2	1.7	(TG)n	2.6	OSSAT2	1.7	ART2A	2.7	AT-rich	4.9
BTLTR1	1.2	C-rich	2.6	BTLTR1	1.2	(TG)n	2.6	Bov-tA1	3.7
ERV2-1-LTR	0.9	(CA)n	2.2	ERV2-1-LTR	0.9	(CA)n	2.2	MIRb	3.3
5S-rRNA	0.9	Bov-tA1	1.9	BTSAT3	0.8	Bov-tA1	2	MIR	2.5
BTSAT3	0.7	MIR	1.9	5S-rRNA	0.7	CHR-2A	1.8	L2a	2.2
LSU-rRNA	0.6	CHR-2A	1.8	LSU-rRNA	0.5	BovB	1.8	L1-2	2.1
G-rich	0.4	G-rich	1.7	BovB	0.4	MIR	1.7	L1	1.8
(CA)n	0.4	BovB	1.6	G-rich	0.4	BOV-A2	1.7	L2c	1.6
(CGGGG)n	0.4	BOV-A2	1.6	(CGGGG)n	0.4	G-rich	1.6	Bov-tA3	1.6
Bov-tA2	0.3	L2b	1.5	(CA)n	0.4	L2b	1.5	BTLTR1	1.5
BovB	0.3	L1-2	1.5	C-rich	0.3	L1-2	1.5	L2b	1.3
(TG)n	0.3	MIR3	1.4	(TG)n	0.2	CHR-2B	1.4	MIRc	1.2
(CGTG)n	0.3	CHR-2B	1.3	(CCCCAG)n	0.2	MIR3	1.4	MIR3	1.2
Bov-tA3	0.3	MLT1B	1.1	(CGG)n	0.2	MIRc	1.1	L1	1.2
C-rich	0.3	MIRc	1.1	L1-3	0.2	MLT1B	1.1	L1-3	1.1

shows that 70,6% of the BTSAT4 base pairs fall within CGIs, 80% of the BTSAT4 repeats being completely or partially located within CGIs (Additional file 13).

Discussion

In this work, the pattern of methylation in high and low motile bull sperm populations was determined using an enrichment step of methyl-CpG sequences combined with bisulfite sequencing.

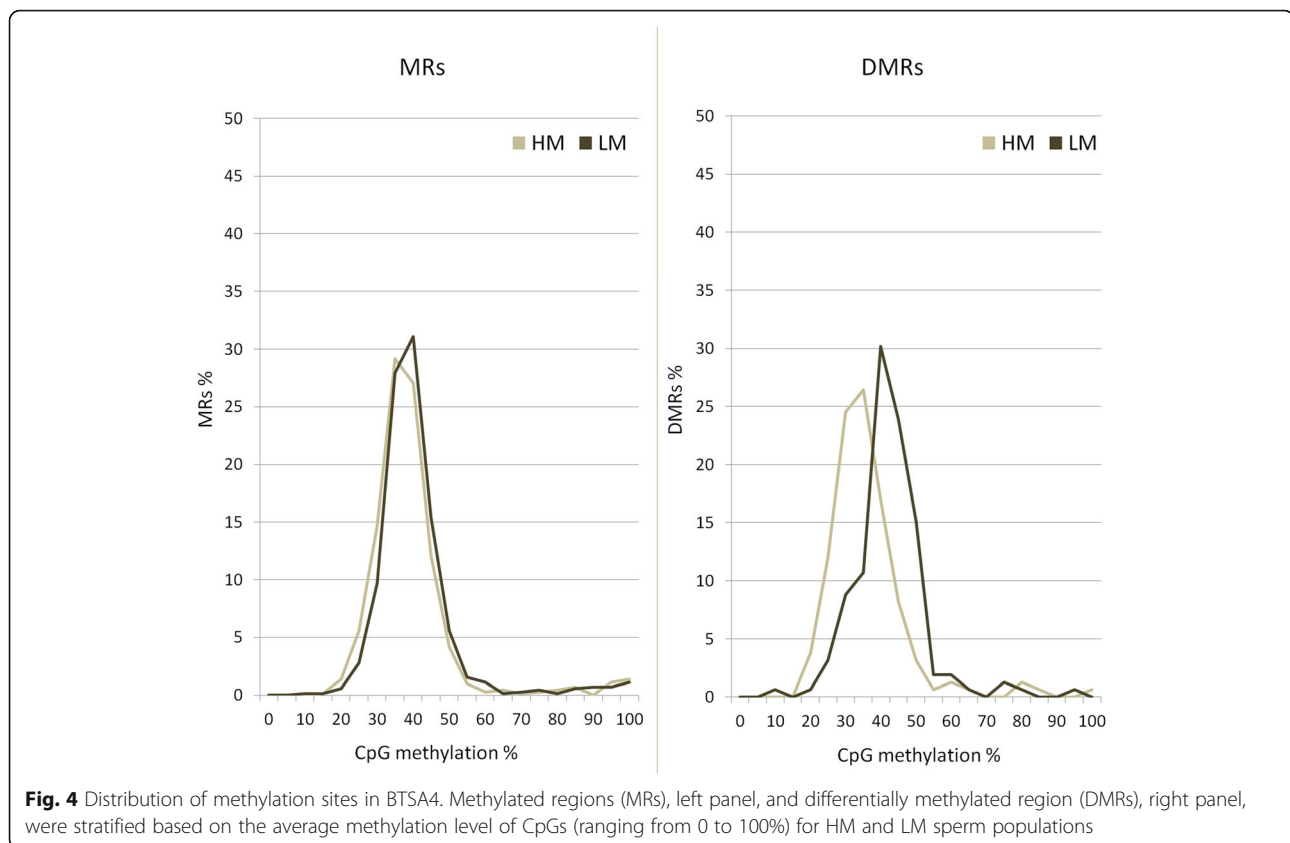
The comparison of different genomic features in HM and LM sperm populations revealed several differentially methylated regions flanking genes with a role in chromatin organization and maintenance. Undoubtedly, a correlation between sperm telomere length and abnormal sperm parameters exists [31]. Interestingly, HM and LM sperm populations showed variation in methylation of telomerase reverse transcriptase (TERT) and telomerase-associated protein 1 (TEP1). Genetic variants in both genes were previously reported to be associated with susceptibility to male infertility [32].

Differential methylation in 3'UTRs was also found in genes (histone lysine methyltransferases 2A (KMT2A), histone lysine demethylases 2A (KDM2A) and nuclear receptor-binding SET Domain 2 (NSD2)/ multiple myeloma

SET domain (MMSET)/ Wolf-Hirschhorn syndrome candidate-1(WHSC1)) influencing chromatin structure by epigenetic mechanisms, such as the regulation of histone H3-K4 methylation. Previous studies reported a strict association between sperm DNA methylation levels and both sperm chromatin condensation and DNA integrity, suggesting that the formation of a compact chromatin and proper DNA methylation are closely related events during spermatogenesis [21].

The NSD family of histone methyltransferase (HMT) comprises three members (NSD1, NSD2/ MMSET/ WHSC1, and NSD3/WHSC1L) that recognise lysine residue of histones H3 and H4 and mediate their methylation [33]. KMT2A (also known as mixed-lineage leukemia 1 (MLL1)) catalyzes the methylation of H3K4 [34, 35]. KDM2A, a Jumonji-C (JmjC)-domain containing histone demethylase (HDM), is a heterochromatin-associated protein that is required to maintain the heterochromatic state, it represses transcription of small non-coding RNAs that are encoded by clusters of satellite repeats at the centromere [36].

Histone Lysine methylation is tightly regulated by distinct families of conserved enzymes, KMTs and KDMs, which add and remove methyl groups at histone lysines



[37]. They play a role in orchestrating methylation of H3K9 and H3K27 in sperm. The methylation increases during meiosis, but the removal of H3K9me at the end of meiosis is essential for the onset of spermiogenesis [38]. In mice, the reduction of MLL2 activity results in a dramatic decrease of the number of spermatocytes by an apoptotic process and prevents spermatogenic differentiation [39]. Lysine-specific histone demethylase 1 (LSD1)/KDM1 is required for spermatogonial differentiation, as well as germ cell survival, in the developing testis [40]. An evolutionarily conserved pathway between histone H3-K9 methylation and DNA methylation exists in mammals, that is likely to be important to reinforce heterochromatic subdomains stability and to protect genome integrity. Suppressor of variegation 3–9 homolog (Suv39h) HMTases (also called KMT1A/B) are required to direct H3-K9 trimethylation and DNA (cytosine-5-)-methyltransferase 3 beta (Dnmt3b)-dependent DNA methylation to major satellite repeats at pericentric region [41].

A concurring variation in methylation of satellite repeats at pericentric region was observed in our dataset. A group of CGIs methylated at intermediate level (20–60%), located within genomic satellite repeats and in particular BTSAT4 Bovine satellite I [29] was observed to be less methylated in HM sperm populations. In the bovine genome BTSAT4 is likely to be the counterpart

of human alpha satellites, because both present in high-copy tandem repeat at centromeric position. Comparative analysis in hundreds of species shows a high variability in size for alpha satellites centromere repeats, e.g. approximately 171-bp in human and 1400-bp in Bovidae [30], which is in agreement with the size of repetitive element that we found to be methylated at intermediate level in bovine sperm. The location of bovine satellite I in all pericentric regions of *Bos taurus* autosomes was also confirmed by fluorescence in situ hybridization [42].

In accordance with our study, a low level of methylation of satellite DNA within pericentric regions was previously observed in primate sperm profiling [13]. Bovine alpha satellite I was observed to have low/intermediate methylation levels in sperm. Embryos obtained by somatic cell nuclear transfer (SCNT) presented a hypermethylation in the bovine alpha satellite I, expected to cause higher chromatin condensation compared to embryos generated by in vitro sperm fertilization (IVF). This may in turn contribute, either immediately or later in development, to the inefficiency of producing live offspring by SCNT [43, 44]. Low methylation levels have been also correlated with the ability to bind cohesin complexes that regulate the separation of chromatids at mitosis [45], suggesting a model in which selective hypomethylation of centromeric satellites might be critical

for accurate chromosome segregation during meiosis. Recently, methylation at satellite repeats throughout the genome has been observed to be increased in obese rat offspring [46]. Although obesity in human is associated with infertility by numerous studies [47], a direct link between satellite repeats methylation and sperm infertility is not yet described.

Conclusion

Methylation profiling in bovine semen revealed differential methylation of the BTSAT4 repetitive element in pericentric regions between HM and LM sperm populations. In addition, many DMRs were enriched in genes often functionally related to sperm DNA organization and maintenance. Together, alteration of methylation in satellite regions within the pericentric regions and in genes associated to lysine histone methylation, highlight that the complex mechanism that regulates DNA condensation during chromosomal packaging in sperm may affect sperm motility.

Methods

Isolation of spermatozoa and evaluation of sperm motility

Frozen semen straws from four mature progeny tested Holstein bulls with satisfactory semen quality were purchased from an Artificial Insemination AI center (INSEME S.P.A., Modena, Italy).

High Motile (HM) and Low Motile (LM) sperm populations were isolated through Percoll gradient as previously described [48]. Total motility and sperm kinetics parameters were assessed by CASA (Computer-Assisted Semen Analysis) system (ISAS[®] v1). Five μ l of semen pellet obtained after Percoll density gradient centrifugation were diluted in 5 μ l Tyrode's albumine lactate pyruvate (TALP) sperm medium [49] pre-warmed at 37 °C. Ten μ l of diluted semen was placed on a pre-warmed (37 °C) Makler chamber. During the analysis, the microscope heating stage was maintained at 37 °C. Using a 10 \times objective in phase contrast, the image was relayed, digitized and analyzed by the ISAS[®] software with user-defined settings as follows: frames acquired, 25; frame rate, 20 Hz; minimum particles area 20 μ m²; maximum particles areas 70 μ m². Spermatozoa speed was assigned to 3 broad categories: rapid (50 μ m/s), medium (25 μ m/s) and slow (10 μ m/s). CASA kinetics parameters were: total motility (MOT TOT, %), progressive motility (PRG, %), curvilinear velocity (VCL, μ m/s), straight-line velocity (VSL, μ m/s), average path velocity (VAP, μ m/s), linearity coefficient (LIN, %, = VSL/VCL \times 100), amplitude of lateral head displacement (ALH, μ m), straightness coefficient (STR, % = VSL/VAP \times 100), wobble coefficient (WOB, % = VAP/VCL \times 100) and beat cross frequency (BCF, Hz).

DNA extraction, library preparation and sequencing

Four HM and four LM sperm samples obtained in previous step were used for DNA extraction. DNA was isolated by NucleoSpin[®] Tissue (Macherey-Nagel) following manufacturer instruction. One μ g of genomic DNA was sonicated to produce DNA fragments of about 350 bp lengths. Methyl-binding domain (MBD) enrichment was performed using the MethylMiner[™] Methylated DNA Enrichment Kit (Thermo Fisher Scientific), following manufacture instruction. Libraries were generated using the TruSeq[®] DNA PCR-Free Library Preparation Kit (Illumina) including a step of bisulfite treatment. After adapters ligation, samples were converted with EpiTect Bisulfite Kits (Qiagen) and finally PCR amplified with KAPA HiFi Uracil+ (Kapa Biosystems) to obtain methyl enriched bisulfite libraries. The eight libraries were used for cluster generation and subsequently sequenced on a single lane of Illumina HiSeq2000.

Statistical analysis and bioinformatics

Data obtained from CASA were analyzed using the SAS[™] package v 9.4 (SAS Institute Inc.). The General Linear Model procedure (PROC GLM) was used to evaluate the efficiency of the sperm separation comparing semen quality parameters at thawing and in the HM population. The model included the fixed effect of the sperm population, and bull as random. Results are given as adjusted least squares means \pm standard error means (LSM \pm SEM).

Preliminary quality control of raw reads was carried out with FastQC (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Illumina raw sequences were then filtered with Trimmomatic [50] to remove adapters and low quality bases at the ends of sequence, using a sliding window approach. Data are available in the Sequence Reads Archive (SRA), (Accession Number SRP119411). Bismark software v.0.17.0 (<https://www.bioinformatics.babraham.ac.uk/projects/bismark/>) was used to align each read to a bisulfite-converted *Bos taurus* genome UMD311 with option -N 1, and methylation calls were extracted using the Bismark methylation_extractor function. Seqmonk software (version 0.34.1) was used for visualization and analysis of the Bismark output (<http://www.bioinformatics.babraham.ac.uk/projects/seqmonk/>). Only position with at least 5 cytosine were recorded in all samples, others were discarded from the data set. Cytosines count for each position ($n = 1,742,816$, count > 5X, present at least in 3 out of 4 in HM and LM samples) was determined and the Edge-R package was used to evaluate if over or under-representation occurs in our dataset between HM and LM groups.

Methylated regions (MRs) were detected genome wide by dividing the genome in 100 bp tiles and analyzing average methylation in a sliding window of 100 bp. MRs were considered if present at least in 3 out of 4 samples

in both HM and LM sperm populations. Methylation was calculated independently for different features: 5' UTR (-2Kb), 3' UTR (+2Kb), gene bodies and CpG islands (CGIs). MRs were also determined per CGI length classes and overlapping BTSAT4 REs. Reads count in BTSAT4 REs was determined and the Edge-R package was used to evaluate if over or under-representation occurs in our dataset between HM and LM groups. Differentially methylated regions (DMRs) between HM and LM populations were calculated using the logistic regression filter in R to assess differential methylation (FDR < 0.05, absolute cut-off of 5%). Hierarchical clustering was produced for DMRs present in CGIs, gene bodies, 5' UTRs and 3' UTRs. The level of methylation was normalized across samples and methylation percentage from a selection of DMRs showing the highest differences in methylation was used for clustering using the Genesis software [51].

Genes included in DMRs at CGIs and different genomic features were submitted to GO analysis. GO classification of the DMRs was performed according to canonical GO categories, using the Cytoscape plug-in ClueGO which integrates GO [52] and enhances biological interpretation of large lists of genes. Evaluation of REs in CGIs was performed by intersecting genomic positions of both features by Bedtools intersect (<http://bedtools.readthedocs.io>), thus frequencies for each RE category were calculated for low/intermediate methylation CGIs (20–60% methylation) and high methylation CGIs (80–100% methylation), in both HM and LM sperm populations and in *Bos taurus* genome.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s12864-019-6317-6>.

Additional file 1. Kinetics parameters evaluated on semen at thawing and in High Motile population: MOT TOT total motility, PRG cells progressive motility, VSL straight-line velocity, VCL curvilinear velocity, VAP average path velocity, LIN linear coefficient, STR straightness coefficient, WOB wobble coefficient, ALH amplitude of lateral head displacement, BCF beat cross-frequency. Results are given as adjusted least squares means \pm standard error means (LSM \pm SEM). a,b values within a row with different superscripts differ significantly at $P < 0.05$.

Additional file 2. Edge-R smear plot representing the Average Log Count Per Millions (CPM) and the abundance differences (logFC = log Fold Change) for the cytosine counts between HM and LM groups. Not significant differences (False Discovery Rate FDR < 0.05) between the two groups were observed (in red).

Additional file 3. Bismark sequencing statistics for the four replicates (1–4) of high motile (HM) and low motile (LM) sperm population.

Additional file 4. Methylated Regions (MRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations overlapping gene bodies. Columns report: Probe name, position (Chromosome, Start, End), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 5. Methylated Regions (MRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations upstream of genes (5'UTR). In each column are reported:

Probe name, position (Chromosome, Start, End), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 6. Methylated Regions (MRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations downstream of genes (3'UTR). In each column are reported: Probe name, position (Chromosome, Start, End), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 7. Methylated Regions (MRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations overlapping CpG islands (CGIs). In each column are reported: Probe name, position (Chromosome, Start, End), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 8. Differentially Methylated Regions (DMRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations overlapping gene bodies. In each column are reported: Probe name, position (Chromosome, Start, End), False Discovery Rate (FDR), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 9. Differentially Methylated Regions (DMRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations upstream of genes (5'UTR). In each column are reported: Probe name, position (Chromosome, Start, End), False Discovery Rate (FDR), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 10. Differentially Methylated Regions (DMRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations downstream of genes (3'UTR). In each column are reported: Probe name, position (Chromosome, Start, End), False Discovery Rate (FDR), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 11. Differentially Methylated Regions (DMRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations overlapping CpG islands (CGIs). In each column are reported: Probe name, position (Chromosome, Start, End), False Discovery Rate (FDR), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM).

Additional file 12. Differentially Methylated Regions (DMRs) found at least in three samples in both high motile (HM) and low motile (LM) sperm populations overlapping BTSAT4 satellite. In each column are reported: Probe name, position (Chromosome, Start, End), False Discovery Rate (FDR), Feature, protein ID, Description, MRs percentage for each sample (1HM, 2HM, 3HM, 4HM, 1LM, 2LM, 3LM, 4LM, average HM, average LM and differences between averages Δ HM-LM).

Additional file 13. Statistics for the distribution of CpG islands (CGIs) and repetitive element BTSAT4 satellite along the *Bos taurus* genome. The intersections of CGIs and BTSAT4 regions in function of length (bps), number (nr) and percentage (%) are reported.

Abbreviations

5-mC: 5 methyl cytosine; ALH: Amplitude of lateral head displacement; BCF: Beat cross frequency; CGIs: CpG islands; CpGs: Cytosine-guanine dinucleotides; DMG: Differentially methylated gene; DMR: Differentially methylated region; FDR: False discovery rate; GO: Gene ontology; HDMase: Histone demethylase; HM: High motile; HMTases: Histone methylases; IVF: In vitro fertilization; KDMs: Histone lysine demethylases; KMTs: Histone lysine methyltransferases; LIN: Linearity coefficient; LINE: Long interspersed elements; LM: Low motile; MBD: Methyl-binding domain; MOT TOT: Total motility; MR: Methylated region; PRG: Progressive motility; RE: Repetitive element; RRBS: Reduced Representation Bisulfite Sequencing; SCNT: Somatic cell nuclear transfer; STR: Straightness coefficient; VAP: Average path velocity; VCL: Curvilinear velocity; VSL: Straight-line velocity; WOB: Wobble coefficient

Acknowledgements

Not applicable.

Authors' contributions

EC, FP, and FT conceived the study. FT Isolated the spermatozoa fractions through Percoll gradient and evaluated sperm characteristics after separation. EC, AML, PC performed the DNA extraction, libraries preparation and sequencing. BL carried out the bioinformatic analysis. EC carried out GO analysis. EC wrote the manuscript and generated the figures. AS, PAM, PC and FP reviewed and approved the final manuscript. All authors have read and approved the final manuscript.

Funding

The research was supported by the Italian Ministry of Education, Universities and Research MIUR GenHome project "Technological Resort for the advancement of animal genomic research" and "Progetto Bandiera INTEROMICS - Sottoprogetto 1: Sviluppo di Infrastrutture di Bioinformatiche per le applicazioni OMICS in Biomedicina", and by the European Union's Horizon 2020 Research and Innovation Programme under the grant agreement n° 677353 IMAGE." The funding bodies did not play any roles in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials

All sequence data are deposited at the NCBI Sequence Read Archive (SRA) (<https://www.ncbi.nlm.nih.gov/sra>) (Accession Number SRP119411).

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Istituto di Biologia e Biotecnologia Agraria, Consiglio Nazionale delle Ricerche, via Einstein, 26900 Lodi, Italy. ²Department of Animal Science, Federal University of Ceará, Fortaleza, Brazil. ³Department of Animal Science, Food and Nutrition - DIANA, Università Cattolica del Sacro Cuore, Piacenza, Italy. ⁴Proteomics and Nutrigenomics Research Center - PRONUTRIGEN, Università Cattolica del Sacro Cuore, Piacenza, Italy.

Received: 17 January 2019 Accepted: 21 November 2019

Published online: 06 December 2019

References

- Gianotten J, Lombardi MP, Zwinderman AH, Lilford RJ, Van der Veen F. Idiopathic impaired spermatogenesis: genetic epidemiology is unlikely to provide a short-cut to better understanding. *Hum Reprod Update*. 2004;10:533–9.
- Tahmasbpour E, Balasubramanian D, Agarwal A. A multi-faceted approach to understanding male infertility: gene mutations, molecular defects and assisted reproductive techniques (ART). *J Assist Reprod Genet*. 2014;31:1115–37.
- Chenoweth PJ. Influence of the male on embryo quality. *Theriogenology*. 2007;68:308–15.
- Parkinson TJ. Evaluation of fertility and infertility in natural service bulls. *Vet J*. 2004;68:215–29.
- De Jonge C. Attributes of fertility spermatozoa: an update. *J Androl*. 1999;20:463–73.
- Dejarnette JM. The effect of semen quality on reproductive efficiency. *Vet Clin North Am Food Anim Pract*. 2005;21:409–18.
- Lewis SE. Is sperm evaluation useful in predicting human fertility? *Reproduction*. 2007;134:31–40.
- Razin A, Shemer R. DNA methylation in early development. *Hum Mol Genet*. 1995;4:1751–5.
- Heard E, Clerc P, Avner P. X-chromosome inactivation in mammals. *Annu Rev Genet*. 1997;31:571–610.
- Gardiner-Garden M, Frommer M. CpG islands in vertebrate genomes. *J Mol Biol*. 1987;196:261–82.
- Jones PA. Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet*. 2012;13:484–92.
- Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet*. 2007;16:2272–80.
- Molaro A, Hodges E, Fang F, Song Q, McCombie WR, Hannon GJ, Smith AD. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. *Cell*. 2011;146:1029–41.
- Marchal R, Chicheportiche A, Dutrillaux B, Bernardino-Sgheri J. DNA methylation in mouse gametogenesis. *Cytogenet Genome Res*. 2004;105:316–24.
- Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM. Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Dev Biol*. 2007;307:368–79.
- Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A. A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature*. 2012;484:339–44.
- Guo H, Zhu P, Yan L, Li R, Hu B, Lian Y, Yan J, Ren X, Lin S, Li J, Jin X, Shi X, Liu P, Wang X, Wang W, Wei Y, Li X, Guo F, Wu X, Fan X, Yong J, Wen L, Xie SX, Tang F, Qiao J. The DNA methylation landscape of human early embryos. *Nature*. 2014;511:606–10.
- Kishigami S, Van Thuan N, Hikichi T, Ohta H, Wakayama S, Mizutani E, Wakayama T. Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. *Dev Biol*. 2006;289:195–205.
- Poplinski A, Tuttmann F, Kanber D, Horsthemke B, Gromoll J. Idiopathic male infertility is strongly associated with aberrant methylation of MEST and IGF2/H19 ICR1. *Int J Androl*. 2010;33:642–9.
- Houshdaran S, Cortessis VK, Siegmund K, Yang A, Laird PW, Sokol RZ. Widespread epigenetic abnormalities suggest a broad DNA methylation erasure defect in abnormal human sperm. *PLoS One*. 2007;2:e1289.
- Montjean D, Zini A, Ravel C, Belloc S, Dalleac A, Copin H, Boyer P, McElreavey K, Benkhalifa M. Sperm global DNA methylation level: association with semen parameters and genome integrity. *Andrology*. 2015;3:235–40.
- Cassuto NG, Montjean D, Siffroi JP, Bouret D, Marzouk F, Copin H, Benkhalifa M. Different levels of DNA methylation detected in human sperms after morphological selection using high magnification microscopy. *Biomed Res Int*. 2016;2016:6372171.
- Jenkins TG, Aston KI, Pflueger C, Cairns BR, Carrell DT. Age-associated sperm DNA methylation alterations: possible implications in offspring disease susceptibility. *PLoS Genet*. 2014;10:e1004458.
- Du Y, Li M, Chen J, Duan Y, Wang X, Qiu Y, Cai Z, Gui Y, Jiang H. Promoter targeted bisulfite sequencing reveals DNA methylation profiles associated with low sperm motility in asthenozoospermia. *Hum Reprod*. 2016;31:24–33.
- Verma A, Rajput S, De S, Kumar R, Chakravarty AK, Datta TK. Genome-wide profiling of sperm DNA methylation in relation to buffalo (*Bubalus bubalis*) bull fertility. *Theriogenology*. 2014;82:750–759.e1.
- Takeda K, Kobayashi E, Akagi S, Nishino K, Kaneda M, Watanabe S. Differentially methylated CpG sites in bull spermatozoa revealed by human DNA methylation arrays and bisulfite analysis. *J Reprod Dev*. 2017;63:279–87.
- Kropp J, Carrillo JA, Namous H, Daniels A, Salih SM, Song J, Khatib H. Male fertility status is associated with DNA methylation signatures in sperm and transcriptomic profiles of bovine preimplantation embryos. *BMC Genomics*. 2017;18:280.
- Yong WS, Hsu FM, Chen PY. Profiling genome-wide DNA methylation. *Epigenetics Chromatin*. 2016;9:26.
- Gaillard C, Doly J, Cortadas J, Bernardi G. The primary structure of bovine satellite 1.715. *Nucleic Acids Res*. 1981;9:6069–82.
- Melters DP, Bradnam KR, Young HA, Telis N, May MR, Ruby JG, Sebra R, Peluso P, Eid J, Rank D, Garcia JF, DeRisi JL, Smith T, Tobias C, Ross-Ibarra J, Korf I, Chan SW. Comparative analysis of tandem repeats from hundreds of species reveals unique insights into centromere evolution. *Genome Biol*. 2013;14:R10.
- Biron-Shental T, Wiser A, Hershko-Klement A, Markovitch O, Amiel A, Berkovitch A. Sub-fertile sperm cells exemplify telomere dysfunction. *J Assist Reprod Genet*. 2018;35:143–8.
- Yan L, Wu S, Zhang S, Ji G, Gu A. Genetic variants in telomerase reverse transcriptase (TERT) and telomerase-associated protein 1 (TEP1) and the risk of male infertility. *Gene*. 2014;534:139–43.

33. Morishita M, Mevius D, di Luccio E. In vitro histone lysine methylation by NSD1, NSD2/MMSET/WHSC1 and NSD3/WHSC1L. *BMC Struct Biol.* 2014;14:25.
34. Milne TA, Briggs SD, Brock HW, Martin ME, Gibbs D, Allis CD, Hess JL. MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell.* 2002;10:1107–17.
35. Dou Y, Milne TA, Ruthenburg AJ, Lee S, Lee JW, Verdine GL, Allis CD, Roeder RG. Regulation of MLL1 H3K4 methyltransferase activity by its core components. *Nat Struct Mol Biol.* 2006;13:713–9.
36. Frescas D, Guardavaccaro D, Kuchay SM, Kato H, Poleshko A, Basrur V, Elenitoba-Johnson KS, Katz RA, Pagano M. KDM2A represses transcription of centromeric satellite repeats and maintains the heterochromatic state. *Cell Cycle.* 2008;7:3539–47.
37. Black JC, Van Rechem C, Whetstone JR. Histone lysine methylation dynamics: establishment, regulation, and biological impact. *Mol Cell.* 2012;48:491–507.
38. Boissonnas CC, Jouannet P, Jammes H. Epigenetic disorders and male subfertility. *Fertil Steril.* 2016;99:624–31.
39. Glaser S, Lubitz S, Loveland KL, Ohbo K, Robb L, Schwenk F, Seibler J, Roellig D, Kranz A, Anastasiadis K, Stewart AF. The histone 3 lysine 4 methyltransferase, MLL2, is only required briefly in development and spermatogenesis. *Epigenetics Chromatin.* 2009;2:5.
40. Myrick DA, Christopher MA, Scott AM, Simon AK, Donlin-Asp PG, Kelly WG, Katz DJ. KDM1A/LSD1 regulates the differentiation and maintenance of spermatogonia in mice. *PLoS One.* 2017;12:e0177473.
41. Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol.* 2003;13:1192–1200.42.
42. Nieddu M, Mezzanotte R, Pichiri G, Coni PP, Dedola GL, Dettori ML, Pazzola M, Vacca GM, Robledo R. Evolution of satellite DNA sequences in two tribes of Bovidae: a cautionary tale. *Genet Mol Biol.* 2015;38:513–8.
43. Couldrey C, Wells DN. DNA methylation at a bovine alpha satellite I repeat CpG site during development following fertilization and somatic cell nuclear transfer. *PLoS One.* 2016;8:e55153.
44. Zhang S, Chen X, Wang F, An X, Tang B, Zhang X, Sun L, Li Z. Aberrant DNA methylation reprogramming in bovine SCNT preimplantation embryos. *Sci Rep.* 2016;6:30345.
45. Parelho V, Hadjur S, Spivakov M, Leleu M, Sauer S, Gregson HC, Jarmuz A, Canzonetta C, Webster Z, Nesterova T, Cobb BS, Yokomori K, Dillon N, Aragon L, Fisher AG, Merckenschlager M. Cohesins functionally associate with CTCF on mammalian chromosome arms. *Cell.* 2008;132:422–33.
46. Youngson NA, Lecomte V, Maloney CA, Leung P, Liu J, Hesson LB, Luciani F, Krause L, Morris MJ. Obesity-induced sperm DNA methylation changes at satellite repeats are reprogrammed in rat offspring. *Asian J Androl.* 2016;18:930–6.
47. Katib A. Mechanisms linking obesity to male infertility. *Cent European J Urol.* 2015;68:79–85.
48. Capra E, Turri F, Lazzari B, Cremonesi P, Gliozzi TM, Fojadelli I, Stella A, Pizzi F. Small RNA sequencing of cryopreserved semen from single bull revealed altered miRNAs and piRNAs expression between high- and low-motile sperm populations. *BMC Genomics.* 2017;18:14.
49. Bavister BD, Leibfried ML, Lieberman G. Development of preimplantation embryos of the golden hamster in a defined culture medium. *Biol Reprod.* 1983;28:235–47.
50. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–20.
51. Sturn A, Quackenbush J, Trajanoski Z. Genesis: cluster analysis of microarray data. *Bioinformatics.* 2002;18:207–8.
52. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pagès F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* 2009;25:1091–3.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

