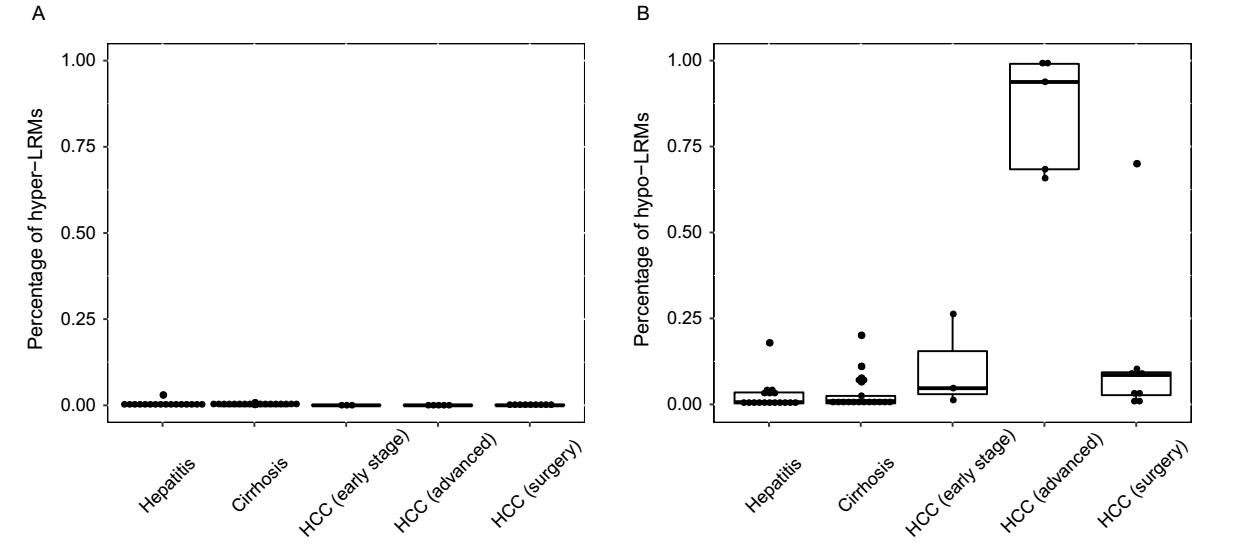
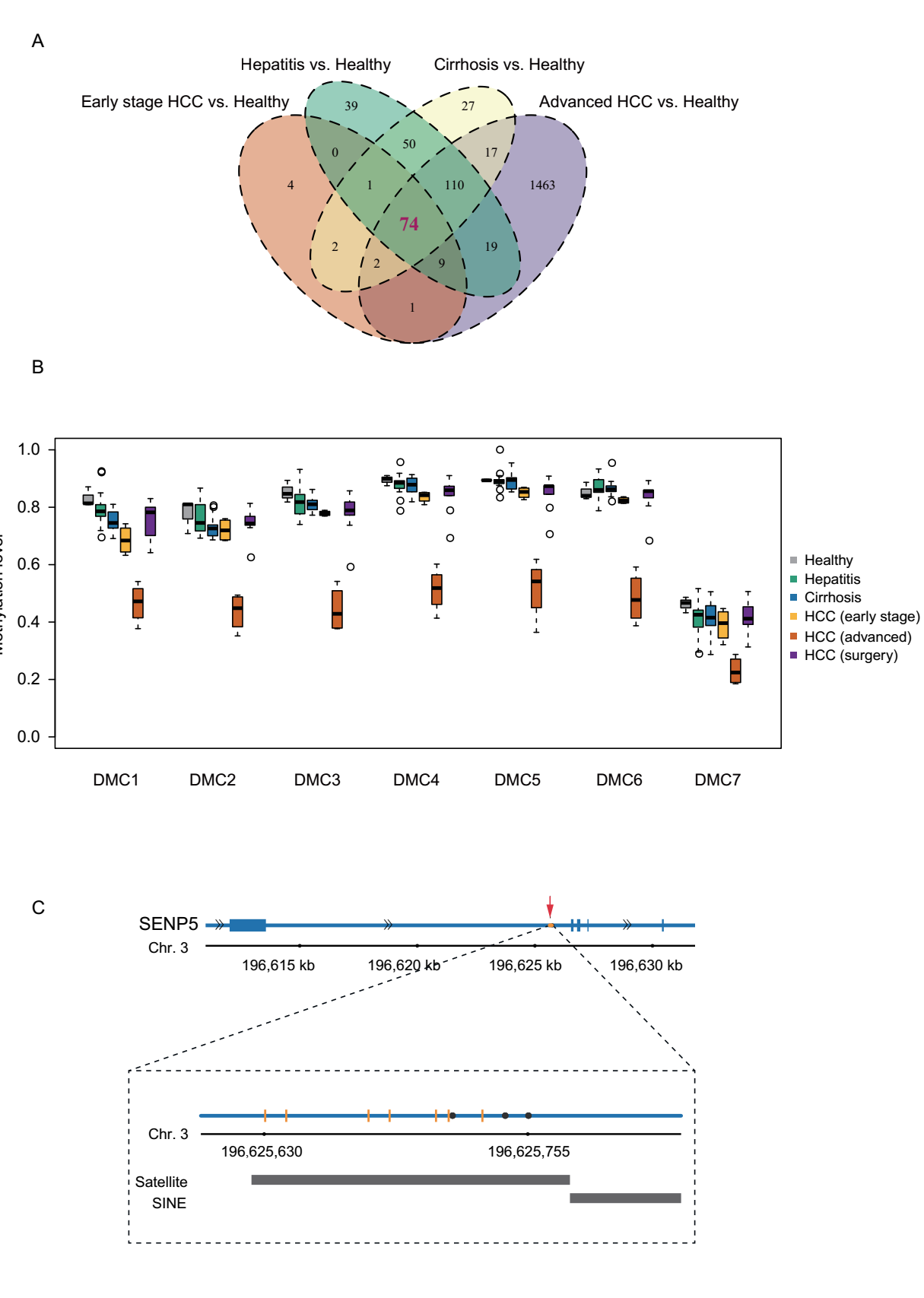


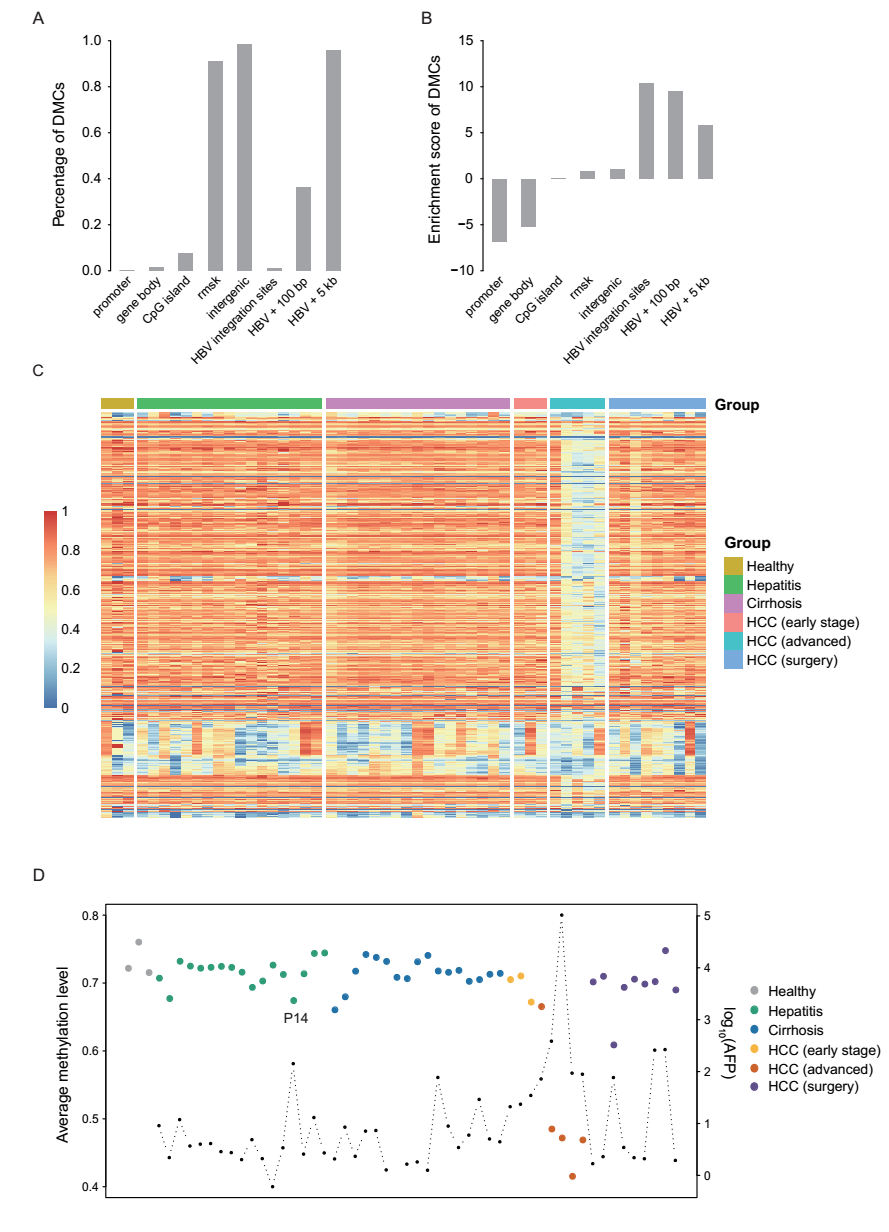
**Fig. 1. The efficiency of re-sampling sequencing depth for low pass WGBS.** Left of the figure showed the correlation coefficient between re-sampling low pass WGBS and total sequencing reads for 100 times from 1M to 10M. Right of the figure showed the coefficient of variation (CV) for 100 correlation coefficient between re-sampling low pass WGBS and total sequencing reads from 1M to 10M.



**Fig. 2 Whole genome-wide changed methylation of all the patients.** (A) The percentage of hyper-methylated long range regions (2-Mb) in chronic hepatitis, cirrhosis and HCC patients. (B) The percentage of hypo-methylated long range regions in chronic hepatitis, cirrhosis and HCC patients.



**Fig. 3. Differentially methylated CpGs (DMCs) identified in all the groups.** (A) Venn diagram showing the overlap of DMCs generated by 2 hypo-methylated chronic hepatitis patients, 1 hypo-methylated cirrhosis patient, 3 early stage HCC patients and 5 advanced HCC patients compared to healthy individuals. (B) Boxplot displays the methylation level of 7 DMCs of SENP5 in 3 healthy individuals, 17 hepatitis, 17 cirrhosis, 3 early stage HCC, 5 advanced HCC and 9 HCC patients after surgery. (D) The locus of 7 DMCs and 3 reported HBV integration sites in intron 2 of SENP5. The black dots represent the HBV integration sites and the orange vertical lines represent the 7 DMCs. The black bar labels in the bottom of the figure represent the locus of repeat marker in this region.



**Fig. 4. DMCs and CpGs are related to HBV integration sites.** (A) The percentage of DMCs located in different genomic elements and regions related to HBV integration sites. (B) The enrichment scores of DMCs in different genomic elements. (C) The heatmap display the methylation level of the CpGs located within 100 bp of the HBV integration sites in all the samples. (D) The average methylation level of all the CpGs located within 100 bp of the HBV integration sites in all the samples. The red arrows showed the examples of P14 patient.

**Table 1. Genes with DMCs between liver disease patients and healthy individuals.**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chr** | **Start** | **End** | **CpG num** | **Position** | **Gene** |
| **Hepatitis vs. healthy** | | | | | |
| chr1 | 156186377 | 156186549 | 18 | gene body | PMF1 |
| chr1 | 156186377 | 156186549 | 18 | gene body | PMF1-BGLAP |
| chr1 | 91853073 | 91853096 | 3 | gene body | HFM1 |
| chr11 | 85195089 | 85195090 | 1 | gene body | DLG2 |
| **Cirrhosis vs. healthy** | | | | | |
| chr1 | 156186492 | 156186493 | 1 | gene body | PMF1 |
| chr1 | 156186492 | 156186493 | 1 | gene body | PMF1-BGLAP |
| chr8 | 70602451 | 70602487 | 2 | gene body | SLCO5A1 |
| **Early stage HCC vs. healthy** | | | | | |
| chr1 | 156186410 | 156186530 | 5 | gene body | PMF1 |
| chr1 | 156186410 | 156186530 | 5 | gene body | PMF1-BGLAP |
| **Advanced HCC vs. healthy** | | | | | |
| chr1 | 91852857 | 91852974 | 2 | gene body | HFM1 |
| chr1 | 156186392 | 156186540 | 10 | gene body | PMF1 |
| chr1 | 156186392 | 156186540 | 12 | gene body | PMF1-BGLAP |
| chr3 | 196625630 | 196625734 | 7 | gene body | SENP5 |
| chr8 | 70602486 | 70602487 | 1 | gene body | SLCO5A1 |
| chr8 | 86572360 | 86572383 | 2 | gene body | REXO1L1P |
| chr11 | 85195100 | 85195101 | 1 | gene body | DLG2 |

**Table 2. The information of hepatitis and cirrhosis patients with lower methylation levels of CpGs located in the 100 bp of HBV integration sites**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Patient** | **Disease** | **Percentage of hypo LRMs** | **MethylHBV** | **AFP (ng/ml)** |
| P2 | chronic hepatitis | 17.87% | 67.69% | 2.2 |
| P14 | cirrhosis | 3.47% | 67.39% | 141.9 |
| P18 | nash-related cirrhosis | 20.04% | 66.04% | 2.07 |
| P19 | alcoholic cirrhosis | 11.00% | 67.96% | 8.46 |