Am J Cancer Res 2022;12(7):3128-3140
www.ajcr.us /ISSN:2156-6976/ajcr0143253

Original Article

SEPTIN9-SDC2-VIM methylation signature as a biomarker for the early diagnosis of colorectal cancer

Rong-Qiang Yuan[3*], Hui Zhao[2*], Yan Wang[1*], Kai Song[4], Jia Yang[3], Wei He[1], Da-Zhuang Miao[1], Qi Wang[1], Yun-He Jia[1]

[1]Harbin Medical University Cancer Hospital Colorectal Cancer Center, Harbin 150086, Heilongjiang, China; [2]The First Affiliated Hospital, Harbin Medical University, Harbin 150001, Heilongjiang, China; [3]Department of Systems Biology, College of Bioinformatics Science and Technology, Harbin Medical University, Harbin 150086, Heilongjiang, China; [4]Zhuhai Interventional Medical Center, Zhuhai Precision Medical Center, Zhuhai People's Hospital, Zhuhai Hospital Affiliated with Jinan University, Jinan University, Zhuhai 519000, Guangdong, China. [*]Equal contributors.

Received April 1, 2022; Accepted June 13, 2022; Epub July 15, 2022; Published July 30, 2022

Abstract: The accurate detection of colorectal cancer (CRC) at its initial stage can reduce mortality. However, the broad application of endoscopy has been limited due to the invasive procedure and patient noncompliance. Liquid biopsy with subsequent mapping of methylation in specific cell-free DNA (cfDNA) may represent an alternative approach for early diagnosis. In this study, we have developed a minimal-invasive blood-based test for detection of precancerous lesions and early-stage CRC. Using TCGA M450K methylation data, we identified candidate methylation sites with the highest Fold Change (FC) for three genes (SEPTIN9, SDC2 and VIM), which were selected from previous studies. Based on logistic regression models, we developed a 3-gene methylation signature for CRC diagnosis with high accuracy (Sensitivity =0.959, Specificity =1, AUC =0.997). Using independent public databases and data from blood samples, this model has demonstrated superior performance. The AUC was 0.919-1 and 0.905-0.916 in public tissue database for CRC and blood sample data, respectively. Thus, our proposed 3-gene methylation signature has a more reliable performance than other methods. Furthermore, signal enhancement effect of 3-gene methylation signature can improve the accuracy of early diagnosis for CRC, which demonstrates the potential for clinical application.

Keywords: DNA methylation, epigenetic biomarker, CRC, early diagnosis, liquid biopsy

Introduction

Colorectal cancer (CRC) ranks as the third most common malignancy and the second leading cause of cancer-related mortality worldwide, with an estimate of 1.93 million new cases and 0.94 million deaths in 2020 [1]. The poor prognosis and high mortality rate of CRC can be attributed to the delayed diagnosis and the insufficient early diagnostic precision. As a result, patients have developed distant metastasis at the time of diagnosis. Therefore, it is crucial to develop high-sensitivity methods for the early diagnosis of CRC.

At present, colonoscopy is the gold standard of CRC diagnosis. However, due to the inconvenience for patients, screening rate using colonoscopy is low. Moreover, the procedure-related risks and patient noncompliance limit the broad application of invasive colonoscopy [2]. Traditionally, non-invasive diagnostic methods include fecal immunochemical testing, fecal occult blood testing and multi-target stool DNA test, while flexible sigmoidoscopy and CT colonoscopy are emerging recently as new and improved non-invasive diagnostic approaches [3]. However, the performance of these methods is compromised by the high false positive rate. Liquid biopsy is a minimally invasive or practically noninvasive method for analyzing a series of tumor-related markers in blood or other body fluids [2]. This method is less harmful and more convenient to patients compared with colonoscopy and is less affected by tumor size. Currently, it is a novel detection method with promising application prospect.

In human cancers, gain- or loss-of-function is mainly caused by genetic and epigenetic abnormalities [4, 5]. During the tumorigenesis of CRC, epigenetic changes, including DNA methylation, histone modification and ncRNA, occur earlier and more frequently than genomic alterations [6-8]. If compared with the phenotypic changes caused by random mutations in the protein coding sequences of genes, epigenetic modifications are more traceable and universal. Therefore, epigenetic modifications may serve as more suitable biomarkers for the early diagnosis of cancer. Indeed, local DNA hypermethylation has been extensively studied in gastrointestinal cancers, and it has been demonstrated to promote cancer initiation via silencing tumor suppressor genes [9-11]. Moreover, the methylation status of specific genes has been suggested to stratify patients with a high risk of developing cancer [12].

At present, early diagnosis methods based on the liquid biopsy technology to detect specific gene methylation have been applied in clinical practice. For example, the ColoSure™ test is focused on VIM gene methylation in feces, with sensitivity of 53% to 83% for CRC and 50% to 84% for adenoma, respectively [13]. The sensitivity of the Epi proColon detection based on SEPTIN9 methylation is 59.6% and 85.7% in patients with stage I and II colorectal cancer, respectively [14, 15]. However, the sensitivity of SEPTIN9 methylation in distinguishing adenoma patients from healthy subjects was low, with only 7.9%-38.7% [16-19]. Recent studies have indicated that using combination of multiple biomarkers may be a more effective strategy to improve the sensitivity for the screening and diagnosis of early-stage cancer. For example, SEPTIN9 combined with SDC2 methylation can increase the sensitivity to 86.5%-88.9% and the specificity to 92.1%-92.8% in the early detection of CRC [20, 21]. However, precancerous lesions, such as adenoma, are excluded from these studies. Earlier identification of such patients is significant as it can effectively reduce the risk of developing CRC and achieve the goal of cancer prevention.

Therefore, in this study, we have constructed an early diagnosis model to distinguish subjects with precancerous lesions and CRC from normal population based on the methylation profiling of specific genes in peripheral blood samples. As shown in **Figure 1**, we selected 3 genes (SEPTIN9, SDC2 and VIM), which were highly related to CRC diagnosis [22], as a set of candidates from previous research. Using TCGA-CRC methylation datasets and the logistic regression (LR) method, we have developed a 3-gene methylation biomarker and verified its performance in independent validation cohorts on both CRC and adenoma patients. The performance of this 3-gene methylation biomarker was proved to be reliable in both training and validation datasets. Furthermore, we used data from peripheral blood samples to build another LR model with different thresholds and achieved a perfect performance. Compared with other early diagnosis biomarkers, this 3-gene methylation biomarker has a comparable performance and thus provides a potential tool for the early diagnosis of CRC.

Material and methods

Data sources

In this study, public tissue data for CRC, adenoma, and colorectal laterally spreading tumors (LSTs) as well as normal samples were obtained from the TCGA [23] (TCGA-COAD and TCGA-READ project) and GEO databases [24]. Additionally, we obtained a whole-blood DNA methylation profile of healthy donors from GSE40279. All the data used were measured by Illumina Human Methylation 450 BeadChip (450K).

We further enrolled 45 CRC and 31 adenoma patients as well as 84 healthy subjects who were biopsied with blood samples at the Harbin Medical University Cancer Hospital from September 1, 2019, to December 31, 2019. The diagnosis of each patient was based on the pathology results from resection specimens by the surgical biopsies. The methylation level of SEPTIN9, SDC2 and VIM for each participant was measured by Luminex 200. This study was reviewed and approved by the ethics committee of Harbin Medical University Cancer Hospital. Each participant provided written informed consent.

Preparation of DNA and bisulfite conversion

For cell-free DNA (cfDNA) extraction, 10 mL peripheral blood samples were collected in K2 EDTA tubes (LAKEbio). Samples were pro-

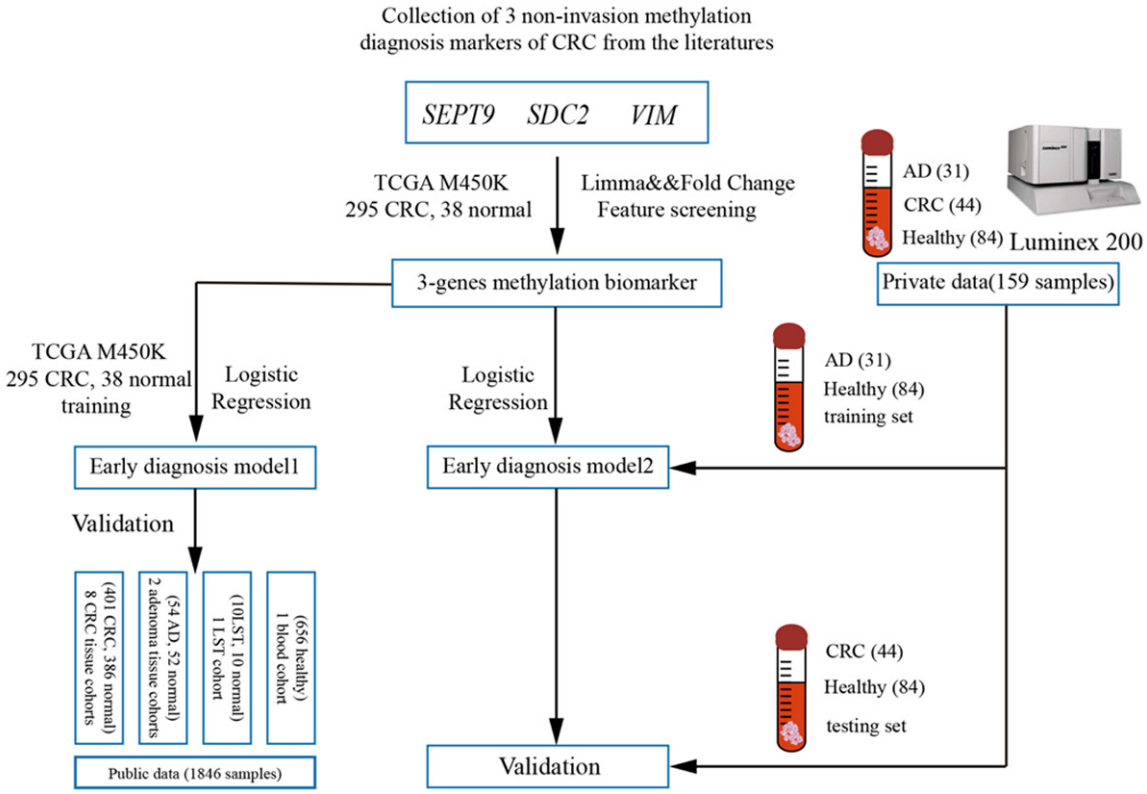

Figure 1. The flowchart for developing early diagnosis model of CRC with two branches.

cessed within 2 h of collection. To separate plasma from cellular components, blood samples were centrifuged at 1350 g for 12 mins at 4°C. Subsequently, the plasma was stored in 1.2 ml aliquots at -40°C until use. To extract DNA from plasma, the frozen plasma aliquots were thawed at room temperature for 1 h and centrifuged at 17900 g for 10 mins. A total of 2 mL of plasma per sample was used for cfDNA extraction. Plasma DNA extraction was proceeded using a plasma DNA extraction kit (Macherey Nagel) following the manufacturer's instruction. Extracted plasma DNA was recovered in a final elution volume of 50 μl. CRC patient samples and healthy control samples were identically processed.

Bisulfite treatment was performed using EZ DNA Methylation-Lightning™ Kit (D5031) with some modifications: conditions for heat treatment were 98°C for 8 mins, 54°C for 50 mins, and subsequently cooled to 4°C in a PCR machine; the elution time was doubled with M-Wash Buffer, and the step of "desufonation were formulated by a manipulator" was omitted. In addition, DNA was resuspended in ddH_2O rather than in TE for optimizing the following multiplex PCR. Human Methylated (Takara), Non-Methylated DNA (Zymo Research) standard, and water were included in each bisulfite conversion batch as positive, negative, and blank controls, respectively. After recovered from bisulfite treatment, the DNA samples were analyzed by multiplex PCR immediately or stored at -20°C until use.

PCR primer and probe design

DNA methylation was assessed by using the methylation-specific PCR (MSP) assay. A pair of primers for bisulfite-treated methylated DNA was designed using Primer 5.0 and Oligonucleotide Properties Calculator [25] to specifically amplify methylated DNA in the first round of PCR. Since the sense and antisense strands are no longer complementary after bisulfite conversion, the PCR primers may be designed for either of the two strands. To increase the specificity, probes longer than primers were designed for the second-round PCR. The probes contained three regions: a universal xTAG sequence (provided by Luminex Corp) at the 5'

end, a gene-specific sequence in the middle, and a biotinylated sequence at the 3' end. A specific region of about 150 bp long was chosen for probe hybridization. Primers and probes were synthesized by Genscript biotechnology. Information about the primers and probes used in this study was shown in Table S8.

Hybridization and Luminex analysis

The second round PCR product was mixed with magnetic beads in 1× TM buffer (0.1 M Tris PH 8.0, 0.2 M NaCl and 0.08% Triton X-100) and transferred to the thermocycler with the following protocol: 96°C for 90 s, followed by 37°C for 15 mins, and cooled at 25°C. To ensure the hybridization of the complementary anti-xTAG sequence covalently bound to the surface of the magnetic beads and amplified DNA in PCR products, each sample was processed by a fixed ratio of 1250 beads to guarantee that 100 values were read by Luminex. After adding the corresponding beads, 2× TM, and diluted PCR products to the 96-well plate of the thermal cycler, we used the program of incubating the samples for 90 s at 95°C, followed by 15 mins at 37°C. After hybridization was completed, 25 μL of SAPE buffer (SAPE and 1× TM) was added for 15 mins at 37°C. Before analysis in a Luminex instrument, the reaction well was washed 5 times with 1× TM. Assays were conducted by two people; one person operated the machine while the other person made sure the accuracy of the result.

The plasma DNA was converted by bisulfite for multiplex PCR, and the target fragments were specifically captured by magnetic beads and fluorescently labeled using SAPE. Finally, the reaction products were measured by Luminex 200 analyzer. The SAPE fluorescent compound that had hybridized to the magnetic beads was measured using xPONENT® Software. Median fluorescence intensity (MFI) of SAPE was used to quantify the amount of DNA bound to the magnetic beads.

Logistic regression model

We constructed early diagnosis models of colorectal cancer based on logistic regression in both the public tissue data and the data we obtained from our blood samples. The formula was as follows:

$$P_1 = 1/(1 + e^{-\alpha}), \alpha = \sum a_i * Gene_i + b \quad (1)$$

P represents the probability of the patient being diagnosed with colorectal cancer, while α represents the sum of the product of gene methylation level and its weight.

Function analysis and statistical analysis

To determine the function of the 3-gene methylation biomarker, a correlation landscape was constructed. For each gene used to construct the biomarker, we utilized the Pearson correlation analysis to calculate the transcriptomics correlations between this particular gene and all the other genes. Then, we utilized gene set enrichment analysis to identify the associated pathways [26]. Briefly, all genes were ranked based on their correlation in expression with a specific marker gene. The ranked gene list was matched to each pathway to explore whether the genes were enriched on the top or at bottom of the list. The enrichment score (ES) and adjusted *P*-value (p. adj) were calculated.

Fisher exact test was used to reveal the influence of clinical factors on the experiment. All the differences among groups in this experiment were calculated by limma package. All statistical analyses were performed using R software [27] version 4.0.3.

Results

Basic information on datasets

We collected 11 tissue cohorts, including eight CRC, two adenoma, and one LSTs cohort, as well as one blood cohort from the public database to validate the predictive performance of our early diagnosis model (**Table 1**). Furthermore, we enrolled 45 CRC and 31 adenoma patients as well as 84 healthy subjects who were biopsied with blood samples at the Harbin Medical University Cancer Hospital from September 1, 2019, to December 31, 2019 (**Tables 1**; S1).

A combinational early diagnosis model could distinguish between disease and normal samples

Among the 174 samples tested by using our 3-gene methylation signature (SEPTIN9, SDC2, VIM) for the diagnosis of CRC, 104 probes were identified (Table S2) with significantly different methylation levels between CRC and normal samples (R, Limma package, adjusted P<0.05).

Table 1. The basic information of datasets in this study

	Data_ID	Disease status	The Number of Normal samples	The Number of Disease samples
Tissue	TCGA	CRC	38	257
	GSE75546	CRC	6	6
	GSE77954ca	CRC	11	13
	GSE77965	CRC	6	6
	GSE42752	CRC	41	22
	GSE48684ca	CRC	41	64
	GSE68060	CRC	36	82
	GSE77718	CRC	96	96
	GSE101764	CRC	149	112
	GSE106556ad	LST	10	10
	GSE48684ad	Adenoma (AD)	41	42
	GSE77954ad	Adenoma (AD)	11	12
Blood	GSE40279	Healthy	656	-
	Self-test	Adenoma (AD)	-	31
		CRC	-	44
		Healthy	84	-

Table 2. Three probes with highest FC in each candidate gene

Probes_ID	UCSC_RefGene Name	T_statistic	*P*-value	Adj.P.Val	Fold_ Change
cg15044248	SEPTIN9	12.688	2.39E-30	2.22E-29	25.431
cg07146119	SDC2	13.637	6.21E-34	7.57E-33	18.771
cg27313572	VIM	15.582	1.78E-41	4.97E-40	13.101

We hypothesized that probes exhibiting bigger difference in the expression level between tumor and normal tissue could distinguish tumor from healthy samples more accurately. Hence, a probe with the highest fold change (FC) value was selected for each gene (**Table 2**). Using methylation profiles of three probes in TCGA data, i.e., cg15044248, cg07146119 and cg27313572, an early diagnosis model for CRC was developed by LR, equation (2). The value of sensitivity, specificity, and AUC in training cohort was 0.949, 1 and 0.997, respectively (**Table 3**).

$$\alpha_1 = 59.187884 * cg15044248 + 283.127031 * cg07146119 + 23.639930 * cg27313572 - 9.752898 \quad (2)$$

With the same threshold (cutoff =0.999999), its overall sensitivity for differentiating CRC, LST and adenoma from normal samples in the public tissue data was 0.86-1, 1 and 0.738-1, respectively (**Table 3**; **Figure 2A**). In addition, this model successfully predicted all 656 blood samples from healthy subjects to be normal. The overall AUC of this LR model was 0.924-1. Finally, this 3-gene methylation signature LR model was compared with each single probe and with the previous 2-gene marker LR model (based on SEPTIN9 and SDC2) [20, 21, 35] in public database. Interestingly, the predictive power of our 3-gene marker LR model and the 2-gene marker model was comparable (Table S3). In addition, our 3-gene methylation biomarker exhibited a stable performance on an independent public tissue data. Thus, the combination of three genes might have a tremendous potential for distinguishing cancer or adenoma patients from healthy subjects using peripheral blood samples.

For the data we collected from our blood samples, we developed another LR model with sensitivity of 0.839 and 0.818 for adenoma and CRC, respectively (cutoff =0.2, **Table 3**; **Figure 2B**), equation (3). More significantly, for those patients whose pathological examination was normal but were predicted to be cancerous by this new model, CRC was diagnosed later during follow-up. Furthermore, the 3-gene marker LR model had a better performance than a single probe and the 2-gene marker LR model in our own blood sample data (**Figure 2B**, **2C**) as well as a more stable performance than a single probe in public data (**Figure 2A**, **2D**; **Table 3**).

$$\alpha_2 = 0.006143975 * B22\_SEPT9 + 0.001456855 * B35\_SDC2 + 0.017228400 * B37\_VIM - 4.981\,966138 \quad (3)$$

Rationale for combinational early diagnosis model

To explore the molecular mechanisms underlying the combinational early diagnosis model, the methylation levels of SEPTIN9, SDC2 and

Table 3. Validation of the early diagnosis model in public data

Data_ID			T	N	Cut_off	Sensitivity	Specificity	AUC
TCGA_M450k	Training set	T	280	15	0.999999	0.949	1	0.997
		N	0	38				
GSE101764	Testing set	T	104	8	0.999999	0.929	0.960	0.97
		N	6	143				
GSE42752		T	21	1	0.999999	0.954	0.975	0.999
		N	1	40				
GSE48684		T	55	9	0.999999	0.86	1	0.979
		N	0	41				
GSE68060		T	74	7	0.999999	0.914	0.889	0.924
		N	4	32				
GSE77718		T	95	1	0.999999	0.989	0.885	0.967
		N	11	85				
GSE77954		T	12	1	0.999999	0.923	0.909	0.937
		N	1	10				
GSE75546		T	6	0	0.999999	1	0.833	1
		N	1	5				
GSE77965		T	6	0	0.999999	1	0.833	1
		N	1	5				
GSE40279*		T	-	-	0.999999	-	1	-
		N	0	656				
Adenoma								
	Testing set		AD	N				
GSE77954AD		AD	12	0	0.999999	1	0.909	0.951
		N	1	10				
GSE106556AD		AD	10	0	0.999999	1	0.8	0.990
		N	2	8				
GSE48684AD		AD	31	11	0.999999	0.738	1	0.975
		N	0	41				
Self-testing								
Adenoma	Training set	AD	26	5	0.2	0.839	0.869	0.916
		N	11	73				
CRC		T	36	8	0.2	0.818	0.869	0.905
		N	11	73				

VIM were analyzed in TCGA database. We found 89.83% tumor samples exhibited hypermethylation in at least one of the three genes, whereas 27.46% in all three genes (**Figure 3A**). Similar results were obtained in the peripheral blood samples, where 88% of tumor samples exhibited hypermethylation in at least one of the three genes whereas 34.67% in all three genes (**Figure 3B**). In both the tissue and the blood samples, these three genes were hypermethylated in patients but hypomethylated in normal samples. Therefore, it is reasonable to identify significantly hypermethylated gene loci from tissue data and verify them in blood samples.

When we compared these three genes, only SDC2 and VIM had a high positive correlation (r=0.901, CI=0.877-0.921, **Table 4**), suggesting a signal enhancement between SDC2 and VIM. We speculated that when the methylation level of one gene was low (at a normal level), signal enhancement by the other gene would correct the final value, and thus making more stable diagnosis. This could explain why multigene biomarker showed more stability than a single gene.

Furthermore, we created gene sets that were sorted according to the correlation among gene

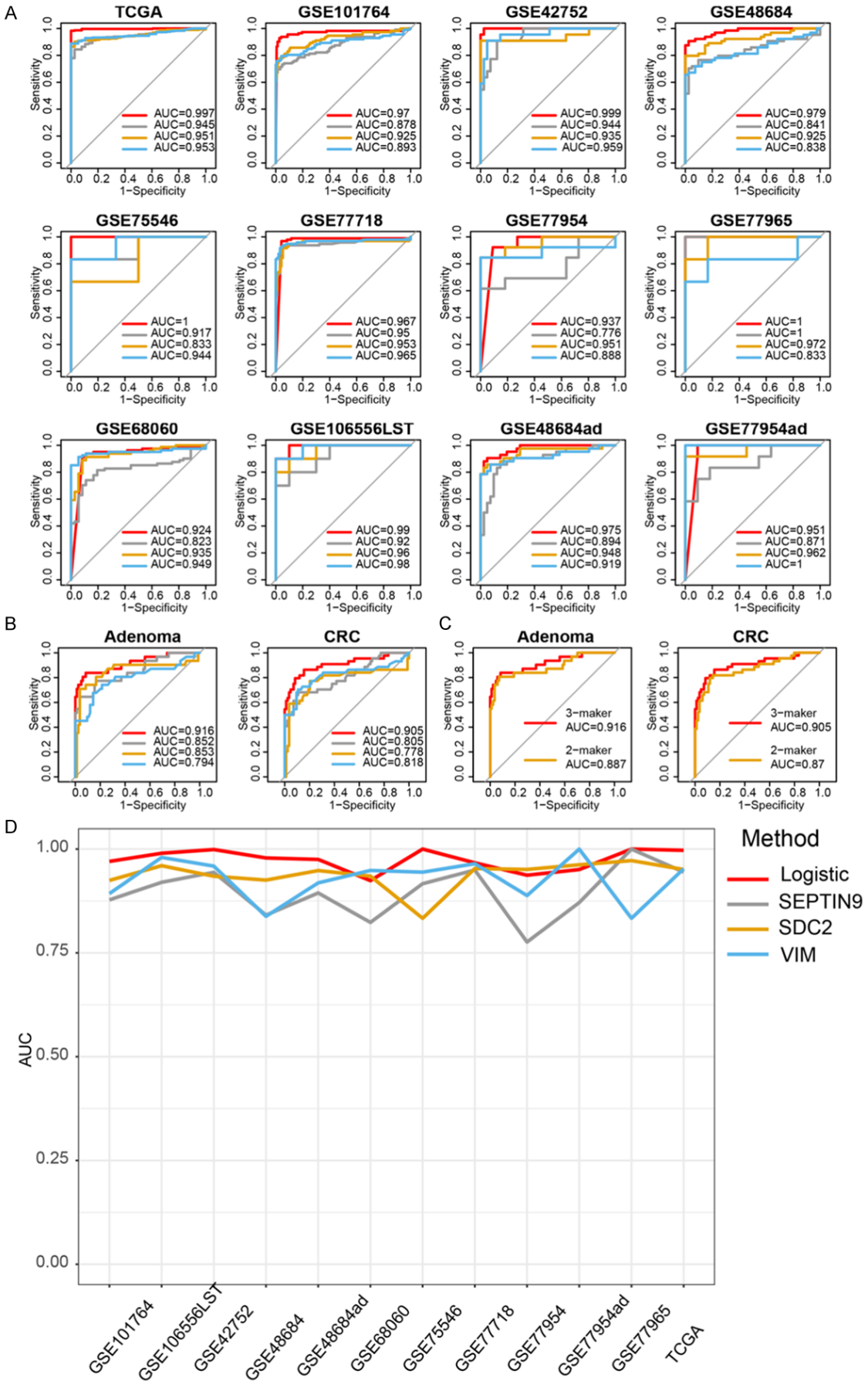
A
TCGA
AUC=0.997
AUC=0.945
AUC=0.951
AUC=0.953
GSE101764
AUC=0.97
AUC=0.878
AUC=0.925
AUC=0.893
GSE42752
AUC=0.999
AUC=0.944
AUC=0.935
AUC=0.959
GSE48684
AUC=0.979
AUC=0.841
AUC=0.925
AUC=0.838
GSE75546
AUC=1
AUC=0.917
AUC=0.833
AUC=0.944
GSE77718
AUC=0.967
AUC=0.95
AUC=0.953
AUC=0.965
GSE77954
AUC=0.937
AUC=0.776
AUC=0.951
AUC=0.888
GSE77965
AUC=1
AUC=1
AUC=0.972
AUC=0.833
GSE68060
AUC=0.924
AUC=0.823
AUC=0.935
AUC=0.949
GSE106556LST
AUC=0.99
AUC=0.92
AUC=0.96
AUC=0.98
GSE48684ad
AUC=0.975
AUC=0.894
AUC=0.948
AUC=0.919
GSE77954ad
AUC=0.951
AUC=0.871
AUC=0.962
AUC=1
Sensitivity
1−Specificity
B
Adenoma
AUC=0.916
AUC=0.852
AUC=0.853
AUC=0.794
CRC
AUC=0.905
AUC=0.805
AUC=0.778
AUC=0.818
C
Adenoma
3−maker
AUC=0.916
2−maker
AUC=0.887
CRC
3−maker
AUC=0.905
2−maker
AUC=0.87
D
AUC
Method
Logistic
SEPTIN9
SDC2
VIM
GSE101764
GSE106556LST
GSE42752
GSE48684
GSE48684ad
GSE68060
GSE75546
GSE77718
GSE77954
GSE77954ad
GSE77965
TCGA

Figure 2. The performance of the early diagnosis LR model in public and private data. A. Validation of this new model in public data with 0.9999 threshold and compared with a single gene. B. Validation of this model in private data with 0.2 threshold and compared with a single gene. C. The AUCs of 3-gene marker and 2-gene marker in private data. D. The AUC line chart of LR model and each probe in public database. A, B and D. Used the same legend in right bottom: method.

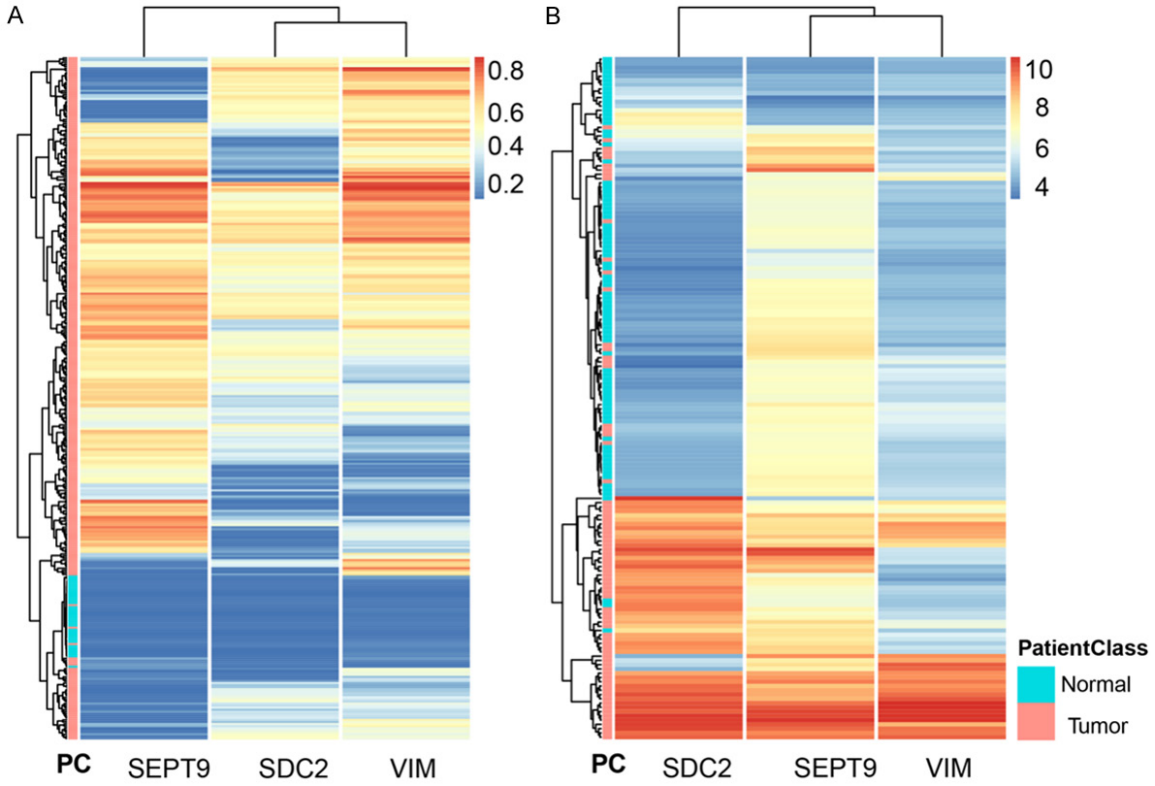

Figure 3. The heatmap of the gene methylation levels in public and private data. Methylation levels of SEPTIN9, SDC2 and VIM in CRC samples from TCGA (A) and in the self-test blood data set (B).

Table 4. The correlation among SEPTIN9, SDC2 and VIM

Gene 1_ID	Gene 2_ID	CI_lower	R	CI_upper	*P*-value	P.adj_value
SDC2	VIM	0.877	0.901	0.921	8.57E-107	5.14E-106
SEPTIN9	VIM	-0.096	0.019	0.133	0.750	0.968
SEPTIN9	SDC2	-0.155	-0.041	0.074	0.484	0.968

panels. Gene Set Enrichment Analysis (GSEA) identified that SEPTIN9, SDC2 and VIM were enriched in 19 cancer hallmark pathways (Tables S4, S5, S6, S7), such as epithelial-mesenchymal transition (EMT) and myogenesis (**Figure 4**). The enrichment score of EMT was the highest to these three genes (**Figure 4**), consistent with the important role of EMT in tumor metastasis [28]. Pathway-based evaluation in early onset CRC suggested the enrichment in focal adhesion and immunosuppression pathways along with EMT [29], which was consistent with our results. Another published study on 22 proteins for a myofibroblastic signature has observed their high expression in mesenchymal-like arrays, rather than in epithelial-like arrays, based on the meta-analysis heat map [30]. This result suggested that both EMT and myogenesis were correlated with tumor growth and metastasis, especially, as SDC2 was closely associated with VIM (23/24, Tables S5, S7).

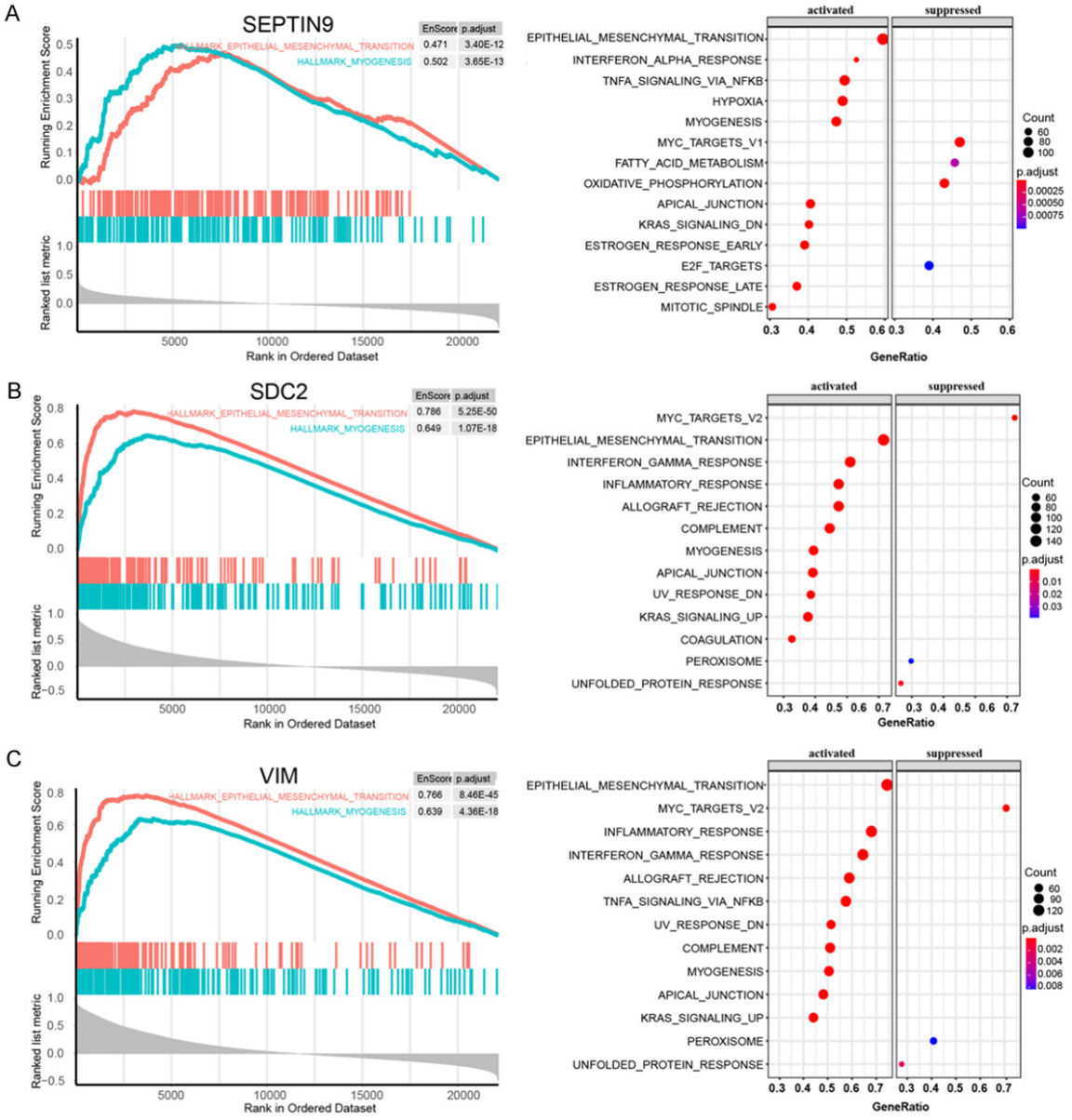

Figure 4. GSEA analysis of SEPTIN9, SDC2 and VIM. Left plots: EMT and myogenesis pathways. The right dot plots described the top ten activated and suppressed pathways related to SEPTIN9 (A), SDC2 (B), and VIM (C).

We also evaluated the effects of different clinical features in TCGA M450K data. We found that such factors as gender, age, tumor stage and location had no influence on the 3-gene methylation pattern (Fisher exact testing, gender P=0.604; age P=0.575; stage P=0.168; location P=0.184). Using Limma package, patients were grouped by sex, age, and tumor stage to assess differences in 3-gene methylation levels among groups. These three genes showed similar correlation to the clinical features (P.adj >0.05) (**Table 5**). Also, the statistical significance of SEPTIN9 methylation level between old and young people was borderline. This result is consistent with the observation that gene methylation increases with age [31, 32]. However, this trend did not affect the performance of 3-gene methylation analysis due to the influence of SDC2 and VIM. Hence, 3-gene methylation biomarker was an independent factor from clinical features; combinational strategy could improve the robustness of early diagnosis under complex clinical conditions. We also set up a nomogram based on the

Table 5. Clinical factors and 3-gene-methylation signature

	Right	Wrong	LR_model Fisher_test_*P*-value	SEPTIN9 P.adj_value	SDC2 P.adj_value	VIM P.adj_value
gender						
male	172	7	0.604	0.774	0.774	0.774
female	144	8				
age						
≥60	215	9	0.575	0.081	0.238	0.655
<60	101	6				
stage						
I-II	150	8	1	0.767	0.767	0.767
III-IV	118	7				
I	39	5	0.168	-	-	-
II	111	3				
III	79	5				
IV	39	2				
Location						
colon	236	12	0.184	-	-	-
cecum	75	2				
rectosigmoid junction	5	1				

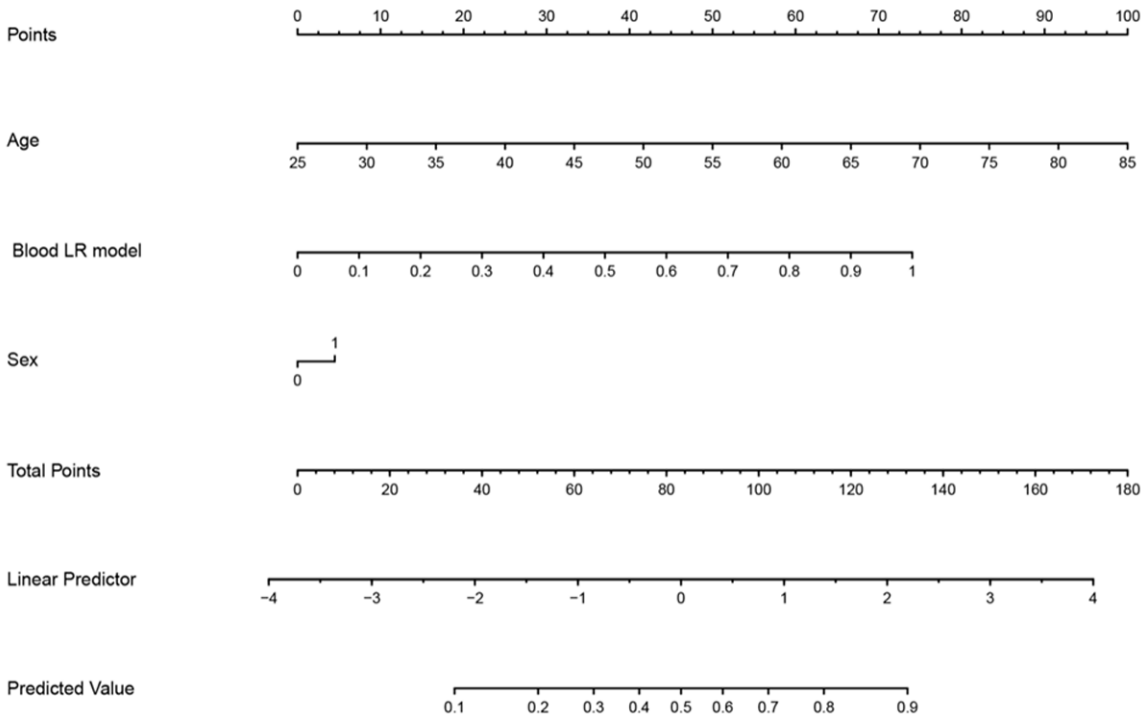

Figure 5. The nomogram based on the blood LR model and baseline information.

blood LR model and baseline information, such as age and sex (**Figure 5**). With the clinical information of patients, the nomogram can help doctors evaluate the probability of patients developing CRC.

Discussion

The lack of reliable, comfortable, and convenient methods for early diagnosis attributes to the high mortality of CRC. Liquid biopsy with

unique advantages is emerging as a non-invasive adjunct or alternative to standard tumor biopsy [33]. Comparing with other methylation-based models for differentiating CRC from healthy controls, our new 3-gene methylation signature has demonstrated a superior performance through combination strategy to successfully distinguish adenoma and LST lesions from benign ones. Although recent studies based on whole cfDNA methylomes [34] have generated a large amount of data valuable for basic research, whole cfDNA methylomes are less suitable for clinical testing as they are costly and time consuming.

In our study, early diagnosis model utilizing 3-gene methylation biomarker was proved to be more stable than single-gene method in multiple independent validation datasets. By comparing with previously reported method which combined SEPTIN9 with SDC2, our model was more efficient in detecting precancerous lesions, such as adenoma. Furthermore, we compared the robustness of our model with traditional antigen biomarkers. For example, we measured the levels of carcinoembryonic antigen (CEA, 0-5 ng/ml) and carbohydrate antigen 199 (CA199, 0-37 U/ml) in patient samples, which could indicate the incidence of cancer by abnormally high levels. The abnormal proportion of CEA and CA199 in blood was 0.468 (30/64) and 0.206 (13/63), respectively. Hence, our LR model has better sensitivity than clinical antigen biomarkers.

We not only developed an early diagnosis model, but also investigated the potential functions of 3-gene methylation biomarker based on 19 shared pathways. The heat map indicated the signal enhancement in multi-gene biomarker approach. Therefore, this combination strategy could improve the sensitivity of diagnosis to make it clinically useful.

We measured the methylation levels of these three genes beyond conventional β-value (see method), so that the LR model that was trained with TCGA M450K data had no effect on self-test blood data. We constructed another LR model using self-test adenoma data with a different threshold and tested it in self-test CRC data. We hypothesized that the methylation level of genes from adenoma patients was lower than that from CRC patients. Our new LR model showed an improved sensitivity to other classification models. Additional large-scale studies of methylation biomarkers for early diagnosis of CRC with cfDNA will benefit its clinical utility and initiate exploration of combinational non-invasive strategy in the future. More blood samples will be analyzed to validate the performance of this 3-gene methylation signature in differentiating CRC, precancerous lesions, and normal tissues.

Conclusions

In summary, we established a predictive blood-based 3-gene methylation biomarker to detect CRC by combining previously reported three non-invasive methylation biomarkers: SEPTIN9, SDC2 and VIM. Our studies suggest that combinational strategy with traditional machine learning algorithm can improve the accuracy of CRC early diagnosis.

Acknowledgements

The authors would like to express their gratitude to EditSprings (https://www.editsprings.cn/) for the expert linguistic services provided. Supported by Haiyan Research Fund of Harbin Medical University Cancer Hospital (JJZD2021-05), Research Fund of Beijing Medical Award Foundation (YXJL-2020-1223-0320), Scientific Research Fund of Heilongjiang Health Commission (2020-069) and the Guangdong Basic and Applied Basic Research Foundation (2021-A1515110238). The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of Harbin Medical University Cancer Hospital (KY2021-30; Jun.20th, 2021).

Disclosure of conflict of interest

None.

Address correspondence to: Yun-He Jia, Harbin Medical University Cancer Hospital Colorectal Cancer Center, Harbin 150086, Heilongjiang, China. E-mail: jiayunhe@hrbmu.edu.cn

References

[1] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2021; 71: 209-249.

[2] Liu SC, Wu J, Xia Q, Liu HD, Li WW, Xia XY and Wang JK. Finding new cancer epigenetic and genetic biomarkers from cell-free DNA by combining SALP-seq and machine learning. Comput Struct Biotechnol J 2020; 18: 1891-1903.

[3] Wolf AMD, Fontham ETH, Church TR, Flowers CR, Guerra CE, LaMonte SJ, Etzioni R, McKenna MT, Oeffinger KC, Shih YT, Walter LC, Andrews KS, Brawley OW, Brooks D, Fedewa SA, Manassaram-Baptiste D, Siegel RL, Wender RC and Smith RA. Colorectal cancer screening for average-risk adults: 2018 guideline update from the American cancer society. CA Cancer J Clin 2018; 68: 250-281.

[4] Ponder BA. Cancer genetics. Nature 2001; 411: 336-341.

[5] Herman JG and Baylin SB. Gene silencing in cancer in association with promoter hypermethylation. N Engl J Med 2003; 349: 2042-2054.

[6] Schuebel KE, Chen W, Cope L, Glockner SC, Suzuki H, Yi JM, Chan TA, Van Neste L, Van Criekinge W, van den Bosch S, van Engeland M, Ting AH, Jair K, Yu W, Toyota M, Imai K, Ahuja N, Herman JG and Baylin SB. Comparing the DNA hypermethylome with gene mutations in human colorectal cancer. PLoS Genet 2007; 3: 1709-1723.

[7] Luo YX, Wong CJ, Kaz AM, Dzieciatkowski S, Carter KT, Morris SM, Wang JP, Willis JE, Makar KW, Ulrich CM, Lutterbaugh JD, Shrubsole MJ, Zheng W, Markowitz SD and Grady WM. Differences in DNA methylation signatures reveal multiple pathways of progression from adenoma to colorectal cancer. Gastroenterology 2014; 147: 418-429, e8.

[8] Wang XK, Wang DW, Zhang HR, Feng MH and Wu XZ. Genome-wide analysis of DNA methylation identifies two CpG sites for the early screening of colorectal cancer. Epigenomics 2020; 12: 37-52.

[9] Grady WM, Willis J, Guilford PJ, Dunbier AK, Toro TT, Lynch H, Wiesner G, Ferguson K, Eng C, Park JG, Kim SJ and Markowitz S. Methylation of the CDH1 promoter as the second genetic hit in hereditary diffuse gastric cancer. Nat Genet 2000; 26: 16-17.

[10] Veigl ML, Kasturi L, Olechnowicz J, Ma AH, Lutterbaugh JD, Periyasamy S, Li GM, Drummond J, Modrich PL, Sedwick WD and Markowitz SD. Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing human MSI cancers. Proc Natl Acad Sci U S A 1998; 95: 8698-8702.

[11] Klutstein M, Nejman D, Greenfield R and Cedar H. DNA methylation in cancer and aging. Cancer Res 2016; 76: 3446-3450.

[12] Luo HY, Zhao Q, Wei W, Zheng LH, Yi SH, Li G, Wang WQ, Sheng H, Pu HY, Mo HY, Zuo ZX, Liu ZX, Li CF, Xie CB, Zeng ZL, Li WM, Hao XK, Liu YY, Cao SM, Liu WL, Gibson S, Zhang K, Xu GL and Xu RH. Circulating tumor DNA methylation profiles enable early diagnosis, prognosis prediction, and screening for colorectal cancer. Sci Transl Med 2020; 12: eaax7533.

[13] Ned RM, Melillo S and Marrone M. Fecal DNA testing for colorectal cancer screening: the ColoSure™ test. PLoS Curr 2011; 3: RRN1220.

[14] He N, Song LL, Kang Q, Jin P, Cai GX, Zhou JF, Zhou GP, Sheng JQ, Cai SJ, Wang JM, Han XL, Nie YZ and Wu KC. The pathological features of colorectal cancer determine the detection performance on blood ctDNA. Technol Cancer Res Treat 2018; 17: 1533033818791794.

[15] Song LL, Jia J, Peng XM, Xiao WH and Li YM. The performance of the SEPT9 gene methylation assay and a comparison with other CRC screening tests: a meta-analysis. Sci Rep 2017; 7: 3032.

[16] Wu D, Zhou GP, Jin P, Zhu JQ, Li SJ, Wu Q, Wang GQ, Sheng JQ, Wang JM, Song LL, Han XL and Qian JM. Detection of colorectal cancer using a simplified SEPT9 gene methylation assay is a reliable method for opportunistic screening. J Mol Diagn 2016; 18: 535-545.

[17] Jung G, Hernandez-Illan E, Moreira L, Balaguer F and Goel A. Epigenetics of colorectal cancer: biomarker and therapeutic potential. Nat Rev Gastroenterol Hepatol 2020; 17: 111-130.

[18] Fu B, Yan P, Zhang S, Lu Y, Pan L, Tang WQ, Chen S, Chen SF, Zhang AQ and Liu W. Cell-Free circulating methylated SEPT9 for noninvasive diagnosis and monitoring of colorectal cancer. Dis Markers 2018; 2018: 6437104.

[19] Song LL, Peng XM, Li YM, Xiao WH, Jia J, Dong C, Gong Y, Zhou GP and Han XL. The SEPT9 gene methylation assay is capable of detecting colorectal adenoma in opportunistic screening. Epigenomics 2017; 9: 599-610.

[20] Chen Y, Wang ZZ, Zhao GD, Sun C, Ma Y, Zhang LY, Zheng MX and Li HC. Performance of a novel blood-based early colorectal cancer screening assay in remaining serum after the blood biochemical test. Dis Markers 2019; 2019: 5232780.

[21] Zhao GD, Li H, Yang ZX, Wang ZZ, Xu MQ, Xiong SM, Li SM, Wu XT, Liu XY, Wang ZW, Zhu Y, Ma Y, Fei SJ and Zheng MX. Multiplex methylated DNA testing in plasma with high sensitivity and specificity for colorectal cancer screening. Cancer Med 2019; 8: 5619-5628.

[22] Grady WM, Yu M and Markowitz SD. Epigenetic alterations in the gastrointestinal tract: current and emerging use for biomarkers of cancer. Gastroenterology 2021; 160: 690-709.

[23] The Cancer Genome Atlas Program [https://tcga-data.nci.nih.gov/tcga/].
[24] Gene Expression Omnibus [https://www.ncbi.nlm.nih.gov/geo/].
[25] Oligo Calc: Oligonucleotide Properties Calculator [http://biotools.nubic.northwestern.edu/OligoCalc.html].
[26] Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES and Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005; 102: 15545-15550.
[27] The R Project for Statistical Computing [http://www.r-project.org/].
[28] Kalluri R. EMT: when epithelial cells decide to become mesenchymal-like cells. J Clin Invest 2009; 119: 1417-1419.
[29] Nam S and Park T. Pathway-based evaluation in early onset colorectal cancer suggests focal adhesion and immunosuppression along with epithelial-mesenchymal transition. PLoS One 2012; 7: e31685.
[30] Karagiannis GS, Petraki C, Prassas I, Saraon P, Musrap N, Dimitromanolakis A and Diamandis EP. Proteomic signatures of the desmoplastic invasion front reveal collagen type XII as a marker of myofibroblastic differentiation during colorectal cancer metastasis. Oncotarget 2012; 3: 267-285.
[31] Horvath S and Raj K. DNA methylation-based biomarkers and the epigenetic clock theory of ageing. Nat Rev Genet 2018; 19: 371-384.
[32] Unnikrishnan A, Freeman WM, Jackson J, Wren JD, Porter H and Richardson A. The role of DNA methylation in epigenetics of aging. Pharmacol Ther 2019; 195: 172-185.
[33] Heitzer E, Haque IS, Roberts CES and Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nat Rev Genet 2019; 20: 71-88.
[34] Liu JQ, Zhao HQ, Huang YK, Xu SP, Zhou Y, Zhang W, Li JQ, Ming Y, Wang XY, Zhao S, Li K, Dong XY, Ma YL, Qian TY, Chen XY, Xing ZY, Zhang Y, Chen HY, Liu ZH, Pang D, Zhou M, Wu ZH, Wang XW, Wang X, Wu N and Su JZ. Genome-wide cell-free DNA methylation analyses improve accuracy of non-invasive diagnostic imaging for early-stage breast cancer. Mol Cancer 2021; 20: 36.
[35] Zhao G, Liu X, Liu Y, Li H, Ma Y, Li S, Zhu Y, Miao J, Xiong S, Fei S and Zheng M. Aberrant DNA methylation of SEPT9 and SDC2 in stool specimens as an integrated biomarker for colorectal cancer early detection. Front Genet 2020; 11: 643.

Table S1. The private data

Class	Sample_ID	B22_SEPT9	B35_SDC2	B37_VIM	B72_ABL*	Quality Control	sex	age	hight (cm)	weight (kg)	CEA (0-5) ng/ml	CA199 (0-37) U/ml	Diabetes	Other diseases	hypertension	family history	History of colitis
Normal	252HP	22	45	30	3679	Pass	female	37									
Normal	249HP	28	30	42.5	3724	Pass	female	44									
Normal	285HP	25	26	36	3013	Pass	male	32									
Normal	271HP	36	51	49	4062	Pass	female	40									
Normal	254HP	33	71	45	3691	Pass	female	48									
Normal	207HP	71	23	32.5	2659	Pass	male	56									
Normal	281HP	176.5	25	40	3628	Pass	female	43									
Normal	220HP	187	31	49	4101	Pass	male	29									
Normal	243HP	182	28	56	4350	Pass	male	35									
Normal	232HP	179	31	49.5	3772	Pass	female	43									
Normal	210HP	112	23	28	3237	Pass	male	58									
Normal	212HP	166.5	18.5	35	3625	Pass	male	67									
Normal	202HP	27	30	36	1837	Pass	female	56									
Normal	248HP	18	98	24	2149	Pass	female	50									
Normal	263HP	160	33	43	3608	Pass	female	27									
Normal	260HP	34	173	41	3575	Pass	female	44									
Normal	201HP	35	42	48	2678	Pass	male	38									
Normal	276HP	225.5	33.5	52	3891	Pass	male	55									
Normal	059HP	182	24	40	3149.5	Pass	male	35									
Normal	233HP	193	30	63	3588	Pass	male	47									
Normal	197HP	127.5	26	31	2341	Pass	male	56									
Normal	269HP	46	76	59	2834	Pass	female	28									
Normal	204HP	217.5	24	64	3682.5	Pass	female	62									
Normal	255HP	223	35	96	3946	Pass	female	59									
Normal	258HP	282.5	23	38	3609	Pass	female	53									
Normal	206HP	216	37	64.5	3846	Pass	male	63									
Normal	244HP	194	32	63	3366.5	Pass	male	41									
Normal	227HP	208	24.5	50	3662	Pass	female	37									
Normal	257HP	37	317	46	4039	Pass	male	48									
Normal	250HP	331	33	50	3844	Pass	female	58									
Normal	242HP	255	32	76	4336.5	Pass	male	41									
Normal	208HP	246.5	33	53	3383.5	Pass	female	39									
Normal	234HP	237.5	27	45	3393.5	Pass	male										
Normal	278HP	288	27	41	3081	Pass	female	61									
Normal	200HP	268	27.5	34	2865.5	Pass	male	32									
Normal	211HP	258	23	42	2995.5	Pass	male	37									
Normal	251HP	35.5	407	44	4091	Pass	female	42									
Normal	205HP	250	34	59	3672	Pass	female	58									
Normal	262HP	303.5	32	52	3187	Pass	male	46									

Normal	267HP	280	29.5	43	2915.5	Pass	female	57				
Normal	284HP	35	324	45	3303.5	Pass	female	40				
Normal	279HP	210.5	28	87	3410	Pass	female	29				
Normal	253HP	261	32.5	113	3174	Pass	male	60				
Normal	236HP	231	39	125.5	2887	Pass	male	45				
Normal	222HP	207	26.5	101	2786	Pass	female	57				
Normal	274HP	298	37	63	2880	Pass	male	47				
Normal	213HP	375	38	106	3937	Pass	male	59				
Normal	221HP	134	23	42.5	3581	Pass	female	52				
Normal	272HP	315	30	70	3718	Pass	female					
Normal	247HP	217	33	51	3212	Pass	male					
Normal	214HP	268.5	33	52.5	3921	Pass	female	60				
Normal	229HP	200.5	23	48	3570	Pass	male	54				
Normal	246HP	453	26	42.5	3021	Pass	female	30				
Normal	224HP	206	23	41	3087.5	Pass	female	55				
Normal	240HP	361.5	29.5	41	2396	Pass	male	39				
Normal	218HP	235.5	33	58	3378	Pass	female					
Normal	282HP	205.5	30	47	3050	Pass	female	29				
Normal	256HP	270	88	55	3309	Pass	female	31				
Normal	264HP	207	23.5	74	3347.5	Pass	female	60				
Normal	230HP	234	26	86.5	3522	Pass	male	37				
Normal	270HP	568.5	510	50	3900	Pass	male	60				
Normal	275HP	320	36	56	4315	Pass	female	58				
Normal	217HP	175	29	36	3629.5	Pass	male	55				
Normal	277HP	248	23	42	3804	Pass	female	62				
Normal	265HP	25.5	61	34	3364	Pass	female	34				
Normal	259HP	590	55	50	2125	Pass	male	66				
Normal	241HP	26.5	27	34	3122	Pass	male	34				
Normal	245HP	218	30	71	3310.5	Pass	female	51				
Adenoma	1	1957	64	65	3765	Pass	female	54	145	64	0.6	6.25
Normal	235HP	342	25.5	51	3054	Pass	male					
Normal	266HP	531	34	58	3306	Pass	female	76				
CRC	2	3326.5	3080.5	3333.5	3408	Pass	female	57	150	49	964.1	145.5
CRC	3	2850	2721	78	2808.5	Pass	male	68	168	65	1.6	6.94
CRC	4	2916	2165	3184.5	3443	Pass	female	70	160	41	17.8	25.56
CRC	5	175	49.5	192	3625	Pass	female	58	160	61	1.77	8.86
CRC	6	1412	35	1410.5	3347.5	Pass	male	67	154	48	14.83	17.98
Adenoma	7	1307.5	1130	72	2759	Pass	female	51	159	60	51.29	39.33
CRC	8	1206.5	32	75	4033	Pass	female	42	160	74	1.33	8.68
Normal	228HP	164	149	42	3195	Pass	female	56				
Normal	216HP	267	33	54	3793.5	Pass	male					
CRC	10	591	27	62.5	3655	Pass	male	62	176	75	1.07	13.89

Adenoma	11	677	1339	78	3809	Pass	female					
Normal	215HP	174.5	29	39	3646.5	Pass	male					
Adenoma	12	222	35	49	3787	Pass	female					
Adenoma	13	572	64	2281.5	3631	Pass	male	68	173	65	13.37	24.5
Adenoma	14	298.5	38	60	3873.5	Pass	female	65	164	66	18.18	30.82
Normal	273HP	452	29	47	3175	Pass	female	55				
CRC	15	1019	2569	1953	4068	Pass	female	70	160	41	17.8	25.56
Adenoma	16	2750	3177	3476.5	3722	Pass	male					
Normal	223HP	220	26	49	3140	Pass	female	58				
CRC	17	378	76	1079.5	2910	Pass	female	74	160	68	1.2	24.97
Normal	231HP	188.5	25.5	36	3105	Pass	male	46				
Adenoma	18	520.5	934	1084	3348	Pass						
CRC	19	1035	1768	79	3399	Pass	male	64	176	64		
Adenoma	21	204.5	1727	63	3819	Pass	male	48	178	90	3.07	10.79
Normal	261HP	239	30	46	3674	Pass	female	28				
Adenoma	22	502	1498	53	3995	Pass	male	85	170	67	32.64	5.92
Adenoma	24	620	1219.5	65	2577.5	Pass						
Adenoma	25	235	1559	249	3495	Pass	male	63	168	74	1.29	10.33
Adenoma	27	303	1626.5	56	3625.5	Pass	male	63	172	76	1.76	5.18
CRC	28	622	999	140.5	3861.5	Pass	female	74	160	72	2.49	6.96
Adenoma	34	1467	1943	1751.5	3602	Pass						
Normal	283HP	162	1505.5	46	2694	Pass	male	56				
Normal	199HP	38	32	50	2325	Pass						
CRC	40	481	892	466	1751.5	Pass	male	76	172	62	19.46	20.26
CRC	42	209	1134	70	2191	Pass	female	59	162	60	1.5	25.22
Adenoma	43	305.5	1112	40	1937	Pass						
Adenoma	46	1632	1464	1760.5	1296	Pass	female	35	168	46	1.64	51.75
Normal	280HP	426	169	47	2978	Pass	female	66				
CRC	47	794	1061.5	1020	1469	Pass	female	60	167	75	50.42	19.25
CRC	49	1082.5	1086	1429.5	2172	Pass	male	76	168	61	72.14	64.9
Adenoma	51	1019.5	1988	1581	2136	Pass	female	35	168	46	1.64	51.75
CRC	52	538	1285.5	1339	2631.5	Pass	female	60	167	75	50.42	19.25
CRC	54	597	1698	1171.5	2554.5	Pass	male	76	168	61	72.14	64.9
CRC	56	913	2095	2475	2200.5	Pass	male	77	160	64	4.89	56.77
CRC	57	194	1593	58	1400.5	Pass	female	31	168	55	5.11	4.82
CRC	59	179.5	1499	486	2654	Pass	male	68	173	62	3.66	11.33
Adenoma	60	998	1457	46	2554	Pass	male	57				
Normal	209HP	275	38	103	4158	Pass	female	30				
Normal	226HP	180	24	46	3442	Pass						
Adenoma	63	177	23.5	36	2103	Pass	female	67	155	52	4.87	13.03
Normal	203HP	170	1628	47	1898	Pass						
Adenoma	64	147	204.5	72	3248	Pass	male	57	160	62.5	2.14	19.79
Normal	268HP	48.5	3410	61	3394	Pass						

CRC	65	233.5	1404	194	2822	Pass	male	72	179	96	15.62	11.68					
CRC	71	234	49	83	3193.5	Pass	male	61	168	64	8.41	7.6					
Adenoma	73	197.5	42	74	3847	Pass	male	59	165	70	46.64	14.91	no	no	no	no	no
CRC	75	2411	2694	3621.5	2743.5	Pass	male	51	173	70	4.75	107.2	no	no	no	no	no
CRC	77	507	55	56	3878.5	Pass	male	69	173	80	3.52	12.36	no	no	no	no	no
CRC	78	417	1135	78	3432	Pass	female	60	158	58	2.34	22.62	yes	no	yes	no	no
CRC	80	337	37	2083	2039.5	Pass	male	58	176	73	3.23	25.47					
CRC	83	208	50	82.5	3961	Pass	female	49	165	65	0.74	2.6					
Adenoma	84	2051	2822	1912	2569.5	Pass							no	no	no	no	no
CRC	85	2418	3106	920.5	2677	Pass	male	70	165	60	15.42	11.15	no	heart disease	no	no	no
CRC	86	384.5	729	155	3865.5	Pass	male	76	172	62	19.46	20.26		Coronary heart disease, fatty liver	yes	no	no
CRC	87	396	1561	710	3608.5	Pass	male	72	179	96	15.62	11.68					
Adenoma	88	288.5	928	67	2564	Pass	male	63	172	76	1.76	5.18		heart disease	yes	no	no
CRC	89	352	1136.5	26.5	3658	Pass	female	66	162	45	7.72	254.4	no	no	no	no	no
Adenoma	93	782	2124	1314	3488.5	Pass	female	79	153	55	30.1	71.76	no	Arrhythmia, heart disease	no	no	no
CRC	94	193.5	16	59.5	2768	Pass	female	71	160	54	5.47	26.58	no	no	no	no	no
Adenoma	95	1726	2181	2146.5	3579	Pass							yes	no	yes	no	no
CRC	96	173.5	1208	33	3346	Pass	male	59	170	72.5	1.04	6.39					
CRC	97	263	20	38	3995.5	Pass	female	83	158	65	1.24	8.25	no		no	no	no
CRC	98	360	100	54	4318	Pass	female	53	165	58	2.77	6.42	no	Heart disease, lumen infarction, hepatitis C	yes	no	no
Adenoma	99	1251	1651	1319.5	3653	Pass	male	61	175	55	6.56	21.39	no	no	yes	no	no
CRC	101	831	676	712	3506	Pass	male	55	169	81	4.65	0.5999	no	no	no	no	no
CRC	103	539	1059.5	98	3841	Pass	female	79	158	58	13.11	0.5999					
CRC	104	806	85	78	3959.5	Pass	male	59	178	95	2.31	8.5					
Adenoma	106	1838	2601	3676	3965	Pass	female	47	165	80	347.4	1000.0001	no	no	no	no	no
CRC	107	220.5	18	129	3917	Pass	male	69	170	53	2.59	11.33	no	Coronary heart disease, cerebral infarction	yes	no	no
CRC	108	168	1290	48	3616.5	Pass	male	33	176	79	0.26	4.56	no	no	yes	no	no
CRC	109	377	19	92.5	4179	Pass	male	63	172	63	2.84	11.15					
CRC	112	134	22	36	4141.5	Pass	male	67	169	78	1.23	9.92	no	Sulfonamide allergy	no	no	no
CRC	113	116.5	22	31	2663	Pass	male	64	170	62	2.52	54.86					
CRC	114	752.5	52	42	3992.5	Pass	male	64	170	70	9.88	27.57	no	no	no	no	no
Adenoma	120	199	21	323	4487.5	Pass					1.13	28,4					
Adenoma	121	530.5	19	31	4161.5	Pass	female	79	153	55	30.1	71.76	yes	no	yes	no	no
Adenoma	122	731.5	2139	36.5	4174.5	Pass	male	61	175	55	6.56	21.39					
Adenoma	219HP	699.5	1991	537	3033	Pass											

B72_ABL*: This is a quality control gene, and only the samples passed the quality control are displayed.

Table S2. The different methylation probes

ID	UCSC_RefGene_Name	logFC	AveExpr	t	*P*.Value	adj.*P*.Val	B
cg20275528	SEPT9	3.269757	-1.17346	23.42865	2.43E-72	4.74E-70	154.2244
cg04261408	SDC2	2.354538	-0.83222	22.46768	1.25E-68	1.22E-66	145.7223
cg25664438	SDC2	2.61509	-1.18268	18.8865	1.49E-54	9.69E-53	113.4604
cg20594401	SDC2	3.543253	-1.88031	18.36976	1.67E-52	8.16E-51	108.7616
cg16935295	SDC2	3.865971	-1.70916	17.44555	7.79E-49	3.04E-47	100.3562
cg14538332	SDC2	2.046011	-1.55072	16.70519	6.68E-46	2.17E-44	93.63439
cg27313572	VIM	3.469989	-2.02147	15.58231	1.78E-41	4.97E-40	83.49186
cg18719750	SDC2	2.361213	-1.49733	15.54057	2.60E-41	6.34E-40	83.11664
cg13096260	SDC2	3.112523	-1.53691	15.09927	1.39E-39	3.01E-38	79.15951
cg12783819	SEPT9	1.670118	-0.91861	15.068	1.84E-39	3.59E-38	78.87985
cg05184938	SEPT9	1.176393	-1.21609	14.5655	1.66E-37	2.94E-36	74.40179
cg10292139	SDC2	2.606294	-1.41502	14.48124	3.52E-37	5.72E-36	73.65397
cg16673702	SDC2	2.042485	-1.89988	14.1387	7.41E-36	1.11E-34	70.62435
cg06848185	SEPT9	3.79982	-2.23369	14.03362	1.88E-35	2.62E-34	69.69841
cg19554255	SEPT9	2.468666	-2.00352	13.66884	4.70E-34	6.11E-33	66.49823
cg07146119	SDC2	3.674454	-2.64289	13.63709	6.21E-34	7.57E-33	66.22084
cg24732574	SDC2	2.346285	-1.98809	13.57301	1.09E-33	1.25E-32	65.66136
cg02320862	SEPT9	0.919192	-0.59679	13.28445	1.36E-32	1.48E-31	63.15163
cg12985929	SEPT9	1.732971	-1.52997	13.26278	1.65E-32	1.69E-31	62.96377
cg14625631	SDC2	2.155366	-2.05443	12.97187	2.06E-31	2.01E-30	60.45167
cg15044248	SEPT9	4.070164	-2.35475	12.68812	2.39E-30	2.22E-29	58.01883
cg26003222	SEPT9	0.108792	-0.145	11.67306	1.30E-26	1.15E-25	49.47887
cg25070637	SDC2	2.246658	-1.93638	11.48126	6.38E-26	5.41E-25	47.89731
cg03804136	SEPT9	2.992025	-2.03104	11.32568	2.30E-25	1.87E-24	46.62275
cg02884239	SEPT9	2.031844	-2.13586	10.88427	8.46E-24	6.60E-23	43.04927
cg23858558	SDC2	1.792841	-1.57643	10.80729	1.57E-23	1.18E-22	42.43281
cg19083265	SEPT9	0.102609	-0.17603	10.14389	3.06E-21	2.21E-20	37.21085
cg26983469	VIM	1.523416	-1.40629	9.953157	1.35E-20	9.40E-20	35.74123
cg23356017	SDC2	-0.98382	-1.07578	-9.31711	1.70E-18	1.15E-17	30.95256
cg10303967	SDC2	-0.91819	-1.02037	-8.97844	2.08E-17	1.35E-16	28.47794
cg06761530	SEPT9	0.074245	-0.12558	8.453798	9.03E-16	5.68E-15	24.75647
cg01320579	SEPT9	0.594359	-0.46709	8.2569	3.58E-15	2.18E-14	23.39688
cg11973177	VIM	2.038894	-2.00424	8.191291	5.64E-15	3.33E-14	22.94853
cg19422205	SEPT9	0.179862	-0.19297	8.036146	1.64E-14	9.40E-14	21.89779
cg01329309	SDC2	-0.96102	-1.23274	-7.34791	1.57E-12	8.77E-12	17.40443
cg22137815	SDC2	-0.52528	-0.66346	-7.19768	4.11E-12	2.23E-11	16.4617
cg27627381	SEPT9	0.104626	-0.30948	6.840825	3.80E-11	2.00E-10	14.27995
cg09481866	SEPT9	0.042632	-0.06639	6.427326	4.51E-10	2.31E-09	11.85716
cg13061373	SEPT9	0.301049	-0.61653	6.350662	7.05E-10	3.52E-09	11.42082
cg00324097	SEPT9	0.831476	-3.35169	6.143093	2.31E-09	1.13E-08	10.26009
cg02746869	VIM	1.609884	-3.22694	6.079967	3.30E-09	1.57E-08	9.913164
cg23051970	SEPT9	-0.30893	-0.6273	-5.86592	1.08E-08	5.01E-08	8.758244
cg18514820	VIM	2.027785	-3.76656	5.855985	1.14E-08	5.17E-08	8.705473
cg26842303	SDC2	-0.65079	-0.7747	-5.72051	2.37E-08	1.05E-07	7.992865
cg20319091	VIM	1.839237	-4.31978	5.470336	8.85E-08	3.84E-07	6.713151
cg23343486	SEPT9	-0.49089	-1.7778	-5.36953	1.49E-07	6.30E-07	6.210986
cg21830368	SEPT9	0.217599	-0.34191	5.358627	1.57E-07	6.52E-07	6.157128

cg03276479	SDC2	0.497221	-1.02417	5.313204	1.98E-07	8.04E-07	5.933793
cg20796999	SEPT9	-0.29887	-0.4191	-5.2123	3.28E-07	1.31E-06	5.443416
cg12190613	SDC2	-0.65099	-1.51938	-5.20075	3.48E-07	1.36E-06	5.38777
cg01003015	VIM	0.997079	-2.95996	5.097012	5.80E-07	2.22E-06	4.892921
cg15293063	SEPT9	0.062962	-0.24548	5.054499	7.14E-07	2.68E-06	4.692576
cg10611580	SEPT9	-0.45933	-1.62936	-4.76326	2.85E-06	1.03E-05	3.359079
cg16962683	SDC2	-0.22566	-0.30935	-4.6932	3.94E-06	1.40E-05	3.048533
cg00446722	SDC2	-0.34406	-0.44566	-4.62928	5.27E-06	1.84E-05	2.76874
cg05151811	VIM	0.818681	-3.67092	4.586074	6.41E-06	2.15E-05	2.581549
cg19111999	VIM	0.678579	-3.16424	4.536378	8.01E-06	2.62E-05	2.368129
cg06513247	SEPT9	0.557058	-3.01605	4.53479	8.06E-06	2.62E-05	2.361344
cg00871371	SEPT9	0.26969	-0.60879	4.493806	9.67E-06	3.09E-05	2.186941
cg18030776	SDC2	-0.35642	-0.44873	-4.25963	2.67E-05	8.40E-05	1.217452
cg27517823	SDC2	-0.49805	-0.84352	-4.16382	3.99E-05	0.000124	0.834202
cg08918274	VIM	1.028488	-3.68199	4.150406	4.22E-05	0.000129	0.781151
cg00146951	VIM	0.738302	-3.30782	4.037419	6.72E-05	0.000198	0.340582
cg14885762	SEPT9	-0.28938	-1.65955	-3.93322	0.000102	0.000297	-0.05595
cg12194864	SEPT9	0.028858	-0.17508	3.906884	0.000113	0.000324	-0.15467
cg17053854	SEPT9	0.02213	-0.08776	3.903587	0.000115	0.000324	-0.16698
cg16686174	SEPT9	0.114944	-0.22602	3.825437	0.000156	0.000434	-0.45616
cg26899651	SEPT9	0.155809	-0.78518	3.7966	0.000174	0.000479	-0.56151
cg21204860	SEPT9	0.347541	-2.46302	3.72017	0.000234	0.000633	-0.83719
cg20198108	VIM	0.489796	-3.48579	3.660767	0.000292	0.000781	-1.04789
cg26674546	SEPT9	-0.31394	-2.17491	-3.61362	0.000349	0.000919	-1.2129
cg17840501	SEPT9	0.111936	-0.56347	3.602538	0.000363	0.000945	-1.2514
cg14183922	SEPT9	0.246277	-1.00687	3.526227	0.000481	0.001234	-1.51352
cg17697835	SEPT9	-0.3204	-1.88387	-3.511	0.000508	0.001287	-1.56519
cg21678736	SEPT9	-0.26452	-1.09524	-3.48555	0.000557	0.001393	-1.65112
cg12098949	SEPT9	0.213696	-1.40629	3.420719	0.000702	0.001734	-1.86735
cg26777303	SDC2	0.093088	-0.26693	3.377547	0.000818	0.001994	-2.00924
cg13342364	SEPT9	-0.21639	-1.49861	-3.25266	0.001261	0.003035	-2.41021
cg02277530	SEPT9	0.105306	-0.58919	3.115786	0.001995	0.00463	-2.83339
cg10755077	SEPT9	-0.14952	-0.84942	-3.09209	0.002156	0.004941	-2.90491
cg06791979	SEPT9	0.063269	-0.1783	3.088845	0.002179	0.004941	-2.91466
cg12858577	SEPT9	0.171116	-0.46306	3.037354	0.002576	0.005707	-3.06814
cg21733927	SEPT9	-0.24042	-1.60141	-3.02279	0.002699	0.005914	-3.11111
cg07827420	SEPT9	0.168528	-1.02451	2.973523	0.003159	0.006803	-3.25501
cg07324245	SEPT9	-0.17249	-1.18107	-2.972	0.003175	0.006803	-3.25943
cg20557159	SEPT9	-0.06563	-0.21552	-2.80739	0.00529	0.011091	-3.72368
cg26063719	VIM	-0.23906	-3.20634	-2.75474	0.006198	0.012857	-3.86687
cg08772789	SEPT9	-0.17694	-0.65479	-2.70055	0.007277	0.014938	-4.01152
cg03236137	SEPT9	0.117218	-0.60989	2.654887	0.008316	0.016891	-4.13128
cg17866352	SEPT9	-0.16939	-1.41112	-2.64611	0.00853	0.017148	-4.15407
cg01154046	VIM	0.63015	-5.21537	2.572244	0.010538	0.020916	-4.34305
cg12203543	SEPT9	0.045769	-0.02667	2.569557	0.010619	0.020916	-4.34983
cg19654743	SEPT9	0.244874	-2.37339	2.565004	0.010756	0.020975	-4.36129
cg16293484	SEPT9	0.091012	-0.6848	2.511554	0.012494	0.024122	-4.49446
cg16317901	SEPT9	-0.2245	-1.59158	-2.50395	0.012761	0.024395	-4.51319
cg23912429	VIM	0.490435	-3.24765	2.469223	0.014043	0.026587	-4.59802

cg05337753	SEPT9		0.075919	-0.27012	2.426078	0.015795	0.029616	-4.70181
cg08034797	SEPT9		0.052125	-0.36741	2.391184	0.01735	0.032222	-4.78446
cg20772590	SEPT9		0.177481	-1.83871	2.35533	0.019087	0.035114	-4.86819
cg14830748	SDC2		-0.08295	-0.20585	-2.31426	0.021263	0.038751	-4.96259
cg03568017	SEPT9		-0.13402	-3.85033	-2.29793	0.022187	0.040059	-4.99967
cg20305489	SEPT9		-0.27297	-0.75401	-2.26233	0.024323	0.043513	-5.07964
cg18278424	SEPT9		-0.14385	-1.64791	-2.25594	0.024724	0.04383	-5.09387
cg14942501	SDC2		-0.04074	-0.11858	-2.21298	0.027579	0.048449	-5.1885

Table S3. The performance of two-marker early diagnosis model

Data_ID			T	N	Cut_off	Sensitivity	Specificity	AUC
TCGA_M450k	Training set	T	276	19	0.999999	0.936	1	0.993
		N	0	38				
GSE101764	Testing set	T	100	12	0.999999	0.893	0.973	0.958
		N	4	145				
GSE42752		T	21	1	0.999999	0.955	0.976	0.999
		N	1	40				
GSE48684		T	54	10	0.999999	0.844	1	0.972
		N	0	41				
GSE68060		T	73	8	0.999999	0.901	0.917	0.92
		N	3	33				
GSE77718		T	93	3	0.999999	0.969	0.885	0.965
		N	11	85				
GSE77954		T	12	1	0.999999	0.923	0.909	0.944
		N	1	10				
GSE75546		T	6	0	0.999999	1	1	1
		N	0	6				
GSE77965		T	6	0	0.999999	1	0.833	1
		N	1	5				
GSE40279*		T	-	-	0.999999	-	1	-
		N	0	656				
Adenoma								
	Testing set		AD	N				
GSE77954AD		AD	12	0	0.999999	1	0.909	0.951
		N	1	10				
GSE106556AD		AD	9	1	0.999999	0.9	0.9	0.97
		N	1	9				
GSE48684AD		AD	30	12	0.999999	0.714	1	0.966
		N	0	41				
Self-testing								
Adenoma	Training set	AD	25	6	0.2	0.806	0.88	0.829
		N	10	74				
CRC		T	34	10	0.2	0.773	0.88	0.87
		N	11	73				

The red colour means that the performance of two-maker model is better than three-maker. The blue means the opposite.

Table S4. The shared Hall_mark pathway of three genes

HALLMARK_ID	SEPT9_ enrich-mentScore	SEPT9_ NES	SEPT9_ pvalue	SEPT9_p. adjust	SEPT9_ qvalues	SDC2_ enrich-mentScore	SDC2_ NES	SDC2_ pvalue	SDC2_p. adjust	SDC2_ qvalues	VIM_ enrich-mentScore	VIM_ NES	VIM_ pvalue	VIM_p. adjust	VIM_ qvalues
HALLMARK_APICAL_JUNCTION	0.519	2.579	1.13E-17	5.64E-16	1.66E-16	0.582	1.743	4.25E-12	2.21E-11	1.05E-11	0.592	1.759	3.16E-13	1.33E-12	5.99E-13
HALLMARK_MYOGENESIS	0.502	2.483	2.19E-14	3.65E-13	1.08E-13	0.649	1.945	5.10E-20	1.07E-18	5.10E-19	0.639	1.898	3.12E-19	4.36E-18	1.97E-18
HALLMARK_EPITHELIAL_MESENCHY-MAL_TRANSITION	0.471	2.346	3.40E-13	3.40E-12	1.00E-12	0.786	2.353	1.25E-51	5.25E-50	2.50E-50	0.766	2.281	2.01E-46	8.46E-45	3.82E-45
HALLMARK_TNFA_SIGNALING_VIA_NFKB	0.421	2.098	1.16E-09	8.32E-09	2.45E-09	0.549	1.645	3.41E-09	1.19E-08	5.68E-09	0.585	1.74	1.37E-13	6.40E-13	2.89E-13
HALLMARK_HYPOXIA	0.426	2.108	1.54E-09	9.63E-09	2.84E-09	0.504	1.509	6.25E-06	1.64E-05	7.81E-06	0.529	1.574	2.72E-07	7.61E-07	3.43E-07
HALLMARK_KRAS_SIGNALING_DN	0.418	2.05	4.27E-08	2.14E-07	6.30E-08	0.469	1.401	7.82E-04	1.82E-03	8.68E-04	0.491	1.456	9.97E-05	2.20E-04	9.94E-05
HALLMARK_INTERFERON_ALPHA_RE-SPONSE	0.47	2.096	1.13E-06	4.73E-06	1.39E-06	0.505	1.469	6.34E-04	1.57E-03	7.45E-04	0.579	1.691	1.07E-06	2.82E-06	1.27E-06
HALLMARK_COAGULATION	0.399	1.882	1.45E-05	5.60E-05	1.65E-05	0.612	1.808	2.65E-11	1.11E-10	5.29E-11	0.595	1.748	9.65E-10	3.12E-09	1.41E-09
HALLMARK_INTERFERON_GAMMA_RE-SPONSE	0.346	1.724	2.15E-05	7.67E-05	2.26E-05	0.582	1.743	4.74E-12	2.21E-11	1.05E-11	0.623	1.855	6.55E-18	5.50E-17	2.48E-17
HALLMARK_IL2_STAT5_SIGNALING	0.34	1.689	4.96E-05	1.65E-04	4.87E-05	0.532	1.593	8.21E-08	2.65E-07	1.26E-07	0.551	1.639	7.39E-10	2.59E-09	1.17E-09
HALLMARK_HEDGEHOG_SIGNALING	0.579	2.072	6.69E-05	2.09E-04	6.16E-05	0.595	1.597	2.69E-03	5.64E-03	2.69E-03	0.62	1.649	1.28E-03	2.62E-03	1.18E-03
HALLMARK_COMPLEMENT	0.321	1.595	3.44E-04	8.60E-04	2.54E-04	0.557	1.668	9.72E-10	3.71E-09	1.77E-09	0.603	1.793	3.28E-15	1.72E-14	7.77E-15
HALLMARK_INFLAMMATORY_RESPONSE	0.297	1.476	1.40E-03	2.91E-03	8.59E-04	0.64	1.917	2.80E-19	3.91E-18	1.86E-18	0.649	1.93	3.42E-21	7.18E-20	3.24E-20
HALLMARK_ANGIOGENESIS	0.505	1.831	1.61E-03	3.23E-03	9.51E-04	0.71	1.915	2.47E-06	6.91E-06	3.29E-06	0.705	1.887	3.55E-06	8.77E-06	3.96E-06
HALLMARK_IL6_JAK_STAT3_SIGNALING	0.354	1.567	5.28E-03	9.11E-03	2.68E-03	0.603	1.741	1.04E-06	3.12E-06	1.48E-06	0.624	1.814	4.38E-08	1.31E-07	5.93E-08
HALLMARK_APOPTOSIS	0.294	1.42	7.16E-03	1.19E-02	3.52E-03	0.453	1.35	3.87E-03	7.74E-03	3.68E-03	0.497	1.471	6.10E-05	1.42E-04	6.42E-05
HALLMARK_TGF_BETA_SIGNALING	0.407	1.635	0.009	0.014	0.004	0.517	1.447	1.50E-02	2.86E-02	1.36E-02	0.531	1.49	6.92E-03	1.26E-02	5.70E-03
HALLMARK_UNFOLDED_PROTEIN_RE-SPONSE	0.324	1.496	0.01	0.015	0.005	-0.286	-1.373	1.18E-03	2.61E-03	1.24E-03	-0.266	-1.471	1.31E-03	2.62E-03	1.18E-03
HALLMARK_UV_RESPONSE_DN	0.299	1.424	0.019	0.027	0.008	0.628	1.863	3.35E-13	2.34E-12	1.12E-12	0.607	1.788	3.80E-11	1.45E-10	6.54E-11

Table S5. The Hall_mark pathway result of SDC2

ID	Description	setSize	enrich-mentScore	NES	pvalue	p.adjust	qvalues	rank	leading_edge	core_enrichment
HALLMARK_EPI-THELIAL_MESEN-CHYMAL_TRANSI-TION	HALLMARK_EPITHELIAL_MESENCHY-MAL_TRAN-SITION	200	0.785755	2.35265	1.25E-51	5.25E-50	2.50E-50	2924	tags=72%, list=13%, signal=63%	11167/10979/4060/800/6678/7431/6444/1634/1462/2200/6591/1842/2191/7070/1281/5159/1009/6695/1290/7058/3624/30008/115908/50863/1296/59/1278/7424/6876/26585/10631/4638/1293/633/5654/4313/10398/4256/5396/1301/8038/6424/10516/4015/1303/1000/1292/1289/8076/10085/4837/5118/2192/25878/2697/2199/7412/3956/1277/1809/2619/3678/4017/2335/4323/10409/22795/131578/6586/1284/9353/5744/4148/8572/1311/5376/7078/3371/1282/7456/2303/1490/2006/2316/3486/2014/5999/4016/813/2669/7169/3690/6422/3685/1004/7057/7040/11010/6387/3491/3915/2882/64175/871/6696/7857/10486/23705/3908/3688/2247/5054/50509/1307/5806/10272/649/333/25890/8325/6445/627/7076/2201/2817/7171/4853/7474/4616/2331/966/3487/3693/7052/1294/4035/7168/57692/5352/7980/3569/1893/6443
HALLMARK_MYO-GENESIS	HALLMARK_MYOGEN-ESIS	188	0.649004	1.945025	5.10E-20	1.07E-18	5.10E-19	3675	tags=49%, list=17%, signal=42%	6678/6444/22808/1281/3490/10278/165/6876/4638/1293/5507/8038/1292/1306/26353/25937/1277/8082/1410/3316/1284/1346/5740/2273/10468/2702/3486/4208/7169/187/2775/7040/9260/4629/8912/1012/845/4046/3908/29970/8877/3479/3688/2247/948/1489/2878/70/6442/6840/6445/11155/3801/5837/347/8736/10290/7136/88/6261/1674/10580/5348/844/4616/1266/4684/3693/58529/4634/4842/5621/3679/26287/4635/1837/2170/4627/2934/6016/3756/4205/6764/4608/1191/4209/1134/5816/3270/6588/7140/89/115
HALLMARK_IN-FLAMMATORY_RE-SPONSE	HALLMARK_INFLAMMA-TORY_RE-SPONSE	197	0.640437	1.917421	2.80E-19	3.91E-18	1.86E-18	5110	tags=57%, list=23%, signal=45%	3624/558/9180/3554/5099/7130/10630/10203/3678/4323/4481/6347/4973/5027/5739/719/2014/728/23601/3690/187/1536/929/1435/6324/366/1880/1240/7096/6401/8685/3937/23533/2357/8877/5054/8578/7097/5996/6354/10316/3587/23166/1441/3383/114548/9435/19/3586/6402/7076/5724/3604/962/6869/5008/3575/1318/6004/970/7439/2015/1902/4049/7432/969/4210/29933/6504/255488/6542/3091/3738/6367/4283/10148/3560/3037/3569/5142/3627/958/10886/1236/131566/10125/6376/6506/3576/6361/3269/7162/60675/3593/6372/684/4693/623/8809/6352/10663/3656/5329/27074/4792/1906/5732/8743/3732/3696/10242/166929/5610
HALLMARK_MYC_TARGETS_V2	HALLMARK_MYC_TAR-GETS_V2	58	-0.71607	-3.27707	1.40E-17	1.47E-16	6.98E-17	3794	tags=72%, list=17%, signal=60%	10196/83743/10171/2193/10514/7965/3099/56915/23223/7374/4173/4869/10244/6949/6652/26354/705/4174/54663/10733/27346/3336/23481/3329/6832/4839/29078/5347/9238/6573/6723/9136/92856/51154/5245/5036/79050/56342/51491/79711/3177/80324

HALLMARK_KRAS_SIGNALING_UP	HALLMARK_KRAS_SIGNALING_UP	197	0.606252	1.815074	6.93E-15	5.82E-14	2.77E-14	4692	tags=48%, list=21%, signal=38%	2982/3624/3075/5178/9945/5396/9358/25960/2791/8404/10457/9935/8829/2995/5175/64123/23643/9590/3290/719/5327/7805/2324/3486/2207/220/80380/3689/51704/4320/4318/5649/6324/2162/6696/5328/1240/2022/50856/55273/3936/6616/51311/2113/1794/4674/7462/83483/3587/4803/951/9734/56729/1438/669/347/9892/2115/101/3575/10320/3481/6004/639/23136/7035/10418/7852/27299/3800/50486/2012/28951/3627/7128/29106/1363/25907/131566/51199/11184/51129/6868/5950/3237/8854/7476/27344/5743/23531/2254/5593/867/5329
HALLMARK_UV_RESPONSE_DN	HALLMARK_UV_RESPONSE_DN	144	0.628318	1.862797	3.35E-13	2.34E-12	1.12E-12	4287	tags=49%, list=19%, signal=39%	6383/6591/1281/5159/1290/10000/2202/1278/23266/1301/10516/4325/3488/273/2697/8829/1277/857/5797/2908/4052/5376/4131/10395/5737/5999/3690/3491/3915/23345/493/10486/29970/5054/8613/3778/627/6310/4853/5577/5293/7035/2534/861/1902/10370/140609/1843/57178/3667/773/2535/3037/9863/3815/2869/10401/1601/11099/7048/23116/50937/11030/10611/323/546/4092/7436/7799/8455
HALLMARK_ALLOGRAFT_REJECTION	HALLMARK_ALLOGRAFT_REJECTION	197	0.590167	1.766917	6.13E-13	3.68E-12	1.75E-12	5468	tags=57%, list=25%, signal=44%	7070/3624/586/2149/6347/2213/9450/2316/2268/1230/942/6688/3689/7040/7903/5552/1435/4318/7042/322/920/7096/10261/3937/3059/2113/8477/7097/6354/729230/3117/23166/3111/3383/114548/5788/6356/7454/5579/3586/940/2533/9734/3603/7076/941/3824/3109/1234/6357/3559/9655/10333/3122/8444/3683/8560/3455/4050/912/6363/3091/3702/1991/6351/6367/4283/3560/3569/56253/5199/958/972/6775/2634/3662/7124/3108/27240/914/3594/973/9607/11184/6772/10225/917/916/50852/3848/10563/4690/3593/3112/3625/4689/7453/919/3596/6352/5551/567/924/3589/915/7321/925/6885/959/3133/3717/3566/1029
HALLMARK_APICAL_JUNCTION	HALLMARK_APICAL_JUNCTION	197	0.582089	1.742733	4.25E-12	2.21E-11	1.05E-11	5356	tags=49%, list=24%, signal=38%	83700/1462/2200/143903/7070/1009/10000/4313/10398/8910/51148/7216/91624/1272/7412/5175/4478/8745/9353/11096/947/9459/257194/7106/942/1004/4318/57555/4628/84552/72/3688/51466/1307/5010/7122/70/3383/649/7450/140885/57863/2771/5788/8515/87/55742/7414/2533/10290/88/347902/794/6464/2318/247/2593/30835/80381/5600/3667/7059/4627/3680/253559/3384/1006/6624/54413/23114/29126/9379/6251/6376/89/10109/1758/4267/7781/9019/147/5310/9672/2770/1013/5880/7791/7087/7185/7082/10487/6237/8935/5728/1432/51776/8517
HALLMARK_INTERFERON_GAMMA_RESPONSE	HALLMARK_INTERFERON_GAMMA_RESPONSE	200	0.581975	1.742506	4.74E-12	2.21E-11	1.05E-11	6399	tags=61%, list=29%, signal=44%	716/3075/715/710/7130/7412/8082/6347/2209/942/9934/7903/26524/1240/3937/6403/2357/1439/6354/3587/3117/3383/10791/10875/55016/9021/669/3434/10068/25939/163351/3455/3433/969/91543/23586/6773/57823/10964/84159/4600/3091/3123/3437/4283/7127/3560/3569/115361/8869/6648/5142/10561/3627/7128/958/952/972/29126/6775/3662/3108/3431/10437/4939/219285/3620/54739/6772/94240/103/684/10628/5292/7453/6375/5743/4261/6352/10906/129607/3663/57169/55008/4792/567/8743/6774/4599/64135/9636/5770/55601/4502/81030/5610/694/8202/51056/10392/6373/3717/3566/317649/3695/11274/4790/837/355/54625/84888/57674/9246/6398/3001/4940/55072/2841/5371/4615/5359/3665

HALLMARK_COAGULATION	HALLMARK_COAGULATION	129	0.612065	1.808163	2.65E-11	1.11E-10	5.29E-11	4031	tags=43%, list=18%, signal=35%	6678/2200/716/1513/3075/715/5654/4313/710/2/5175/2335/4323/7056/4973/1311/7078/5327/1397/718/3690/7057/4320/10404/4318/5155/5328/341/5054/5341/1465/712/649/7450/7076/1519/4317/2157/1508/2534/301/8609/4035/2934/7980/2151/3426/1191/344/7044/3920/8237/4312/9104/5641
HALLMARK_COMPLEMENT	HALLMARK_COMPLEMENT	198	0.556889	1.668062	9.72E-10	3.71E-09	1.77E-09	6426	tags=55%, list=29%, signal=39%	716/59345/7077/3075/715/6935/23414/710/2335/4323/1284/4973/3684/5327/718/2207/1514/54331/10404/55619/2219/4855/5155/1012/7941/3937/23533/341/948/5054/714/9732/5341/3988/1465/712/308/4322/2771/7454/1378/5294/2793/7076/1519/4317/2157/5547/1508/88/10019/2625/966/1356/2534/9806/8942/60489/1368/51440/4035/3003/4321/5290/7980/3569/7128/29106/1191/30818/1509/10125/6283/2153/3920/868/5641/6280/1075/5292/3303/11216/6352/5329/23348/5055/1847/51056/959/3717/3697/5251/4057/1051/1512/64145/837/9958/151/2885/200316/2811/3001/1520/3932/5359/3665/9101
HALLMARK_TNFA_SIGNALING_VIA_NFKB	HALLMARK_TNFA_SIGNALING_VIA_NFKB	200	0.549313	1.644711	3.41E-09	1.19E-08	5.68E-09	6146	tags=53%, list=28%, signal=39%	3624/9945/7130/9242/6347/4973/3371/80176/2669/1959/3491/6515/1435/1880/5328/8877/5054/8613/5341/7097/597/5806/3572/23258/3383/11080/19/57007/9021/3604/941/55332/3575/11182/9308/4616/10957/8013/687/604/3433/10365/6446/969/23586/1843/8942/24145/6351/7127/3569/50486/6648/5142/1827/4082/3627/7128/8303/24147/7124/9516/9120/5187/10611/3593/6372/5209/10769/56937/5971/1960/4814/4791/8837/5743/4794/329/6352/1326/4783/7071/5329/4792/1906/4170/5055/64135/7185/6776/330/694/1847/6373/1051/4790/25816/2683/1437/5966/2355/388/10010/8553/5791/1958
HALLMARK_IL2_STAT5_SIGNALING	HALLMARK_IL2_STAT5_SIGNALING	198	0.531887	1.593173	8.21E-08	2.65E-07	1.26E-07	5705	tags=49%, list=26%, signal=37%	2983/23208/84525/8829/1291/4135/126669/942/3685/6335/51284/6515/84230/1435/6696/152007/6403/8477/5179/3587/6095/7293/5745/8611/23258/333/8728/10875/57045/3586/943/6402/57157/80760/3563/659/55509/23710/3604/962/3559/1233/9308/4616/7074/6004/399/58528/4050/11230/5621/5583/8600/974/23075/7052/9173/29851/1493/3560/2151/115361/1893/2012/3627/1522/2323/3662/1028/23413/79026/64375/8784/2623/8530/8320/5292/8809/10538/3596/1326/3480/4783/5732/8743/22807/7185/8202/51429/26018/8835/3566/5033/3482/975/92749/596/1437
HALLMARK_IL6_JAK_STAT3_SIGNALING	HALLMARK_IL6_JAK_STAT3_SIGNALING	86	0.602681	1.740848	1.04E-06	3.12E-06	1.48E-06	5981	tags=59%, list=27%, signal=43%	56034/9180/3554/2/1230/3676/3690/929/7040/1435/23533/948/1439/7097/6354/3572/1441/3563/1438/9021/3559/3953/4050/3162/64109/6773/4283/10148/3569/3627/952/7124/3594/6772/10563/5292/8809/3581/23765/1326/867/6774/5770/6373/3566/1437/3597/355/7132/3454/2885
HALLMARK_ANGIOGENESIS	HALLMARK_ANGIOGENESIS	36	0.709934	1.915457	2.47E-06	6.91E-06	3.29E-06	3575	tags=56%, list=16%, signal=47%	11167/4060/1462/1281/1290/10631/8829/2260/7056/3764/4973/3685/6696/6781/7076/6275/4023/7448/5104/6578

HALLMARK_HYPOXIA	HALLMARK_HYPOXIA	196	0.503921	1.508765	6.25E-06	1.64E-05	7.81E-06	4756	tags=39%, list=21%, signal=31%	6383/1634/633/5507/4015/1289/284119/7043/857/2908/9590/1490/3486/9469/112464/55076/10570/3491/6515/8406/5155/1466/5054/2113/3798/6095/9435/6781/57007/5837/8839/2817/5066/5228/3340/10957/3423/1356/3036/7852/58528/3162/55577/10370/8609/6518/1843/7052/1837/6275/4627/3569/7128/272/2651/2645/1028/6533/51129/7162/56925/5209/7436/5292/8277/538/2632/9672/8459/26136/2131/665/4783/5329/3073/3516/901
HALLMARK_INTERFERON_ALPHA_RESPONSE	HALLMARK_INTERFERON_ALPHA_RESPONSE	97	0.505486	1.469369	0.000634	0.001565	0.000745	6399	tags=51%, list=29%, signal=36%	716/1435/55281/6402/3433/91543/6773/10964/3437/115361/10561/3627/972/2634/3431/10437/219285/94240/103/684/10628/7453/83666/10906/10161/129607/54809/55008/27074/567/4599/1267/64135/9636/1997/55601/5610/51056/6373/3566/11274/54625/85363/9246/3107/55072/5359/51667/3665
HALLMARK_KRAS_SIGNALING_DN	HALLMARK_KRAS_SIGNALING_DN	174	0.468837	1.401312	0.000782	0.001824	0.000868	4924	tags=34%, list=22%, signal=27%	51226/55190/3779/5737/3699/80303/7042/83439/658/57451/8900/5158/1237/70/9435/11346/55220/941/9194/64137/6261/3351/2593/3750/5020/6446/4842/2044/91543/4210/10964/7704/6262/9499/3780/9746/8303/6474/2668/6620/83714/3848/3593/4908/6414/57089/26548/29903/7781/1630/5133/525/29799/1906/5071/8115/4599/55315/5055
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	HALLMARK_UNFOLDED_PROTEIN_RESPONSE	112	-0.28551	-1.37316	0.00118	0.002609	0.001242	4459	tags=37%, list=20%, signal=29%	1977/1938/30827/1973/50628/8570/27230/1662/3376/9797/8140/7422/58477/10897/54512/8602/55272/2521/55466/8815/9775/468/56915/4869/10797/440/6208/1163/79094/28972/29968/1054/5393/5394/1978/3313/55651/9136/23404/25804/3014
HALLMARK_HEDGEHOG_SIGNALING	HALLMARK_HEDGEHOG_SIGNALING	34	0.5953	1.597134	0.002686	0.00564	0.002686	6899	tags=68%, list=31%, signal=47%	7070/2735/8828/8829/1400/23493/7857/6252/23462/8633/4627/3897/8861/4897/7436/4983/6585/7090/5727/1808/5371/1271/43
HALLMARK_APOPTOSIS	HALLMARK_APOPTOSIS	160	0.453251	1.350263	0.003868	0.007736	0.003684	5934	tags=40%, list=27%, signal=29%	4060/1634/5159/9638/7077/633/4313/2149/9890/857/3489/7078/5327/4744/3082/929/1806/8682/7042/552/8900/2878/7076/51176/1540/4616/3162/301/969/5538/90427/2934/3569/2012/6648/952/1191/7124/118/914/6093/5330/1387/4092/10628/1960/8837/5914/665/5551/4170/8743/330/10018/2026/3482/599/7049/6041/837/355/351/356/388
HALLMARK_TGF_BETA_SIGNALING	HALLMARK_TGF_BETA_SIGNALING	54	0.51667	1.447499	0.014997	0.02863	0.013633	5792	tags=50%, list=26%, signal=37%	11031/25937/4053/7057/7040/2022/7046/5054/90/9241/659/6497/1028/7044/51742/4092/56937/64750/2280/7071/4086/7082/9612/6885/387/324/57154
HALLMARK_PEROXISOME	HALLMARK_PEROXISOME	103	-0.2506	-1.25552	0.021401	0.039079	0.018609	4882	tags=40%, list=22%, signal=31%	23417/7153/5825/10455/8915/6342/54884/23457/2224/25824/2172/26986/1022/291/3417/1503/1962/4436/10478/5052/1891/6820/3295/1718/5195/8856/3155/11001/30/60386/6647/390916/8800/1487/3265/2053/373156/3418/3291/706/54982

Table S6. The Hall_mark pathway result of SEPTIN9

ID	Description	setSize	enrich-mentScore	NES	*p*value	p.adjust	qvalues	rank	leading_edge	core_enrichment
HALLMARK_APICAL_JUNC-TION	HALLMARK_APICAL_JUNCTION	197	0.519018	2.5788	1.13E-17	5.64E-16	1.66E-16	3909	tags=41%, list=18%, signal=34%	3728/3691/71/81/10458/3675/208/5097/8189/2886/2962/80381/3757/7410/9344/6714/4902/3384/2771/5819/3914/7450/54509/140885/8449/1495/2318/7283/57863/5880/6624/4763/5310/10290/3265/70/6464/7082/5522/9672/7122/4628/7791/4627/5600/5871/9379/51754/8517/5335/87/4267/999/8751/10174/3383/3918/4597/478/1384/147/6810/1364/2683/3636/947/3667/253559/7106/72/357/57555/24146/10398/794/649/4192/1500/58/92359
HALLMARK_OXIDATIVE_PHOSPHORY-LATION	HALLMARK_OXIDATIVE_PHOSPHORY-LATION	200	-0.50767	-2.45606	8.45E-16	2.11E-14	6.23E-15	4061	tags=43%, list=18%, signal=35%	9377/7385/3052/521/2271/64981/517/4694/529/4701/26521/1743/7417/4724/4702/1892/6391/4552/4329/2395/4714/4719/1349/2108/29796/4729/54704/5250/840/8802/38/1738/56898/10632/64960/4700/3032/23479/6390/1355/2110/51660/4704/7386/4718/56993/4697/36/1350/4190/1345/29088/51318/4709/2746/1737/9550/7419/9167/4712/5264/10440/34/3417/10884/522/51382/81689/539/7381/10935/2230/509/56945/6832/10651/5162/7388/515/4711/4707/6392/26520/54205/1666/4698
HALLMARK_MYOGENESIS	HALLMARK_MYOGEN-ESIS	188	0.501974	2.483213	2.19E-14	3.65E-13	1.08E-13	5446	tags=47%, list=25%, signal=36%	3691/208/4851/3757/3801/115/23654/43/816/5664/10014/4209/8557/375790/1284/4641/7004/5091/7249/1191/6764/2775/10290/3693/2548/1674/537/1292/274/70/6442/4622/9260/9144/4627/8912/3679/6709/25970/4629/10174/1277/2065/3756/624/4625/203/187/1384/5837/9172/7138/6535/140465/845/9215/89797/2314/6261/1728/7135/4303/7040/11155/6840/58/8877/1293/1760/4842/6649/1158/22933/10278/26353/1306/4616/688/351/8497/5740/165/8736/4046/10580/3479/1026/6300/1675
HALLMARK_ESTROGEN_RESPONSE_EARLY	HALLMARK_ESTROGEN_RESPONSE_EARLY	197	0.48168	2.393281	5.07E-14	6.34E-13	1.87E-13	4488	tags=39%, list=20%, signal=31%	9368/5914/2194/9862/9612/124583/7466/3880/8165/60598/6513/3866/115/25987/8416/399665/23030/64699/84187/55806/27134/3480/221/7869/11057/8140/347733/23286/7052/9687/3856/10160/23541/3898/3487/595/3625/5655/26018/60436/916883/23223/1952/6478/11145/10439/10140/29984/55793/267/949/18/25800/6196/8839/2683/3705/9620/79754/9094/89797/5174/56603/23254/64284/4137/794/6840/23329/2296/6542/9145/1717/8645/26353/54020/10514
HALLMARK_EPITHELIAL_MESENCHY-MAL_TRANSI-TION	HALLMARK_EPITHELIAL_MESENCHY-MAL_TRAN-SITION	200	0.470972	2.346123	3.40E-13	3.40E-12	1.00E-12	7613	tags=60%, list=34%, signal=39%	8985/5351/4035/7052/50509/333/1284/3693/9235/5768/2817/1292/5817/2316/3487/3371/290/3690/4017/1294/6385/8572/3915/6586/388/6382/10272/3678/56937/1277/7078/3918/7049/1282/4323/22795/6535/79709/1289/7980/4853/10398/7040/649/64175/1314/5645/5217/1293/6422/22943/5054/6641/4312/667/4314/5159/8325/2669/2006/4616/1278/7128/1893/3569/2192/30008/2199/2335/5118/4176/3673/4313/3909/1490/3725/6372/10516/3576/1647/7857/5329/5654/57692/3688/2026/7422/1281/2014/7076/51330/4016/633/2619/1311/3486/3956/4638/284217/6695/1290/1307/1303/2331/7070/4487/7045/966/871/2201/7474/3491/25878/59/2697/2919/7171/5479/6876

HALLMARK_ MITOTIC_ SPINDLE	HALLMARK_ MITOTIC_ SPINDLE	199	0.438441	2.181648	2.98E-10	2.48E-09	7.32E-10	2734	tags=31%, list=12%, signal=27%	10801/396/23580/1453/6904/22994/9564/81/140735/55201/4296/9344/7461/201176/8440/3619/5829/25777/10300/54509/10426/3996/25/613/57787/6780/6624/4763/10160/23332/64857/274/9371/2316/6711/8243/5347/4628/89941/9826/11190/4627/11346/10844/10163/57580/4650/2017/10435/6709/7204/8874/4926/3797/9700/64411/85464/7430/10013/55722/79658
HALLMARK_ TNFA_SIGNAL-ING_VIA_ NFKB	HALLMARK_ TNFA_SIG-NALING_VIA_ NFKB	200	0.421113	2.097753	1.16E-09	8.32E-09	2.45E-09	6525	tags=50%, list=29%, signal=35%	4088/10318/8878/5606/10938/5970/3949/5971/3914/9334/6776/7127/5209/7262/1051/79155/8303/604/23135/3371/4791/2355/595/23529/4814/6385/11182/388/7280/1052/5187/56937/4790/3164/3383/23645/4794/8660/3976/2683/602/24147/9021/22822/1958/3433/1435/8613/10209/3726/3659/10365/51561/9592/8877/5054/3601/23258/127544/8744/7538/4929/1827/6890/2669/4616/8013/7128/3569/5743/23764/8061/1437/80149/80176/5328/6515/1026/57007/3725/6372/9242/1647/1942/5329/2354/23308/9945/7422/1843/7071/7185/1846/1960/8553/50486/24145/5055/330
HALLMARK_ HYPOXIA	HALLMARK_ HYPOXIA	196	0.426073	2.107693	1.54E-09	9.63E-09	2.84E-09	6359	tags=49%, list=29%, signal=35%	23210/63827/6513/5578/3669/229/5155/3340/3036/2584/1028/126792/54541/5209/226/7052/2548/2817/2645/2355/9672/3069/3484/51129/5211/4627/284119/6385/5214/2997/6478/9590/133/3798/2651/4214/25819/949/2023/10370/23645/8987/8660/6781/9957/5837/8839/23327/4502/9469/5313/3098/9215/3099/1289/1356/112464/302/10570/230/3309/2821/2309/2131/5054/3623/7538/9435/7128/3569/23764/2745/8497/6576/4282/6515/55818/1026/57007/1490/3725/6533/5292/3219/10397/58528/1942/5329/1837/5105/2026/7422/1843/2239/633/8553
HALLMARK_ MYC_TAR-GETS_V1	HALLMARK_ MYC_TAR-GETS_V1	200	-0.42724	-2.06692	3.43E-09	1.90E-08	5.61E-09	5697	tags=47%, list=26%, signal=35%	10549/10528/6175/7514/52/890/11331/23016/6950/4999/3251/1973/5901/57819/7913/6418/10728/10146/9868/26135/7531/1977/5687/7398/9377/6188/3015/6434/5984/26156/6059/10399/4673/10213/4085/51020/6427/708/3336/3475/7411/3066/7298/328/5111/6632/6732/6629/10575/8664/5250/6637/6146/10576/6741/5682/5685/4686/1176/3608/10054/6634/1854/51491/1964/6742/54107/3183/6128/4869/1933/6164/1207/6432/6204/5634/7419/5478/7555/26354/9184/2079/1965/11137/6194/6428/10935/11335/5683/10987/7324/2739/9045/5706
HALLMARK_ KRAS_SIG-NALING_DN	HALLMARK_ KRAS_SIG-NALING_DN	174	0.41822	2.050146	4.27E-08	2.14E-07	6.30E-08	5319	tags=40%, list=24%, signal=31%	5164/6927/6531/3866/56848/152/55890/6712/23031/5795/3785/8303/64137/10991/9328/70/8085/6620/4599/4524/83439/3352/11346/1952/3779/4868/5650/29903/55315/10913/1271/11202/11254/4625/79085/27076/27006/3425/9746/29842/1280/2813/23676/6261/9127/54842/5133/51156/57089/9499/2668/7704/242/4842/6528/2811/1158/2044/7137/4929/9435/3780/91543/51384/2261/5409/5020/80303/525/55220
HALLMARK_ ESTROGEN_ RESPONSE_ LATE	HALLMARK_ ESTROGEN_ RESPONSE_ LATE	197	0.387824	1.926945	1.47E-07	6.70E-07	1.97E-07	5139	tags=37%, list=23%, signal=29%	9368/9612/7466/5104/3993/3880/6768/9231/51181/11187/8416/3669/399665/64699/51599/55806/27134/23650/221/7869/11057/8140/2109/10160/12/3487/595/5655/3418/1952/6478/2852/7177/11145/10439/999/10140/267/949/2030/4057/3918/5364/6196/5265/8839/3705/3547/9215/6280/5174/56603/7980/57834/4137/4192/2707/2296/11012/64377/1717/8645/7538/26353/2810/6509/2261/7054/4665/760/55540/51195/5777

HALLMARK_ INTERFERON_ ALPHA_RE- SPONSE	HALLMARK_ INTERFER- ON_ALPHA_ RESPONSE	97	0.46957	2.095818	1.13E-06	4.73E-06	1.39E-06	6077	tags=53%, list=27%, signal=38%	1267/3959/11054/85441/3430/79132/3669/103/55281/3665/7706/9636/27074/2766/4599/3107/3566/9830/64761/4061/93343/10581/55072/7726/10410/3433/10906/1435/54809/5698/129607/9246/3659/116071/55601/54625/6890/55008/91543/3429/4343/51667/3437/3431/972/5610/4938/83666/23070/1997/7318
HALLMARK_ COAGULATION	HALLMARK_ COAGULA- TION	129	0.398887	1.882163	1.45E-05	5.60E-05	1.65E-05	6772	tags=50%, list=31%, signal=35%	5045/3172/9948/2159/5155/7450/4035/8237/6271/1191/4324/10603/3690/2783/2160/1397/11202/7078/4323/5265/2147/1510/629/1369/11072/335/7980/649/7056/5054/4312/8754/2811/22933/1803/4314/2534/5327/10404/23764/2335/824/5328/3673/4313/1675/1512/4319/2161/4318/7044/2934/5654/9104/726/715/7076/1311/4317/5055/3249/717/718/4224/22921
HALLMARK_ INTERFERON_ GAMMA_RE- SPONSE	HALLMARK_ INTERFER- ON_GAMMA_ RESPONSE	200	0.346172	1.724441	2.15E-05	7.67E-05	2.26E-05	4848	tags=34%, list=22%, signal=27%	3959/11054/85441/3430/57674/6774/9961/5371/79132/3669/81030/5770/103/4940/26524/7127/8651/57169/3665/6398/7706/6892/9636/4599/3566/9830/64761/5214/4600/3663/4061/10581/55072/4790/5690/3383/5699/7726/4502/629/9021/10410/3433/10906/5698/3105/129607/9246/3659/116071/958/10616/3601/55024/55601/54625/6890/1439/6403/8202/7128/55008/3569/5743/10791/91543/3429/3560/4615
HALLMARK_ IL2_STAT5_ SIGNALING	HALLMARK_ IL2_STAT5_ SIGNALING	198	0.34007	1.68899	4.96E-05	0.000165	4.87E-05	5294	tags=36%, list=24%, signal=28%	598/5045/5339/6510/3482/3480/1028/1522/6271/6050/8651/7133/7052/333/4641/896/79971/28960/1291/23413/9144/8784/26018/3566/64866/10075/2118/23191/11230/388/51429/2323/9261/3976/81848/7293/975/4135/4643/22822/10410/943/92749/1435/5727/3099/8530/200734/27242/2184/28996/7371/23258/2623/6403/4616/79026/5745/3596/8202/1893/23764/1437/3560/64375/80324/760/191/6515/898/5583
HALLMARK_ HEDGEHOG_ SIGNALING	HALLMARK_ HEDGEHOG_ SIGNALING	34	0.578517	2.071561	6.69E-05	0.000209	6.16E-05	7163	tags=68%, list=32%, signal=46%	7090/3897/43/5371/9289/4763/6585/8851/4627/1271/9620/5727/6469/1808/8861/7857/1400/23493/7422/7070/5921/8829/7088
HALLMARK_ P53_PATHWAY	HALLMARK_ P53_PATH- WAY	199	0.333343	1.658692	0.00011	0.000324	9.56E-05	4197	tags=30%, list=19%, signal=24%	79651/3691/10458/4851/3732/29927/6768/57799/6256/23654/3654/51181/9883/1969/54541/2872/8651/29948/7421/896/64787/100/1509/3265/23303/1029/3625/6520/9618/4814/8851/10614/55367/2232/6382/11202/10140/91/3872/23645/203/18/10608/3976/7508/975/7039/10906/5900/7135/6236/7040/64782/8877/2309/7704/5784/127544/8744
HALLMARK_ WNT_BETA_ CATENIN_SIG- NALING	HALLMARK_ WNT_BETA_ CATENIN_ SIGNALING	42	0.533442	1.991911	0.000143	0.000397	0.000117	5570	tags=60%, list=25%, signal=45%	9612/4851/5664/8312/10014/6932/5467/2648/4855/9794/85407/7475/7471/28514/23385/5727/79885/81029/22943/10023/8325/3714/1856/27121/8313
HALLMARK_ FATTY_ACID_ METABOLISM	HALLMARK_ FATTY_ACID_ METABOLISM	151	-0.3641	-1.70455	0.000233	0.000613	0.000181	6793	tags=46%, list=31%, signal=32%	5160/10961/5720/39/51478/3320/8803/9588/2326/6611/1371/51703/1376/874/7358/2168/2180/223/1429/364/6281/658/7503/23173/3295/3157/158/10455/3015/3052/2169/7533/2271/1743/51170/1892/6391/328/7390/4683/4128/3033/8802/5092/1738/670/3032/10965/2110/4190/3422/5910/6296/26275/51109/8801/3158/34/58505/3417/131076/84263/5805/594/549/5162/6392/1666/84693

HALLMARK_ COMPLEMENT	HALLMARK_ COMPLE-MENT	198	0.321219	1.595366	0.000344	0.00086	0.000254	5811	tags=38%, list=26%, signal=29%	2885/3172/808/6714/10938/2771/2159/2155/391/5155/4035/1051/1284/3665/30818/1191/4324/1509/4855/978/1608/2783/8851/151/23348/4057/6810/4323/5265/2147/1380/629/7077/51734/9986/6280/1356/5698/7980/5531/29106/3659/11108/3309/2219/2153/5054/308/8754/146547/2811/1803/2534/5327/10404/9806/7128/3569/23764/2625/2335/3684/27445/760/6047/5321/200316/7454/10019/51440/5580/5292/1512/27154/5646/5329
HALLMARK_ E2F_TARGETS	HALLMARK_ E2F_TAR-GETS	200	-0.33468	-1.61913	0.000414	0.000986	0.000291	5893	tags=39%, list=27%, signal=29%	9125/29127/1027/10549/10528/9837/7514/1111/7465/83461/55646/5411/4999/9738/253714/11340/993/1062/5901/7913/10051/4678/204/7374/3070/55148/7398/64858/3015/4361/6434/9183/4673/4085/5931/6427/3364/5111/29980/4683/56655/7884/9232/57405/4292/5885/79075/11168/1633/23594/9833/57122/6119/580/1854/79677/5511/3161/84844/3925/54556/7112/29089/8726/9787/6118/4436/983/5558/1033/5810/11200/55635/1965/10635/84312/1164/25842
HALLMARK_ APICAL_SUR-FACE	HALLMARK_ APICAL_SUR-FACE	43	0.500373	1.895975	0.000525	0.001193	0.000352	7145	tags=65%, list=32%, signal=44%	2319/55959/23054/2050/3315/9696/51754/27076/2683/357/266727/131566/80274/351/6376/22854/3560/2625/142680/4118/672/5329/2619/1946/51458/6517/7070/50617
HALLMARK_ XENOBIOTIC_ METABOLISM	HALLMARK_ XENOBIOTIC_ METABOLISM	189	0.314997	1.561193	0.000746	0.001622	0.000478	3877	tags=26%, list=18%, signal=22%	3728/4836/9564/53/26007/5831/3172/10237/6510/51181/9361/2159/54498/4355/488/1969/5467/5447/6560/7448/5091/1543/55748/978/3487/65018/4814/3484/3083/7132/2160/10653/2651/50/2646/1593/1244/5699/4502/629/8310/1586/1728/6539/5909/4864/8714/51/2184
HALLMARK_ INFLAMMA-TORY_RE-SPONSE	HALLMARK_ INFLAMMA-TORY_RE-SPONSE	197	0.297101	1.476177	0.001398	0.002913	0.000859	6846	tags=41%, list=31%, signal=29%	3732/5970/11047/3949/391/488/6541/7133/3665/6892/5817/27074/64127/3690/3566/6367/10507/133/5739/4061/91/3678/4790/3383/187/3976/4323/23095/1440/5724/1435/3659/3656/958/23166/8877/5054/3601/6542/7439/8744/2811/558/9435/13 1566/3569/6376/4049/3560/7432/1902/6869/623/1441/1026/6372/3576/1236/7855/5008/5329/3759/8578/5610/23308/4210/6401/2014/7076/9180/2769/970/7098/8685/3249/728/4224/8671/6361/10630/6324
HALLMARK_ ANGIOGEN-ESIS	HALLMARK_ ANGIOGEN-ESIS	36	0.505432	1.83062	0.001613	0.003225	0.000951	7155	tags=61%, list=32%, signal=41%	5104/7410/7448/4043/6781/2260/27242/7056/3714/351/6578/5154/6372/7422/1281/7076/350/5553/1290/4487/8993/8829
HALL-MARK_UV_RE-SPONSE_UP	HALLMARK_ UV_RE-SPONSE_UP	155	0.305752	1.464019	0.003799	0.007306	0.002153	5040	tags=32%, list=23%, signal=25%	1267/10948/2011/5045/3757/8878/23542/58525/1028/2907/226/11333/23467/896/1543/9392/4245/1742/388/10494/10439/8817/10963/3164/3383/1212/10652/5566/4323/409/1181/6535/1280/8692/6539/3726/3659/6236/1175/682/6890/3134/3569/997/4885/760/27161/10018/6257
HALLMARK_ GLYCOLYSIS	HALLMARK_ GLYCOLYSIS	196	0.293096	1.449883	0.004047	0.007494	0.002209	3768	tags=28%, list=17%, signal=23%	64132/9862/9123/2539/79586/54480/3669/10331/229/5351/26229/5315/5834/2584/126792/54541/226/6662/375790/1468/5091/5768/2817/9672/5704/3069/51129/5214/8912/2997/55501/6382/9514/2023/54982/10370/29926/8660/6781/56052/8534/2132/2683/3425/29925/90161/7039/4351/3099/1289/1365/10570/3309/4191/10616
HALLMARK_ PI3K_AKT_ MTOR_SIG-NALING	HALLMARK_ PI3K_AKT_ MTOR_SIG-NALING	104	0.353287	1.592927	0.004332	0.007736	0.00228	2318	tags=22%, list=10%, signal=20%	396/57521/2885/10298/1845/156/6513/8878/5606/4793/84335/1213/2872/31/7249/7186/3265/1072/7132/1869/5335/811/5300

HALLMARK_IL6_JAK_STAT3_SIGNALING	HALLMARK_IL6_JAK_STAT3_SIGNALING	86	0.354368	1.566838	0.005283	0.009109	0.002685	5718	tags=36%, list=26%, signal=27%	2885/6774/5770/8651/7133/23765/3690/7132/3566/5320/7297/1271/91/4055/9021/1435/3659/7040/27242/3601/1439/867/3569/1437/4615/55540/1441/3725/5781/5292/5967
HALLMARK_APOPTOSIS	HALLMARK_APOPTOSIS	160	0.293868	1.419971	0.007157	0.011929	0.003516	6439	tags=40%, list=29%, signal=29%	5914/598/2064/1387/8567/3482/8878/5970/5664/118/3669/4000/10134/1191/5611/595/5551/89941/3315/6709/388/2232/2065/7078/4092/7049/2147/994/1611/4744/7077/10410/3659/3000/5159/5526/6890/4616/5327/3569/351/3489/10018/4313/1026/3725/6574/672/1647/581/6304/2934/3875/2026/7076/4739/1960/51330/633/1499/7153/552/4255/2769
HALLMARK_TGF_BETA_SIGNALING	HALLMARK_TGF_BETA_SIGNALING	54	0.406766	1.634639	0.009214	0.014405	0.004246	6480	tags=57%, list=29%, signal=41%	9612/4088/5045/1028/6497/7082/6711/2022/57154/999/56937/90/4092/23645/4091/4086/3726/7040/28996/5054/9241/1025/64750/6574/3397/5499/7044/7071/25937/1499/4053
HALLMARK_NOTCH_SIGNALING	HALLMARK_NOTCH_SIGNALING	32	0.469852	1.652034	0.009219	0.014405	0.004246	3552	tags=34%, list=16%, signal=29%	4851/5664/5578/5467/4854/2648/595/23291/113878/28514/4853
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	HALLMARK_UNFOLDED_PROTEIN_RESPONSE	112	0.324302	1.496083	0.010139	0.015363	0.004528	2271	tags=22%, list=10%, signal=20%	7466/9570/2081/1639/55738/1981/10525/23644/54541/6734/7709/8140/1051/5611/8570/6464/8602/9114/3484/1938/30827/9775/79094/811/7280
HALLMARK_UV_RESPONSE_DN	HALLMARK_UV_RESPONSE_DN	144	0.299496	1.423573	0.018578	0.02732	0.008052	6564	tags=42%, list=30%, signal=30%	2064/4088/7410/6256/57178/5578/3949/6541/3480/5002/2869/4363/7082/3690/2535/5195/3915/1628/50937/8871/4790/1277/5293/10370/4092/3488/7049/8405/3667/9112/8613/302/773/4853/11030/5054/2534/5159/861/1278/4131/8476/5581/5797/1902/273/493/10516/3397/1281/1843/5607/8455/11044/8553/7035/4255/22884/10395/10401/11343

Table S7. The Hall_mark pathway result of VIM

ID	Description	setSize	enrichmentScore	NES	*p*value	p.adjust	qvalues	rank	leading_edge	core_enrichment
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	200	0.766387	2.281349	2.01E-46	8.46E-45	3.82E-45	3665	tags=74%, list=17%, signal=63%	7431/7424/6678/11167/5159/2200/2191/1292/1281/6695/1290/1634/10979/633/800/1278/7058/1293/7070/4060/1462/30008/3678/8038/4837/4313/1296/115908/5396/5654/6591/2619/26585/1009/1277/59/10516/1842/6876/10631/3624/1289/7412/2199/4015/8076/50863/6444/4256/5118/10409/1000/6586/25878/10398/4323/1301/2335/22795/3956/1303/2014/1284/2192/1809/4638/7456/2697/4017/1490/6424/5376/9353/8572/2316/3371/5744/7040/4016/1282/1311/6422/3491/7078/5054/131578/2669/5999/64175/6387/7057/10085/3690/50509/6696/7169/871/2303/4148/3685/1307/333/2006/11010/3486/649/7857/10272/3915/3908/25890/2247/2882/23705/813/5806/10486/7076/1004/4616/6445/3688/4853/3487/2331/2817/3693/8325/7052/627/7171/1294/7128/3569/966/4035/5351/2201/284217/7980/1893/26577/7168/7474/57692/667/4312/6443/5329

HALLMARK_INFLAMMATORY_RESPONSE	HALLMARK_INFLAMMATORY_RESPONSE	197	0.649146	1.929984	3.42E-21	7.18E-20	3.24E-20	5991	tags=68%, list=27%, signal=50%	558/3678/3554/9180/3624/5739/7130/4323/10630/5027/5099/4973/719/2014/728/929/4481/10203/6347/1435/23533/23601/187/1240/1536/8877/3937/3587/8685/5054/366/8578/3383/3690/1880/6324/23166/2357/7096/114548/1441/7097/9435/6401/5724/6354/6402/5996/3604/3586/962/7076/19/1318/6869/4049/3560/6004/3575/969/5008/10316/6504/4210/2015/4283/3738/29933/255488/7439/970/10148/6367/958/1902/1236/6542/3627/3091/5142/10125/6352/3569/6376/7432/6361/684/10663/8809/3593/5329/7162/131566/6506/8807/623/3732/27074/3576/3932/4792/5732/6373/60675/391/3566/3037/3269/6372/3665/10886/2867/2811/166929/8743/7133/3656/5791/4693/4790/2769/196/5610/1316/3600/3696/1906/3553/3659/4061/64127/5970/4067/3454
HALLMARK_MYOGENESIS	HALLMARK_MYOGENESIS	188	0.639044	1.897752	3.12E-19	4.36E-18	1.97E-18	4035	tags=51%, list=18%, signal=42%	22808/6678/165/1292/1281/1293/8038/10278/1306/1277/6876/6444/26353/1284/4638/3490/5740/2702/25937/4046/3316/7040/187/1410/8082/5507/8877/1346/10468/2273/9260/4208/2775/2878/7169/948/3479/3486/4629/8912/3908/2247/3801/845/29970/6442/1012/7136/1489/6261/4616/6445/5837/70/3688/8736/11155/10290/6840/3679/2934/88/3693/1674/347/58529/4627/4684/5816/26287/5348/10580/6764/844/7140/5621/1191/4205/6588/6300/4608/4209/1266/115/4635/4634/2170/4842/1837/3270/89/3756/9659/6016/816
HALLMARK_ALLOGRAFT_REJECTION	HALLMARK_ALLOGRAFT_REJECTION	197	0.633317	1.882923	1.04E-18	1.09E-17	4.92E-18	4591	tags=59%, list=21%, signal=47%	586/7070/3624/2213/2268/6688/3689/9450/942/6347/1435/3059/2316/1230/7040/5552/920/2149/3937/7903/7454/2113/4318/3117/3383/3111/23166/7096/114548/10261/322/7097/5579/729230/8477/5788/3603/6354/3824/3683/2533/1234/3586/3109/6357/940/3559/10333/3122/7076/941/7042/4050/3560/6351/9734/3702/6363/6356/56253/8444/972/4283/5199/11184/2634/6367/958/916/3108/3594/8560/912/10225/3455/914/3091/6775/3662/6352/10563/3569/6772/917/919/9655/50852/973/27240/924/5551/7453/1991/7124/3112/925/915/3593/3133/4689/8807/3848/9437/7535/3932/3001/356/3596/3717/3566/3625/567/9607/3665/959/3458
HALLMARK_INTERFERON_GAMMA_RESPONSE	HALLMARK_INTERFERON_GAMMA_RESPONSE	200	0.623138	1.85493	6.55E-18	5.50E-17	2.48E-17	5763	tags=64%, list=26%, signal=48%	716/715/3075/710/7412/7130/2209/942/6347/1240/8082/3937/3587/7903/1439/3117/3383/2357/6403/26524/9934/10068/6354/10791/9021/10875/55016/25939/6773/3560/969/3123/7127/163351/3434/972/115361/3433/4283/57823/29126/23586/10964/958/91543/3437/4600/3108/6648/3627/952/219285/3455/3091/6775/5142/3662/669/3431/7128/4939/10437/54739/6352/4261/3569/84159/6772/10561/3620/7453/8869/4502/684/5292/103/10628/10906/6398/3695/4599/129607/6375/9246/3001/4792/54625/9636/5371/94240/51056/6373/3717/81030/55601/5743/3566/567/694/11274/57169/64135/6774/3663/55008/3665/317649/57674/4940/8743/2841/55072/3669/10392/355/4790/4615/8737/55024/5777/6890/5610/22914/3600/837/80830/3659/4061/3106
HALLMARK_MYC_TARGETS_V2	HALLMARK_MYC_TARGETS_V2	58	-0.70109	-3.45075	5.83E-16	4.08E-15	1.84E-15	3678	tags=71%, list=17%, signal=59%	4173/11335/1019/10199/23160/79077/7374/23481/6652/5347/10733/10171/6723/705/6573/4839/54663/64216/4869/7965/10196/56915/56342/92856/6832/10244/26354/51154/79050/3329/27346/9136/5036/79711/3177/3336/9238/80324/5245/29078/51491

HALLMARK_KRAS_SIGNALING_UP	HALLMARK_KRAS_SIGNALING_UP	197	0.608945	1.810463	2.44E-15	1.47E-14	6.61E-15	4084	tags=44%, list=18%, signal=36%	9945/3075/9935/2995/10457/8829/5396/5175/2982/3624/8404/7805/25960/23643/5178/3689/2791/719/80380/64123/2324/2207/3290/9358/9590/1240/2022/1794/951/3587/3936/2113/4318/51311/6324/6696/5327/50856/5328/7462/2162/83483/51704/220/3486/101/10320/6616/5649/4674/55273/7852/56729/6004/4803/3575/9734/4320/639/9892/10418/1438/23136/28951/11184/2115/7035/3627/347/669/7128/3481/3800/50486/51129/27299/2012/29106/51199/8854/5329/131566/7185/55107/1363/330/25907
HALLMARK_COMPLEMENT	HALLMARK_COMPLEMENT	198	0.602762	1.792737	3.28E-15	1.72E-14	7.77E-15	5164	tags=51%, list=23%, signal=39%	716/715/59345/3075/710/7077/23414/3684/4323/6935/2335/4973/1284/2207/718/23533/1514/2219/55619/3937/5054/54331/712/7454/5155/4855/714/5341/7941/2771/5327/341/1378/948/2793/9732/10404/5294/1508/10019/7076/1012/2625/2534/1465/9806/308/3988/60489/3003/1356/4317/1519/4322/8942/88/1509/2157/7128/4321/5547/10125/6352/3569/51440/966/4035/1368/1191/5641/7980/5290/29106/2153/5292/6283/5329/868/6280/3932/3001/1847/3697/51056/3717/30818/391/5055/3665/959/2811/1075/1380/3303/4057/23348/151/2885/11216/10938/200316
HALLMARK_TNFA_SIGNALING_VIA_NFKB	HALLMARK_TNFA_SIGNALING_VIA_NFKB	200	0.584685	1.740468	1.37E-13	6.40E-13	2.89E-13	5917	tags=57%, list=27%, signal=43%	9945/9242/3624/7130/4973/6347/1435/3371/1959/8877/6515/80176/3491/5054/2669/5341/3383/1880/597/5328/7097/9021/3604/3572/5806/23258/604/11182/8613/941/19/4616/11080/3575/969/55332/9308/6351/687/8013/7127/10365/6446/3433/8942/57007/23586/6648/3627/1843/10957/5187/5142/7128/6352/3569/50486/4791/24145/8303/1827/7124/9120/5971/5209/9516/10769/1960/3593/24147/5329/4794/10611/7185/4814/8837/6776/330/4792/2355/4170/1847/4082/6373/5743/4783/1326/25816/694/64135/6372/5055/23135/2354/2683/10318/5791/10938/4790/23529/8553/1437/7071/5271/1958/388/6890/1316/1906/3553/3659/329/5966/5970/56937
HALLMARK_APICAL_JUNCTION	HALLMARK_APICAL_JUNCTION	197	0.59154	1.758714	3.16E-13	1.33E-12	5.99E-13	5297	tags=48%, list=24%, signal=37%	83700/2200/143903/7070/1462/10000/4313/5175/51148/1009/4478/7216/7412/91624/10398/1272/8910/942/7106/947/9459/9353/8745/51466/11096/57555/4318/3383/7450/2771/7122/4628/57863/5788/257194/1307/84552/649/2533/5600/8515/247/140885/5010/1004/87/30835/794/72/70/3688/7414/10290/2318/6624/29126/55742/88/2593/80381/7059/6464/4627/3667/347902/6376/3384/7087/253559/54413/5880/5310/9672/89/7185/23114/7781/1006/3680/7791/1758/9379/1013/60/7082/10109/2683/147/8503/10487/51208/4267/2770/478/8517
HALLMARK_UV_RESPONSE_DN	HALLMARK_UV_RESPONSE_DN	144	0.606756	1.788003	3.80E-11	1.45E-10	6.54E-11	5168	tags=51%, list=23%, signal=40%	5159/6383/1281/1290/1278/10000/8829/6591/1277/10516/23266/2202/1301/857/3488/5797/273/2697/2908/5376/4131/10395/4325/23345/3491/5054/5999/5737/3690/4052/493/5293/3915/3778/29970/10486/8613/2534/4853/10370/5577/6310/1902/7035/2535/1843/627/861/3667/10401/57178/2869/773/50937/11099/9863/10611/3815/140609/323/3037/4092/7048/11030/7799/3480/1601/8455/7082/7049/8503/4790/7436/8553

HALLMARK_IL2_STAT5_SIGNALING	HALLMARK_IL2_STAT5_SIGNALING	198	0.551082	1.63903	7.39E-10	2.59E-09	1.17E-09	5551	tags=52%, list=25%, signal=39%	23208/8829/2983/1291/84525/126669/942/1435/84230/6515/3587/4135/152007/943/7293/6403/51284/6696/8477/3685/333/6402/5745/6335/55509/3604/3586/962/10875/3559/399/23258/6095/23710/4616/7074/4050/3560/5179/8728/6004/57045/3563/9308/974/1233/5583/1493/115361/8611/29851/8784/80760/23075/57157/2323/3627/7052/659/3662/58528/9173/11230/8530/1522/8320/79026/5621/1893/2012/64375/5292/8809/23413/10538/7185/2623/8600/1028/5732/22807/3596/596/4783/3566/1326/3480/2151/8743/9261/7133/975/8553/1437/4641/26018/5339/196/388/51429/1316/8835
HALLMARK_COAGULATION	HALLMARK_COAGULATION	129	0.594674	1.74813	9.65E-10	3.12E-09	1.41E-09	3045	tags=36%, list=14%, signal=32%	6678/716/2200/715/3075/710/4313/5654/5175/1513/2/4323/2335/4973/7056/718/1397/1311/7078/5054/712/5155/4318/5341/7057/3690/7450/5327/341/5328/649/10404/1508/7076/2534/1465/4320/4317/1519/301/2934/2157/4035/1191/5641/7980/3426
HALLMARK_IL6_JAK_STAT3_SIGNALING	HALLMARK_IL6_JAK_STAT3_SIGNALING	86	0.624047	1.813628	4.38E-08	1.31E-07	5.93E-08	5991	tags=64%, list=27%, signal=47%	3554/9180/2/56034/929/1435/23533/1230/7040/1439/3690/1441/3676/948/7097/6354/9021/3572/3559/4050/6773/3563/1438/3162/4283/10148/3627/3594/952/64109/10563/3569/6772/3581/3953/7124/5292/8809/23765/6373/3566/1326/867/6774/7132/7133/2885/355/1437/4615/3553/3659/1271/5770/3454
HALLMARK_HYPOXIA	HALLMARK_HYPOXIA	196	0.529267	1.573523	2.72E-07	7.61E-07	3.43E-07	5839	tags=47%, list=26%, signal=35%	6383/1634/284119/633/1289/4015/7043/857/2908/1490/9590/5507/6515/3491/5054/5155/2113/9469/55076/10570/112464/8406/9435/3486/6781/1466/3798/6095/8839/7852/6518/5837/3036/1356/3340/3162/10370/55577/57007/5228/2817/5066/1843/7052/10957/4627/7128/58528/3569/3423/51129/2651/4502/6275/272/5209/8609/1837/5292/5329/7162/9672/2026/2645/56925/2355/1028/3073/596/4783/694/665/8277/3516/2131/2632/3669/901/6533/7436/8509/8553/8459/4214/1316/3098/5260/8497/25819/2548/3623/9215
HALLMARK_INTERFERON_ALPHA_RESPONSE	HALLMARK_INTERFERON_ALPHA_RESPONSE	97	0.57886	1.690966	1.07E-06	2.82E-06	1.27E-06	6985	tags=65%, list=32%, signal=45%	716/1435/55281/6402/6773/972/115361/3433/10964/2634/91543/3437/3627/219285/3431/10437/10561/7453/684/83666/103/10628/10906/54809/4599/27074/129607/9246/54625/9636/94240/51056/6373/55601/3566/567/1267/11274/64135/3107/55008/3665/55072/3669/10161/51667/6890/5610/55603/3600/3659/4061/85363/79132/3660/64108/3430/5698/85441/64761/6737/10379/10581
HALLMARK_ANGIOGENESIS	HALLMARK_ANGIOGENESIS	36	0.705099	1.887372	3.55E-06	8.77E-06	3.96E-06	3504	tags=56%, list=16%, signal=47%	11167/1281/1290/4060/1462/8829/10631/2260/4973/7056/3764/6696/3685/6781/7076/7448/6275/6578/5104/4023
HALLMARK_APOPTOSIS	HALLMARK_APOPTOSIS	160	0.496714	1.471499	6.10E-05	0.000142	6.42E-05	5911	tags=45%, list=27%, signal=33%	5159/1634/7077/633/4060/4313/9638/857/929/3489/1806/4744/9890/2149/552/7078/5327/2878/3082/8682/7076/8900/7042/4616/1540/969/3162/301/2934/6648/952/914/3569/5330/118/51176/90427/5551/1191/7124/5538/2012/1960/6093/10628/8837/2026/10018/1387/330/4170/356/4092/665/5914/8743/3669/7049/355/599/2769/6709/388/6890/5526/837/3553/3659/3482/4739/6574/5970

HALLMARK_KRAS_SIGNALING_DN	HALLMARK_KRAS_SIGNALING_DN	174	0.491152	1.45588	9.97E-05	0.00022	9.94E-05	6051	tags=41%, list=27%, signal=30%	51226/55190/3779/5737/5158/80303/9435/1237/3699/83439/11346/658/941/8900/7042/6261/9194/70/57451/64137/55220/4210/6446/3351/3750/2593/7704/10964/91543/3780/9746/5133/2044/8303/6474/6620/5020/9499/4842/6262/3593/8115/3848/7781/4599/2668/6559/83714/57089/26548/525/4908/5055/959/2811/3458/1630/29903/3852/29799/6414/54855/51384/10734/55315/1906/1271/50489/5795/8085/6531
HALLMARK_HEDGEHOG_SIGNALING	HALLMARK_HEDGEHOG_SIGNALING	34	0.620253	1.648788	0.001278	0.002622	0.001183	5769	tags=62%, list=26%, signal=46%	7070/8829/8828/2735/1400/6252/7857/23493/23462/8633/4627/3897/8861/4897/5371/7090/1808/6585/7436/43/1271
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	HALLMARK_UNFOLDED_PROTEIN_RESPONSE	112	-0.26551	-1.47101	0.001311	0.002622	0.001183	3311	tags=28%, list=15%, signal=24%	1736/9775/10897/51013/8602/54512/3376/468/5394/55272/1977/8815/29968/58477/1662/27230/440/5393/1978/4869/1163/6208/56915/1054/28972/3014/9136/25804/3313/23404/55651
HALLMARK_PEROXISOME	HALLMARK_PEROXISOME	103	-0.26714	-1.27646	0.004349	0.008302	0.003745	4533	tags=41%, list=20%, signal=33%	1891/223/8443/10005/7153/23417/2224/5052/3265/1718/1962/1487/291/5194/6342/4436/51/10478/6820/3155/60386/5825/3422/51703/26986/30/3417/6647/11001/54982/2172/25824/1022/3295/706/3418/390916/8856/8800/373156/2053/3291
HALLMARK_TGF_BETA_SIGNALING	HALLMARK_TGF_BETA_SIGNALING	54	0.531099	1.489911	0.00692	0.012637	0.005701	6845	tags=61%, list=31%, signal=42%	11031/4053/25937/7040/2022/5054/7057/9241/7046/90/6497/659/1028/7044/4092/51742/9612/4086/7082/7071/324/387/64750/2280/57154/6574/56937/4088/5045/6885/8412/3726/6711
HALLMARK_APICAL_SURFACE	HALLMARK_APICAL_SURFACE	43	0.524472	1.436085	0.022893	0.040062	0.018073	6500	tags=49%, list=29%, signal=35%	7070/2619/266727/8406/2625/3560/4118/22854/55959/6376/5329/131566/51738/2683/1946/146760/84632/4067/27076/3315/11126

Table S8. The information of the Primers and probes

Target Genes		Sequence (5'-3')
SEPT9	Primer F	CGTTGTTTATTAGTTATTATGTCG
	Primer R	AATAATCCCATCCAACTACG
	Probe	CAAACAAACATTCAAATATCAATCTAGTTATTATGTCGGATTTCG
SDC2	Primer F	AAATTGAATTTCGGTACGG
	Primer R	CACTTATTAATTTCTACACTCCC
	Probe	CATCTTCATATCAATTCTCTTATTTCGGTACGGGAAAGGAGTT
VIM	Primer F	ATGTTACGCGTTTTTTTGTC
	Primer R	TCCTACAACTCCACCTTCTC
	Probe	TACAACATCTCATTAACATATACATACGCGTTTTTTTGTCGTG
ABL	Primer F	GGGATGGAAAGAATTTTATGT
	Primer R	GTGACAAACTAACCCCAAATTCTCT
	Probe	CTATCATTTATCTCTTTCTCAATTGAAAGAATTTTATGTTTGGTGA

Note: Primer F = forward primer; Primer R = reverse primer; Probe = second-round primer.