Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Targeted DNA Methylation Analysis by High-Throughput Sequencing in Hepatic Carcinoma Tissue and Adjacent Normal Tissues

Guangmou Zhang
Xinxiang Medical University
Huigen Feng
Xinxiang Medical University
Zhiqing Yuan (yz5269@yeah.net)
Xinxiang Medical University

Research

Keywords: DNA methylation, High-throughput sequencing, Hepatic carcinoma

Posted Date: October 8th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-78102/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Abstract

Background: The relationship between epigenetic abnormalities and tumorigenesi has been investigated in the past decade and made major advances, particularly the abnormal expression of small RNAs, DNA methylation, and histone modification in cancer. In many tumor-related studies, the regulatory changes in DNA methylation during cancer development and the development of resistance to anticancer drugs have show that DNA methylation can be used as a biomarker for cancer diagnosis and concomitant diagnosis, but there is a lack of clinically useful biomarkers associated with hepatic carcinoma. Using high-throughput sequencing technology, appropriate testing and validation can be carried out in large samples. The relationship between DNA methylation and tumor development can be explored, contributing to clinical diagnosis and personalized treatment of hepatic carcinoma.

Methods: In this study, we implemented and evaluated the effectiveness of high-throughput sequencing for DNA methylation analysis in hepatic carcinoma. For the relationship between DNA methylation and gene expression, Pearson correlation analysis was used to evaluate the correlation. Twenty-five isolated genomic regions were amplified by PCR using bisulfite-transformed liver cancer tissue (Ca) and paracancer tissue (T) as template DNA. PCR final product sequence information was obtained by sequence analysis using Illumina Hiseq/Miseq platform.

Results: The average depth of coverage across all amplicons was 30,548 for T and 29,346 for Ca, with a maximum of 3,675 at the ARID1A amplicon and a minimum of 65 at the PTEN amplicon. Methylation spectra were obtained for each genomic locus of the two groups of samples, and the results showed that methylation was significantly different at the X target loci and slightly different at the Y target loci. Cluster analysis showed that all T tissues were clustered in one group (except tissues T2 and T3), while Ca tissues were clustered on the other side. The results showed that DNA methylation at the loci may be closely related to liver cancer, providing references for the research and development of biomarkers in clinical diagnosis.

Conclusions: The study demonstrates that high-throughput sequencing technology is a powerful and cost-effective method for methylation analysis of target DNA in cancer tissues.

Background

The liver is one of the important organs in the human body and is involved in several critical functions such as detoxification, metabolism of protein, glucose and fat, glycogen storage and decomposition of red blood cells. Abnormal liver functions may lead to several complications and diseases. A recent study identified that along with reduced folate levels, individuals with type 2 diabetes exhibited changes in DNA methylation and gene expression in the liver, supporting that epigenetic modifications may contribute to liver function[1].

DNA methylation at cytosines of CpG dinucleotides is the most commonly studied epigenetic mark, and numerous environmental factors were found to alter the epigenome. As DNA methylation can easily be

investigated in the laboratory, it can be applied for epigenome-wide association studies and used as epigenetic biomarkers for the prediction of complex diseases. Age-related changes in DNA methylation have been observed in several types of human tissues such as the brain, skeletal muscle, adipose tissue, and pancreatic islets[2, 3]. Epigenetic alterations in blood mirror age-associated DNA methylation and gene expression changes in human liver[4].

DNA methylation is usually associated with diseases such as cancer and autoimmune disorders, occurring with environmental exposure and other biological phenomena. Strong correlations between DNA methylation patterns and clinical phenotypes can serve as biomarkers for disease diagnosis and treatment guidance. For example, DNA methylation biomarkers have been shown to support clinical decisions in various cancers and are also used in non-invasive prenatal testing, quality control of cultured cells, and forensic applications[5].

To investigate the tumorigenesi and the corresponding treatment options, the study of genetic and epigenetic background information related to liver cancer will promote the understanding of the pathogenesis of liver cancer. As the one of the most widely used sequencing platforms, the application of MiSeq in this study makes it possible to analyze DNA methylation of single base pairs and contributes to the understanding of more detailed regulatory mechanisms. As an important epigenetic marker, DNA methylation is involved in a series of functions, such as suppression or activation of gene expression, nucleosome positioning transcription factor binding, and splicing[6]. Methylation polymorphism is an important reason for individual phenotypic difference, and methylation variation may also often lead to individual phenotypic difference. A large number of studies have shown that abnormal methylation is closely related to cancer, diabetes, Parkinson's disease and other complex diseases, methylation detection has great significance foro biology and translational medicine.In this study, we used multiple PCR technology combined with Illumina Hiseq/Miseq second-generation sequencing platform to capture and sequence multiple specific CpG islands at the same time, and accurately calculate the methylation level of each CpG site with high-depth sequencing data. By analyzing the the expression and methylation status of 25 cancer-related genes in liver cancer and its adjacent tissues to explore the correlation between DNA methylation and the development of liver cancer

Methods

Preparation of DNA samples

Twenty-six pairs of fresh liver cancer tissue and adjacent normal liver tissue samples were obtained from the hospital, as approved by Xinxiang Medical University (XYLL-2016025). All samples were immediately frozen in liquid nitrogen after surgery and stored at -80℃. Genomic DNA was isolated from these samples via the phenol-chloroform method, after incubation with proteinase K and DNA lysis buffer (50 mM Tris-HCl at pH 8.0, 100 mM ethylenediaminetetraacetic acid at pH 8.0, 100 mM NaCl, 1% sodium dodecyl sulfate). The quality of genomic DNA of each sample was good. The cleaning bands were detected by agarose gel electrophoresis, and there was no obvious degradation or RNA contamination.

The quality and quantity of the DNA were determined by NanoDrop 2000 spectrophotometer (NanoDrop Technologies, DE, USA). Total DNA concentration was ≥ 20 ng/ml and the total content was ≥ 1 mg. The DNA was stored at -20 °C until use.

Genomic DNA methylation

According to the manufacturer's instructions, both liver cancer and adjacent normal liver tissue samples DNA were subjected to bisulfite treatment using the QiagenEpiTect Bisulfite Kit (QIAGEN, Hilden, Germany). The EpiTect Spin Colums was utilized to purify the methylated DNA.

Sulfites treat DNA:

1) Defrost the DNA and dilute the Bisulfite Mix with 800 µl Rnase-free water, swirling and oscillating until all the Bisulfite Mix is completely dissolved.
2) DNA was stabilized to 20 µl by Rnase-free water for sulfite reaction, the reaction system as follows:

DNA: 20 µl

Bisulfite Mix: 85 µl

DNA Protect Buffer: 35 µl

Total: 140 µl

3) At room temperature, the mixture will change from green to blue with the addition of the DNA Protect Buffer.
4) Sulfite reaction was performed using a PCR instrument.The reaction procedure is as follows: 95℃ 5min→60℃ 25min→95℃ 5min→60℃ 85min→95℃ 5min→60℃ 175min→20℃ Forever.

Target area enrichment

According the following list of Cancer-related genes (Table S1), we designed forward and reverse primers for methylated and unmethylated statuses with MethPrimer (Li Lab, China, Table S2). The volume of each PCR reaction mix is 20 µl, which includes 10.0 µl of KAPA2G Robust Hot Start ReadyMix, 1.0 µl of forward primer, 1.0 µl reverse primer, > 5 ng of bisulfite-treated genomic DNA, and topped with double-distilled water to 20 µl. The PCR reaction mix was then subjected to thermal cycling for target amplification. PCR was carried out with the following conditions: 3 min of initial enzyme activation; 45 cycles of 94 °C at 30 s, 55 °C at 30 s, and 72 °C at 40 s; and final extension at 72 °C for 10 min. PCR products of the same template were mixed by the same volume and purified by backman agtampure XP magnetic beads. The products were subjected to subsequent tests after product concentration detection.

DNA methylation analysis

Illumina plate-compatible specific tag sequences were introduced to the end of the library by PCR amplification using a primer with Index sequence. The Index PCR amplification products of all samples were equally mixed and the final MethylTarget sequencing library was obtained after glue cutting recovery. After accurate quantization of library molality, the Illumina Hiseq/Mis EQ platform was finally used for high-throughput sequencing in a 2 × 150 bp/2 × 250 bp double-end sequencing mode to obtain FastQ data.

Statistical analysis

Statistical analysis was performed using the statistical software package SPSS21 (SPSS, Chicago, USA). We used a chi-square test to compare the abundances of the three gene methylation statuses (methylated, unmethylated, and partially methylated). Two-tailed chi-square (χ^2) test or Fisher's exact test was performed to analyze the correlation between gene expression/methylation and clinicopathological parameters. The Chi2 test was conducted to compare the genomic distribution of significant CpG sites.

Results And Discussion

DNA methylation analysis

The main principles of DNA methylation analysis be classified into various types based on pretreatment (sodium bisulfite, concentrated enzyme digestion, and affinity) and analytical method (next-generation sequencing-based, locus-specific, gel-based, and array-based analyses) [7]. Among them, sodium bisulfite-modified genomic DNA is one of the most commonly used techniques to evaluate CpG methylation status by modifying genomic unmethylated cytosine to uracil [8, 9]. As the gold standard of DNA methylation analysis, bisulfite-based DNA methylation offers the advantage of quantitative assessment, detection sensitivity, and ability to analyze a wide variety of samples [10]. However, it is limited by the amount of DNA because the transformation of sodium bisulfite requires DNA denaturation before treatment, and the removal of sodium bisulfite after purification results in a large amount of degraded DNA. Most published studies on DNA methylation have used low-throughput, conventional bisulfite (Sanger) sequencing for the analysis of mammalian samples. With advances in nucleic acid sequencing technologies, next-generation sequencing (also known as deep or high-throughput sequencing) makes it possible to generate millions of base pairs of sequence information in a matter of hours [11]. At present, most reports on DNA methylation based on high-throughput sequencing are focused on genome-wide epigenomic analysis [12, 13] and is very expensive. However, because of the large size of the mammalian genome, statistics on the ratio of methylated to non-methylated cytosine in the genome are not reliable, and it is difficult to ensure sufficient coverage depth for each CpG site. Absolute DNA methylation assays provide a quantitative measurement of DNA methylation levels at single-CpG resolution. In this study, 25 discrete genomic regions were amplified by PCR using bisulfite-converted genomic DNA derived from hepatic carcinoma and adjacent normal tissues as template DNA. The resulting PCR products were subjected to high-throughput sequencing using the Illumina Genome Analyzer IIx platform.

Although the accumulation of genetic and epigenetic abnormalities has been shown to play an important role in carcinogenesis, the molecular mechanism of hepatic carcinoma remains to be further studied[14]. Epigenetic effects such as microRNAs, DNA methylation, histone modification, non-coding Rnas, and nucleosome localization have similar functions, whilch can alter gene expression without altering DNA sequence [15–19]. In the past few decades, it has become increasingly clear that altered epigenetic regulation plays a key role in many diseases, particularly cancers[18]. As the most widely studied epigenetic mechanism, abnormal expression of DNA methylation has been found to be related to tumor development in many studies. For example, in the study of gastric carcinogenesis, it was found that the tumor suppressor genes RUNX3 and cancer-related genes CDH1 and MLH1 were regulated by promoter hypermethylation and silenced in the early stage of carcinogenesis[16]. In this study, we selected 25 tumor-related genes based on comprehensive literature review. Next-generation sequencing has the ability to detect small amounts of methylated DNA. To explore the distribution of CpGs in the selected genes, DNA methylation can serve as a molecular biomarker in GC in a variety of samples, such as serum, plasma, and tissue. In total, we applied 33 samples, among which 26 were used for high-throughput sequencing (13 pairs of hepatic carcinoma and adjacent normal tissues).

Sequence information and chromosomal coordinates for these 25 tumor-related genes are shown in Table 1. Genomic DNA was collected from homologous hepatic carcinoma tissue (Ca) and adjacent normal tissues (T) and subjected to bisulfite conversion, providing a mark to differentiate unmethylated from methylated cytosine bases within the nucleotide sequences evaluated. A single PCR amplicon from each of the 25 distinct tumor-related genes was generated using bisulfite-converted DNA from hepatic carcinoma and adjacent normal tissues. The average amplicon length was 279.4 base pairs (bp; range 120–650 bp), covering an average of 36.8 CpG dinucleotides per amplicon (range 2-145 CpG sites). The optimized multiple PCR primer panel was used to perform multiple PCR amplification using the transformed sample genome as the template.

Table 1
Summary of PCR products and number of Illumina reads aligned for each CpG island studied

PCR product	Length (bp)	#CpG	T Reads	Ca Reads
STK11	443	59	24050	36279
PTEN	438	57	5704	6147
MET-1	258	40	23909	22260
MET-2	282	24	775256	738751
NFE2L2-1	237	20	337	187
NFE2L2-2	233	36	98322	69172
KEAP1-1	215	28	1533	1526
KEAP1-2	181	18	809359	744181
CTNNB1-1	214	22	195809	175086
CTNNB1-2	172	36	56113	47358
FGFR3-1	198	10	61834	58389
FGFR3-2	247	52	28233	44704
FBXW7	203	20	2958	2772
PIK3CA-1	175	30	141784	131981
PIK3CA-2	285	60	32522	9634
MAP2K1-1	214	6	31852	22975
MAP2K1-2	199	17	45349	22565
EGFR-1	208	17	135042	80566
EGFR-2	241	46	50172	37268
ERBB2-1	268	46	39532	17938
ERBB2-2	246	24	25729	16240
ARID1A-1	358	46	37157	36329
ARID1A-2	120	22	14002	6387
CBL-1	650	145	1173	3745
CBL-2	274	56	10773	9594
ARID2-1	346	54	216472	58561

PCR product	Length (bp)	#CpG	T Reads	Ca Reads
ARID2-2	130	12	262008	254750
ERCC2-1	247	22	48719	31427
ERCC2-2	201	28	5279	1181
EP300-1	187	24	34433	21379
EP300-2	180	44	12620	8345
NOTCH3	259	26	320072	262190
NOTCH2	466	94	7209	6077
ERCC1-1	259	22	536945	477879
ERCC1-2	187	22	103036	93346
RB1	441	89	45692	18105
ABCA3-1	406	2	34039	27829
ABCA3-2	313	40	16364	24719
ABCA3-3	367	43	13424	16205
PRDM14-1	343	26	198822	41791
PRDM14-2	401	39	41885	31167
PRDM14-3	442	69	14448	17588
DAB2IP	277	28	1531	2272
RAB31-1	264	16	308870	235449
RAB31-2	298	19	1399534	1126432
Total			6269906	5098726

Gene coverage depth analysis

A total of 11,368,632 sequencing readings were aligned and overall, the number of readings of T sample is significantly higher than that of Ca sample. Table 1 summarizes the statistics of the number of reads aligned and used for downstream methylation analyses for each CpG island queried. Although more readings were obtained in the T sample, the differences between the readings for each gene in the T sample were small and followed the same trajectory (except for T14 sample). Similarly, the distribution of each gene in the Ca sample was uniform. Figure 1A and Fig. 1B show the coverage depth of each gene in four Ca samples and four T samples, respectively. And Fig. 1C and Fig. 1D show the coverage depth of four tumor suppressant genes and cancer-related genes in different Ca samples and T samples, respectively. Because read distribution bias was extreme for one sample (T14; other genes are not shown)

and large gaps in coverage were observed, this genomic locus was not evaluated further. Although each gene has different coverage depth in the same sample, the locus of read distribution of each gene is consistent in all samples (Fig. 1A, Fig. 1B). The average coverage depth across all genes in sample T was 30,548, Whole Coverage was 96.16%, while the average coverage depth across all genes in sample Ca was 29346 and Whole Coverage was 98.13%. Notably, the sequencing results of T14 samples showed that only 74.83% of bases could completely cover the reads of their target sequence, which was significantly lower than the overall coverage of other samples. Therefore, the samples were removed in the subsequent analysis. ARID1A had the highest average coverage depth of all genes, reaching 3841 in Ca samples and 3995 in T samples, respectively.The average coverage depth of PTEN was the lowest, with only 47 and 71 amplitudes detected in Ca and T samples.

Analysis of differences in methylation ratio

Using a sophisticated single nucleotide polymorphism identification algorithm, we can accurately determine the position of each methylation site, the sequencing depth, and the methylation ratio in the 25 amplicons of a single sample. The difference in methylation ratio of 25 amplicons (and their subtypes) between the two groups of samples was obtained through data analysis. Three methylated types (CG, CGH, and CHH) were analyzed, and their methylation ratios of these three methylated types in the amplicons of Ca and T were calculated. Sequencing results showed that CG-type methylation was present in all amplicons, including 45 of the 46 subtypes. CHG and CHH were only present in 4 amplicons (nine subtypes). The comparison of methylation data is shown in Fig. 2. In Fig. 2A, we observed that among the 45 amplicon subtypes, the methylation ratio of 3 amplicons (DAB2IP, 15.44% difference; Prdm14-1, 16.34% difference; Rab31-1, 14.36% difference) showed a significant difference between Ca and T (greater than 10%), whereas 19 amplicons showed no significant difference (the methylation patterns were greater than 1% and less than 10%) and the remaining 23 amplicons only showed slight differences (less than 1%). The methylation ratios of 20 amplicons were up-regulated while those of 25 were down-regulated in Ca. It is worth noting that the methylation ratio of all 25 down-regulated amplicons was less than 4% between Ca and T, and most were less than 1%. In Fig. 2B, 3 amplicons showed insignificant differences in methylation ratio between Ca and T, and 6 showed extremely slight differences. Among the 9 amplicons, the methylation ratio of 3 amplicons was up-regulated in Ca, while that of 6 amplicons was down-regulated. Analysis of CHH methylation mode (Fig. 2C) revealed that 2 amplicons showed insignificant methylation ratio difference between Ca and T, and 7 showed extremely slight differences. Only one amplicon showed an up-regulated methylation ratio in Ca, while the others were down-regulated.

Since there is no clear threshold for DNA methylation, it is not clear whether these subtle differences have physiological significance. However, there is evidence that very small changes in DNA methylation patterns may also affect the physiology of cells and organisms [23–25]. Therefore, any change in methylation pattern need to be taken seriously. Physiological changes can be caused by changes in methylation of more than 10%, which can be regarded as the key analysis direction. For example, Waterland et al. found that a change in methylation of only 10% was enough to cause phenotypic

variation in live yellow aguti mice [26]. In view of the obvious changes in methylation ratio of DAB2IP, prdm14-1, and rab31-1, it is speculated that the methylation abnormalities of these 3 amplicons may be related to the growth and development of hepatic carcinoma cells. The specific relationship still needs further exploration.

Haplotype type analysis

According to analysis by genome-wide Association Studies (GWASs), the statistical data indicate that single nucleotide polymorphisms (SNPS) are the predominant form of DNA sequence variation.The occurrence of SNPS may be due to individual differences or disease susceptibility. Haplotype-dependent allele-specific methylation (hap-ASM) is also thought to impact disease susceptibility, but maps of this phenomenon using stringent criteria in disease-relevant tissues remain sparse. According to the research resultsIn of Mirabello et al., the HPV16 methyl haplotype determined by a new-generation sequencing method is considered to be associated with pre-cervical cancer [20]. Jo et al. found that haplotypes can affect DNA methylation levels and promoter activity in breast carcinomas [21]. In the study on the mechanism of DNA methylation and disease association, methyl-SeQ was used to sequence cells and tissues such as brain, T lymphocytes, glial cells, purified neurons and placenta, and a large number of unreported quantitative trait loci (mQTLs) and differential methylated regions (DMRs) were finally analyzed. The results showed that hap-ASM is highly tissue-specific; an important trans-acting regulator of genomic imprinting is regulated by this phenomenon; variations in CTCF and TF binding sites are an underlying mechanism, and maps of hap-ASM and mQTLs reveal regulatory sequences underlying supra- and sub-threshold GWAS peaks in immunological and neurological disorders [22].

Allele-specific methylation influences allele-specific expression. Therefore, the methylation state of haplotypes within genetically associated regions can determine epigenetic differences with potential functional effects. DNA methylation data and association-determined risk and non-risk haplotypes can be compared by haplotype-specific methylation analysis. With the help of high-throughput sequencing, we can carry out multi-dimensional methylation analysis. In this study, the amplicon was used as the unit of analysis to analyze the haplotype of CpG loci, and the results are listed in the site haplotype (Table 2).

Table 2
Haplotype type statistics

Target	Haplotype	Depth	Ca10	Ca12
ARID1A_1_	ttttttttttt	287238	0.571429	0.608743
ARID1A_1_	cttttttttttt	37373	0.099065	0.111211
ARID1A_1_	ttctttttttt	17717	0.063523	0.046122
ARID1A_1_	tttctttttttt	11133	0.021312	0.021255
ARID1A_1_	ttttttttcttt	10476	0.013474	0.016317
ARID1A_1_	tttttttctttt	8603	0.020624	0.017702
ARID1A_1_	ttttttctttttt	6722	0.014368	0.009995
ARID1A_1_	ctcttttctttt	1179	0.000894	0
ARID1A_2_	ttttttttttt	149929	0.773132	0.822343
ARID1A_2_	ttctttttttt	5491	0.025589	0.018645

Analysis of the difference in haplotype abundance

T-test, u-test, and logistic regression models were used to analyze the difference in haplotype abundance between the two groups of samples. Table 3 demonstrates that only a few haplotypes showed significant differences in abundance between the two groups of samples. For example, among the 8 haplotypes of the ARID1A_1_ sequence, only TTCTTTTT showed statistical significance after t-test analysis ($p \leq 0.5$). Distinct from genomic imprinting, where the methylation of an allele is determined by its parent-of-origin, for loci with hap-ASM, the local sequence context (haplotype) acts in cis to dictate the methylation status of local CpGs[22]. Abundance differences in haplotypes such as TTCTTT between Ca and T samples may affect tumor development by indicating the methylation status of local CpGs.

Table 3
Abundance differences in haplotype

Target	Haplotype	P-value (T-test)	P-value (U-test)	P-value (Logistic)
ARID1A_1_	ttttttttttttt	0.77824	0.9361618	0.7717829
ARID1A_1_	cttttttttttt	0.94024	0.9786974	0.9370238
ARID1A_1_	ttctttttttttt	0.033058	0.0768398	0.05901794
ARID1A_1_	tttttttttcttt	0.35586	0.8937907	0.4012169
ARID1A_1_	ttttttttctttt	0.74136	0.8517193	0.7187338
ARID1A_1_	tttttttctttttt	0.90396	0.7283391	0.8952096
ARID1A_1_	ctcttttctttt	0.35548	0.6046477	0.5704589
ARID1A_2_	ttttttttttt	0.22912	0.503319	0.2271915
ARID1A_2_	ttctttttttt	0.23392	0.3203219	0.2982954
ARID1A_2_	tttttttttttc	0.4623	0.8517193	0.4533269

Principal component analysis (PCA) and cluster analysis

PCA was performed on all samples based on the DNA methylation profiles to better differentiate the performance of methylated DNA. PCA showed that almost all Ca samples are closely clustered together and distinct from T samples (Fig. 3). This indicates that all Ca samples have the same trend of methylation status change, which may lead to changes in the physiological properties of normal samples.

To evaluate the classification power of the identified differentially methylated DNA regions, we performed cluster analysis according to the methylation level of CpG loci in the amplicon. The results showed that all samples were divided into two clusters (Fig. 4). T tissues were clustered in one group (except T2 and T3), while Ca tissues were clustered on the other side. Heatmap intuitively shows the relative methylation level of CpG sites in each sample. Figure 4 shows that almost all T samples showed low methylation levels in blue, while Ca samples on both sides showed higher methylation levels, indicated by orange or red. On the whole, this reveals that the methylation level of Ca samples is improved compared with that in homologous T samples.

Conclusion

High-throughput sequencing technology offers a powerful method for methylation analysis of target DNA in hepatic carcinoma tissue and adjacent normal tissues. Importantly, the cost of high-throughput sequencing is no higher than that of conventional bisulfite sequencing. In this study, high-throughput sequencing provided an average coverage depth of 30548 for T samples and 29346 for Ca samples. It allowed us to compare the methylation levels of homologous samples at each methylation site and

provided sufficient abundance for haplotype analysis, PCA, and cluster analysis. The results of homologous sample methylation comparison and haplotype analysis showed that there are some significant methylation sites and potential methylation regulatory units between the two groups. The difference in haploid and methylated sites between hepatic carcinoma and adjacent normal tissues may provide references for the research and development of biomarkers in clinical diagnosis. On the other hand, PCA and cluster analysis further revealed the difference in methylation level between the two groups of samples, which would result in different performance. This difference in methylation between hepatic carcinoma and adjacent normal tissues may be related to the growth and development of hepatic carcinoma.

Declarations

Ethics approval

This study is approved by Xinxiang Medical University (XYLL-2016025)

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The analyzed data sets generated during the present study are available from the corresponding author on reasonable request.

Funding

This study is supported by the grant from Xinxiang science and technology major project

Authors' contributions

ZGM and YZQ conceptualized and designed the study. ZGM and FHG all of the data and prepared the manuscript. YZQ revised the manuscript. All authors have read and approved the final version of this manuscript.

References

1. Nilsson E, Matte A, Perfilyev A, D de Mello V, Käkelä P, Pihlajamäki J, et al. Epigenetic alterations in human liver from subjects with Type 2 diabetes in parallel with reduced folate levels. J Clin Endocrinol Metab. 2015;100(11):E1491–501.

2. Day K, Waite LL, Thalacker MA, West A, Bamman MM, Brooks JD, et al. Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. Genome Biol. 2013; R102 .
3. Yang JY, Yu L, Gaiteri C, Srivastava GP, Chibnik LB, Leurgans SE, et al. Association of DNA methylation in the brain with age in older persons is confounded by common neuropathologies. Int J Biochem Cell Biol. 2015;67:58–64.
4. Bacos K, Gillberg L, Volkov P, Olsson AH, Hansen T, Pedesen O, et al. Blood-based biomarkers of age-associated epigenetic changes in human islets associate with insulin secretion and diabetes. Nat Commun. 2015;10:1–13.
5. The BLUEPRINT consortium. Quantitative comparison of DNA methylation assays for biomarker development and clinical applications. Nat Biotechnol. 2016;34:726–37.
6. Matsushita J, Okamura K, Nakabayashi K, Suzuki T, Horibe Y, Kawai T, et al. The DNA methylation profile of liver tumors in C3H mice and identification of differentially methylated regions involved in the regulation of tumorigenic genes. BMC Cancer. 2018;18:317–31.
7. Laird PW. Principles and challenges of genomewide DNA methylation analysis. Nat Rev Genet. 2010;11(3):191–203.
8. Jones PA, Baylin SB. The fundamental role of epigenetic events in cancer. Nat Rev Genet. 2002;3(6):415–28.
9. Jones PA, Baylin SB. The epigenomics of cancer. Cell. 2007;128:683–92.
10. Greenberg ES, Chong KK, Huynh KT, Tanaka R, Hoon DSB. Epigenetic Biomarkers in Skin Cancer. Cancer Let. 2014;342(2):170–7.
11. Morrill BH, Cox L, Ward A, Heywood S, Prather RS, Som SC. Targeted DNA Methylation Analysis by High Throughput Sequencing in Porcine Peri-attachment Embryos. J Reprod Dev. 2013;59(3):314–20.
12. Boerno ST, Grimm C, Lehrach H, Schweiger MR. Next-generation sequencing technologies for DNA methylation analyses in cancer genomics. Epigenomics. 2010;2(2):199–207.
13. Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Filippini JT, et al. Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. 2009;462(7271):315–22.
14. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou S, Diaz LA Jr, Kinzler KW. Cancer Genome Landscapes Science. 2013;339(6127):1546–58.
15. Ushijima T, Asada K. Aberrant DNA methylation in contrast with mutations. Cancer Sci. 2010;101(2):300–5.
16. Calcagno DQ, Gigek CO, Chen ES, Burbano RR, Smith M. de AC. DNA and histone methylation in gastric carcinogenesis. World J Gastroenterol. 2013;19(8):1182–92.
17. Tsai KW, Wu CW, Hu LY, Li SC, Liao YL, Lai CH, et al. Epigenetic regulation of miR-34b and miR-129 expression in gastric cancer. Int J Cancer. 2011;129(11):2600–10.

18. Ziogas D, Roukos D. Epigenetics in Gastric Cancer: Challenges for Clinical Implications. Ann Surg Oncol. 2009;16(7):2077–8.
19. Wolffe AP, Matzke MA. Epigenetics: regulation through repression. Science. 1999;286(5439):481–6.
20. Mirabello L, Frimer M, Harari A, McAndrew T, Smith B, Chen Z, et al. HPV16 methyl-haplotypes determined by a novel next-generation sequencing method are associated with cervical precancer. Int J Cancer. 2015;136(4):E146–53.
21. Rønneberg JA, Tost J, Solvang HK, Alnaes GIG, Johansen FE, Brendeford EM, et al. GSTP1 Promoter Haplotypes Affect DNA Methylation Levels and Promoter Activity in Breast Carcinomas. Cancer Res. 2008;68(14):5562–71.
22. Do C, Lang CF, Lin J, Darbary H, Krupska I, Gaba A, et al. Mechanisms and Disease Associations of Haplotype-Dependent Allele-Specific DNA Methylation. Am J Hum Genet. 2016;98(5):934–55.
23. Couldrey C, Lee RS. DNA methylation patterns in tissues from mid-gestation bovine foetuses produced by somatic cell nuclear transfer show subtle abnormalities in nuclear reprogramming. BMC Dev Bio. 2010;10:27.
24. Breton CV, Salam MT, Gilliland FD. Heritability and role for the environment in DNA methylation in AXL receptor tyrosine kinase. Epigenetics. 2011;6(7):895–8.
25. Downing C, Johnson TE, Larson C, Leakey TI, Siegfried RN, Rafferty TM, et al. Subtle decreases in DNA methylation and gene expression at the mouse Igf2 locus following prenatal alcohol exposure: effects of a methyl-supplemented diet. Alcohol. 2011;45(1):65–71.
26. Waterland RA, Jirtle RL. Transposable elements: targets for early nutritional effects on epigenetic gene regulation. Mol Cell Biol. 2003;23(15):5293–300.

Figures

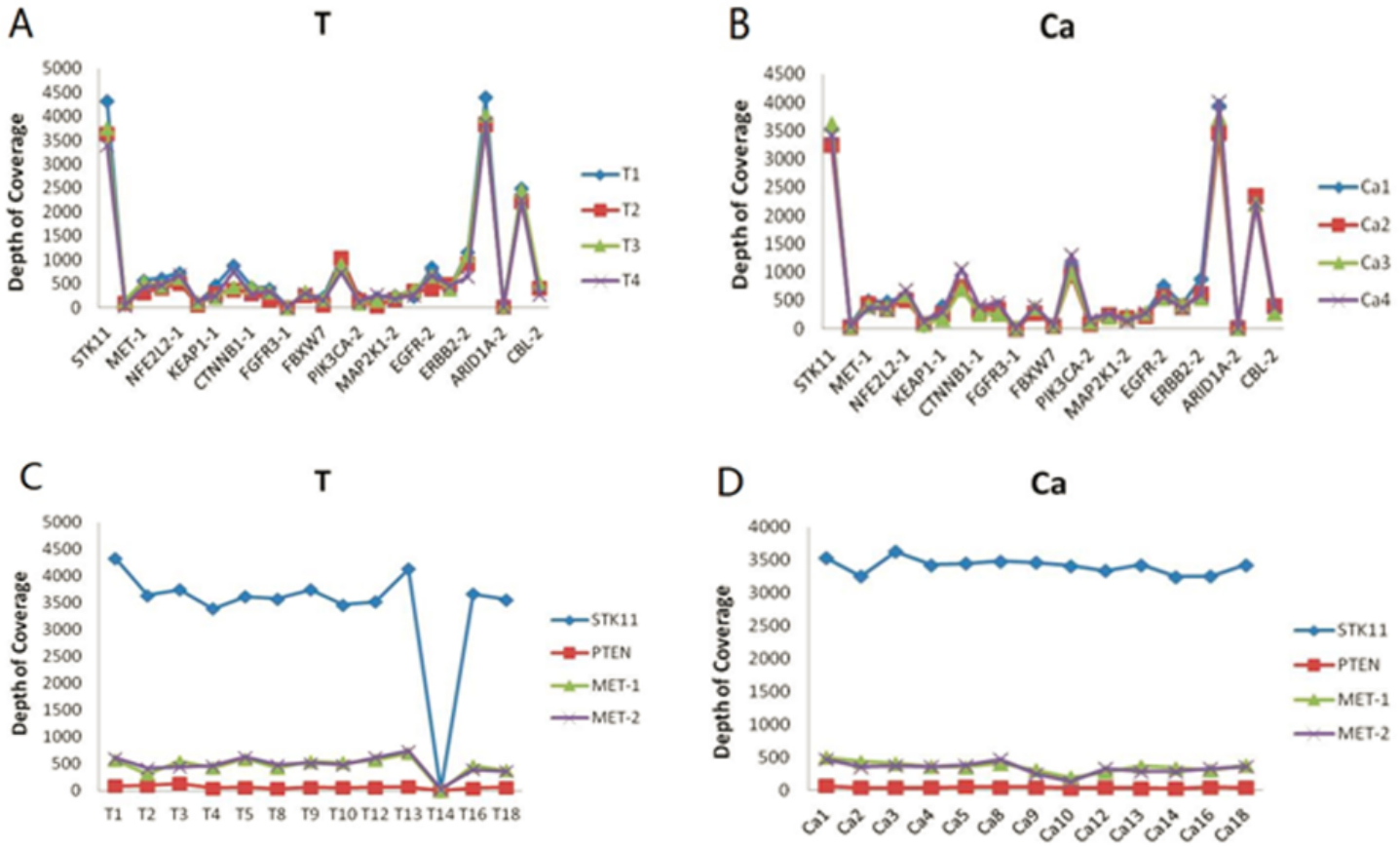


Figure 1

Depth of coverage obtained using high-throughput bisulfite sequencing (A and B) Representative traces plotting absolute depth of coverage (Y-axis) against genes (X-axis) along the entire length of the amplicon for 8 of the 26 sample PCR products sequenced and aligned. The samples correspond to 4 T samples and 4 Ca samples. The traces of all the samples are almost identical in shape, with only slight variations in depth of coverage at each location. (C, D) Representative traces plotting absolute depth of coverage (Y-axis) against samples (X-axis) along the entire length of the amplicon for four of the 25 PCR products sequenced and aligned.

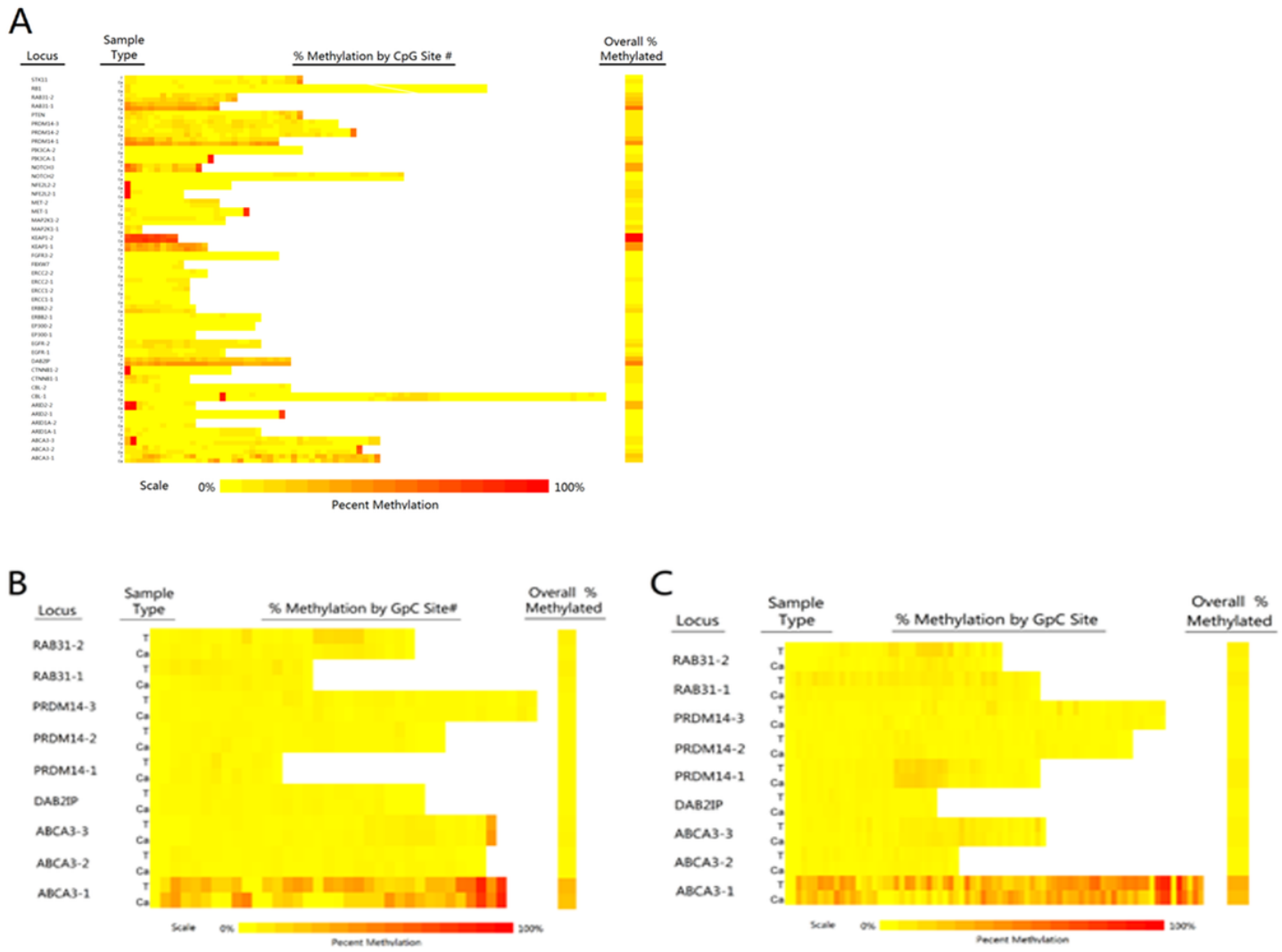


Figure 2

Analysis of difference in methylation ratio (A) Comparative analysis of DNA methylation patterns (CG) in hepatic carcinoma (Ca) and adjacent normal tissues (T). The percentages of methylated cytosines at each potential methylation site for each of 25 distinct PCR products are depicted. (B) Comparative analysis of DNA methylation patterns (CHG) in Ca and T. The percentages of methylated cytosines at each potential methylation site for each of 9 distinct PCR products are depicted. (C) Comparative analysis of DNA methylation patterns (CHH) in Ca and T. The percentages of methylated cytosines at each potential methylation site for each of 9 distinct PCR products are depicted. Deep red indicates high levels of methylation; yellow indicates low levels of methylation.

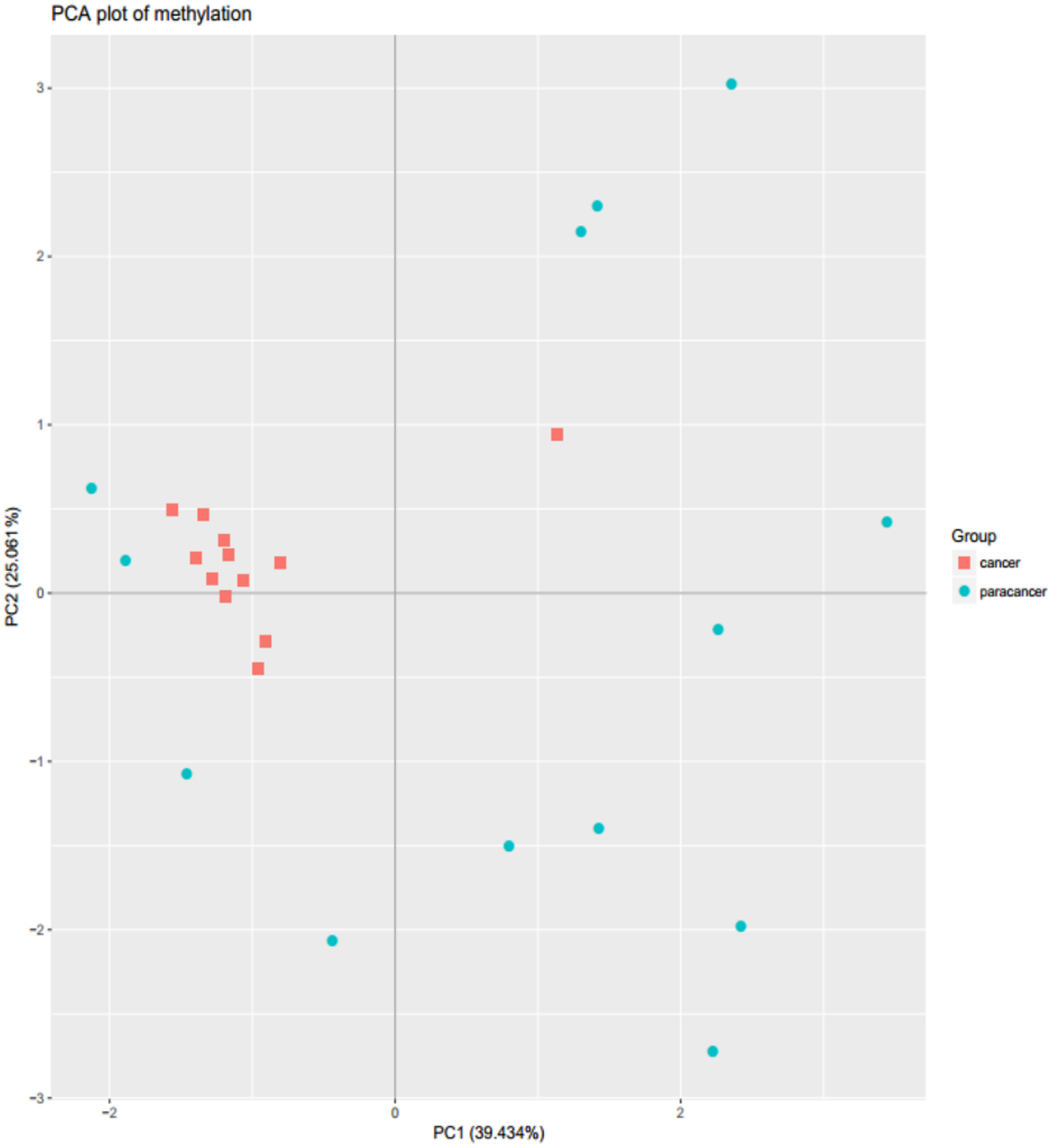


Figure 3

Principle component analysis Color discrimination of sample group on PCA; The X and Y axes respectively correspond to the indicators that best reflect the true species composition of samples.

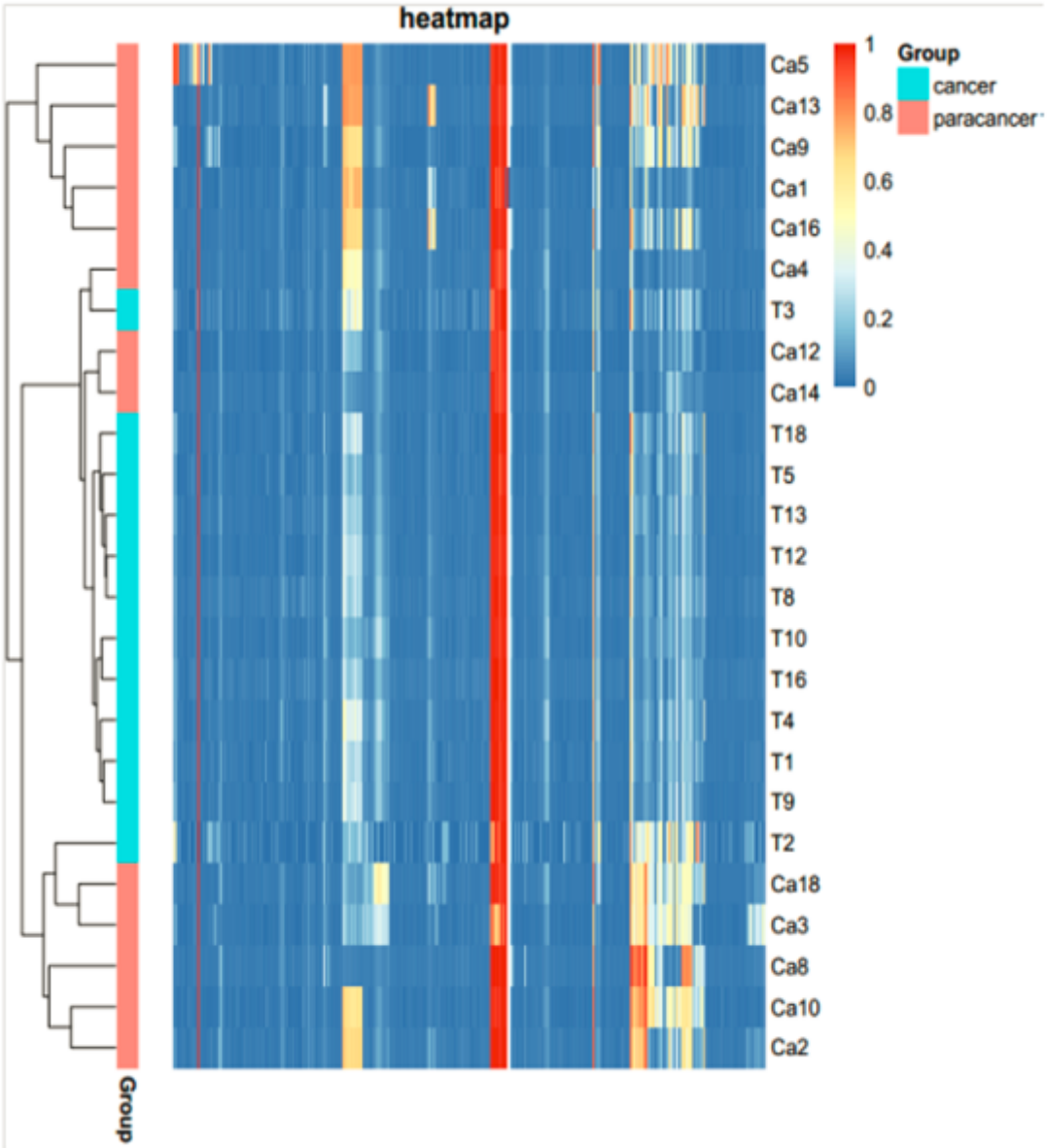


Figure 4

Cluster analysis. Each cell of the heatmap represents the relative methylation level of CpG sites in the corresponding row of samples and reflects the relative methylation level of the sites by color gradient. Blue and red represent low and high levels of methylation, respectively. The similarity of methylation level is shown by row arrangement order. Adjacent rows indicate higher overall similarity of sample methylation level. The tree on the left of the figure systematically describes the degree of similarity.