doi: 10.17221/78/2015-CJAS

Genome-wide analysis of CpG islands in some livestock genomes and their relationship with genomic features

A. Barazandeh[1,2], M.R. Mohammadabadi[1], M. Ghaderi-Zefrehei[3], H. Nezamabadi-pour[4]

[1]Department of Animal Science, Shahid Bahonar University of Kerman, Kerman, Iran
[2]Department of Animal Science, University of Jiroft, Jiroft, Iran
[3]Department of Animal Science, University of Yasouj, Yasouj, Iran
[4]Department of Electrical Engineering, Shahid Bahonar University of Kerman, Kerman, Iran

ABSTRACT: CpG islands (CGIs) are an important group of CpG dinucleotides in the guanine- and cytosine-rich regions as they harbour functionally relevant epigenetic loci for whole genome studies. As a matter of fact, since there has not been a formal comparative analysis of CGIs in domestic even-toed ungulate genomes, this study was performed to serve this comparison. The Hidden Markov Model was used to detect CGIs in the genomes. The results indicated that the CGIs number and CGI densities had scant variations across genomes. The goat genome had the highest number of CGIs (99 070), whereas the alpaca genome had the highest CGI density (43.39 CGIs/Mb). Significant positive correlations were observed among CGI densities with chromosome pair number, observed CpG/expected CpG, recombination rate, and gene density. When the size of chromosomes increased, the CGI densities decreased and a trend of higher CGI densities in the telomeric regions was observed. Only 10.96% of CGIs were methylated underscoring this postulation that the majority of CGIs remains to be unmethylated. The highest amount of the methylated CGIs was observed in the introns, intergenic, and coding (CDS) regions and the lowest amount of the methylated CGIs was observed in the promoter regions, implying that the DNA methylation of CGIs may control gene expression at the genome level. Detected differences between even-toed ungulate and other vertebrate genomes showed that CGI densities varied greatly among the genomes. These findings would contribute to better understanding the even-toed ungulate (epi) genomes, the role of CGIs in epigenomic functions and molecular evolution.

Keywords: even-toed ungulate; (epi) genomic; Hidden Markov Model; DNA methylation

INTRODUCTION

Since ancient times, domestic even-toed ungulates have played pivotal roles for man, being exploited for meat and milk, for fibre production, as beasts of burden for transport in agricultural/rural oriented community; they were even worshipped. These animals have served for traditional technologies since the very early era of domestication (Elsik et al. 2009; Archibald et al. 2010; Dong et al. 2013; Wu et al. 2014). In addition, some of domestic even-toed ungulates are now used as animal models in biomedical researches to investigate the genetic basis of complex traits, to produce transgenic peptide medicines, and to come up with some special form of antibody offering particular advantages in various medical and biotechnological applications (Kastelic et al. 2009; Jirimutu et al. 2012). DNA methylation is a type of chemical modification of DNA that can be inherited without changing the backbone of DNA sequence. It involves the addition of a methyl group to DNA, typically occurring at CpG dinucleotide, which can affect gene regulation (Irizarry et al. 2009; Wu et al. 2010; Jia et al. 2016). The DNA of most vertebrates, especially

doi: 10.17221/78/2015-CJAS

mammals, is depleted in CpG dinucleotides. The remaining CpGs clustering in DNA regions are generally referred to as CpG islands (CGIs). There has been growing interest in CGI because they are enriched with the promoters of genes (Hackenberg et al. 2010) and by altering DNA methylation in CGIs, they play important roles in the regulation of gene expression and gene silencing in biological processes such as X-chromosome inactivation, imprinting, silencing of intragenomic parasites (Su et al. 2010) and considerably, they might help discover the epigenetic causes of cancer (Han and Zhao 2009; Wu et al. 2010; Koh et al. 2016). Because of their crucial roles, multiple algorithms(either specific species or general purpose) have been developed to identify CGIs in the genomes. The first algorithm was proposed by Gardiner-Garden and Frommer (1987). In this algorithm, CGI is featured as a region of at least 200 bp, with a guanine and cytosine (GC) content greater than 50% and observed-to-expected CpG ratio (Obs_{CpG}/Exp_{CpG}) greater than 0.6 (Gardiner-Garden and Frommer 1987). Later on, many intelligent algorithms have been designed to circumvent disadvantages of the preceding algorithm (Wang and Leung 2004). These algorithms were mainly developed for studying the human genome, depended on cut-off values and left out important CGI clusters associated with epigenetic markers that were relevant to development and disease; therefore, they were not directly applicable to non-vertebrate genomes (Irizarry et al. 2009). To resolve these problems, Wu et al. (2010) have proposed an alternative Hidden Markov Model (HMM)-based approach that is flexible in the definition of a CGI and creation of CGI lists for other species. From 2009 onwards, the results of many whole genome sequencing projects in domestic even-toed ungulates have been released (Elsik et al. 2009; Archibald et al. 2010; Jirimutu et al. 2012; Dong et al. 2013; Wu et al. 2014). There have been some reports in terms of comparisons of CGIs and their correlation with genomic features in some mammalian genomes (mostly human and mouse) and non-mammalian genomes (fish) in the literature (Han et al. 2008; Han and Zhao 2008; Irizarry et al. 2009; Wu et al. 2010), but to the best of our knowledge, there has not been a thorough research on the detection of CGIs and their correlation with genomic features in domestic even-toed ungulates. Therefore, the main objective of this study was to analyze and compare CGIs and their correlation with genomic features in domestic even-toed ungulates. In this way, a systematic survey of CGIs in six domestic even-toed ungulate genomes was undertaken, including cow, goat, sheep, alpaca, dromedary, and bactrian camel. Also, the results on even-toed ungulates were compared with those on six other vertebrates (human, mouse, dog, horse, chicken, and zebrafish).

MATERIAL AND METHODS

Genome sequences and other genome information. Assembled genome sequences of even-toed ungulate genomes (cow, goat, sheep, alpaca, dromedary, and bactrian camel) were downloaded from the National Center for Biotechnology Information (NCBI) (the referenced accession numbers for each species were: GCF_000003205.7 (http://www.ncbi.nlm.nih.gov/assembly/GCF_000003205.7) (cow), GCF_000317765.1 (http://www.ncbi.nlm.nih.gov/assembly/GCF_000317765.1/) (goat), GCF_000298735.2 (http://www.ncbi.nlm.nih.gov/assembly/GCF_000298735.2) (sheep), JEMW00000000 (http://www.ncbi.nlm.nih.gov/nuccore/JEMW00000000) (alpaca), JDVD00000000 (http://www.ncbi.nlm.nih.gov/nuccore/JDVD00000000) (dromedary), and JARL00000000 (http://www.ncbi.nlm.nih.gov/nuccore/JARL00000000) (bactrian camel)). The genome information is given in Table 1. CGIs analysis of other vertebrate genomes (human, mouse, dog, horse, chicken, and zebrafish) was downloaded from the website (rafalab.Jhsph.edu/CGI/) which was predicted by Wu et al. (2010).

Algorithm for detection of CGIs. CGIs were identified based on HMM (Irizarry et al. 2009; Wu et al. 2010). HMM was implemented by applying an R software package called makeCGI. We first divided each genome into non-overlapping segments of length L (in bp), then we used the number of cytosine (C), guanine (G), and CpG in a segment of length L as model parameters and the hidden state $Y(s)$ for segment s with states $Y(s) = 1$ as CGI and $Y(s) = 0$ as baseline. Finally, we obtained posterior probabilities for genomic regions being CpG islands. The length segment L = 16 was chosen based on the association of CGI with epigenetic markers.

CGI mapping of different genomic regions. The method of Han et al. (2008) was used to identify CGIs in different genomic regions (genes, inter-

doi: 10.17221/78/2015-CJAS

genic regions, intragenic regions, and transcription start site (TSS) regions). Concisely, locations of CGIs were compared with the coordinates of different genomic regions based on the cow gene annotation information from the NCBI database by BEDTools (Quinlan and Hall 2010). CGIs overlapped with any genes were classified as gene-associated CGIs; CGIs whose total sequences were in intergenic regions were classified as intergenic CGIs; CGIs whose sequences were in gene regions were classified as intragenic CGIs, and CGIs overlapped with TSSs were classified as TSS CGIs.

Methylation data. To assess the methylation status of CpG islands, the MeDIP-seq methylation data of placental tissue downloaded from Su et al. (2014) study was used. Here, highly methylated regions (HMRs), which are also called peaks, were used to evaluate the methylation status in CGIs. 145 218 HMRs were obtained in the placenta. To detect methylated CGIs, the overlapping of CGIs with HMRs was calculated by BEDTools (Quinlan and Hall 2010) and those CGIs that overlapped with the HMRs were considered as methylated CpG islands.

RESULTS AND DISCUSSION

CGIs and CGI density in domestic even-toed ungulate genomes. Along with six domestic even-toed ungulate genomes, CGIs were detected by HMM. The genome information and statistics of the detected CGIs are summarized in Table 1. Due to the fact that the genome size varied across genomes, the CGI density was calculated. The number of predicted CGIs and correspondingly CGI densities varied between the genomes. The goat genome had the highest number of CGIs (99 070) and the alpaca genome had the highest CGI density (43.39 CGIs/Mb). Figure 1 shows box plots of the lengths of the detected CGIs for all genomes. The average length and variability of CGI across species were low. The goat genome had the shortest length and variety of CGI across genomes. Han et al. (2008) have analyzed CGIs in 9 mammalian genomes. They have found low diversity across mammalian genomes in terms of predicted CGIs and the CGI density. Our results are consistent with their results. The lower number of CGIs in camels may be due to lower genomes size in this than in other species (Table 1). According to the previous studies, there is a significant similarity between aforementioned genomes like a proximity in terms of phylogeny, syntenic maps, and orthologous genes (Elsik et al. 2009; Archibald et al. 2010; Jirimutu et al. 2012; Dong

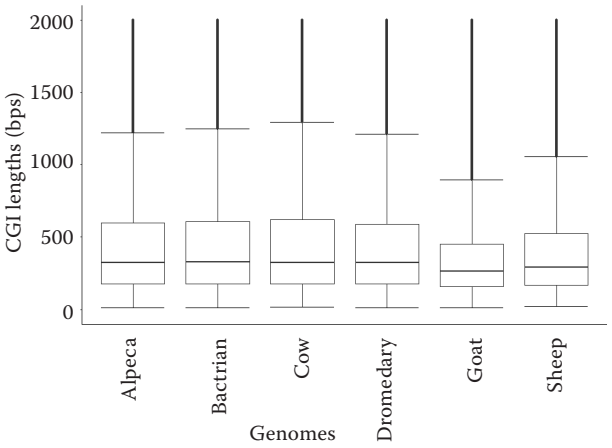

Figure 1. Lengths of CpG islands (CGIs) of domestic even-toed ungulate genomes by the Hidden Markov Model. Average length and variance of CGI between species changed fairly slowly

Table 1. CpG islands (CGI) and other genomic features in the domestic even-toed ungulate genomes

Genomes					CpG islands				
Species	size (Gb)	number of chromosome pairs	GC content[1] (%)	ObsCpG/ExpCpG[2]	number of CGIs	CGI density (Mb)	average length (bp)	GC content[1] (%)	ObsCpG/ExpCpG[2]
Cow	2.63	30	41.7	0.251	90 668	34.47	485	63.3	0.749
Sheep	2.59	27	41.8	0.253	95 542	36.89	420	62.2	0.729
Goat	2.52	30	41.8	0.251	99 070	39.31	361	61.2	0.730
Bactrian	2.01	37	41.3	0.266	86 417	42.99	488	60.1	0.685
Dromedary	2.01	37	41.2	0.259	84 854	42.22	482	59.7	0.686
Alpaca	2.05	37	41.4	0.261	88 949	43.39	489	60.6	0.694

[1]guanine and cytocine content

[2]observed/expected CpG ratio

doi: 10.17221/78/2015-CJAS

et al. 2013; Wu et al. 2014). Presumably, this is the reason of an almost identical number of CGIs among these genomes. Previous studies (Han et al. 2008; Weidner et al. 2014; Jung and Pfeifer 2015) have revealed that lifespan, body temperature, and body mass are related to CGIs. Thus, the records of body temperature, body mass, and lifespan retrieved from AnAge database (http://genomics.senescence.info/species/) as well as the correlation between CGI density and these traits were calculated for even-toed ungulate genomes. However, a significant correlation was not observed between CGI densities with lifespan (r = 0.75, P = 0.08), body temperature (r = –0.11, P = 0.83), and body mass (r = 0.78, P = 0.14). A high correlation was observed between CGI density and lifespan but it was not significant. This may be because of low divergence of lifespan among even-toed ungulate livestock.

Correlation between CGI density and other genomic features. A significant positive correlation (r = 0.87, P = 0.02) was discovered between CGI densities and the number of chromosome pairs (Figure 2A). Also, negative significant (P = 0.01) and non-significant (P = 0.07) correlation was found between CGI density with genome size (r = –0.94) and GC content (r = –0.79), respectively. The correlation between CGI density and Obs_{CpG}/Exp_{CpG} was positive and significant (r = 0.85, P = 0.03). The pattern of significant correlation between CGI densities in these genomes with some genomic features such as genome size, the number of chromosome pairs, GC content, and Obs_{CpG}/Exp_{CpG} was in agreement with Han et al. (2008). The CGI densities reported by Han et al. (2008) were quite lower than those detected in the current study, which might be due to different approaches used for CGI prediction. As a

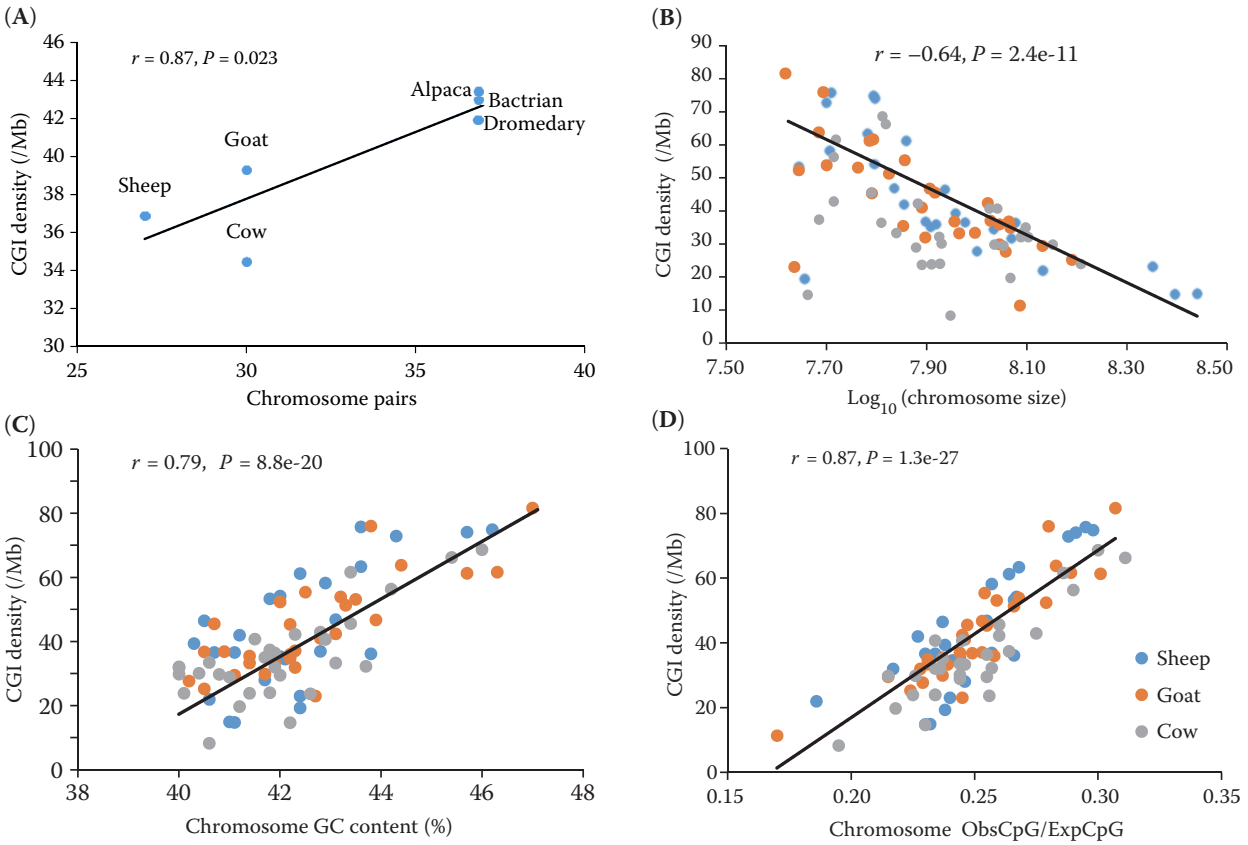

Figure 2. Correlation between CpG islands (CGI) densities with genomic features in even-toed ungulate genomes (**A**) CGI density (per Mb) vs the number of chromosomes pairs, (**B**) CGI density (per Mb) vs log10 (chromosome size), (**C**) CGI density (per Mb) vs chromosome guanine and cytosine content (%), (**D**) CGI density (per Mb) vs chromosome observed/expected CpG ratio

Substantial overlaps are seen between various species. Genomes of camels were excluded because of no data available at chromosome level for this species

doi: 10.17221/78/2015-CJAS

Table 2. CpG islands (CGI) densities in chromosomes grouped according to sizes in the domestic even-toed ungulate (cow, sheep, goat) genomes

Chromosome size (Mb)	Number of chromosomes	Mean of CGI density (Mb ± SD)
< 50	11	58.4 ± 31.9
50–100	47	47.2 ± 15.5
100–150	24	31.9 ± 6.9
> 150	5	20.6 ± 5.0
Total	87	42.84 ± 19.1

matter of fact that in this study we investigated the camel genomes in the form of the contig and scaffold format (data at chromosomal level was not available), the calculation of correlation between CGI density and other genomic features, e.g. the size of chromosome, was performed only for cow, goat, and sheep genomes. An extremely significant negative ($P = 2.4 \times 10^{-11}$) correlation (–0.64) was observed between CGI density and $\log_{10}$ (chromosome size) (Figure 2B). Also, an extremely significant positive correlation was discovered between CGI density with GC content ($r = 0.79$, $P = 8.8 \times 10^{-20}$) and Obs_{CpG}/Exp_{CpG} ($r = 0.87$, $P = 1.3 \times 10^{-27}$) of the chromosomes (Figure 2C, D, respectively). These very strong correlations observed between CGI densities and some genomic features at the chromosome level were in accordance with other studies (Han et al. 2008; Han and Zhao 2008). CGI density was significantly positively correlated with gene density ($r = 0.51$, $P = 5.4 \times 10^{-7}$), which was in agreement with the research of Han and Zhao (2009) in the dog genome ($r = 0.63$, $P = 8.0 \times 10^{-6}$). Table 2 shows the mean of CGI densities in grouped chromosomes based on sizes (< 50, 50–100, 100–150, and >150 Mb). When the size of chromosomes increased, the CGI densities decreased. The CGI density in the smallest chromosome group (< 50 Mb) was about three times greater than in the largest chromosome group (> 150 Mb). Previous studies on GC content and CGIs in the chicken genomes (Hillier et al. 2004; Rao et al. 2013) and in some mammalian genomes (Han et al. 2008) have revealed a high density of CGIs on microchromosomes and smaller chromosome groups, respectively. The results of the current research were consistent with these studies. An extremely significant correlation was discovered between gene-associated CGI densities with $\log_{10}$ (chromosome size) ($r = –0.54$, $P = 5.4 \times 10^{-8}$), GC content of the chromosomes ($r = 0.64$, $P = 3.8 \times 10^{-11}$), and Obs_{CpG}/Exp_{CpG} of the chromosomes ($r = 0.70$, $P = 4.6 \times 10^{-14}$). But the correlation between intergenic CGI densities with $\log_{10}$ (chromosome size) ($r = –0.28$, $P = 0.01$), GC content of the chromosomes ($r = 0.44$, $P = 1.8 \times 10^{-5}$), and Obs_{CpG}/Exp_{CpG} of the chromosomes ($r = 0.45$, $P = 9.9 \times 10^{-6}$) was quite low and less significant compared to gene-associated CGI densities. These findings support the motivated postulation that CGIs would function as markers of genes. The significant positive correlation between CGI density with GC content and gene density indicated that CGIs depend on both local genomic features and gene number (Han and Zhao 2009). Because of the high quality of cow's genes annotation with respect to other genomes, cow data were used to search for CGIs in different regions of the genome. Significant correlation between CGI densities in all genomic regions of cow and genomic features ($\log_{10}$ (chromosome size), GC content, and Obs_{CpG}/Exp_{CpG}) at the chromosome level are presented in Table 3. According to gene annotation in the NCBI database, 61 095, 29 613, 48 759, and 12 315 CGIs were overlapped with genes (gene-associated CGIs), intergenic regions (intergenic CGIs), intragenic regions (intragenic

Table 3. Correlation between CpG islands (CGI) density and genomic features

Genomic features	Gene-associated CGIs (n = 61 095)		Intergenic CGIs (n = 29 613)		Intragenic CGIs (n = 48 795)		Transcription start site CGIs (n = 12 315)	
	r	P	r	P	r	P	r	P
$\log_{10}$ (chromosome size)	–0.61	2.9×10^{-5}	–0.39	0.03	–0.48	0.01	–0.53	0.003
GC content[1]	0.91	4.1×10^{-12}	0.47	0.01	0.85	2.3×10^{-9}	0.87	3.2×10^{-10}
Obs_{CpG}/Exp_{CpG}[2]	0.94	3.1×10^{-14}	0.65	8.5×10^{-5}	0.83	1.8×10^{-8}	0.86	8.6×10^{-10}

[1]guanine and cytocine content

[2]observed/expected CpG ratio

doi: 10.17221/78/2015-CJAS

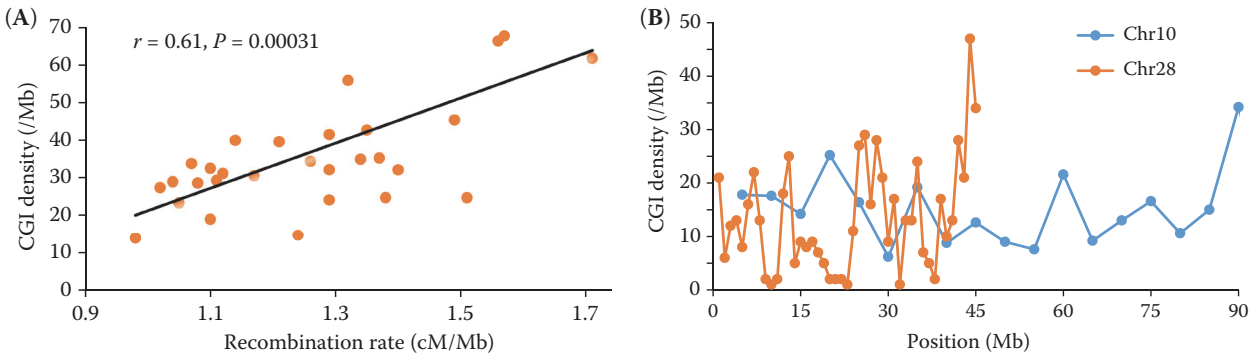

Figure 3. Correlation between CpG islands (CGI) density and recombination rate (cM/Mb) in the cow genome (**A**), distribution of CGI density (per Mb) on a long cow chromosome (Chr 10) and a short chromosome (Chr 28) (**B**). The data demonstrate a trend of higher CGI density in telomeric regions. A similar trend was found for other chromosomes

CGIs), and transcriptional start sites (TSS CGIs), respectively. Number of CGIs and corresponding CGI density in the intergenic region was remarkably lower than in intragenic and genes regions; in addition, they have shown lower significant correlation with other genomic features. These observations indicate that CGIs are considerable gene features in even-toed ungulate genomes. Other researchers have found similar results for CGIs in various genomic regions (Han et al. 2008; Han and Zhao 2009; Medvedeva et al. 2010).

CGI density and recombination rate. A set of recombination rate data on cow (window size, 1 Mb) was obtained from Weng et al. (2014). A meaningful correlation (r = 0.61, P = 0.0003) was detected between CGI density and recombination rate (Figure 3A). Because of recombination rate increases from centromeric toward telomeric regions (Han et al. 2008; Poissant et al. 2010; Weng et al. 2014), a trend of CGI density along chromosomes was obtained. Attractively, a trend of higher CGI density in the telomeric regions was observed (Figure 3B). This feature may be the reason of positive correlation between CGI density and recombination rate. Several GC related measures (GC-rich repeats, CpG dinucleotide sites, and CpG islands) were positively correlated with recombination rate in previous studies (Han et al. 2008; Tortereau et al. 2012; Rao et al. 2013). The same result was obtained in this study. Also, recombination rates are related to the distance from the centromere (Romiguier et al. 2010; Paape et al. 2012); in the current study, a trend of higher CGI density in the telomeric regions was observed. It is difficult to reveal the reason of the observed relationships between GC related measures such as CGIs and recombination rate. Further analysis of the mechanisms underlying recombination is needed to identify the related molecular mechanism.

CGIs and DNA methylation. Among a total of 90 668 CGIs detected in cow genome, only 9942 (10.96%) were methylated in the placenta. These results specified that most CGIs are unmethylated. In the study of Su et al. (2014) in cattle, 20.12% of CGIs were methylated, which could be probably due to the different approaches used to detect CGIs. When comparing the status of methylated CGIs investigated in different regions of the genome, most of the methylated CGIs were present in the introns, intergenic and CDS regions and the lowest number of the methylated CGIs existed in the promoter regions (Figure 4). A genome-wide profile of DNA methylation in CGIs has been reported for many organisms (Du et al. 2012; Kwak et al. 2014), which confirms the obtained results in the cur-

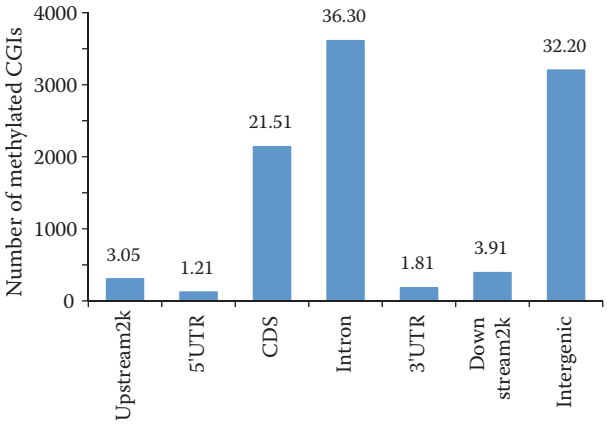

Figure 4. Methylated CpG islands (CGIs) distribution in different regions of cow genome. Percentage of methylated CGIs in each region is presented in the columns

doi: 10.17221/78/2015-CJAS

rent study. The CGIs in the gene body (CDS plus introns) are relatively highly methylated, whereas CGIs in the promoter (around TSSs) remain hypomethylated, a finding that is consistent with human (Du et al. 2012) and chicken (Hu et al. 2013). The results of Hu et al. (2013) have shown that the average methylation density was the lowest in CGIs followed by promoters and within the gene body, however, the methylation density of introns was high in comparison with the regions. Also, methylated CGIs were prominently distributed in the intergenic regions. Our results suggested that the DNA methylation of CGIs may control gene expression at the genome level.

Comparison of predicted CGIs in even-toed ungulate genomes and other vertebrate genomes. To detect the information on the discrepancy of CGI density between even-toed ungulate and other vertebrate genomes, CGI density was scanned in six vertebrate genomes, including the human, mouse, horse, dog, chicken, and zebrafish genomes (Figure 5). CGI density ranged from 23.05 (human) to 85.17 (zebrafish). Variation of CGI density between even-toed ungulates and chicken and other mammals was not substantial. In mammals, the dog genome had the largest CGI density in comparison with other mammals and especially even-toed ungulates. Han and Zhao (2009) studied the contradictory features of CpG islands in the promoter and other regions in the dog genome. They revealed a remarkably higher CGI density in the dog genome than in the human and mouse genomes. They showed that dog genome had fewer promoter-associated CGIs than the human and

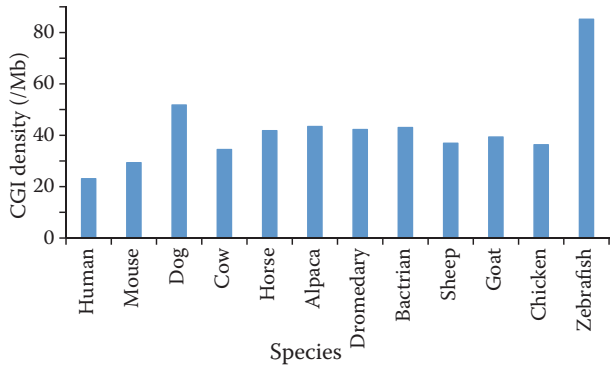

Figure 5. Comparison of CpG islands (CGI) density (per Mb) in various vertebrate genomes. Results of six even-toed ungulates obtained from this study and results of other six vertebrates retrieved from website (rafalab.Jhsph.edu/CGI/). CGI density varied between genomes

mouse; also, the abundance of CGIs in the dog genome was largely supported by the noncoding regions including the intergenic and intronic regions. The human genome had the lowest CGI density and the zebrafish genome had the largest CGI density in comparison with even-toed ungulate genomes. CGIs are progressively depleted from fish to mammals. This is mostly consistent with the idea of a greater degree of CGIs losses with increasing organism complexity (Irizarry et al. 2009). Variation in the number of CGIs and CGI density in warm-blooded vertebrates such as mammals is low, but in cold-blooded ones such as fish it is quite high (Han et al. 2008). Fish genome had very different CGIs number and CGI density. In the study of Han and Zhao (2008) on fish genomes, it was found that the number of CGIs and the CGI density varied greatly across four fish genomes. They concluded that this feature might be caused by genetics (sequence composition evolution) and environmental factors such as water temperature, the speed of flow, the light intensity at different water depths during the long evolutionary period after the divergence of the common ancestor of fishes.

CONCLUSION

A systematic comparative genomic analysis of CGIs and CGI density was conducted in the domestic even-toed ungulate genomes. The number of CGIs and the CGI density showed low variations in these genomes. This may be due to their phylogenetic proximity. This study has shown significant correlations between CGI density and genomic features such as the number of chromosome pairs, chromosome size, GC content, ObsCpG/ExpCpG ratio, gene density, and recombination rate. These results indicate the effects of these genomic features on the evolution of CGIs during domestic even-toed ungulate evolution. The majority of CGIs were unmethylated; most of the methylated CGIs were present in introns, intergenic and CDS regions, and the lowest number of methylated CGIs existed in the promoter regions. This finding confirmed that the DNA methylation of CGIs may regulate gene expression at the genome level. When predicted CGIs in even-toed ungulates were compared with other vertebrates, it was observed that in warm-blooded vertebrates (such as even-toed ungulates) the variation in the number of

doi: 10.17221/78/2015-CJAS

CGIs and in CGI density was lower than in cold-blooded vertebrates, such as fish. This is mostly consistent with the idea of a greater degree of CGIs losses with increasing organism complexity.

REFERENCES

Archibald A.L., Cockett N.E., Dalrymple B.P., Faraut T., Kijas J.W., Maddox J.F., McEwan J.C., Hutton Oddy V., Raadsma H.W., Wade C., Wang J., Wang W., Xun X. (2010): The sheep genome reference sequence: a work in progress. Animal Genetics, 41, 449–453.

Dong Y., Xie M., Jiang Y., Xiao N., Du X., Zhang W., Tosser-Klopp G., Wang J., Yang S., Liang J., et al. (2013): Sequencing and automated whole-genome optical mapping of the genome of a domestic goat (Capra hircus). Nature Biotechnology, 31, 135–141.

Du X., Han L., Guo A.Y., Zhao Z. (2012): Features of methylation and gene expression in the promoter-associated CpG islands using human methylome data. Comparative and Functional Genomics, 2012, ID 598987.

Elsik C.G., Tellam R.L., Worley K.C., Gibbs R.A., Muzny D.M., Weinstock G.M., Adelson D.L., Eichler E.E., Elnitski L., Guigo R., et al. (2009): The genome sequence of taurine cattle: a window to ruminant biology and evolution. Science, 324, 522–528.

Gardiner-Garden M., Frommer M. (1987): CpG islands in vertebrate genomes. Journal of Molecular Biology, 196, 261–282.

Hackenberg M., Barturen G., Carpena P., Luque-Escamilla P.L., Previti C., Oliver J.L. (2010): Prediction of CpG-island function: CpG clustering vs. sliding-window methods. BMC Genomics, 11: 327.

Han L., Zhao Z. (2008): Comparative analysis of CpG islands in four fish genomes. Comparative and Functional Genomics, 2008, ID 565631.

Han L., Zhao Z. (2009): Contrast features of CpG islands in the promoter and other regions in the dog genome. Genomics, 94, 117–124.

Han L., Su B., Li W.H., Zhao Z. (2008): CpG island density and its correlations with genomic features in mammalian genomes. Genome Biology, 9, R79.

Hillier L.W., Miller W., Birney E., Warren W., Hardison R.C., Ponting C.P., Bork P., Burt D.W., Groenen M.A.M., Delany M.E. (2004): Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature, 432, 695–716.

Hu Y., Xu H., Li Z., Zheng X., Jia X., Nie Q., Zhang X. (2013): Comparison of the genome-wide DNA methylation profiles between fast-growing and slow-growing broilers. PLoS ONE, 8, e56411.

Irizarry R.A., Wu H., Feinberg A.P. (2009): A species-generalized probabilistic model-based definition of CpG islands. Mammalian Genome, 20, 674–680.

Jia M., Gao X., Zhang Y., Hoffmeister M., Brenner H. (2016): Different definitions of CpG island methylator phenotype and outcomes of colorectal cancer: a systematic review. Clinical Epigenetics, 8: 25.

Jirimutu, Wang Z., Ding G., Chen G., Sun Y., Sun Z., Zhang H., Wang L., Hasi S., Zhang Y., et al. (2012): Genome sequences of wild and domestic bactrian camels. Nature Communications, 3, Article No. 1202.

Jung M., Pfeifer G.P. (2015): Aging and DNA methylation. BMC Biology, 13: 7.

Kastelic D., Frkovic-Grazio S., Baty D., Truan G., Komel R., Pompon D. (2009): A single-step procedure of recombinant library construction for the selection of efficiently produced llama VH binders directed against cancer markers. Journal of Immunological Methods, 350, 54–62.

Koh Y.W., Chun S.-M., Park Y.-S., Song J.S., Lee G.K., Khang S.K., Jang S.J. (2016): Association between the CpG island methylator phenotype and its prognostic significance in primary pulmonary adenocarcinoma. Tumor Biology, 37, 1–10.

Kwak W., Kim J., Kim D., Hong J.S., Jeong J.H. (2014): Genome-wide DNA methylation profiles of small intestine and liver in fast-growing and slow-growing weaning piglets. Asian-Australasian Journal of Animal Sciences, 27, 1532–1539.

Medvedeva Y.A., Fridman M.V., Oparina N.J., Malko D.B., Ermakova E.O., Kulakovskiy I.V., Heinzel A., Makeev V.J. (2010): Intergenic, gene terminal, and intragenic CpG islands in the human genome. BMC Genomics, 11: 48.

Paape T., Zhou P., Branca A., Briskine R., Young N., Tiffin P. (2012): Fine-scale population recombination rates, hotspots, and correlates of recombination in the medicago truncatula genome. Genome Biology and Evolution, 4, 726–737.

Poissant J., Hogg J.T., Davis C.S., Miller J.M., Maddox J.F., Coltman D.W. (2010): Genetic linkage map of a wild genome: genomic structure, recombination and sexual dimorphism in bighorn sheep. BMC Genomics, 11: 524.

Quinlan A.R., Hall I.M. (2010): BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics, 26, 841–842.

Rao Y.S., Chai X.W., Wang Z.F., Nie Q.H., Zhang X.Q. (2013): Impact of GC content on gene expression pattern in chicken. Genetics Selection Evolution, 45: 9.

Romiguier J., Ranwez V., Douzery E.J.P., Galtier N. (2010): Contrasting GC-content dynamics across 33 mammalian genomes: relationship with life-history traits and chromosome sizes. Genome Research, 20, 1001–1009.

Su J., Zhang Y., Lv J., Liu H., Tang X., Wang F., Qi Y., Feng Y., Li X. (2010): CpG_MI: a novel approach for identifying

doi: 10.17221/78/2015-CJAS

functional CpG islands in mammalian genomes. Nucleic Acids Research, 38, e6.

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

Tortereau F., Servin B., Frantz L., Megens H.-J., Milan D., Rohrer G., Wiedmann R., Beever J., Archibald A.L., Schook L.B., Groenen M.A.M. (2012): A high density recombination map of the pig reveals a correlation between sex-specific recombination and GC content. BMC Genomics, 13: 586.

Wang Y., Leung F.C.C. (2004): An evaluation of new criteria for CpG islands in the human genome as gene markers. Bioinformatics, 20, 1170–1177.

Weidner C.I., Lin Q., Koch C.M., Eisele L., Beier F., Ziegler P., Bauerschlag D.O., Jockel K.-H., Erbel R., Muhleisen T.W., Zenke M., Brummendorf T.H., Wagner W. (2014): Aging of blood can be tracked by DNA methylation changes at just three CpG sites. Genome Biology, 15, R24.

Weng Z.-Q., Saatchi M., Schnabel R.D., Taylor J.F., Garrick D.J. (2014): Recombination locations and rates in beef cattle assessed from parent-offspring pairs. Genetics Selection Evolution, 46: 34.

Wu H., Caffo B., Jaffee H.A., Irizarry R.A., Feinberg A.P. (2010): Redefining CpG islands using hidden Markov models. Biostatistics, 11, 499–514.

Wu H., Guang X., Al-Fageeh M.B., Cao J., Pan S., Zhou H., Zhang L., Abutarboush M.H., Xing Y., Xie Z., et al. (2014): Camelid genomes reveal evolution and adaptation to desert environments. Nature Communications, 5, Article No. 5188.

Received: 2015–10–13
Accepted after corrections: 2016–07–26

Corresponding Author

Dr. Mohammad Reza Mohammadabadi, Associate Professor, Shahid Bahonar University of Kerman, Agricultural College, Department of Animal Sciences, 22 Bahman Blvd, Kerman, 7616914111 Iran
Phone: +989 133 987 534, e-mail: mmohammadabadi@yahoo.ca