Genome-wide analysis of DNA methylation reveals selection signatures of the grass carp during domestication

Yichao Li[a,1], Bing Fu[a,1], Junming Zhang[a,1], Jun Xie[a], Guangjun Wang[a], Peng Jiang[a], Jingjing Tian[a], Hongyang Li[a], Yun Xia[a], Zhifei Li[a], Ermeng Yu[a,*]

[a] *Pearl River Fisheries Research Institute, Chinese Academy of Fishery Sciences, Guangzhou 510380, China*

* Corresponding author. *E-mail* address: yem@prfri.ac.cn (E. Yu).

[1] These authors contributed equally to this work.

ABSTRACT

With the rapid development of aquaculture, many fish species are domesticated and brought into cultivation. In the process of domestication, the domesticated fish undergone intense selection pressures and develop some adaptations and phenotypic traits, namely selection signatures, such as growth and metabolism, immunity, foraging and learning behaviors. However, how this selection signatures emerges is still not clear and the knowledge of molecular epigenetic mechanisms underlying selection signatures in fish is still in its infancy. Thus, we used a farmed fish, grass carp (*Ctenopharyngodon idellus*), as model species to detect these selection signatures and identify the candidate differentially methylated genes that are closely associated with these selection signatures at the level of whole genome, investigating the role of DNA methylation in the emergence of selection signatures during domestication. Our results showed that domesticated grass carp demonstrated four selection signatures, including growth and metabolism, immunity, foraging and learning behaviors, and 38 candidate genes were found associated with these traits. 16 genes are significant candidate genes which play major roles in the growth and metabolism, such as IGF-1 , GK , GYS1, etc. 11 genes are related to immunity, including . The GRM1, TAS1R1 and TAS1R3 genes are essential for the adaptation of domesticated grass carp to commercial feed in artificial rearing condition. The C-FOS, POMC and CBP genes may be responsible for the acquisition of novel feeding habits and contribute to faster growth indirectly by enhancing food intake. The findings here in will provide new insights to expand our understanding about the role of epigenetic modifications in shaping physiological phenotypes in this and other teleost models, which can contribute to efficient breeding of aquaculture stocks and restocking programmes.

Keywords:

DNA methylation

grass carp

selection signatures

1. Introduction

As an economically important food source of humans, fish is one of the most important domesticated species (Teletchea ,et al. 2012). In modern agricultural industry, fish have undergone strong long-term artificial selection and developed a range of adaptations and phenotypic traits, namely selection signatures (Milla et al.,2021), such as higher flexibility in diet, rapid growth, less stress susceptibility, a more socially tolerant disposition and enhanced prolificacy (Pasquet et al., 2018), which may distinguish those domesticated breeds from their wild counterparts. These phenotypic differences make fish highly suitable for animal agriculture and comparative studies.

Epigenetic changes is increasingly recognized to contribute to the emergence of phenotypic differences (Goldberg et al., 2007). DNA methylation, the most widely studied epigenetic regulatory mechanisms, can respond to environmental change and, at the same time, be stable enough to be maintained throughout lifetime, even across generations (Anastasiadi et al., 2019). DNA methylation has been associated with many biological process, including gene expression regulation, genomic imprinting, X chromosome inactivation, embryonic development, the alteration of chromatin structure and transposon inactivation, etc (Wilson et al., 2012; Jones et al., 2012). Multiple approaches have been developed to analyze DNA methylation profiles at the genome-wide level, including comprehensive high-throughput arrays for relative methylation (CHARM), methylated DNA immunoprecipitation (MeDIP), methylation array, reduced-representation bisulfite sequencing (RRBS), whole-genome bisulfite sequencing (WGBS), single-cell reduced-representation bisulfite sequencing (scRRBS), single-cell whole-genome bisulfite sequencing (scWGBS), TET-assisted bisulfite sequencing (TAB-seq), etc (Laird et al., 2010, Yong et al., 2016). Compared to other approaches, WGBS has two major advantages: (1) it is able to assess the methylation state of nearly every CpG site, including low-CpG-density

89 regions, such as intergenic 'gene deserts', partially methylated domains and distal regulatory
90 elements, (2) it can determine absolute DNA methylation level and reveal methylation sequence
91 context (Yong et al., 2016). This method has been broadly employed to analyze the genome-wide
92 methylation profiles of many animals, including sea cucumber (Yang et al., 2020), pig (Yang et al.,
93 2016), broilers (Hu et al., 2013), Chinese perch (Pan et., 2021).

94 In recent years, several studies have focused on the role of DNA methylation regulation on
95 fish during domestication. However, most studies focus on the role of DNA methylation in the
96 changes of one particular phenotypic characteristics, such as growth (Li et al., 2017), reproduction
97 (Tomasz Podgorniak et al., 2019), morphological changes (Zhang et al., 2017), learning behaviors
98 (Dou et al., 2018). Thus, until now, no systematic studies have investigated the role of DNA
99 methylation in the emergence of different selection signatures in the process of fish domestication,
100 including growth and metabolism, immunity, foraging and learning behaviors.

101 Grass carp (*Ctenopharyngodon idellus*) is one of the most important freshwater aquaculture
102 species, and its global production reaches 5.704 million tons in 2018, the highest in fish
103 production worldwide, providing low-cost, high-quality animal protein especially for developing
104 and underdeveloped regions (FAO, 2020). Although some farms have carried out breeding for
105 several generations, others still rely on wild-caught fish for broodstock to maintain genetic
106 diversity (Huang et al., 2015; Fu et al, 2015). In addition, farmed grass carp also shares some of
107 the characteristics of the domestication syndrome, for instance, changes in morphological features
108 (Zhao et al., 2020), rapid growth (Ashraf et al., 2011), anti-predator behavior (Tang et al, 2017),
109 indicate that this species could be a suitable model species for studying the role DNA methylation,
110 one of key molecular epigenetic pathways, in the formation of selection signatures in this and

111 other teleost species.

112 In this study, we used WGBS to carry out genome-wide DNA methylation analysis of blood
113 from domesticated and wild carp with four selection signatures differences: growth and
114 metabolism, immunity, foraging and learning behaviors. We obtained comprehensive DNA
115 methylation profiles for the two groups and identified differentially methylated genes (DMGs) that
116 might contribute to the emergence of selection signatures between domesticated and wild carp. To
117 the best of our knowledge, this study is the first one to report the role of epigenetic modifications
118 in the emergence of different selection signatures in fish. Thus, the findings here in will provide
119 new insights to expand our understanding about the role of epigenetic modifications in shaping
120 physiological phenotypes in this and other teleost models
121

122 ## 2. Materials and Methods

123 *2.1 Domesticated trial*

124 A total of 15 wild sub-adult grass carp (initial weight 250 ± 20 g) were captured using traps
125 and gill nets along the Xijiang river in Zhaoqing, Guangdong Province, China, within a 1000 m
126 radius from the following coordinates (latitude: 23.1034.8 °N; longitude: 112.4554.0 °E).
127 Subsequently, the fish were transported to breeding base and randomly allocated into three ponds
128 (1.5 m × 1.5 m × 1 m), 5 individuals in each pond for subsequent domesticated experiments in the
129 Pearl River Fisheries Research Institute. The fish were fed commercial feed at 8:00 am and 4:00
130 pm each day for 180 days, and the feed amount for each day was 2~3% of fish weight. The water
131 temperature was kept at 22~27 °C, pH was 6.5~7.5, and dissolved oxygen was above 5.0 mg/L.

132 *2.2 Sample Collection and Preparation*

133 Three domesticated grass carp after six month of domestication were collected from ponds

(one fish each pond, with a body weight of 720 ± 50 g). Wild grass carp with a body weight of 550 ± 50 g were collected from same place where wild sub-adult grass carp were captured, and DGC and WGC were at approximately the same age and ontogeny stage (i.e.,puberty), which was determined by annulus characteristics of scale (Cai et al., 2020). Whole blood was collected from were collected from the domesticated grass carp (sample DGC1 - DGC3) and wild grass carp (sample WGC1 - WGC3) groups. Subsequently, blood samples were treated with EDTAK2 (Sanli, Liuyang, China) and then centrifuged at 12,000 rpms for 3 min to separate red blood cells (RBCs) from serum.

The experimental protocols used in this study were approved by the Laboratory Animal Ethics Committee of Pearl River Fisheries Research Institute, CAFS, China, under permit number LAEC-PRFRI-2021-06-03.

2.3 Genomic DNA Extraction and Whole-Genome Bisulfite Sequencing

Genomic DNA was extracted from RBCs of grass carp using DNeasy Blood and Tissue Kit following the manufacturer' s protocol. Genomic DNA of RBCs from domesticated grass carp and wild grass carp groups was sent to BGI (BGI Tech Solutions Co., Ltd., Shenzhen, China) for whole-genome bisulfite sequencing. For normal WGBS library constructing, the DNA was fragmented by sonication using a Bioruptor (Diagenode, Belgium) to a mean size of approximately 250 bp, followed by the blunt-ending, dA addition to 3'-end, finally, adaptor ligation(in this case of methylated adaptors to protect from bisulfite conversion), essentially according to the manufacturer's instructions. Ligated DNA was bisulfite converted using the EZ DNA Methylation-Gold kit (ZYMO). Different Insertsize fragments were excised from the same lane of a 2% TAE agarose gel. Products were purified by using QIAquick Gel Extraction kit

(Qiagen) and amplified by PCR. At last, Sequencing was performed using the HighSeq4000 or other Illumina platforms.

2.4 Data Filtering and Reads Alignment

After sequencing data was delivered, the raw reads were filtered by removing adaptor sequences, contamination and low-quality reads from raw reads. Low-quality reads include two types,and the reads meet any one of the two conditions will be removed: (1) Unknown bases are more than10%; (2) The ratio of bases whose quality was less than 20was over 10%. After filtering, the Clean data was then mapped to the reference genome of grass carp (Wang et al., 2015) by BSMAP, and then remove the duplication reads and merge the mapping results according to each library. The BSMAP script was BSMAP -a filename_1.clean.fq.gz -b filename_2.clean.fq.gz -o filename.sam - d ref.fa -u -v 8 -z 33 -p4 -n 0 -w20 -s 16 -f 10 -L 100. The sam files were converted to bam files using scripts (samtools view -S-b-o filename.bam filename.sam; samtools sort-m 2000000000 filename.bam filename.sort; samtools index filename.sort.bam). The mapping rate and bisulfite conversion rate of each sample were calculated.

2.5 Identification of Differentially Methylated Regions

The methylation level was determined by dividing the number of reads covering each mC by the total reads covering that cytosine (Lister R., 2009), which was also equal the mC/C ratio at each reference cytosine (Xiang et al., 2010). The formula is Rmaverge = Nmall/Nmall+Nnmall. Nm represents the reads number of mC, while Nnm represents the reads number of non-methylation reads. Putative DMRs were identified by comparison of the sample DGC and sample WGC methylomes using windows that contained at least 5 CpG (CHG or CHH) sites with a 2- fold change in methylation level and Fisher test p value < = 0.05.

178 *2.6 Gene Ontology and Pathway Enrichment of DMRs*

179 GO enrichment analysis provides all GO terms that significantly enriched in a list of
180 differentially methylated genes, comparing to a genome background, and filter the differentially
181 methylated genes that correspond to specific biological functions.This method firstly maps all
182 differentially methylated genes to GO terms in the database (http://www.geneontology.org/),
183 calculating gene numbers for every term, then uses hypergeometric test to find significantly
184 enriched GO terms in the input list of differentially methylated genes, based on ' GO ::
185 TermFinder' (http://www.yeastgenome.org/help/analyze/go-term-finder). KEGG pathway
186 enrichment analysis helps to identify significantly enriched metabolic pathways or signal
187 transduction pathways in differentially methylated genes comparing with the whole genome
188 background. The calculating formula is the same as that in GO analysis.
189
190
191
192

193 ## 3. Results

194 *3.1 Global mapping and statistical analysis of the WGBS reads*

195 We conducted whole-genome bisulfite sequencing of grass carp blood from domesticated
196 grass carp group and wild grass carp group. After quality control of filtering, a total of 2.36 billion
197 clean reads were generated, consisting of 378.12 million, 387.63 million, and 321.7 million reads
198 for each domesticated grass carp sample and 371.48 million, 493.82 million, and 410.42 million
199 reads for each wild grass carp sample (Table 1). The bisulfite conversion rate (%) of all
200 sequencing libraries ranges from 99.16% to 99.28%. After read alignment, clean reads were
201 mapped to the reference genome of grass carp with mapping rates ranging from 88.03 % to

92.11 % (Table 1).

3.2 Global DNA methylation patterns of domesticated and wild grass carp.

The methylation levels of the whole genom were listed in Table 2. In each group, approximately 9% of all genomic C sites were methylated. Methylation in grass carp was found to exist in three sequence contexts: CG, CHG (where H is A, C, or T), and CHH. The average methylation levels of CG, CHG, and CHH at the whole genome levels were 80.3%, 1.09%, and 1.08% in the domesticated grass carp group, and 75.04%, 1.07%, and 1.13% in the wild grass carp group. The average methylation of CG showed a significantly increased within genome in the domesticated grass carp group compared with the wild grass carp group ($p < 0.05$). The methylation level of CHG and CHH showed no differences between domesticated grass carp and wild grass carp groups ($p > 0.05$).

Methylation status of CG, CHG and CHH of grass carp genome showed the methylome' s overall characteristics (Figure 1). The methylation levels of approximately 25% of all mCG were hypermethylated (methylation level > 90%). However, only about 3% of mCHH and mCHG were hypermethylated (methylation level > 90%) compared with mCG. The methylated Cs mostly occur in the form of mCG, followed by mCHH and mCHG. The methylation level distribution of mC and mCG were alike. The methylated Cs mostly occur in the form of mCG; approximately 96% of all detected mCs (Table 3). Proportion of mCHG range from 0.75% to 1.15%, and proportion of mCHH range from 2.7% to 3.05% (Table 3). The proportion of mCG, mCHG and mCHH showed no significant difference between domesticated grass carp group and wild grass carp group ($p > 0.05$).

3.3 DNA methylation levels of different genomic features

225 Different genomic features show distinct methylation patterns which are related with distinct
226 regulation functions (Cokus et al, 2008). The heatmap presents the methylation landscape in
227 different genomic features (whole genome, CGI, downstream 2 kb, upstream 2 kb, mRNA, repeat,
228 CDS, exon), providing additional information as well as a global assessment of some of these
229 components (Figure2). CpG islands contained the highest numbers of CpG sites (approximately
230 10 - 20 CpG sites in a 200 bp window) compared with other genomic features. About 60% of CpG
231 sites in CpG islands were hypermethylated (methylation levels >90%) in the heatmap (Figure2
232 and Supplementary Material Figure S2). The other genomic features generally contained 0 - 10
233 CpG sites in the 200 bp window. A lower proportion of CpG sites within whole genome, mRNA
234 and repeat were hypomethylated (methylation levels < 10%) than CpG sites within remaining
235 genomic features (CGI, downstream 2 kb, upstream 2 kb, CDS, exon).

236 *3.4 DNA Methylation Patterns Across the Entire Transcriptional Units at Whole Genome Level*

237 In order to reveal the relationship between DNA methylation profiles and genes expression in
238 detail, Canonical DNA methylation profiles of the entire transcriptional units were divided into
239 distinct functional elements (Li et al., 2010) to study the changes of methylation levels in different
240 features (Figure 3). Methylation differences between CG and non-CpG methylation (CHG and
241 CHH) are visible (Figure 3), as methylation levels of CG are higher than those of CHG and CHH
242 across the entire transcriptional units. Another feature is a modest elevation in methylation level at
243 internal exons and internal intron during transcriptional unit scanning. The lowest methylation
244 level occurs in the first intron, followed by the first exons and down stream.

245 *3.5 Identification and Enrichment Analysis of Differential Methylated Regions*

246 To characterize the differences of genome methylation levels between domesticated grass

247 carp and wild grass carp groups, DMRs and differentially methylated genes (DMGs) were
248 detected. For CG context methylation, a total of 533,235 DMRs were identified between the two
249 groups, which corresponded to 20,010 DMGs in promoter regions and 27,016 DMGs in gene
250 body.

251 GO enrichment analysis of DMGs was performed to provide significantly enriched GO
252 terms corresponding to specific biological process, cellular component, and molecular function
253 (Figure 4). In promoter regions, the over-represented GO terms in the biological process are
254 cellular process, metabolic process, and biological regulation. The top enriched GO terms in the
255 cellular component are cellular anatomical entity, intracellular and protein−containing complex. In
256 terms of molecular function, the top enriched GO terms are binding, catalytic activity, and
257 molecular transducer activity. In gene body region, the over-represented GO terms in the
258 biological process are cellular process, metabolic process and biological regulation. The top
259 enriched GO terms in the cellular component are cellular anatomical entity, intracellular and
260 protein-containing complex. In terms of molecular function, the top enriched GO terms are
261 binding, catalytic activity and transporter activity. KEGG pathway analysis, which is an
262 alternative approach to categorize gene functions, was also conducted for the DMGs in promoter
263 and gene body regions. In promoter region, DMGs were significantly enriched in pathways in
264 cancer, neuroactive ligand−receptor interaction and PI3K−Akt signaling pathway. In gene body
265 region, DMGs were significantly enriched in metabolic pathways, pathways in cancer and MAPK
266 signaling pathway (Figure 4).

267 *3.6 Candidate DMGs associated with selection signatures.*

268 Many transcriptome and association studies have explored the molecular mechanisms

underlying selection signatures in fish and other species, laying a foundation for our investigation of the involvement of DMGs in selection signatures. Beside, it is reported that DNA methylation, especially in the promoter regions, usually affects gene expression by different modes (Moore et al, 2012). 38 candidate DMGs in promoter region were identified according to the following criteria: (1) genes were differentially methylated in grass carp and wild grass carp groups; (2) genes were enriched in pathways related to selection signatures; (3) genes were differentially expressed in fish and other species or related with selection traits reported by previous studies. However, KEGG pathway analysis showed that these methylated genes were not significantly enriched in predicted process and signaling pathway (Figure 5). The protein-protein interaction network analysis indicated that these DMGs were highly correlated with each other (Figure 6).

4. Discussion

4.1 DNA methylation profiles in domesticated and wild grass carp

Domestication is a process by which humans select some phenotypes of wild animal species (i.e., morphological traits or growth). During the domestication process, some phenotypic traits of animal could be altered by the artificial selection to help domesticated animal species adapt to new environmental conditions (Sylvain et al, 2021). DNA methylation, one of the most important and stable epigenetic modifications in eukaryotes, can lead to heritable phenotypic and transcriptomic changes (Höglund et al., 2020). However, in aquatic species, there is still limited research that illustrate the effects of DNA methylation on selection signatures by comparing genome-wide methylation profiles between domesticated and wild fish species. In this study, we identified some key differentially methylated genes related to selection signatures (including growth and metabolism, immunity, foraging behaviors, social behaviors) in domesticated and wild

291 grass carp to uncover its genetic characters.

292 In the present study, we reported a genome-wide examination of DNA methylation in
293 domesticated grass carp and wild grass carp. The mapping rates of clean reads ranged from
294 86.25% to 92.11% (Table 1), which is above the average value reported in other fish species (Pan
295 et al., 2021). However, the WGBS results in this study are still consistent because bisulfite
296 conversion reduces the complexity of the genomic sequence and the ability of most computational
297 programs to align sequences onto the reference genome (Yang et al., 2020).

298 Further DNA methylation profile analysis was also conducted in our present study across the
299 distinct genomic features and entire transcriptional units. Our results showed that the
300 genome-wide methylation patterns were similar between two groups. Approximately 9% of all
301 total cytosines and 80% of CG were methylated and only a small proportions of methylation at
302 non-CpG methylation (CHG and CHH) was observed within all regions, which is consistent with
303 previous study (Cai et al., 2020). The high level of DNA methylation in CG content was a specific
304 characteristic of animals. DNA methylation in CHH and CHG patterns is a major characteristic of
305 plant methylomes and largely absent in animal methylomes (Zemach et al., 2010).

306 A modest elevation in methylation level at internal exons and intron was observed during
307 transcriptional unit scanning. The lowest methylation level occurs in the first exons, followed by
308 the the first intron and last exon. These results was consistent with previous findings in grass carp
309 (Cai et al., 2020), suggesting that mutagenic effects appear to occur at the first exons and internal
310 exon and intron tend to be influenced by the regulatory impact of DNA methylation.

311 To better understand the functional classification of DMGs during domestication, GO
312 enrichment analysis was conducted to gain insight into the biological processes in which the

313 DMGs might be involved. The top enriched GO terms of DMRs identified in the present study
314 were cellular process, biological regulation and metabolic process, suggesting that DNA
315 methylation may play a critical role in regulating global transcription during domestication.
316 Furthermore, to further explored the molecular mechanisms underlying selection signatures
317 in grass carp, 38 candidate DMGs in promoter region were identified, including IGF-1
318 (Insulin-like growth factor 1), GK (Glycerol kinase), GYS1 (Glycogen synthase 1), etc. But these
319 methylated genes were not significantly enriched in predicted process and signaling pathway of
320 KEGG.
321
322 *4.2 Key differentially methylated genes related to growth and metabolism*
323 Domestication is a long process which forces animals to adapt to captivity by modifying an
324 animal' s growth and metabolism, immune response, foraging and social behaviors. In the
325 process of domestication, fish undergo great change in the growth and metabolism due to the
326 difference of environmental conditions and food resource (Teletchea et al., 2016), which is closely
327 associated with the differential expression of genes involved in growth and metabolism (Shen et
328 al., 2021). In this study, several genes involved in growth and metabolism, such as IGF-1
329 (Insulin-like growth factor 1), GK (Glycerol kinase), GYS1 (Glycogen synthase 1), FASN (Fatty
330 acid synthase), etc, exhibited hypomethylation in domesticated grass carp group compared to wild
331 grass carp group. This result indicated that the expression of these genes might be upregulated. In
332 contrast, some genes related to growth and metabolism exhibited hypermethylation in
333 domesticated grass carp group, including G6PC (Glucose-6-phosphatase),
334 PCK1(Phosphoenolpyruvate carboxykinase), FBP1 (Fructose 1,6 bisphosphatase 1),
335 FOXO1(Forkhead box protein O1), ACAA1 (Acetyl-Coenzyme A acyltransferase 1), CPT1

(Carnitine palmitoyltransferase I), suggesting the expression of these genes might be downregulated.

IGF-1, an important growth hormone, mediates the anabolic and linear growth promoting effect of pituitary GH protein and has been well demonstrated to play an important role in the skeletal muscle development (Laron et al., 2001). Thus, IGF-1 is thought to be a candidate gene for regulating muscle growth.

GK is a key enzyme that catalyzes the first step in glycolysis (Enes et al., 2009). FOXO1, G6PC, FBP1 and PCK1 are key enzymes of gluconeogenesis (Lu et al., 2018). Our results indicated that, compared to wild grass carp, the glycolytic capacity is enhanced and the capacity for gluconeogenesis is inhibited in domesticated grass carp feeding commercial diet. These results was consistent with the previous finding that high-carbohydrate diet could induce glycolysis and inhibit gluconeogenesis (Liu et al., 2021).

GSY1 is a rate-limiting enzyme in glycogen synthesis and plays an important role in the synthesis of glycogen in the muscle (Martins et al., 2013). The present results suggested that the capacity for muscle glycogen storage in domesticated grass carp maybe increase compared with that of wild grass carp.

FASN, which catalyzes the last step in fatty acid biosynthesis, is believed to be the central enzyme of hepatic lipid accumulation (Dorn et al, 2010). ACAA1, CPT1 gene play a certain role in the lipid degradation (Song et al., 2019). In the present study, our results suggested that, in domesticated grass carp, the ability to synthesize fat was enhanced and lipid degradation was inhibited compared to wild grass carp, which may account for elevated lipid levels of muscle in domesticated grass carp feeding commercial feed (Ashraf et al.,, 2011). According to the results,

an explanation for the enhanced fat deposition in fish might be excessive energy intake in commercial diet (Wu et al., 2021). Overall, DNA methylation is likely to play an important role in regulating growth and metabolism of domesticated grass carp by influencing these gene expression.

4.3 Key differentially methylated genes related to immunity.

To maximize profitability, domesticated fish were cultured in intensive farming conditions with limited space, high density and other stressors, which could influence fish immune response negatively and even result in large-scale disease (Lin et al., 2018, Chen et al.,2017). It was reported that, in the process of domestication, the immune status of domesticated fish could be affected negatively due to increased chronic stress by confinement (Mandiki et al., 2011), which would make fish more susceptible to pathogens and ultimately impair fish survival.

In this study, several immune related genes exhibited hypermethylation in domesticated grass carp group compared to wild grass carp group, including MHCI (Major histocompatibility complex class I), MHCII (Major histocompatibility complex class II), C1QA (Complement C1q subcomponent subunit A), C3 (Complement C3), C4A (Complement C4-A), C5 (Complement C5), IFN-γ (Interferon gamma), etc, which suggested the expression of these genes might be downregulated. MHCI and MHCII are two cell surface proteins essential for the acquired immune system for antigen presentation to recognize foreign molecules in vertebrates and can promote the development and expansion of T cells (Huang et al., 2005). The complement system, an essential part of both innate and adaptive immunity in teleosts, is initiated by one or a combination of three pathways, the alternative, lectin and classical (Yang et al., 2016). IFN-γ, one of critical antiviral cytokines, modulate functions of the immune system by up-regulate major histocompatibility

complex molecules (MHC I and MHC II) and directly activate other immune cells, such as macrophages and natural killer cells (Borden et al., 2007). In our results, these immune related genes mentioned above identified in domesticated grass carp exhibited hypermethylation, indicating a poor immune performance in domesticated domesticated grass carp during long-term domestication.

4.4 Key differentially methylated genes related to foraging behaviors

During the domestication process, fish change their foraging habits to obtain food in captive conditions (Pasquet et al., 2018). In this study, foraging behaviors-related genes were also found to be under strong selection in domesticated carp. The genes of GRM1(Metabotropic glutamate receptor 1), TAS1R1 (Taste receptor type 1 member 1) and TAS1R3 (Taste receptor type 1 member 3) exhibited hypomethylation in domesticated grass carp group compared to wild grass carp group, suggesting the expression of these genes might be upregulated. The TAS1R1 and TAS1R3 heterodimer receptor functions as an umami receptor, responding to L-amino acid binding, especially L-glutamate (Nelson et al., 2001). GRM1, which is widely expressed throughout the central nervous system and regulates synaptic signaling, is another L-glutamate receptor (Gabriel et al., 2009). Generally, these taste receptors play an important roles in perception of L-amino acids and feeding behavior (Cai et al., 2021). Thus, these genes with hypomethylation are likely to be upregulated in domesticated carp, facilitating the adaptation of domesticated grass carp to commercial feed in artificial rearing condition.

4.5 Key differentially methylated genes related to learning behaviors

Learning and memory could enable the organism to plastically respond to the changing environment. Increasing research has investigated the learning (cognitive) and memory

403 characteristics of fish in the past few decades, including spatial cognition, learned recognition,
404 social learning, foraging activity, etc (Dou et al., 2018). In this study, through domestication,
405 social learning-related genes have also been under selection pressure. The gene POMC
406 (Pro-opiomelanocortin) exhibited hypermethylation. In contrast, the genes of C-FOS
407 (Proto-oncogene C-FOS) and CBP (CREB-binding protein) exhibited hypomethylation in
408 domesticated grass carp group compared to wild grass carp group, suggesting the expression of
409 these genes might be upregulated. C-FOS is necessary for consolidation of non-spatial
410 hippocampal-dependent memory (Countryman et al., 2005), and C-FOS mRNA expressions are
411 up-regulated in response to a variety of neuronal activation protocols, including behavioral
412 training (Smulders et al., 2000) and long-term protentiation (Alberini et al., 2009). CBP is a
413 coactivator of transcription that play an essential role in memory consolidation (Korzus et al.,
414 2004). It was reported that gene expression of C-FOS and CBP facilitated the acquisition of novel
415 feeding habits in fish through the memory formation (Peng et al., 2019; Dou et al., 2018). Beside,
416 the C-FOS gene might be an important transcriptional factor to inhibit the expression of the
417 anorexigenic gene POMC, resulting in an increase of the food intake of dead prey fish in mandarin
418 fish (Peng et al., 2019). Therefore, it tempts us to speculate that the learning-related genes of
419 C-FOS and CBP may be responsible for the acquisition of novel feeding habits in domesticated
420 grass carp, which considerably increase the rate of domestication. Furthermore, individual food
421 intake is enhanced through the interaction between the learning gene C-FOS and the appetite
422 control gene POMC, contributing to the faster growth of domesticated grass carp indirectly.

423

424

425

Reference

Alberini, C. M., 2009. Transcription factors in long-term memory and synaptic plasticity. Physiol Rev. 89(1),121-145.

Ashraf, M., Zafa, A., Rauf, A., Mehboob, S., Qureshi, N. A., 2011. Nutritional values of wild and cultivated silver carp (hypophthalmichthys molitrix) and grass carp (*ctenopharyngodon idella*). Int J Agric Biol. 13(2), 1560-8530.

Borden, E.C., Sen, G.C., Uze, G., Silverman, R. H., Ransohoff, R.M., Foster, G.R., 2007. Interferons at age 50: past, current and future impact on biomedicine. Nat Rev Drug Discov.

Cai, Z.L., Duan, Y.J., Xu, J.C., He, Y., Geng J.Z., Liu, X., Han, M.X., Liu, Y.H., 2020 Comparison of five structure materials in age determination of Amur ide Leuciscus waleckii in Dali Lake. Journal Of Dalian Ocean University. 35, 1. DOI: 10．16535 /j．cnki．dlhyxb．2019-183

Cai, W.J., Liang, X.F., Yuan, X.C., Li, A.X., He, S., 2020. Changes of DNA methylation pattern in metabolic pathways induced by high-carbohydrate diet contribute to hyperglycemia and fat deposition in grass carp (*ctenopharyngodon idellus*). Front Endocrinol. 11, 398.

Cai, W.J., Li, J., Li, L., Chen, X., Liang, X.F., 2021. Knockout of t1r1 gene in zebrafish (*danio rerio*) by crispr/cas9 reveals its roles in regulating feeding behavior. Aquaculture, 545, 737189.

Chen, X., Wang, J., Long, Q., Gaughan, S., Wang, C., 2017. Domestication drive the changes of immune and digestive system of eurasian perch (*perca fluviatilis*). PLoS One. 12(3), e0172903.

Christoph, D., Marc-Oliver, R., Georgi, K., Michael, S., Kathrin, S., Weiss, T.S., 2010. Expression of fatty acid synthase in nonalcoholic fatty liver disease. Int J Clin Exp Patho, 3(5), 505.

448 Countryman, R.A., Kaban. N.L., Colombo, P.J. 2005. Hippocampal c-fos is necessary for
449 long-term memory of a socially transmitted food preference. Neurobiol Learn Mem. 84(3),
450 175-183.
451 Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., 2008. Shotgun
452 bisulphite sequencing of the arabidopsis genome reveals DNA methylation patterning. Nature.
453 452(7184), 215-9.
454 Dafni, A ., Francesc, P., 2019. Epimutations in developmental genes underlie the onset of
455 domestication in farmed European sea bass. Mol Biol Evol. 10, 10.
456 Dou, Y., He, S., Liang, X. F., Cai, W., Wang, J., Shi, L., 2018. Memory function in feeding habit
457 transformation of mandarin fish (*Siniperca chuatsi*). Int J Mol Sci, 19(4), 1254.
458 Doi, A., Park, I.H., Wen, B., Murakami, P., Aryee, M.J., Irizarry, R., 2009. Differential
459 methylation of tissue- and cancer-specific CpG island shores distinguishes human induced
460 pluripotent stem cells, embryonic stem cells and fibroblasts. Nat Genet, 41(12), 1350.
461 Enes, P., Panserat, S., Kaushik, S., Oliva-Teles, A. A., 2008. Nutritional regulation of hepatic
462 glucose metabolism in fish. Fish Physiol Biochem, 35(3), 519-539.
463 FAO, 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in Action. Rome.
464 Fu, J., Shen, Y., Xu, X., Liu, C., Li, J., 2015. Genetic parameter estimates and genotype by
465 environment interaction analyses for early growth traits in grass carp (*ctenopharyngodon*
466 *idella*). Aquacult Int.
467 Gabriel, A.S., Maekawa, T., Uneyama, H., Torii, K. 2009. Metabotropic glutamate receptor type 1
468 in taste tissue. Am J Clin Nutr. 90(3), 743S-746S.
469 Goldberg, A.D., Allis, C,D., Bernstein, E., 2007. Epigenetics: a landscape takes shape. Cell.

128(4), 635-638.

Pasquet, A., 2018. Effects of Domestication on Fish Behaviour. DOI: 10.5772/intechopen.78752.

Höglund, A., Henriksen, R., Fogelholm, J., Churcher, A.M., Guerrero-Bosagna, C.M., Martinez-Barrio, A., 2020. The methylation landscape and its role in domestication and gene regulation in the chicken. Nat Ecol Evol. 4, 1713-1724.

Hu, Y.S., Xu, H,P., Li, Z.H., Zheng, X.J., Jia, X.Z., Nie, Q.H., Zhang, X.Q., 2013. Comparison of the genome-wide DNA methylation profiles between fast-growing and slow-growing broilers. PLoS One, 8(2), e56411.

Huang, S.X, Gilfillan, S.S, Cella., Marina., 2005. Evidence for mr1 antigen presentation to mucosal-associated invariant t ceils. J Biol Chem, 280(22), 21183-21193.

Huang, R., Sun, J.X., Luo, Q., He, L.B., Liao, L.J., Li, Y.M., Guo, F.H., Zhu, Z.Y., Wang, W.P. 2015. Genetic variations of body weight and GCRV resistance in a random mating population of grass carp. Oncotarget. 6(34), 35433-35442.

Jones., Peter, A., 2012. Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet, 13(7), 484-492.

Korzus, E., Rosenfeld, M.G., Mayford, M., 2004. CBP histone acetyltransferase activity is a critical component of memory consolidation. Neuron, 42(6), 961-972.

Laird, P.W., 2010. Principles and challenges of genome-wide DNA methylation analysis. Nat Rev Genet.

Laron, Z., 2001. Insulin-like growth factor 1 (igf-1): a growth hormone. Mol Pathol MP. 54(5), 311.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., 2009. Human

492 DNA methylomes at base resolution show widespread epigenomic differences. Nature,
493 462(7271), 315-22.
494 Li, Y., Zhu, J., Geng, T., Li, N., Zhang, X., 2010. The DNA methylome of human peripheral blood
495 mononuclear cells. PLoS Biol. 8(11), e1000533.
496 Li, S.P., He, F., Wen, H.S., Li J.F., Si, Y.F., Liu, M.Y., He H.W., Huang, Z.J., 2017. Analysis of
497 DNA Methylation Level by Methylation-Sensitive Amplification Polymorphism in Half
498 Smooth Tongue Sole (*Cynoglossus semilaevis*) Subjected to Salinity Stress. J. Ocean Univ.
499 China. 16 (2), 269-278.
500 Li, C,L., Zuo, W.D., Tong, X.L., Han, M.J., Gao, R., Hu, H., Lu, K.P., Luan, Y., Zhang, B, L.,
501 Liu,Y.Y., Dai, F.Y., 2020. Whole-genome resequencing reveals loci under selection during
502 silkworm improvement. 138, 278-290.
503 Lin, W., Li, L., Chen, J., Li, D.P., Hou, J., Guo, H.H., Shen, J.Z., 2018. Long-term crowding stress
504 causes compromised nonspecific immunity and increases apoptosis of spleen in grass carp
505 (*Ctenopharyngodon idella*). 80, 540-545.
506 Fang, L., Guo, X.Z., Liang, X.F., 2021. First feeding of grass carp (*Ctenopharyngodon idellus*)
507 with a high-carbohydrate diet:the effect on glucose metabolism in juveniles. 21, 100830.
508 Lu, S., Wu, X., Gao, Y., Gatlin, D.M., Wu, M., Yao, W., 2018. Effects of dietary carbohydrate
509 sources on growth, digestive enzyme activity, gene expression of hepatic gluts and key
510 enzymes involved in glycolysis-gluconeogenesis of giant grouper epinephelus lanceolatus
511 larvae. Aquaculture. 343-350.
512 Martins, F.O., Rito, J., Jarak, I., Viegas, I., Jones, J. G., 2013. Disposition of [U-2H7] glucose into
513 hepatic glycogen in rat and in seabass. CBPA, 166(2).

514 Mandiki, S., Henrotte E., Milla S., Douxfils J., Wang N., Rougeot C., 2011. How Physiological
515 Status and Immune Defense are Affected by Photo-thermal Regimes and Domestication
516 Process in Captive Eurasian Perch? ISEE. 2011.4(S8), 298–299.
517 Milla, S., Pasquet, A.,Mohajer, L, El., Fontaine, P., 2021. How domestication alters fish
518 phenotypes. 13, 388-405.
519 Moore, L.D., Le, T., Fan, G., 2012. DNA methylation and its basic function.
520 Neuropsychopharmacology. 38(1), 23-38.
521 Nelson, G., Hoon. M.A., Chandrashekar, J., Zhang, Y.F., Nicholas, J.P.R., Zuke, C.S., 2001.
522 Mammalian sweet taste receptors. Cell. 106, 381-390.
523 Pan, Y.X., Chen, L., Cheng, J., Zhu, X., Wu, P., Bao, L.S., Chu, W.Y., He, S., Liang, X.F., Zhang,
524 J.S. 2021. Genome-wide DNA methylation profiles provide insight into epigenetic regulation
525 of red and white muscle development in Chinese perch *Siniperca chuatsi.* 256, 110647
526 Pasquet, A. 2018. Effects of Domestication on Fish Behaviour. DOI: 10.5772/intechopen.78752.
527 Song, M., Zhong, Y., Guo, J., Chen, Y., Luo, L., Lin, S., 2019. Effects of dietary lipid levels on the
528 expression of cpt1 in *Micropterus salmoides*. Journal of Fisheries of China. 43, 2166-2174.
529 Smulders, T.V., Devoogd, T.J., 2000. Expression of immediate early genes in the hippocampal
530 formation of the black-capped chickadee (*poecile atricapillus*) during a food-hoarding task.
531 Behav Brain Res, 114(1), 39-49.
532 Shen, Y.W., Li, H.Y., Zhao, J.L., Tang, S.J., Chen, X.W., 2021. The digestive system of mandarin
533 fish (*siniperca chuatsi*) can adapt to domestication by feeding with artificial diet. Aquaculture,
534 538, 736546.
535 Tang, L.J., Jacquin, L., Lek, S., Liu, H.Z., Li, Z.J., Liu, J.S., Zhang, T.L., 2017. Differences in

536 anti-predator behavior and survival rate between hatchery-reared and wild grass carp
537 (*Ctenopharyngodon idellus*). Ann Limnol Int. 53, 361-367.
538 Teletchea, F., Fontaine, P., 2012. Levels of domestication in fish: implications for the sustainable
539 future of aquaculture. Fish Fish, 15(2), 181-195.
540 Teletchea, F., 2016. Is fish domestication going too fast? Nat Resour, 07(6), 399-404.
541 Podgorniak, T., Brockmann, S., Konstantinidis, I.,Fernandes., J.MO., 2019. Differences in the fast
542 muscle methylome provide insight into sex-specific epigenetic regulation of growth in Nile
543 tilapia during early stages of domestication. Epigenetics. 14(8).
544 Wilson, G.A., Pawandeep, D., Andrew, F., Daniel, C., Yuka, S., Reiner, S., Schär, P., Beck, S.,
545 2012. Resources for methylome analysis suitable for gene knockout studies of potential
546 epigenome modifiers. Gigascience, 1(1), 3-3.
547 Wu, B.L., Huang, L., Chen, X.J., Zhang, Y., Chen, X.J., Wu, C.C. Deng, X.J., Gao, J., He, J.X.,
548 2021. Effects of feeding frequency on growth performance, feed intake, metabolism and
549 expression of fgf21 in grass carp (*ctenopharyngodon idellus*). Aquaculture, 545(15), 737196.
550 Wang, Y.P., Lu, Y., Zhang, Y., Ning, Z.M., Li, Y., Zhao, Q., Lu,H.Y., 2015. The draft genome of
551 the grass carp (*Ctenopharyngodon idellus*) provides insights into its evolution and vegetarian
552 adaptation. Nat Genet. 47, 625-631.
553 Xiang, H., Zhu, J.D., Chen, Q., Dai, F.Y., Li, X., Wang, J., 2010. Single base-resolution
554 methylome of the silkworm reveals a sparse epigenomic map. Nat Biotechnol. 28(5), 516-20.
555 Yang, Y.J., Zheng, Y.Q., Sun, L.N., Chen, M.Y. 2020. Genome-wide DNA methylation signatures
556 of sea cucumber *apostichopus japonicus* during environmental induced aestivation. Genes,
557 11(9), 1020.

558 Yang, Y.L., Zhou, R., Mu, Y.L., Hou, X.H., Tang, Z.L., Li, K., 2016. Genome-wide analysis of
559 DNA methylation in obese, lean, and miniature pig breeds. Sci Rep-UK. 6, 30160.
560 Yang, Y., Yu, H., Li, H., Wang, A,L., Yu, H.Y., 2016. Effect of high temperature on immune
561 response of grass carp (*ctenopharyngodon idellus*) by transcriptome analysis. Fish Shellfish
562 Immun. 58, 89-95.
563 Yong, W.S., Hsu, F.M., Chen, P.Y., 2016. Profiling genome-wide DNA methylation. Epigenet
564 Chromatin. 9(1), 26.
565 Zemach, A., McDaniel, I.E, Silva, P., Zilberman, D., 2010. Genome-wide evolutionary analysis of
566 eukaryotic DNA methylation. Science. 328, 916.
567 Zhang, Y.Q., Liu, J.H., Fu, W., Xu, W.T., Zhang, H.Q., Chen, S,J., Xiao, Y.M., 2017. Comparative
568 transcriptome and DNA methylation analyses of the molecular mechanisms underlying skin
569 color variations in Crucian carp (*Carassius carassius L*). BMC Genet. 18(1), 95.
570 Zhao, Y.M., Zhang, L., Wang, C.F., Xie, C.X., 2020. Biology and ecology of grass carp in china: a
571 review and synthesis. N Am J Fish Manage. 40(6), 1379-1399.
572
573
574
575
576
577
578
579

Table 1

Alignment statistics with reference genome

Sample ID	Clean Reads	Mapped Reads	Mapping Rate (%)	Uniquely Mapped Reads	Uniquely Mapping Rate (%)	Bisulfite Conversion Rate (%)
DGC1	378,121,760	347,944,087	92.02	329,583,800	87.16	99.28
DGC2	387,627,636	353,434,830	91.18	335,393,994	86.52	99.23
DGC3	321,704,534	296,306,210	92.11	281,433,956	87.48	99.16
WGC1	371,478,196	321,457,393	86.53	300,678,599	80.94	99.16
WGC2	493,819,092	425,920,306	86.25	398,434,532	80.68	99.24
WGC3	410,419,590	361,301,972	88.03	339,621,985	82.75	99.22

Note: DGC, domesticated grass carp; WGC, wild grass carp

Table 2

Average methylation levels of different genomic regions.

Groups	Regions	C (%)	CG (%)	CHG (%)	CHH (%)
DGC1	Genome	9.76	80.38	0.99	1.01
DGC2	Genome	9.69	80.55	1.18	1.06
DGC3	Genome	9.68	79.98	1.11	1.17
WGC1	Genome	9.21	74.7	1.12	1.18
WGC2	Genome	9.19	75.04	1.02	1.09
WGC3	Genome	9.13	75.39	1.07	1.12

Note: DGC, domesticated grass carp; WGC, wild grass carp. CG, CHG, and CHH (where H is A, C, or T)

Table 3

Proportion of CG, CHG and CHH in all Methyl-cytosine

		mCG	mCHG	mCHH
DGC1	mC number	26,890,172	208,311	750,652
	proportion (%)	96.557	0.748	2.695
DGC2	mC number	26,953,558	323,600	784,048
	proportion (%)	96.053	1.153	2.794
DGC3	mC number	26,520,365	219,138	811,594
	proportion (%)	96.259	0.795	2.946
WGC1	mC number	26,700,829	228,299	848,145
	proportion (%)	96.125	0.822	3.053
WGC2	mC number	27,258,571	220,189	831,299
	proportion (%)	96.286	0.778	2.936
WGC3	mC number	26,934,911	222,172	821,714
	proportion (%)	96.269	0.794	2.937

Note: DGC, domesticated grass carp; WGC, wild grass carp. CG, CHG, and CHH (where H is A, C, or T)

Figure 1. Distribution of methylation level of mC in each sequence context

The x-axis was defined as the percentage of reads showing mC at a reference cytosine site. The y-axis indicates the fraction of total mC calculated within bins of 10%. DGC, domesticated grass carp; WGC, wild grass carp. (A) Sample DGC 1; (B) Sample DGC2; (C) Sample DGC3; (D) Sample WGC1; (E) Sample WGC2; (F) Sample WGC3.

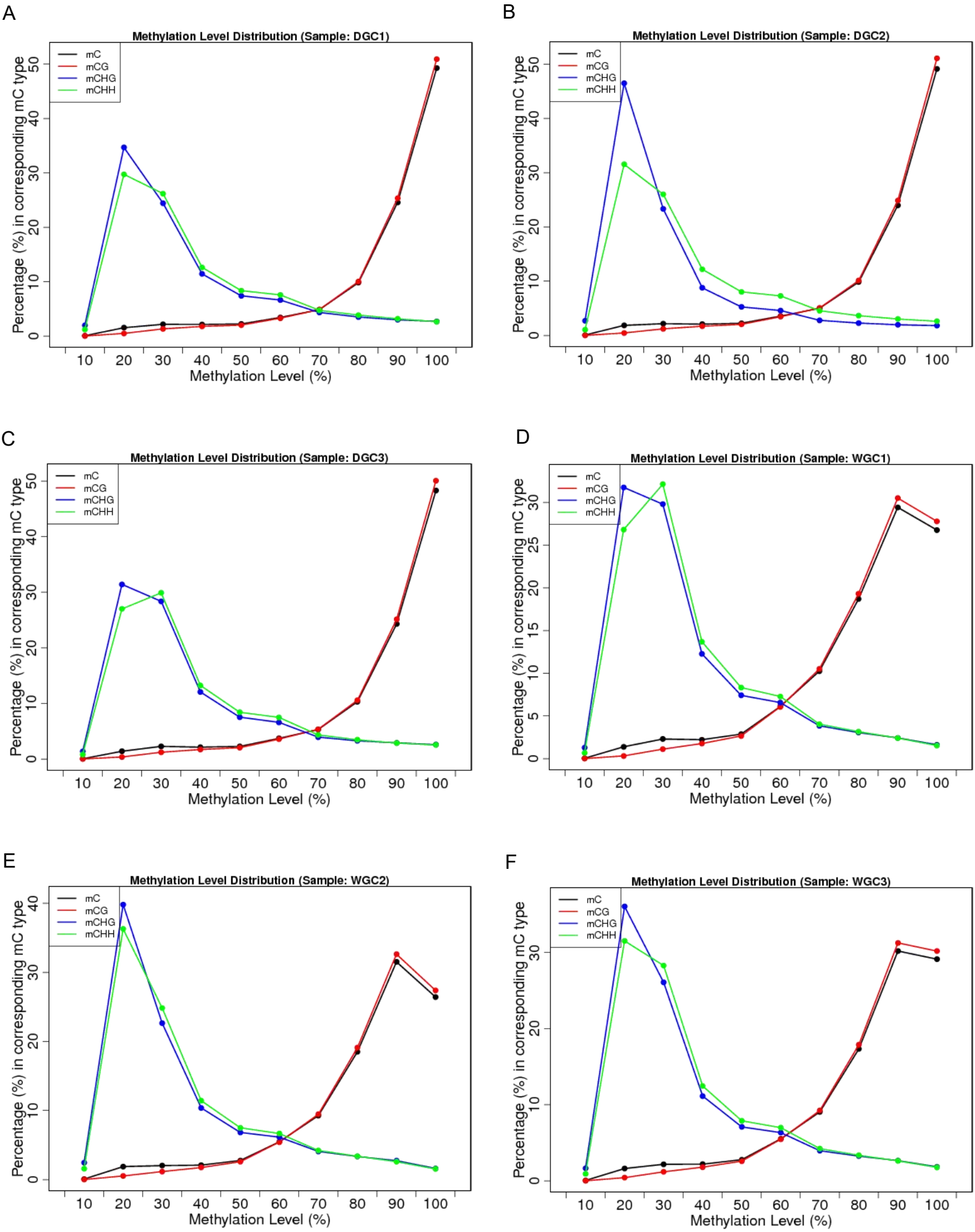
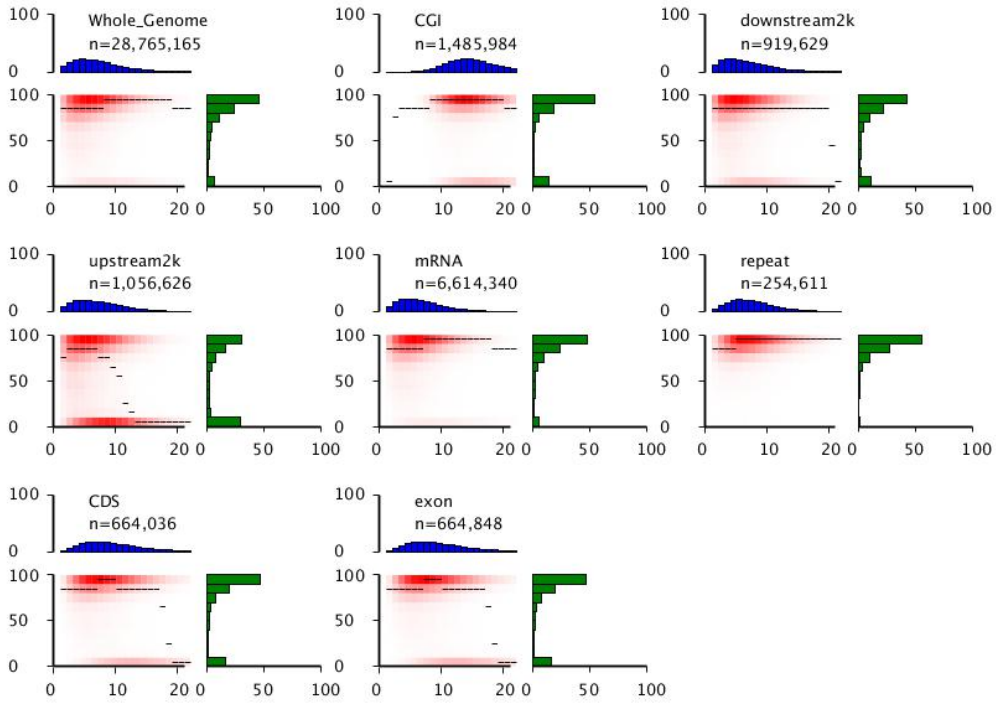


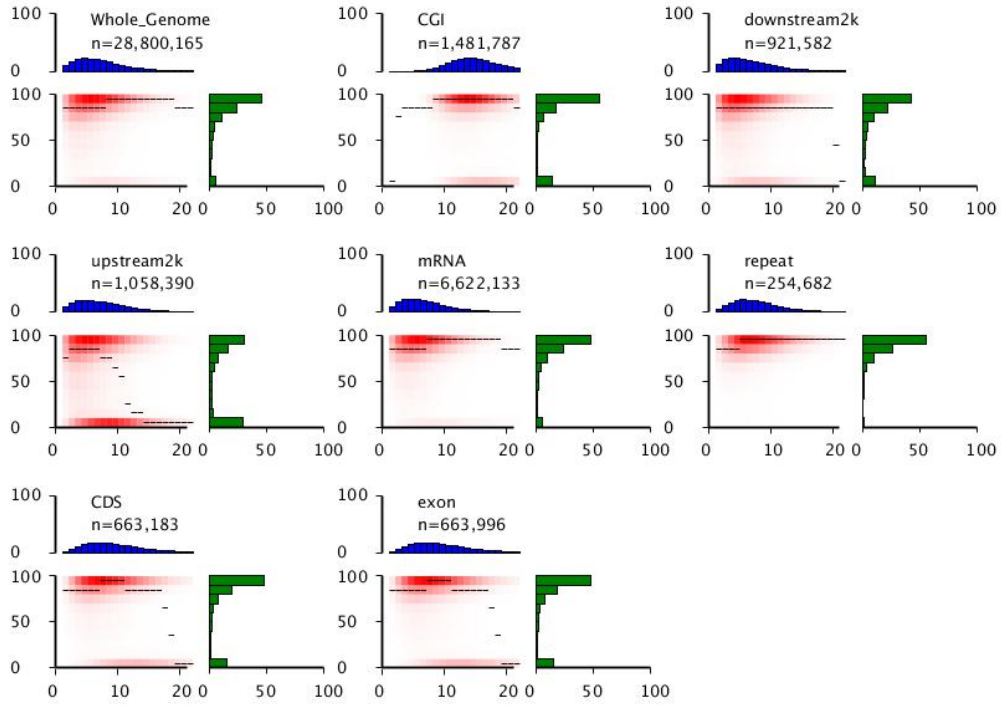
Figure 2. Heat maps show distinct methylation and CpG density patterns.

CpG density (x-axis) is defined as numbers of CpG dinucleotides in 200bp windows. Methylation level (y- axis) is defined as average methylation level of cytosines in CpGs. The thin black lines within each heat map denote the median methylation level of CpGs at the given local density. The red color gradient indicates abundance of CpGs that fall into bins of given methylation levels and CpG densities. The blue bar charts above each heat map show the distribution of CpG densities, projected onto the x-axis of the heat maps. The green bar charts to the right of the heat maps show the distribution of methylation levels, projected onto the y-axis of the heat maps. DGC, domesticated grass carp; WGC, wild grass carp. (A) Sample DGC 1; (B) Sample DGC2; (C) Sample DGC3; (D) Sample WGC1; (E) Sample WGC2; (F) Sample WGC3.

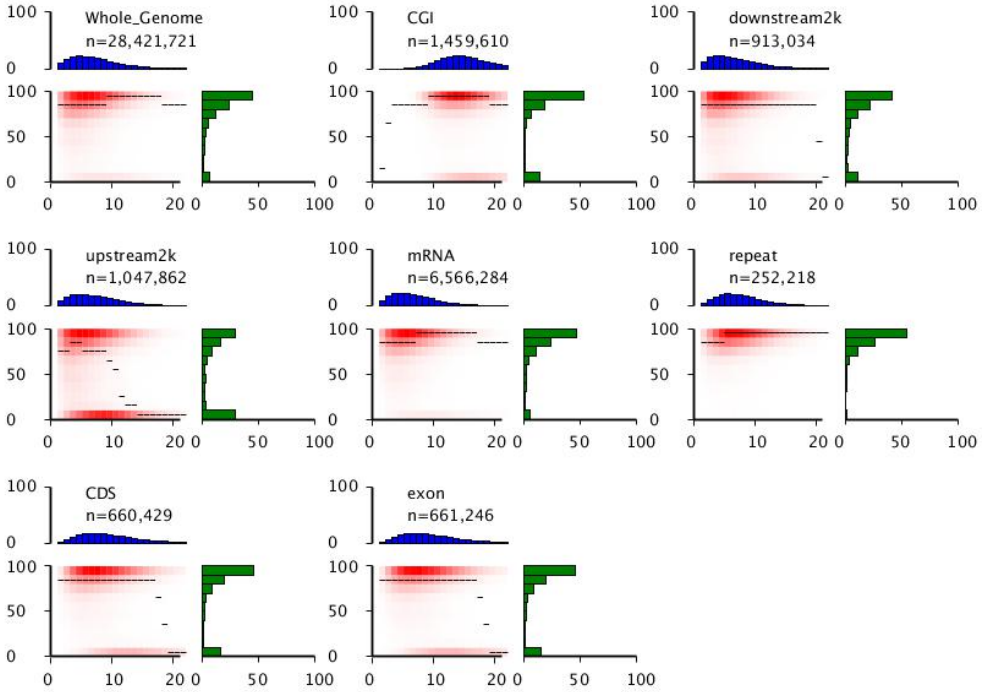
A



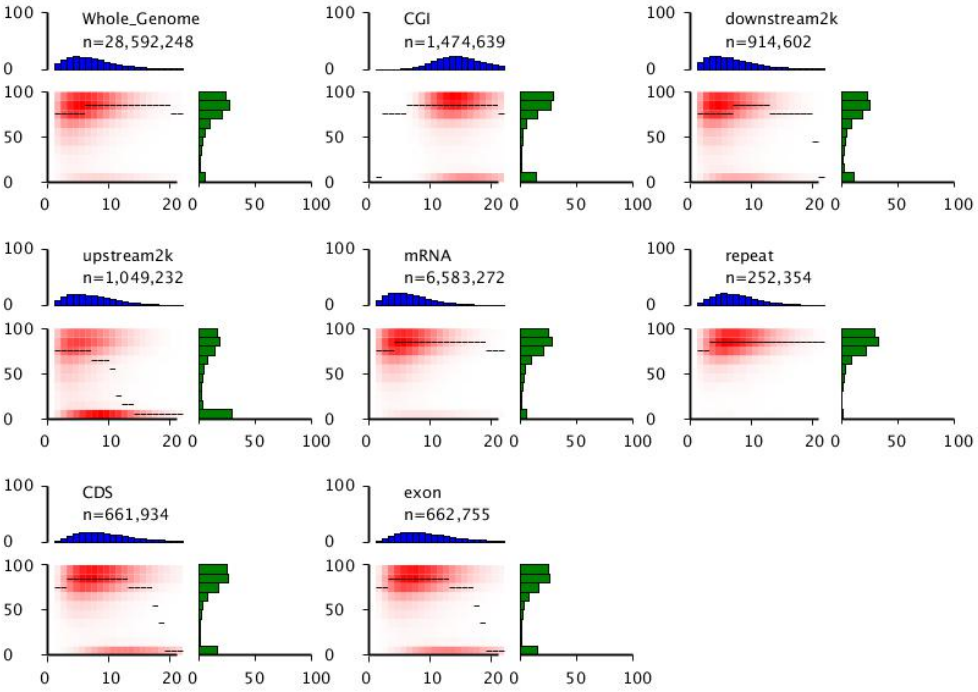
B



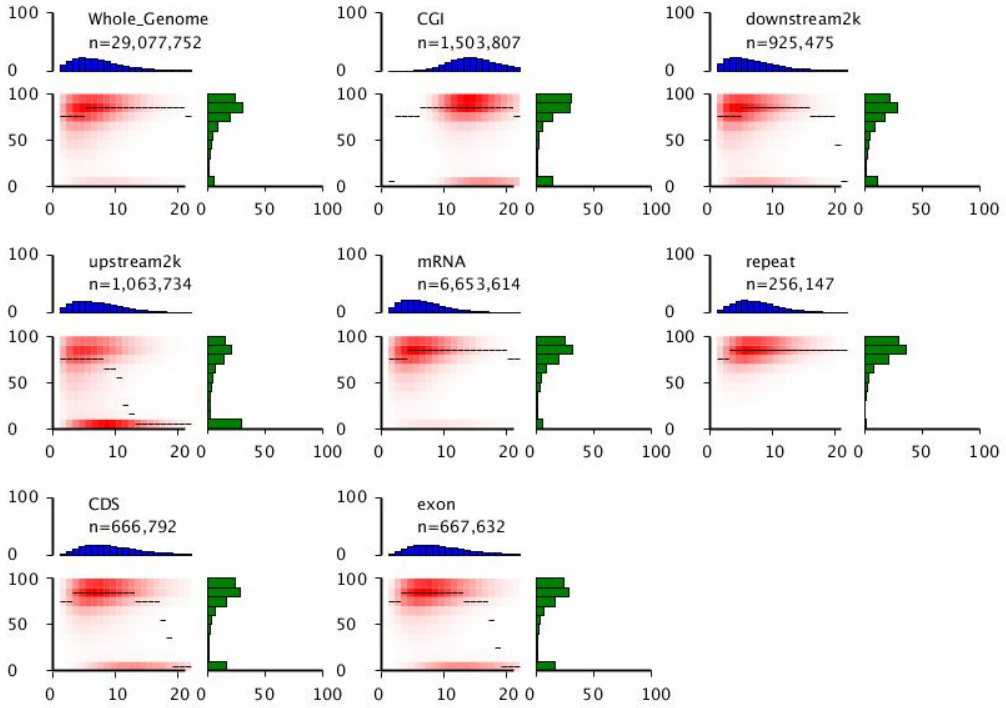
C



D



E



F

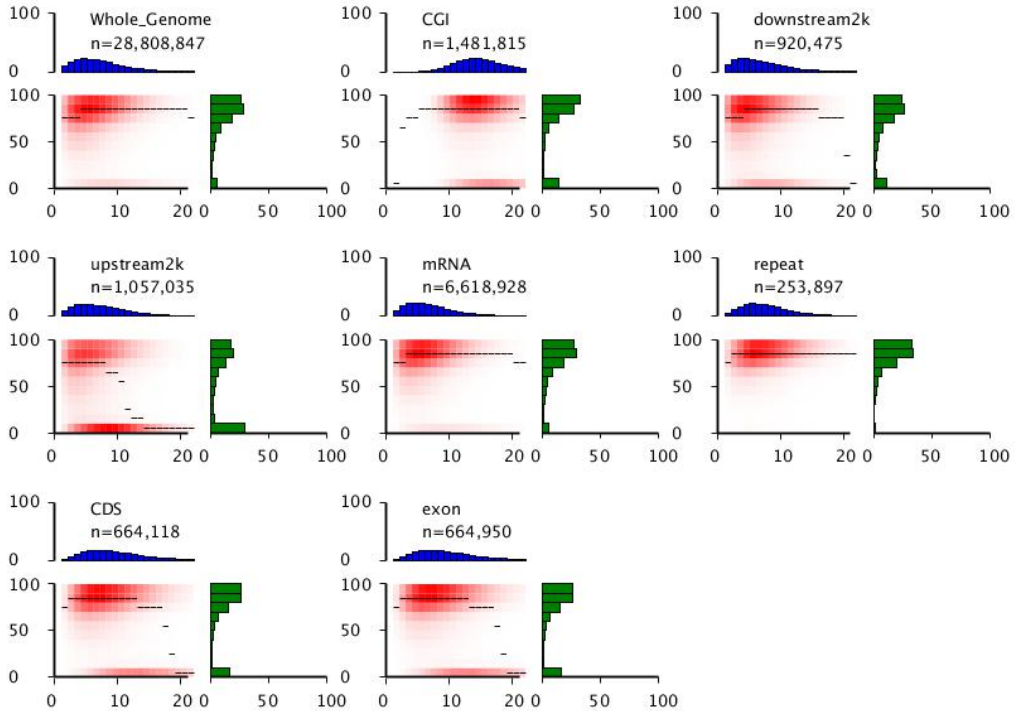


Figure 3. DNA Methylation Patterns Across the Entire Transcriptional Units at Whole Genome Level of sample DGC1. The canonical gene structure is defined by 7 different features, denoted by the x-axis. The length of each feature was normalized and divided into equal numbers of bins. Each dot denotes the mean methylation level per bin and the respective lines denote the 5-bin moving average. Each feature was analyzed separately for the numbers listed in the table below the figure. The green vertical line indicates the mean location of the transcription start sites. DGC, domesticated grass carp; WGC, wild grass carp. (A) Sample DGC 1; (B) Sample DGC2; (C) Sample DGC3; (D) Sample WGC1; (E) Sample WGC2; (F) Sample WGC3.

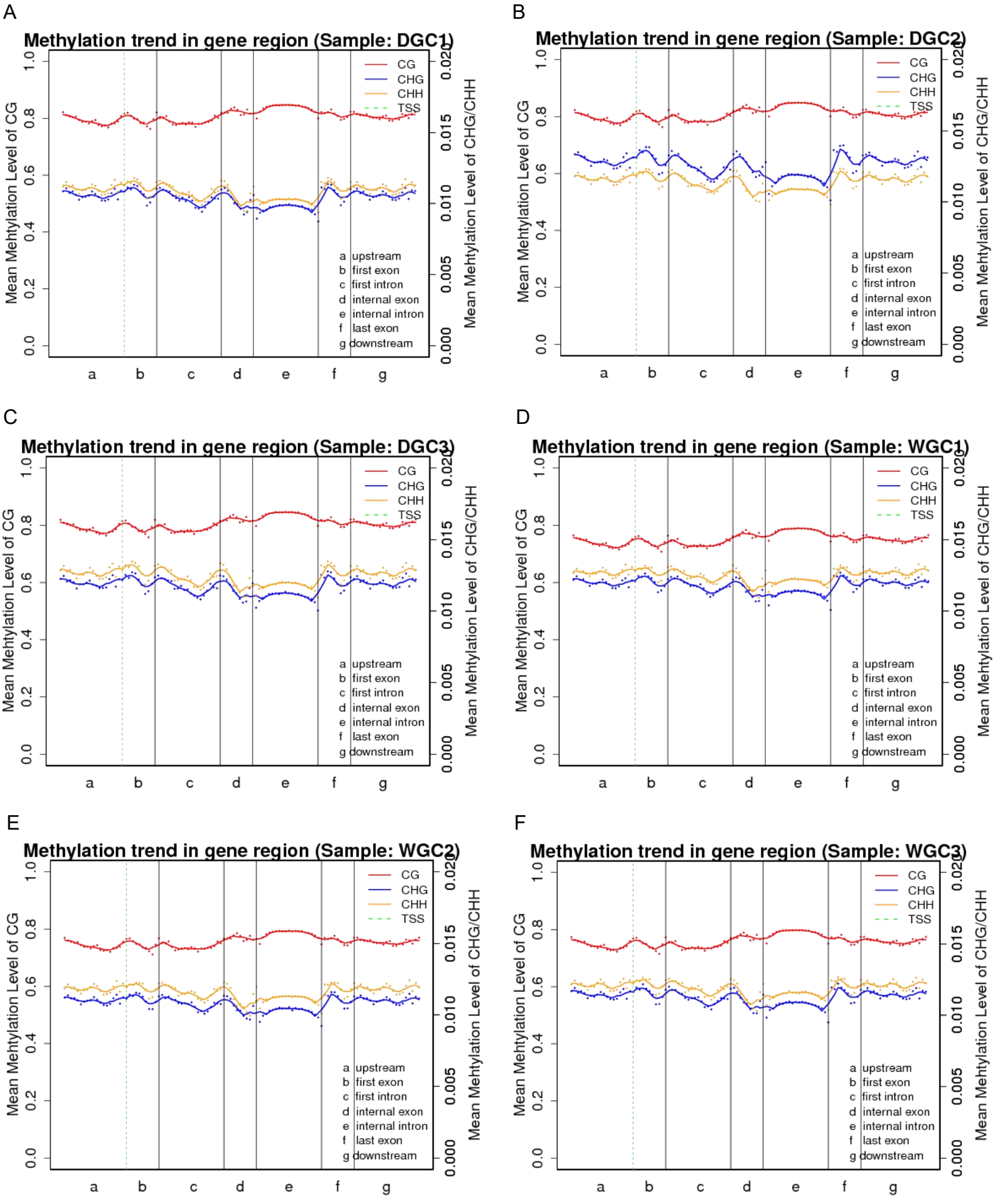
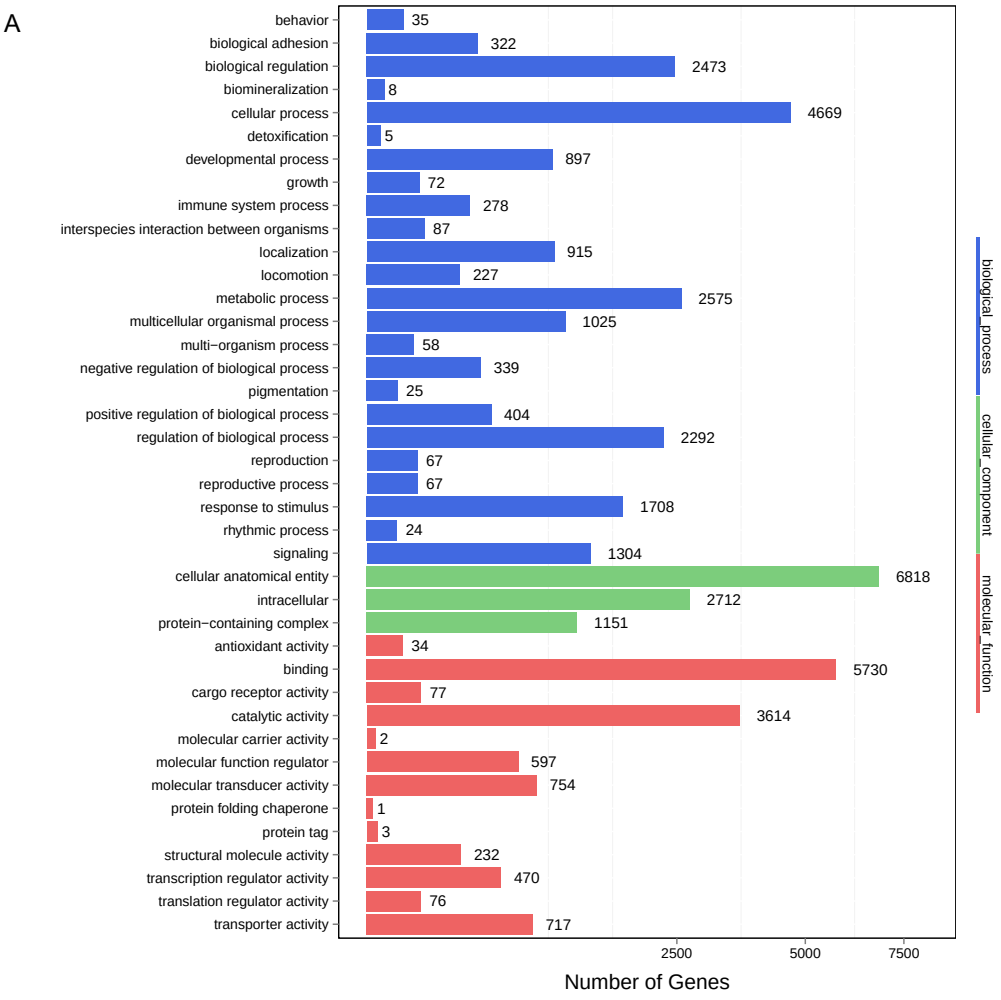


Figure 4. GO analysis of differentially methylated regions (DMRs)-related genes

The x-axis represents three domains of GO while the y-axis represents the gene number in every pathway and processes.

(A).GO analysis of differentially methylated regions (DMRs)-related genes in promoter region

(B).GO analysis of differentially methylated regions (DMRs)-related genes in gene body region



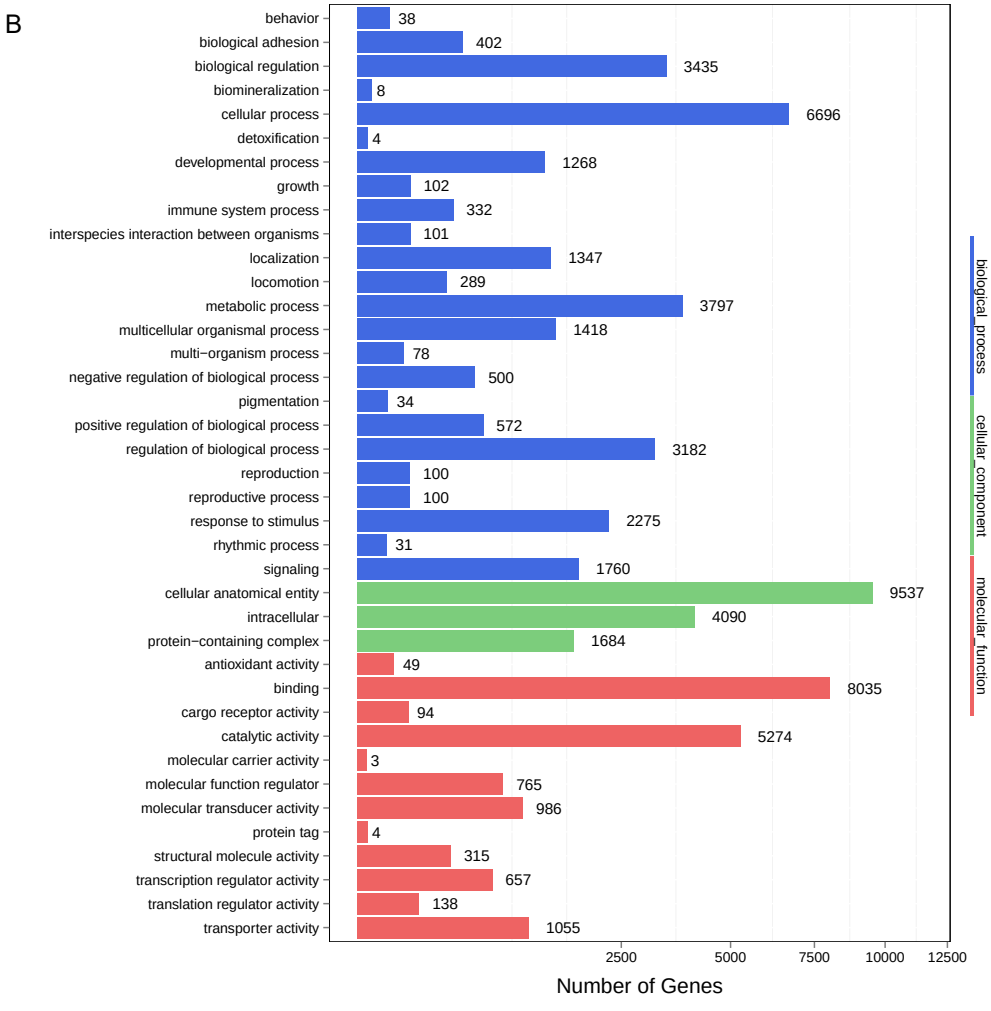
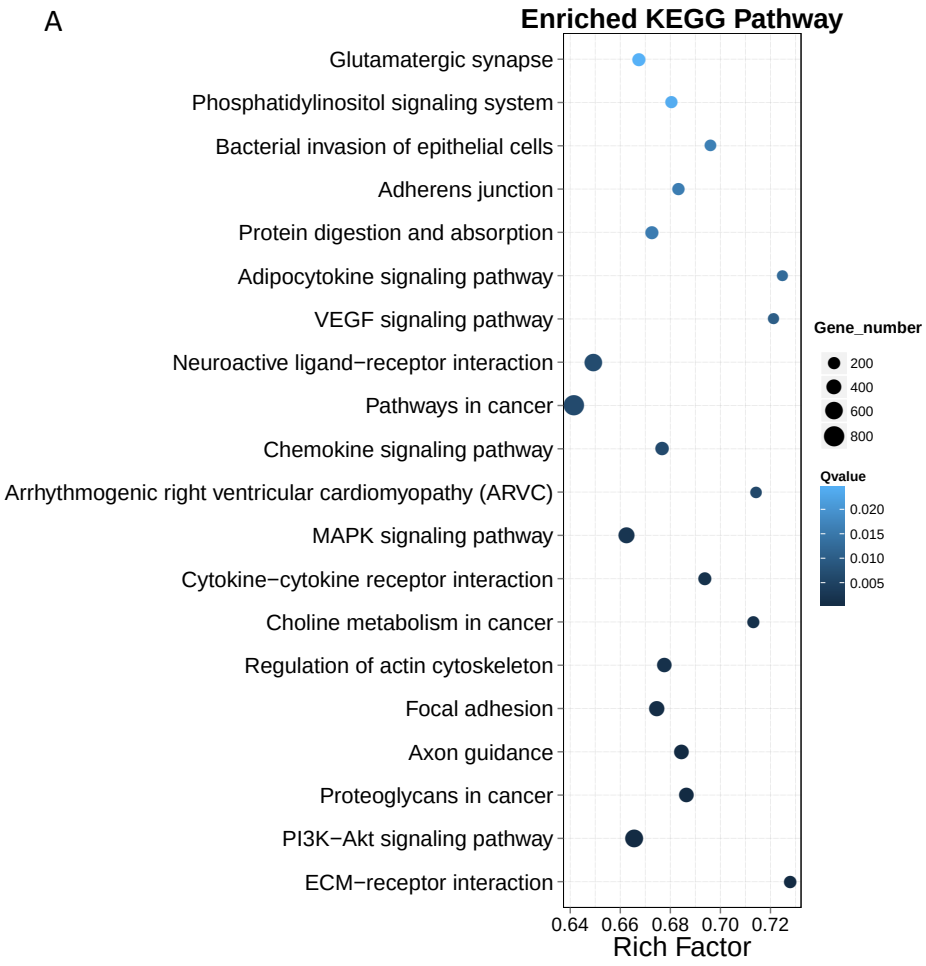


Figure 5. Pathway analysis of DMRs-related genes

The abscissa represent the richness factor, the ordinate represent the enriched pathway terms. Q-value represents the corrected P, and a small Q-value indicates high significance.

(A)Pathway analysis of DMRs-related genes in promoter region; (B) Pathway analysis of DMRs-related genes in gene body region.



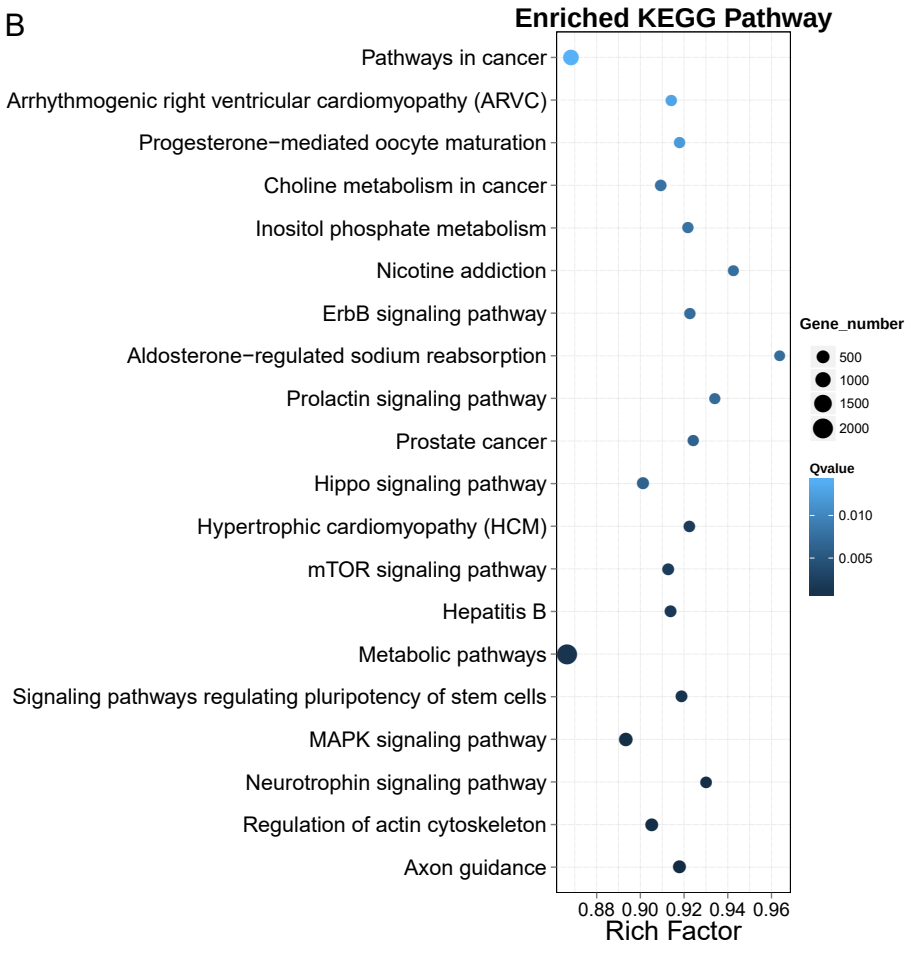


Figure 6. Protein-protein interaction network analysis of DMGs in different selection signatures.