Pancreatic Cancer Cell Fraction Estimation in a DNA Sample

Pancreatic cancer cell fraction estimation in a DNA sample

Hiroki Ishihara[1, 2], Satoshi Yamashita[1], Ryosuke Amano[3], Kenjiro Kimura[3], Kosei Hirakawa[3], Takako Ueda[4], Yoshiki Murakami[4], Akihiro Tamori[4], Kazunari Tanabe[2], Norifumi Kawada[4], Atsushi Hagihara[4], Toshikazu Ushijima[1*]

[1]Division of Epigenomics, National Cancer Center Research Institute, Tokyo, Japan

[2]Department of Urology, Tokyo Women's Medical University, Tokyo, Japan

[3]Department of Surgical Oncology, Osaka City University Graduate School of Medicine, Osaka, Japan

[4]Department of Hepatology, Osaka City University Graduate School of Medicine, Osaka, Japan

Running head: Pancreatic cancer cell fraction marker

*Corresponding author: Toshikazu Ushijima, Division of Epigenomics, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan, Tel: +81 3 3547 5240, Fax: +81 3 5565 1753, E-mail address: tushijim@ncc.go.jp

Keywords:

DNA methylation; pancreatic cancer; cancer cell fraction; epigenetics; IPMN

Abstract

Objective: Pancreatic cancers are characterized by dense stroma. To estimate the degree of interference by co-existing non-cancerous cells in molecular analyses, we aimed to develop a DNA methylation marker that assesses a cancer cell fraction in DNA samples.

Methods: Microarray data of 22 pancreatic cancer tissues from TCGA and nine non-cancerous tissues were used for genome-wide screening. 31 surgical tumor samples (10 intraductal papillary mucinous neoplasms [IPMNs] and 21 pancreatic cancers), 4 normal, and 26 non-tumorous samples were used for validation. Gene-specific methylation analysis was conducted by bisulfite pyrosequencing.

Results: Genome-wide screening isolated *SIM1*, *MIR129-2*, *NR1I2*, and *HOXB-AS4*, as specifically methylated in pancreatic cancer cells. Bisulfite pyrosequencing validated that one or more of three genes (*SIM1*, *MIR129-2*, and *NR1I2*) were methylated in 22 (71.0%) tumor samples (8 IPMNs and 14 cancers), and all showed low levels of methylation in 26 (86.7%) normal and non-tumorous samples. Therefore, the three genes collectively constituted one marker for a pancreatic cancer cell fraction. The cancer cell fraction estimated by the marker was highly correlated with that estimated using the *KRAS* mutant allele frequency ($R = 0.79$).

Conclusion: The DNA methylation marker is useful to estimate the pancreatic cancer cell fraction in DNA samples.

Introduction

Pancreatic cancer is a highly lethal disease with few modalities for early diagnosis and few effective systematic therapies for advanced cases [1-4]. To develop effective diagnostic and therapeutic strategies, a deeper understanding of the underlying molecular pathophysiology is needed. What hampers this understanding is the large stromal component comprising the bulk of the tumor mass while pancreatic epithelial cells are a minor constituent of the mass [5]. The co-existing normal cells can seriously affect the results of molecular analyses of pancreatic cancers, such as next-generation sequencing and gene expression analyses [6-8]. Thus, accurate estimation of the cancer cell fraction in pancreatic tumor samples is a critical step in the molecular analysis of pancreatic cancers.

To estimate the cancer cell fraction in DNA samples, we previously established a unique method using cancer cell-specific DNA methylation in esophageal squamous cell carcinomas and gastric cancers [9, 10]. The method utilizes a panel of a limited number of genes, typically three, specifically methylated in cancer cells and unmethylated in non-cancerous cells, such as normal epithelial cells, fibroblasts, and blood cells. Therefore, the methylation levels of the genes reflect the cancer cell fraction in DNA samples [9]. Since DNA methylation can be analyzed in DNA samples, one of the advantages of using a DNA methylation marker is that tissue sections or frozen cells are not needed. At the same time, DNA methylation patterns are specific to individual cell types and thus to individual tissues [11-14], and a panel of genes specific to pancreatic cancers must therefore be isolated.

In this study, we aimed to establish a cancer cell fraction marker based on DNA

methylation profiles in pancreatic cancers.

Materials and Methods

Pancreatic surgical specimens, blood samples, and pancreatic cancer cell lines

A total of 60 pancreatic surgical specimens were obtained at Osaka City University Hospital from 31 patients (eight intraductal papillary mucinous neoplasm [IPMN] patients, 20 pancreatic cancer patients, and three patients with benign pancreatic disorders) (Supplementary Table 1). From the 20 patients with pancreatic cancer, 21 tumor samples and 19 non-tumorous samples were obtained. From the eight patients with IPMNs, 10 tumor samples and seven non-tumorous samples were obtained. From the three patients with benign disorders, three normal pancreatic tissue samples were obtained. In addition, genomic DNA from one normal pancreatic tissue from a healthy 33-year-old male (D1234188) was purchased (Biochain, Hayward, CA, USA). In total, four normal samples (normal tissues from non-cancer individuals), 26 non-tumorous samples (normal tissues from cancer and IPMN patients), and 31 tumor samples (10 IPMNs and 21 pancreatic cancers) were used as pancreatic specimens in this study. Eighteen patients (56.3%) were male, and the mean age was 65.3 years (range: 33-87). Normal, non-tumorous, and tumor samples were macroscopically harvested from the surgical specimens, and stored at -80 °C. The present study was conducted with the approval of the Institutional Review Board of the National Cancer Center and Osaka City University Hospital. Written informed consent was obtained from all participants.

Four peripheral leucocyte samples were obtained from four heathy volunteers. Seven human pancreatic cancer cell lines (Capan-2, HPAC, HPAF-II, Hs 766T, MIA PaCa-2, PANC-1, and BxPC-3) were purchased from the American Type Culture Collection (Manassas, VA, USA). Genomic DNA was extracted by the phenol-chloroform and

ethanol precipitation method.

Genome-wide DNA methylation analysis

Genome-wide DNA methylation analysis was performed using an Infinium HumanMehylation450 BeadChip array (Illumina, San Diego, CA, USA) as previously reported [15]. Additionally, we downloaded HumanMethylation450 data of 22 pancreatic cancerous tissues and nine non-cancerous tissues, which were randomly selected from 184 pancreatic tissue samples available in the TCGA database (Supplementary Table 2).

The DNA methylation level of an individual probe was obtained as a β value that ranged from 0 (unmethylated) to 1 (fully methylated). From all of the 482,421 CpG sites, we excluded 11,421 CpG sites on the sex chromosomes and 5,077 CpG sites at genomic positions that could not be specified according to the human genome assembly hg 38. The remaining 465,923 CpG sites on autosomes were evaluated in this analysis.

Measurement of DNA methylation levels of specific genomic regions

Gene-specific DNA methylation levels were measured by bisulfite pyrosequencing. Bisulfite modification was performed using 1 μg of *Bam*HI-digested genomic DNA as previously reported [16]. The modified DNA was suspended in 40 μl of TE buffer, and an aliquot of 1μl was used for bisulfite pyrosequencing. A target region was amplified by biotinylated primers, as previously reported [17]. The PCR product was annealed to a 0.2 μM pyrosequencing primer. Pyrosequencing was carried out using the PSQ 96 Pyrosequencing System (QIAGEN, Valencia, CA, USA), and a methylation level was obtained using PSQ Assay Design software (QIAGEN).

Somatic mutant allele frequency of *KRAS* and other cancer-related genes

Mutation status of *KRAS* and other cancer-related genes were analyzed by targeted deep sequencing using an Ion Proton next-generation sequencer (Thermo Fischer Scientific, Waltham, MA, USA) and a panel of genes (291 regions of 55 cancer-related genes, covering 48,005 base positions) [18]. A sequence library was prepared by multiplex PCR using Ion AmpliSeq Library Kits 2.0 (Thermo Fischer Scientific). A sequence variant was considered as a somatic mutation when the variant allele frequency was ≥ 3.0% and the variant was found in at least two forward and two reverse reads using CLC Genomics Workbench (QIAGEN). Additionally, only single base substitution mutations were analyzed, and deletions and insertions were neglected. A *KRAS* mutation was confirmed by conventional Sanger sequencing using primer sets for exon 2 of *KRAS* [19].

To estimate the cancer cell fraction, the *KRAS* mutant allele frequency was doubled because a *KRAS* mutation is a gain-of-function mutation and is present in one of the two alleles when copy number alterations are absent [20-23].

Statistical analysis

The correlation analysis was performed using the Pearson’s product-moment correlation coefficients. The analyses were performed using PASW statistics version 18.0 (SPSS Japan Inc., Tokyo, Japan).

Results

Isolation of genomic regions specifically methylated in pancreatic cancer cells

Genomic regions specifically methylated in pancreatic cancer cells but unmethylated in non-cancerous cells, including normal pancreatic epithelial cells and stromal cells, were isolated. Specifically, genome-wide DNA methylation data of peripheral leucocyte samples and cell lines were obtained by their beadarray analysis with Infinium Human Methylation450, and those of cancerous and non-cancerous pancreatic tissues were extracted from The Cancer Genome Atlas (TCGA) database (Supplementary Table 2). First, from probes located at 465,923 CpG sites on autosomes, we selected 143,241 CpG sites unmethylated (β value ≤ 0.2) in four peripheral leucocyte samples from four healthy volunteers and in nine non-cancerous tissues from nine patients (Figure 1).

We further selected 799 CpG sites methylated in six or more of seven pancreatic cancer cell lines ($\beta \geq 0.8$) and in 14 or more of 22 pancreatic cancer tissues ($\beta \geq 0.3$). To isolate genomic regions that were highly likely to be specifically methylated, we selected those with three or more flanking probes of consistent values, as previously reported [24, 25], and finally isolated 24 genes corresponding to 29 genomic regions. Alternatively, we isolated genomic regions with a higher incidence of methylation but with less chance of validation. From the 143,241 CpG sites, we selected 281 CpG methylated sites in all of the seven cell lines ($\beta \geq 0.8$) and in 16 or more of the 22 cancer tissues ($\beta \geq 0.3$). To isolate genomic regions likely to be specifically methylated, we selected those with two or more flanking probes of consistent values, and isolated 27 genes corresponding to 31 genomic regions. Nine genes were overlapping in the two algorithms, and 42 genes were finally isolated as candidate marker genes (Figure 1 and Supplementary Table 3).

Selection of a panel of fraction marker genes with broad coverage across patients

To select candidate fraction marker genes, we screened genes that showed a low frequency of copy number alternations (CNAs), which can influence the methylation levels of the genes [26]. From the 42 candidate marker genes, we selected 34 genes with low frequencies of CNAs in pancreatic cancers [27, 28] (Supplementary Table 3). Next, to identify marker genes methylated in different sets of patients, we conducted a hierarchical clustering analysis using the 34 candidate marker genes, and observed four second-level clusters (I, II, III, and IV) (Figure 2A). From each of the four second-level clusters, we selected genes which had broad coverage across patients and for which we successfully designed high-quality primers for bisulfite pyrosequencing (Figure 3). Consequently, four genes, namely *SIM1*, *MIR129-2*, *NR1I2*, and *HOXB-AS4*, met the two conditions as candidate marker genes (Table 1, and Supplementary Table 4). As expected, the four genes covered different sets of patients and collectively had broad coverages across pancreatic cancer patients and cell lines (Figure 2B, and Supplementary Figure 1).

Methylation levels in pancreatic surgical specimens

To evaluate whether the four candidate marker genes could accurately estimate cancer cell fractions, methylation levels of the four genes were validated by bisulfite pyrosequencing in 61 pancreatic surgical specimens, consisting of four normal samples, 26 non-tumorous samples, 10 IPMN samples, and 21 pancreatic cancer samples (Supplementary Table 1). Unexpectedly, *HOXB-AS4* had relatively high methylation levels even in normal samples and non-tumorous samples (11 of 30 with methylation levels $\geq$ 10%) (Supplementary Figure 2). Therefore, excluding *HOXB-AS4*, we adopted a

panel of the remaining three genes, *SIM1*, *MIR129-2*, and *NR1I2*, as the final candidate marker genes for pancreatic cancer cell fraction.

Among the 10 IPMN samples and 21 pancreatic cancer samples, at least one of the three genes, *SIM1*, *MIR129-2*, and *NR1I2*, was highly methylated in eight (80.0%) IPMNs and 14 (66.7%) pancreatic cancers respectively. At the same time, the three genes had little methylation among the four normal samples and 22 of 26 (84.6%) non-tumorous samples (Figures 4A and 4B). Among the four non-tumorous samples (20N, 23N, 42N, and 60N) with methylation levels ≥ 10%, sample 60N had a high allele frequency of a *KRAS* mutation (Supplementary Table 5). However, in the remaining three samples, somatic mutations were not detected in the following mutation analysis (Supplementary Tables 5 and 6). Overall, the panel of the three genes, *SIM1*, *MIR129-2*, and *NR1I2*, was considered capable of estimating cancer cell fractions with a broad coverage of patients.

Cancer cell-specific mutations were additionally used to estimate the fraction of pancreatic cancer cells. Mutation allele frequencies of 55 cancer-related genes including *KRAS*, which is known to be mutated in approximately 90% of pancreatic cancers [20, 29-31], were analyzed by target deep sequencing. In the 19 tumor samples with a *KRAS* mutation (seven IPMNs, 12 pancreatic cancers), substantial levels of methylation of the fraction marker genes (methylation level ≥ 10%) were also observed in all of them (Supplementary Table 5). In the 12 tumor samples without a *KRAS* mutation (three IPMNs, nine pancreatic cancers), two (22T and 23T) had substantial methylation levels (24.1% in 22T and 36.2% in 23T) (Figures 4A and 4B). Accordingly, these two samples had somatic mutations in other genes (*CDKN2A* in 22T; *ALK*, *SMO*, and *CCND1* in 23T), confirming that they indeed contained cancer cells (Supplementary Table 6). As discussed previously, sample 60N had a variant allele frequency of 4.2 % for a *KRAS* mutation.

Correlation between the cancer cell fraction estimated by the DNA methylation marker and that estimated using the *KRAS* mutant allele frequency

To assess the accuracy of the cancer cell fraction marker, we evaluated the correlation between the cancer cell fraction estimated by the panel of the three DNA methylation marker genes and that estimated using the *KRAS* mutant allele frequency. As shown in Figure 5, we found the two to be highly correlated (R = 0.79).

Discussion

The pancreatic cancer cell fraction was successfully estimated using a panel of three genes, namely *SIM1*, *MIR129-2*, and *NR1I2*, that was extensively selected based upon their lack of methylation in non-tumorous samples, their high incidence of methylation in tumorous samples, and their broad coverage across different patients. The high performance of the panel was shown by a high correlation (R = 0.79) between the cancer cell fraction estimated by the panel of DNA methylation marker genes and that estimated by the *KRAS* mutant allele frequency. Notably, the fraction marker was informative not only in pancreatic cancers, but also in IPMNs, which are known risk factors for pancreatic cancers [32]. Functionally, *MIR129-2* is reported to be a tumor-suppressor gene [33, 34], having the potential to be a driver gene in pancreatic carcinogenesis. On the other hand, neither *NR1I2* nor *SIM1* were expressed in normal pancreatic tissues, suggesting that they are passenger genes.

The cancer cell fraction can be estimated using the mutant allele frequency obtained by next-generation sequencing. Although *KRAS* mutation is useful for this purpose in pancreatic cancers, the DNA methylation marker had a broader coverage across patients. As shown in Figure 4A, substantial methylation levels were obtained even in the two samples without *KRAS* mutations (22T and 23T). Importantly, when these two samples were included in the analysis of correlation of cancer cell fractions, the correlation coefficient remained high (R = 0.79) (Supplementary Figure 3). Furthermore, the fraction marker genes were methylated in *KRAS*-wild-type pancreatic cancer cell lines (Hs 766T and BxPC-3) as shown in Figure 2B. Notably, the DNA methylation marker developed here detected the presence of pre-malignant cells or cancer cells in a "normal" sample

(60N). This sample harbored a *KRAS* mutation, which can be present in pancreatic cancer or pre-malignant cells [35].

Furthermore, cancer cell fractions in early-stage pancreatic cancers were successfully estimated in eight samples. This suggests that the DNA methylation fraction marker may be used for early detection of pancreatic cancers using pancreatic juice or peripheral blood samples. Also, the use of cancer cell fraction markers has been reported to improve the performance of molecular analyses of cancers. For example, in esophageal squamous cell carcinomas, the predictive power of a biomarker was improved after the correction of the cancer cell fraction using a similar fraction marker [25]. Moreover, in lung cancers, correction of the cancer cell fraction allowed for more accurate estimations of copy number alterations from sequencing data [7]. Thus, estimation and correction of pancreatic cancer cell fraction will be useful for many aspects of pancreatic cancer research.

The use of DNA methylation for a cancer cell fraction marker has multiple advantages. The methylation marker can be used in DNA samples, without the need for histological sections and cell counting. In our previous study, we demonstrated that the accuracy of estimation of the cancer cell fraction by DNA methylation marker was equivalent to that of the pathological measurement [9]. Although the pathological approach has been established as the gold standard to estimate the cancer cell fraction, it is time-consuming and prone to technical difficulties when distinguishing cancer cells from co-existing stromal cells, especially when cancer cells infiltrate their surrounding non-cancerous tissue, as is the case with pancreatic cancers [36]. In addition, quantitative methylation analysis is much more cost-effective compared with next-generation sequencing, and can be useful in cancers without measurable mutations.

Three non-tumorous samples (20N, 23N, and 42N) had relatively high levels of methylation (≥ 10%), but no evidence of the presence of tumor cells was obtained by mutation analysis. Since 6.67% of pancreatic cancers without *KRAS* mutations have mutations of other driver genes [27], there remains a possibility that undetected tumor cells were contaminated in these samples as was the case for 60N.

Some limitations remain in this study. Firstly, the sample number was relatively small. Secondly, the sensitivity (71.0%) and specificity (86.7%) of the fraction marker may be insufficient for some analyses that require high accuracy. Thirdly, our present criterion for the isolation of fraction marker genes might have been too strict. Indeed, although *BNC1*, *ADMTS1*, and *CD1D* are reported to be accurate in the detection of pancreatic cancers [37, 38], these genes were excluded from the candidate marker genes in this study. Specifically, *ADMTS1* showed a low incidence of methylation ($\beta > 0.3$) in the cancerous tissues (0-2 of 22 tissues). Three of the 33 *BNC1* probes had a high incidence of methylation in cancer tissues (14-18 of 22), and one of the 15 *CD1D* probes also had a high incidence (18 of 22). However, our screening criterion for the methylation of consecutive probes hampered isolation of these two genes.

In conclusion, we established a pancreatic cancer cell fraction marker using DNA methylation that can be used in DNA samples. This marker will facilitate a variety of molecular analyses of pancreatic cancers by providing an accurate estimation of cancer cell fractions in pancreatic tissue samples.

Abbreviations

TCGA, The Cancer Genome Atlas; CNAs, copy number alternations; Intraductal papillary mucinous neoplasms, IPMNs

Author contributions

A.H., T.U., and S.Y. conceived the study and designed the experiments. R.A., K.K., K.H., N.K., Y.M., A.T., T.U., and A.H. collected tissue samples. H.I. conducted the experiments and data analysis. H.I., S.Y., and T.U. interpreted data and wrote the manuscript. A. H. and T.U. provided support for the studies. All authors have read and approved the submission of the manuscript.

Acknowledgements

The authors are grateful to Drs. K. Ichimura, Y. Matsushita, and M. Kitahara of the Division of Brain Tumor Translational Research in the National Cancer Center Research Institute for their technical assistance with the experiments.

Conflicts of interest

The authors declare no conflicts of interest.

Funding

The research described in this report was funded by the Practical Research for Innovative Cancer Control (grant number; 15ck0106030h0002) from the Japan Agency for Medical Research and development, AMED.

References

1 Siegel R, Ma J, Zou Z, Jemal A: Cancer statistics, 2014. CA: a cancer journal for clinicians 2014;64:9-29.

2 Jemal A, Siegel R, Ward E, Hao Y, Xu J, Murray T, Thun MJ: Cancer statistics, 2008. CA: a cancer journal for clinicians 2008;58:71-96.

3 Li D, Xie K, Wolff R, Abbruzzese JL: Pancreatic cancer. Lancet (London, England) 2004;363:1049-1057.

4 Makohon-Moore A, Iacobuzio-Donahue CA: Pancreatic cancer biology and genetics from an evolutionary perspective. Nature reviews Cancer 2016;16:553-565.

5 Rath N, Olson MF: Regulation of pancreatic cancer aggressiveness by stromal stiffening. Nature medicine 2016;22:462-463.

6 Meyerson M, Gabriel S, Getz G: Advances in understanding cancer genomes through second-generation sequencing. Nature reviews Genetics 2010;11:685-696.

7 Gusnanto A, Wood HM, Pawitan Y, Rabbitts P, Berri S: Correcting for cancer genome size and tumour cell content enables better estimation of copy number alterations from next-generation sequence data. Bioinformatics (Oxford, England) 2012;28:40-47.

8 Roma C, Esposito C, Rachiglio AM, Pasquale R, Iannaccone A, Chicchinelli N, Franco R, Mancini R, Pisconti S, De Luca A, Botti G, Morabito A, Normanno N: Detection of EGFR mutations by TaqMan mutation detection assays powered by competitive allele-specific TaqMan PCR technology. BioMed research international 2013;2013:385087.

9 Takahashi T, Matsuda Y, Yamashita S, Hattori N, Kushima R, Lee YC, Igaki H, Tachimori Y, Nagino M, Ushijima T: Estimation of the fraction of cancer cells in a tumor

DNA sample using DNA methylation. PloS one 2013;8:e82302.

10 Zong L, Hattori N, Yoda Y, Yamashita S, Takeshima H, Takahashi T, Maeda M, Katai H, Nanjo S, Ando T, Seto Y, Ushijima T: Establishment of a DNA methylation marker to evaluate cancer cell fraction in gastric cancer. Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association 2016;19:361-369.

11 Yi JM, Dhir M, Guzzetta AA, Iacobuzio-Donahue CA, Heo K, Yang KM, Suzuki H, Toyota M, Kim HM, Ahuja N: DNA methylation biomarker candidates for early detection of colon cancer. Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine 2012;33:363-372.

12 Hoque MO, Prencipe M, Poeta ML, Barbano R, Valori VM, Copetti M, Gallo AP, Brait M, Maiello E, Apicella A, Rossiello R, Zito F, Stefania T, Paradiso A, Carella M, Dallapiccola B, Murgo R, Carosi I, Bisceglia M, Fazio VM, Sidransky D, Parrella P: Changes in CpG islands promoter methylation patterns during ductal breast carcinoma progression. Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology 2009;18:2694-2700.

13 Okamoto Y, Sawaki A, Ito S, Nishida T, Takahashi T, Toyota M, Suzuki H, Shinomura Y, Takeuchi I, Shinjo K, An B, Ito H, Yamao K, Fujii M, Murakami H, Osada H, Kataoka H, Joh T, Sekido Y, Kondo Y: Aberrant DNA methylation associated with aggressiveness of gastrointestinal stromal tumour. Gut 2012;61:392-401.

14 Harada T, Yamamoto E, Yamano HO, Nojima M, Maruyama R, Kumegawa K, Ashida M, Yoshikawa K, Kimura T, Harada E, Takagi R, Tanaka Y, Aoki H, Nishizono M, Nakaoka M, Tsuyada A, Niinuma T, Kai M, Shimoda K, Shinomura Y, Sugai T, Imai

K, Suzuki H: Analysis of DNA methylation in bowel lavage fluid for detection of colorectal cancer. Cancer prevention research (Philadelphia, Pa) 2014;7:1002-1010.

15 Shigematsu Y, Niwa T, Yamashita S, Taniguchi H, Kushima R, Katai H, Ito S, Tsukamoto T, Ichinose M, Ushijima T: Identification of a DNA methylation marker that detects the presence of lymph node metastases of gastric cancers. Oncology letters 2012;4:268-274.

16 Yamashita S, Takahashi S, McDonell N, Watanabe N, Niwa T, Hosoya K, Tsujino Y, Shirai T, Ushijima T: Methylation silencing of transforming growth factor-beta receptor type II in rat prostate cancers. Cancer research 2008;68:2112-2121.

17 Yoshida T, Yamashita S, Takamura-Enya T, Niwa T, Ando T, Enomoto S, Maekita T, Nakazawa K, Tatematsu M, Ichinose M, Ushijima T: Alu and Satalpha hypomethylation in Helicobacter pylori-infected gastric mucosae. International journal of cancer 2011;128:33-39.

18 Yamashita S, Iida N, Takeshima H, Hattori N, Maeda M, Kishino T, Nagano R, Shimazu T, Tsugane S, Ushijima T: A novel method to quantify base substitution mutations at the 10-6 per bp level in DNA samples. Cancer letters 2017;403:152-158.

19 Matsubara A, Ogawa R, Suzuki H, Oda I, Taniguchi H, Kanai Y, Kushima R, Sekine S: Activating GNAS and KRAS mutations in gastric foveolar metaplasia, gastric heterotopia, and adenocarcinoma of the duodenum. British journal of cancer 2015;112:1398-1404.

20 Rozenblum E, Schutte M, Goggins M, Hahn SA, Panzer S, Zahurak M, Goodman SN, Sohn TA, Hruban RH, Yeo CJ, Kern SE: Tumor-suppressive pathways in pancreatic carcinoma. Cancer research 1997;57:1731-1734.

21 Hezel AF, Kimmelman AC, Stanger BZ, Bardeesy N, Depinho RA: Genetics and

biology of pancreatic ductal adenocarcinoma. Genes & development 2006;20:1218-1249.

22 Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, Miller DK, Wilson PJ, Patch AM, Wu J, Chang DK, Cowley MJ, Gardiner BB, Song S, Harliwong I, Idrisoglu S, Nourse C, Nourbakhsh E, Manning S, Wani S, Gongora M, Pajic M, Scarlett CJ, Gill AJ, Pinho AV, Rooman I, Anderson M, Holmes O, Leonard C, Taylor D, Wood S, Xu Q, Nones K, Fink JL, Christ A, Bruxner T, Cloonan N, Kolle G, Newell F, Pinese M, Mead RS, Humphris JL, Kaplan W, Jones MD, Colvin EK, Nagrial AM, Humphrey ES, Chou A, Chin VT, Chantrill LA, Mawson A, Samra JS, Kench JG, Lovell JA, Daly RJ, Merrett ND, Toon C, Epari K, Nguyen NQ, Barbour A, Zeps N, Kakkar N, Zhao F, Wu YQ, Wang M, Muzny DM, Fisher WE, Brunicardi FC, Hodges SE, Reid JG, Drummond J, Chang K, Han Y, Lewis LR, Dinh H, Buhay CJ, Beck T, Timms L, Sam M, Begley K, Brown A, Pai D, Panchal A, Buchner N, De Borja R, Denroche RE, Yung CK, Serra S, Onetto N, Mukhopadhyay D, Tsao MS, Shaw PA, Petersen GM, Gallinger S, Hruban RH, Maitra A, Iacobuzio-Donahue CA, Schulick RD, Wolfgang CL, Morgan RA, Lawlor RT, Capelli P, Corbo V, Scardoni M, Tortora G, Tempero MA, Mann KM, Jenkins NA, Perez-Mancera PA, Adams DJ, Largaespada DA, Wessels LF, Rust AG, Stein LD, Tuveson DA, Copeland NG, Musgrove EA, Scarpa A, Eshleman JR, Hudson TJ, Sutherland RL, Wheeler DA, Pearson JV, McPherson JD, Gibbs RA, Grimmond SM: Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 2012;491:399-405.

23 Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, Johns AL, Miller D, Nones K, Quek K, Quinn MC, Robertson AJ, Fadlullah MZ, Bruxner TJ, Christ AN, Harliwong I, Idrisoglu S, Manning S, Nourse C, Nourbakhsh E, Wani S, Wilson PJ, Markham E, Cloonan N, Anderson MJ, Fink JL, Holmes O, Kazakoff SH, Leonard C,

Newell F, Poudel B, Song S, Taylor D, Waddell N, Wood S, Xu Q, Wu J, Pinese M, Cowley MJ, Lee HC, Jones MD, Nagrial AM, Humphris J, Chantrill LA, Chin V, Steinmann AM, Mawson A, Humphrey ES, Colvin EK, Chou A, Scarlett CJ, Pinho AV, Giry-Laterriere M, Rooman I, Samra JS, Kench JG, Pettitt JA, Merrett ND, Toon C, Epari K, Nguyen NQ, Barbour A, Zeps N, Jamieson NB, Graham JS, Niclou SP, Bjerkvig R, Grutzmann R, Aust D, Hruban RH, Maitra A, Iacobuzio-Donahue CA, Wolfgang CL, Morgan RA, Lawlor RT, Corbo V, Bassi C, Falconi M, Zamboni G, Tortora G, Tempero MA, Gill AJ, Eshleman JR, Pilarsky C, Scarpa A, Musgrove EA, Pearson JV, Biankin AV, Grimmond SM: Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 2015;518:495-501.

24 Gyobu K, Yamashita S, Matsuda Y, Igaki H, Niwa T, Oka D, Kushima R, Osugi H, Lee S, Suehiro S, Ushijima T: Identification and validation of DNA methylation markers to predict lymph node metastasis of esophageal squamous cell carcinomas. Annals of surgical oncology 2011;18:1185-1194.

25 Takahashi T, Yamahsita S, Matsuda Y, Kishino T, Nakajima T, Kushima R, Kato K, Igaki H, Tachimori Y, Osugi H, Nagino M, Ushijima T: ZNF695 methylation predicts a response of esophageal squamous cell carcinoma to definitive chemoradiotherapy. Journal of cancer research and clinical oncology 2015;141:453-463.

26 Robinson MD, Stirzaker C, Statham AL, Coolen MW, Song JZ, Nair SS, Strbenac D, Speed TP, Clark SJ: Evaluation of affinity-based genome-wide DNA methylation data: effects of CpG density, amplification bias, and copy number variation. Genome research 2010;20:1719-1729.

27 Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer cell 2017;32:185-203.e113.

28 Liang JW, Shi ZZ, Shen TY, Che X, Wang Z, Shi SS, Xu X, Cai Y, Zhao P, Wang CF, Zhou ZX, Wang MR: Identification of genomic alterations in pancreatic cancer using array-based comparative genomic hybridization. PloS one 2014;9:e114616.

29 Hidalgo M: Pancreatic cancer. The New England journal of medicine 2010;362:1605-1617.

30 Wood LD, Hruban RH: Pathology and molecular genetics of pancreatic neoplasms. Cancer journal (Sudbury, Mass) 2012;18:492-501.

31 Zhou L, Baba Y, Kitano Y, Miyake K, Zhang X, Yamamura K, Kosumi K, Kaida T, Arima K, Taki K, Higashi T, Imai K, Hashimoto D, Yamashita Y, Chikamoto A, Beppu T, Tan X, Baba H: KRAS, BRAF, and PIK3CA mutations, and patient prognosis in 126 pancreatic cancers: pyrosequencing technology and literature review. Medical oncology (Northwood, London, England) 2016;33:32.

32 Kamisawa T, Wood LD, Itoi T, Takaori K: Pancreatic cancer. Lancet (London, England) 2016;388:73-85.

33 Dyrskjot L, Ostenfeld MS, Bramsen JB, Silahtaroglu AN, Lamy P, Ramanathan R, Fristrup N, Jensen JL, Andersen CL, Zieger K, Kauppinen S, Ulhoi BP, Kjems J, Borre M, Orntoft TF: Genomic profiling of microRNAs in bladder cancer: miR-129 is associated with poor outcome and promotes cell death in vitro. Cancer research 2009;69:4851-4860.

34 Wong KY, Yim RL, Kwong YL, Leung CY, Hui PK, Cheung F, Liang R, Jin DY, Chim CS: Epigenetic inactivation of the MIR129-2 in hematological malignancies. Journal of hematology & oncology 2013;6:16.

35 Kanda M, Matthaei H, Wu J, Hong SM, Yu J, Borges M, Hruban RH, Maitra A, Kinzler K, Vogelstein B, Goggins M: Presence of somatic mutations in most early-stage

pancreatic intraepithelial neoplasia. Gastroenterology 2012;142:730-733.e739.

36 Geiersbach K, Adey N, Welker N, Elsberry D, Malmberg E, Edwards S, Downs-Kelly E, Salama M, Bronner M: Digitally guided microdissection aids somatic mutation detection in difficult to dissect tumors. Cancer genetics 2016;209:42-49.

37 Kisiel JB, Raimondo M, Taylor WR, Yab TC, Mahoney DW, Sun Z, Middha S, Baheti S, Zou H, Smyrk TC, Boardman LA, Petersen GM, Ahlquist DA: New DNA Methylation Markers for Pancreatic Cancer: Discovery, Tissue Validation, and Pilot Testing in Pancreatic Juice. Clinical cancer research : an official journal of the American Association for Cancer Research 2015;21:4473-4481.

38 Yi JM, Guzzetta AA, Bailey VJ, Downing SR, Van Neste L, Chiappinelli KB, Keeley BP, Stark A, Herrera A, Wolfgang C, Pappou EP, Iacobuzio-Donahue CA, Goggins MG, Herman JG, Wang TH, Baylin SB, Ahuja N: Novel methylation biomarker panel for the early detection of pancreatic cancer. Clinical cancer research : an official journal of the American Association for Cancer Research 2013;19:6544-6555.

Figure legends

Figure 1. Genome-wide isolation of genomic regions specifically methylated in pancreatic cancer cells

The genome-wide DNA methylation data were obtained from i) our own analysis of peripheral leucocytes and pancreatic cancer cell lines and ii) the TCGA database (cancerous and non-cancerous tissues). Probes unmethylated in non-cancerous cells were first selected, and then those methylated in cancer cells were selected using three- and two-consecutive-probe approaches. Candidate genomic regions were converted to candidate genes, and 24 and 27 genes were selected as candidate fraction markers. Nine genes were overlapping in the two approaches, and 42 genes were finally isolated as candidate marker genes.

Figure 2. Selection of a panel of fraction marker genes using a hierarchical cluster analysis

(A) A hierarchical cluster analysis was conducted using the 121 CpG sites in 34 genes with low frequencies of CNAs and 22 pancreatic cancerous tissues. Consequently, four second-level clusters (I, II, III, and IV clusters), as shown by blue bars on the left side, were obtained. (B) Four candidate genes (*SIM1*, *MIR129-2*, *NR1I2*, and *HOXB-AS4*) selected from the four second-level clusters in panel A had broad and different coverage sets of patients and cancer cell lines. Red and blue cells show samples with β values ≥ 0.3 and ≥ 0.8, respectively.

Figure 3. Genomic structures of the four candidate marker genes

Genomic structure and the location of a CpG island are shown at the top. A CpG map around the target CpG sites is shown in the bottom. Vertical lines show individual CpG sites. Arrows show the targeted CpG locations of bisulfite pyrosequencing. The triangle shows the CpG site where the methylation level was measured by bisulfite pyrosequencing.

Figure 4. Methylation levels of *SIM1, MIR129-2,* and *NR1I2* in pancreatic surgical specimens

Methylation levels of the three marker genes, *SIM1*, *MIR129-2*, and *NR1I2*, were validated by bisulfite pyrosequencing using 61 pancreatic surgical specimens. Four normal samples, 26 non-tumorous samples, 10 IPMN samples, and 21 pancreatic cancer samples were used. Among the three genes, one or more were highly methylated in 22 tumor samples (eight IPMNs, 14 pancreatic cancers). Samples with a *KRAS* mutation are shown by asterisks. Only a small amount of methylation was detected in the normal and non-tumorous samples. (A) Individual methylation levels of the three genes. (B) Collective presentation of the highest methylation level among the three genes.

Figure 5. Correlation between the cancer cell fraction estimated by the DNA methylation marker and that estimated using the *KRAS* mutant allele frequency

The cancer cell fraction estimated by the DNA methylation marker had a high correlation with that estimated using the *KRAS* mutant allele frequency (R = 0.79). Using the *KRAS* mutant allele frequency, the cancer cell fraction was estimated by doubling its frequency.

Supplementary Figure legends

Supplementary Figure 1. An analysis of the patient coverage by the four candidate marker genes

A hierarchical clustering analysis was conducted using the four candidate CpG sites (four genes) and 22 cancerous samples. The four genes covered all samples (β value ≥ 0.3).

Supplementary Figure 2. Methylation levels of *HOXB-AS4* in pancreatic surgical specimens

The methylation level of *HOXB-AS4* was validated by bisulfite pyrosequencing using 61 pancreatic surgical specimens. High methylation levels (≥ 10%) were observed even in normal and non-tumorous samples (11 of 30 samples), in contrast with the other three genes, *SIM1*, *MIR129-2*, and *NR1I2*.

Supplementary Figure 3. Correlation between the cancer cell fraction estimated by the DNA methylation marker and that estimated using the somatic mutant allele frequency of *KRAS* and other cancer-related genes

The cancer cell fraction estimated by the DNA methylation marker had a high correlation with that estimated using the somatic mutant allele frequency of *KRAS* or other cancer-related genes (R = 0.79). Using somatic mutations of oncogenes, the cancer cell fraction was estimated by doubling their mutant allele frequencies. Black circles show the tumor samples with *KRAS* mutations, the blue triangle shows the 22T sample with a *CDKN2A* mutation, and the red diamond shows the 23T sample with *ALK*, *SMO*, and *CCND1* mutations. In the 23T sample, the cancer cell fraction was estimated using the *ALK*

mutation frequency because its frequency was the highest among that of the three genes.

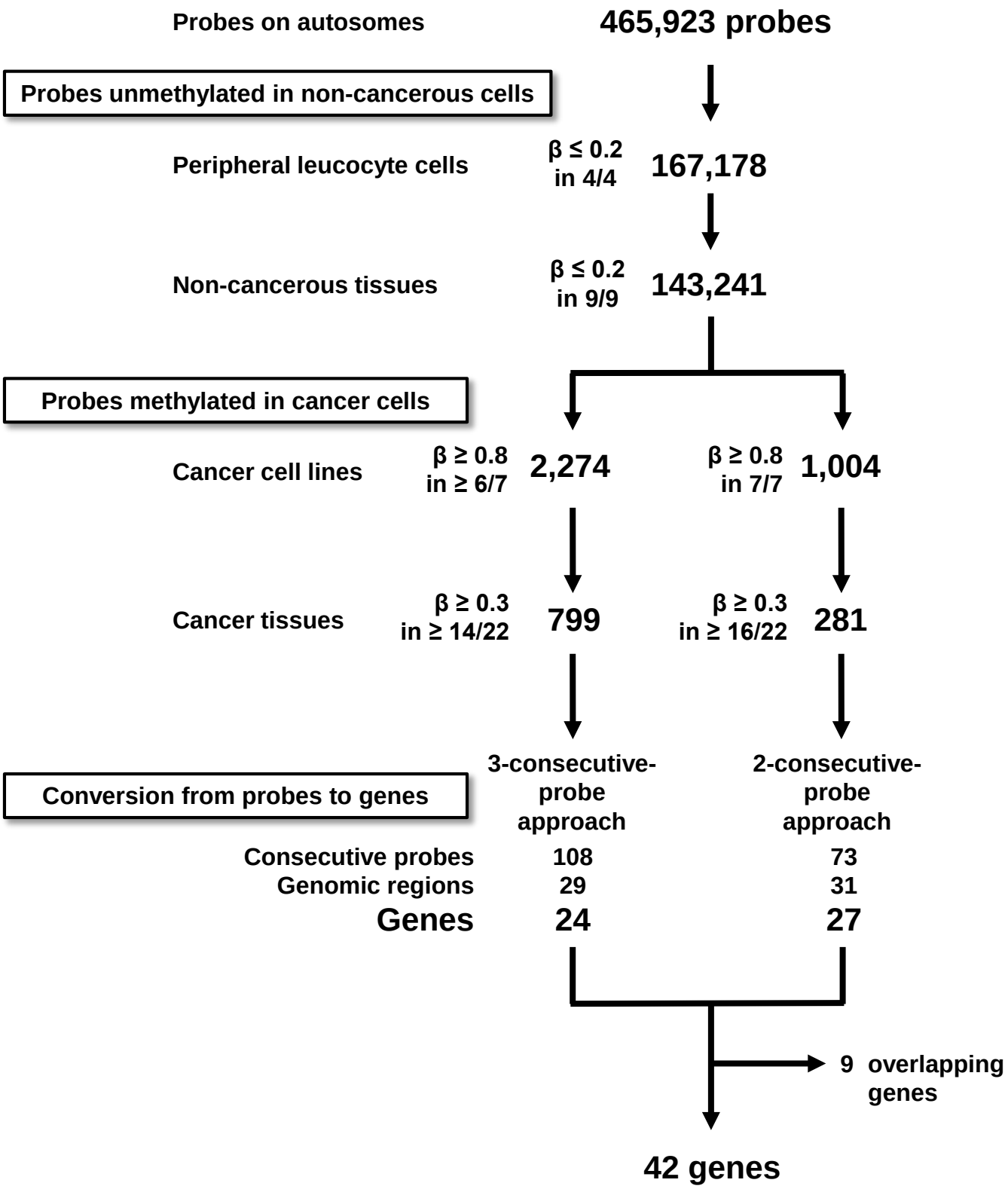
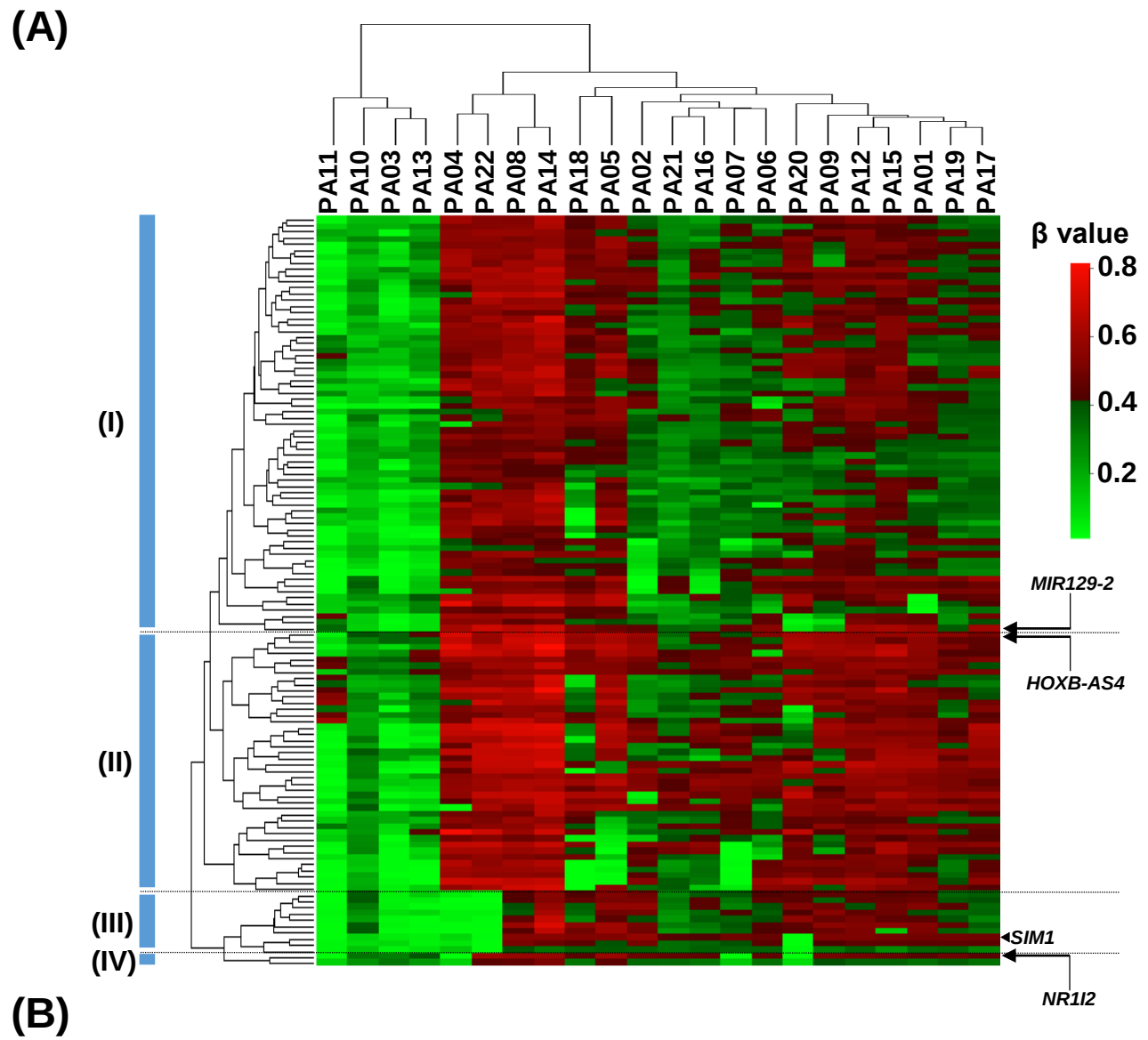

Probes on autosomes
465,923 probes
Probes unmethylated in non-cancerous cells
Peripheral leucocyte cells
β ≤ 0.2 in 4/4
167,178
Non-cancerous tissues
β ≤ 0.2 in 9/9
143,241
Probes methylated in cancer cells
Cancer cell lines
β ≥ 0.8 in ≥ 6/7
2,274
β ≥ 0.8 in 7/7
1,004
Cancer tissues
β ≥ 0.3 in ≥ 14/22
799
β ≥ 0.3 in ≥ 16/22
281
Conversion from probes to genes
3-consecutive-probe approach
2-consecutive-probe approach
Consecutive probes
108
73
Genomic regions
29
31
Genes
24
27
9 overlapping genes
42 genes

Figure 1



(B)

	PA 01	PA 02	PA 03	PA 04	PA 05	PA 06	PA 07	PA 08	PA 09	PA 10	PA 11	PA 12	PA 13	PA 14	PA 15	PA 16	PA 17	PA 18	PA 19	PA 20	PA 21	PA 22	Capan-2	HPAC	HPAF-II	Hs 766T	MIA PaCa-2	PANC-1	BxPC-3
SIM1	■	■			■	■	■	■	■			■		■	■	■	■	■		■	■		■	■	■		■	■	■
MIR 129-2	■	■		■	■	■	■	■			■	■		■	■	■	■	■		■	■	■	■	■	■	■	■	■	■
NR1I2	■	■	■		■	■		■	■	■		■	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■
HOXB-AS4	■	■	■	■	■	■	■	■	■	■		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■

Cancerous tissues

Cell lines

Figure 2

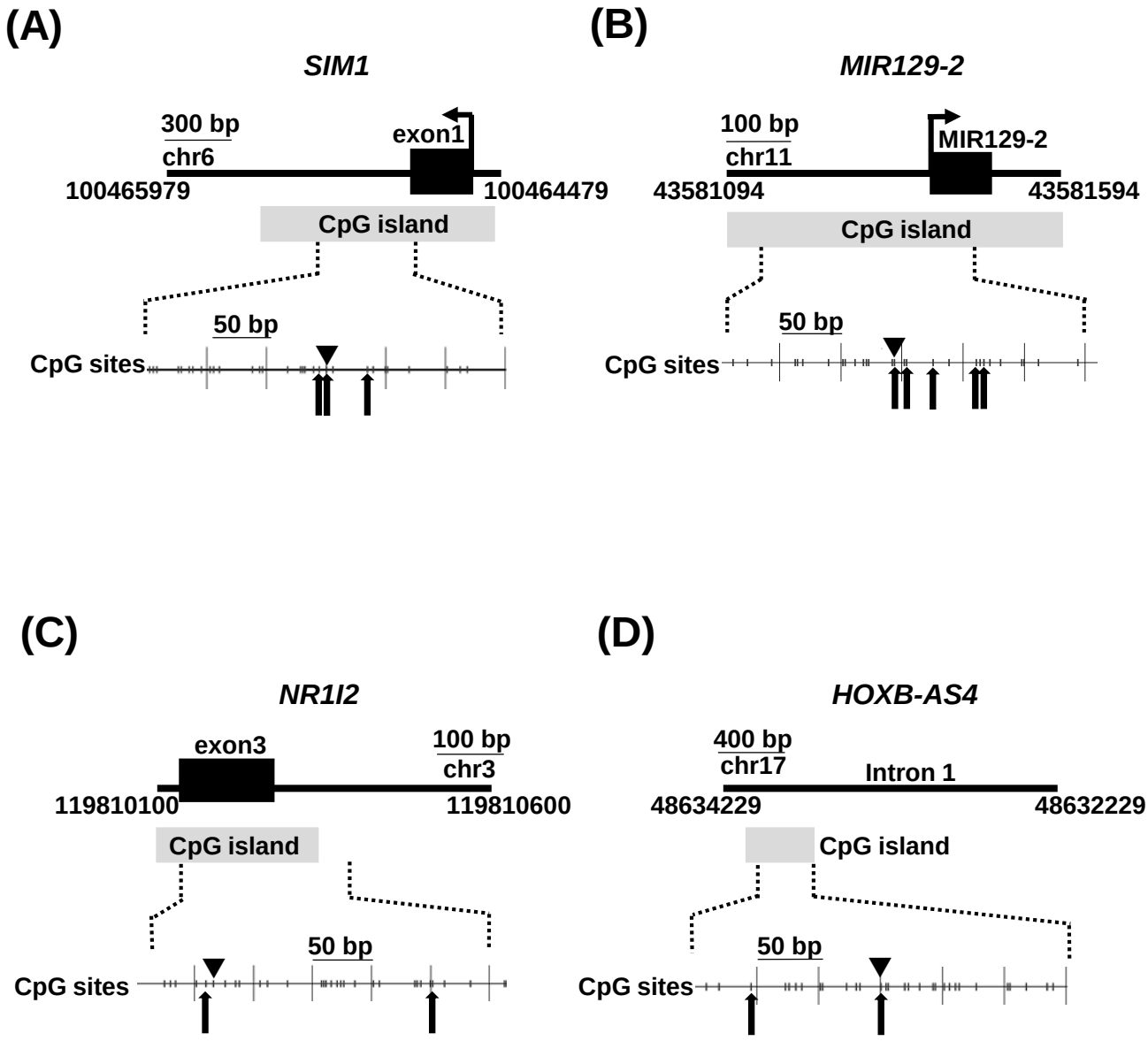
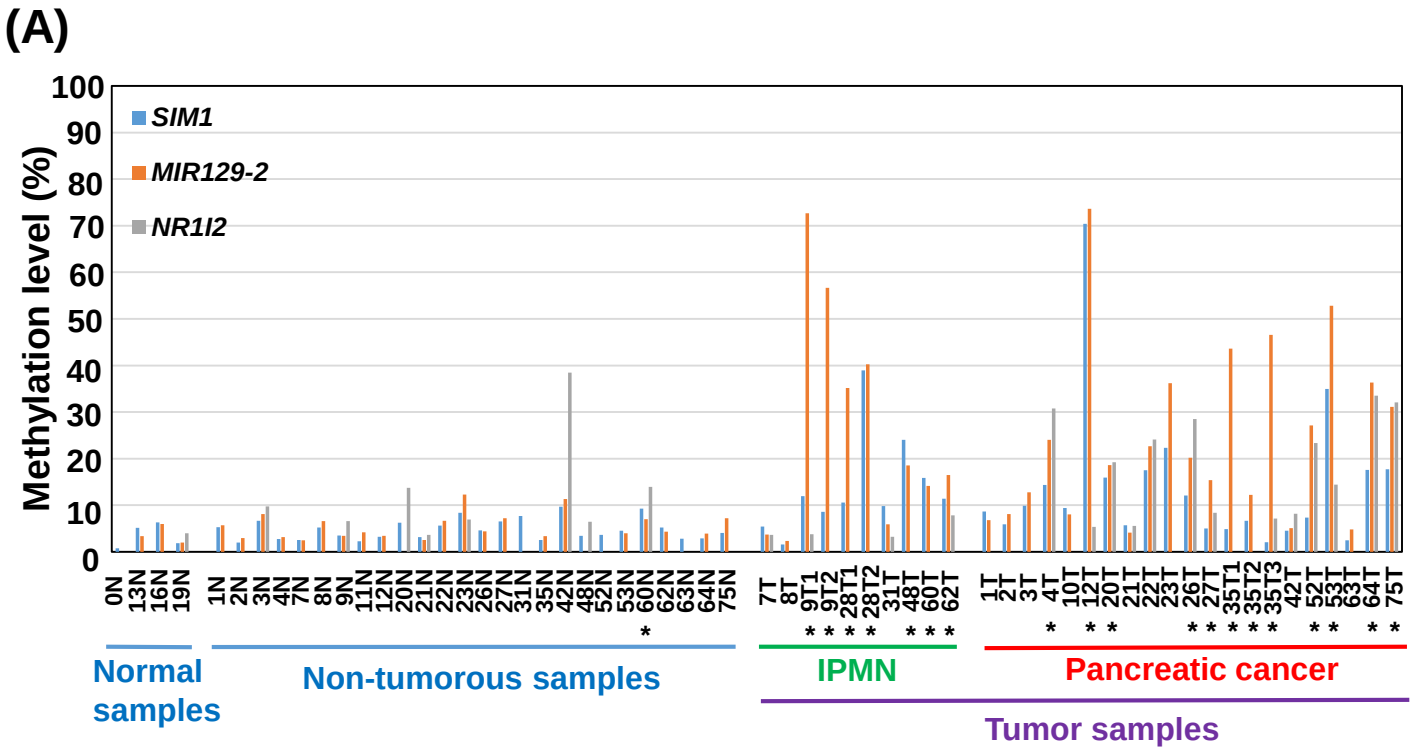

(A)
SIM1
300 bp
chr6
exon1
100465979
100464479
CpG island
50 bp
CpG sites
(B)
MIR129-2
100 bp
chr11
MIR129-2
43581094
43581594
CpG island
50 bp
CpG sites
(C)
NR1I2
exon3
100 bp
chr3
119810100
119810600
CpG island
50 bp
CpG sites
(D)
HOXB-AS4
400 bp
chr17
Intron 1
48634229
48632229
CpG island
50 bp
CpG sites

Figure 3


(A)
SIM1
MIR129-2
NR1I2
Methylation level (%)
0 10 20 30 40 50 60 70 80 90 100
0N 13N 16N 19N 1N 2N 3N 4N 7N 8N 9N 11N 12N 20N 21N 22N 23N 26N 27N 31N 35N 42N 48N 52N 53N 60N 62N 63N 64N 75N 7T 8T 9T1 9T2 28T1 28T2 31T 48T 60T 62T 1T 2T 3T 4T 10T 12T 20T 21T 22T 23T 26T 27T 35T1 35T2 35T3 42T 52T 53T 63T 64T 75T
Normal samples
Non-tumorous samples
IPMN
Pancreatic cancer
Tumor samples

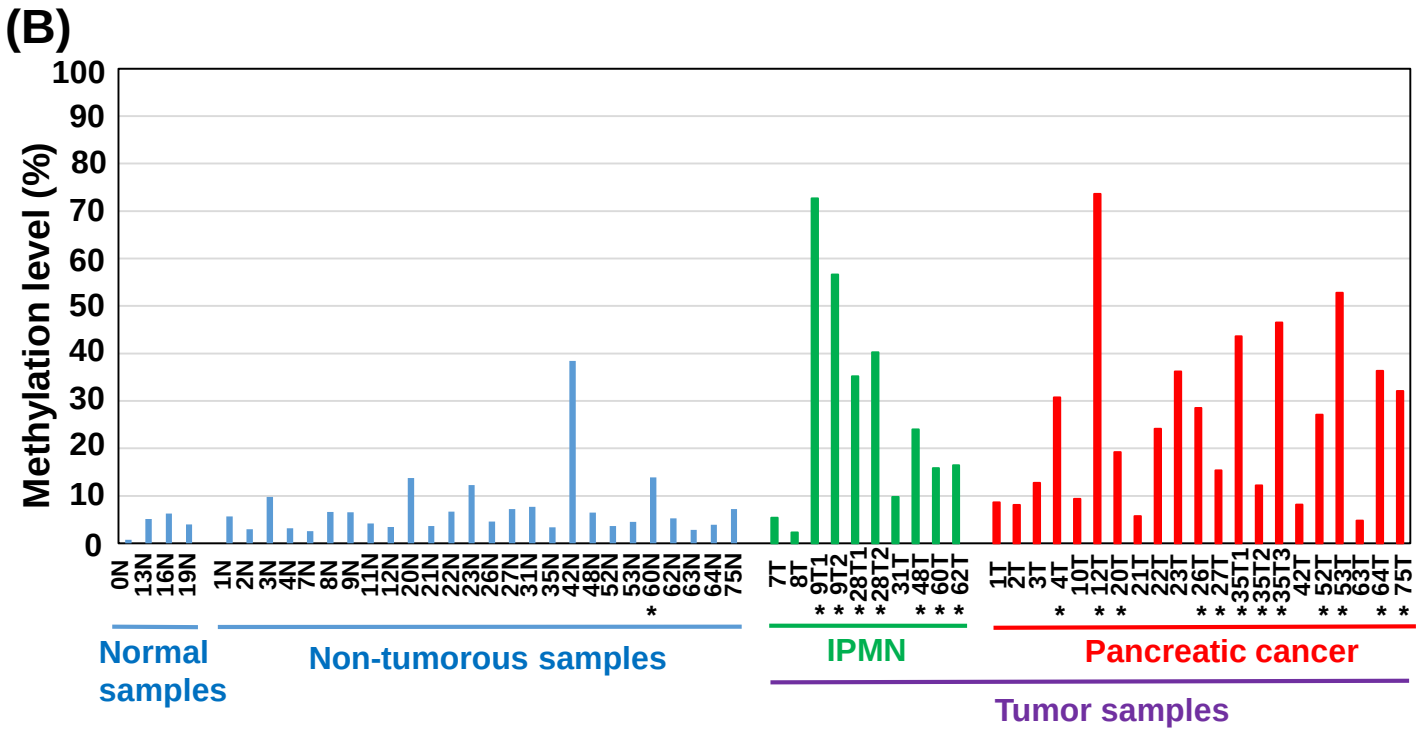

(B)
Methylation level (%)
0 10 20 30 40 50 60 70 80 90 100
0N 13N 16N 19N 1N 2N 3N 4N 7N 8N 9N 11N 12N 20N 21N 22N 23N 26N 27N 31N 35N 42N 48N 52N 53N 60N 62N 63N 64N 75N 7T 8T 9T1 9T2 28T1 28T2 31T 48T 60T 62T 1T 2T 3T 4T 10T 12T 20T 21T 22T 23T 26T 27T 35T1 35T2 35T3 42T 52T 53T 63T 64T 75T
Normal samples
Non-tumorous samples
IPMN
Pancreatic cancer
Tumor samples

Figure 4

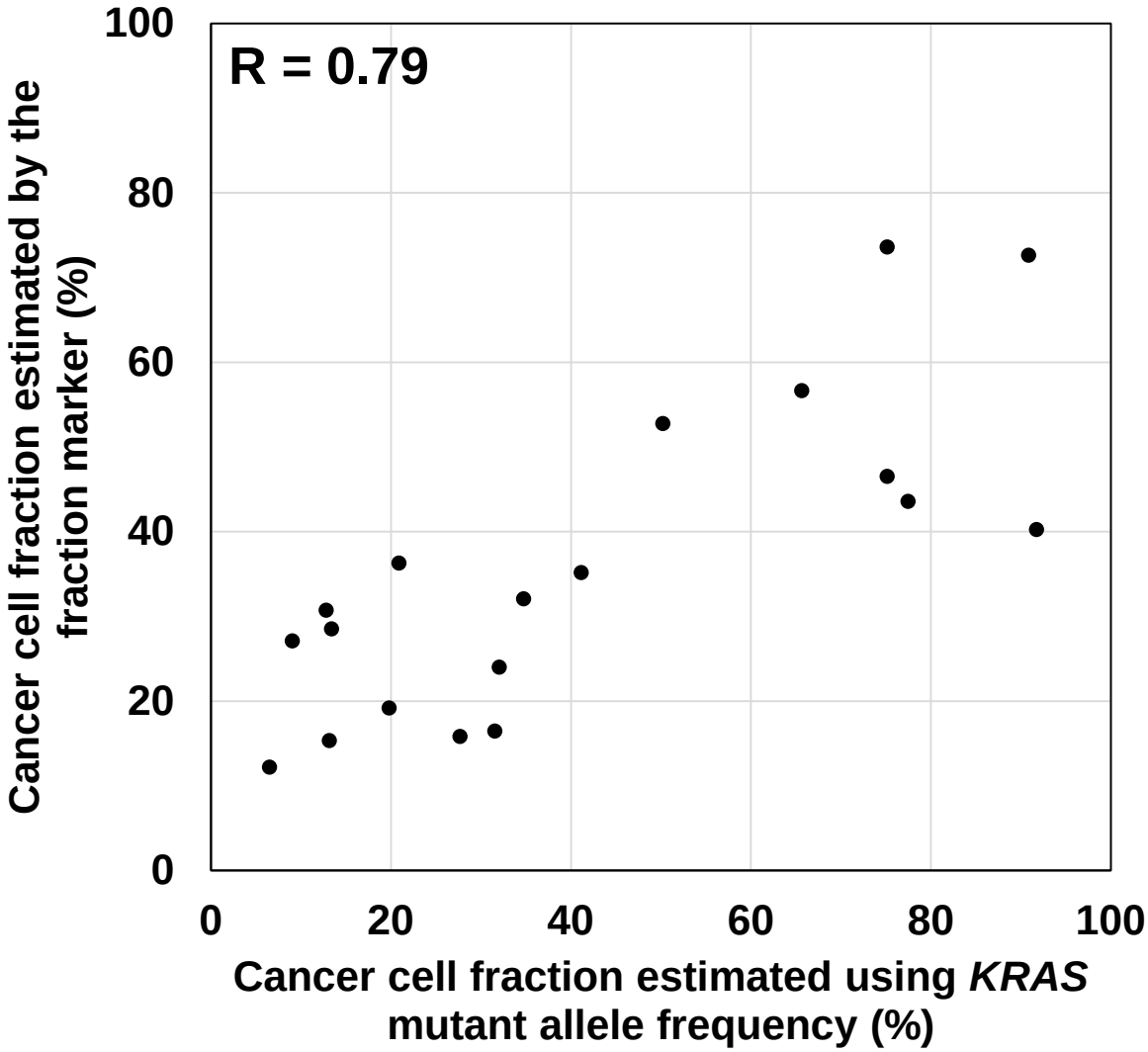
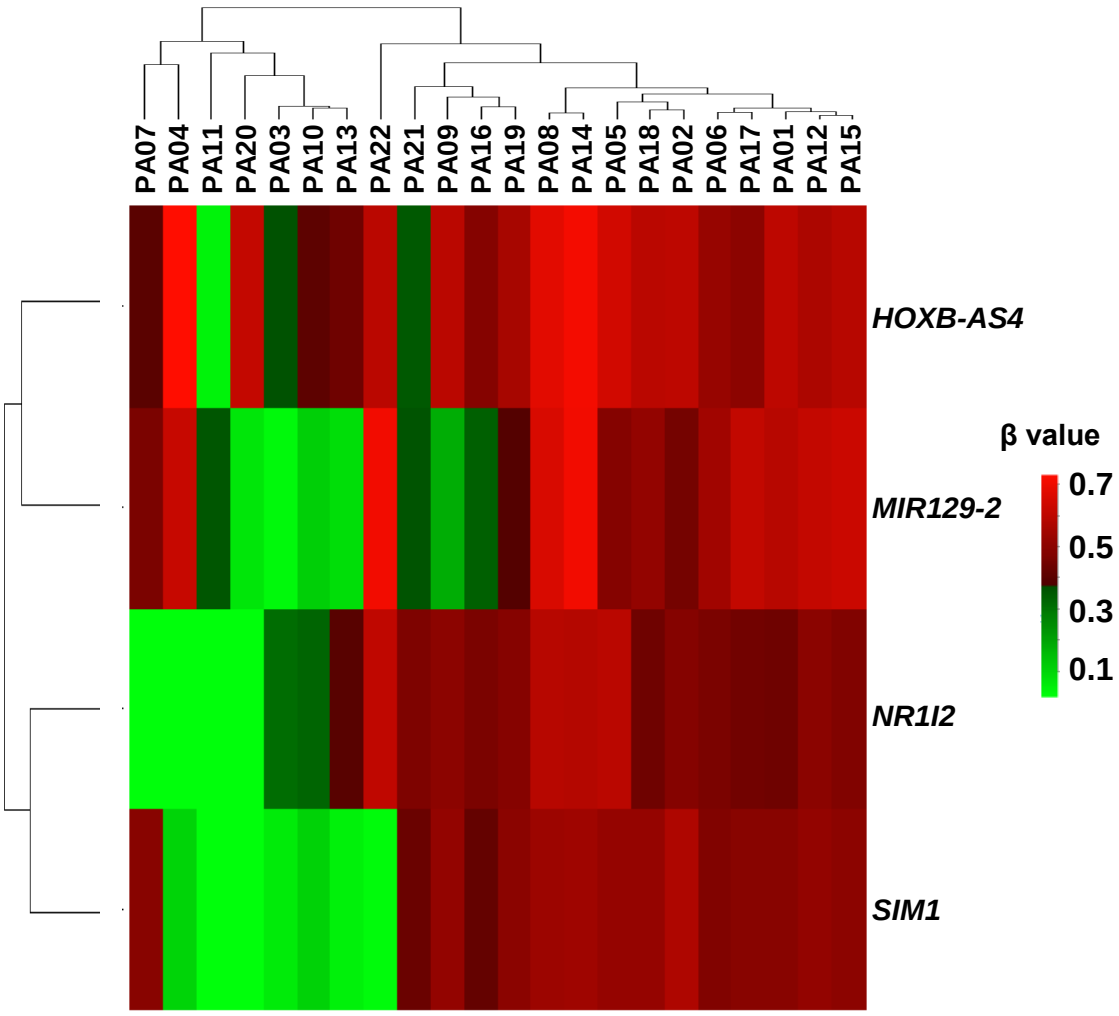
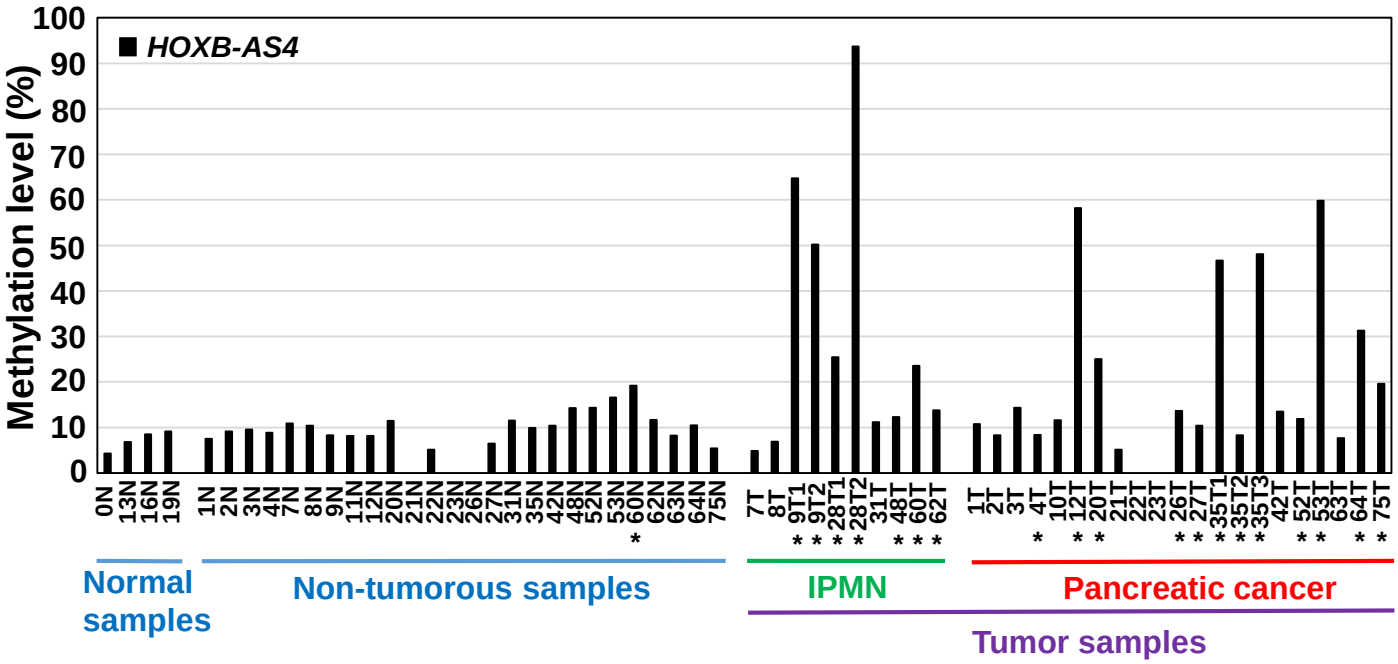

R = 0.79
Cancer cell fraction estimated by the fraction marker (%)
Cancer cell fraction estimated using *KRAS* mutant allele frequency (%)
0
20
40
60
80
100

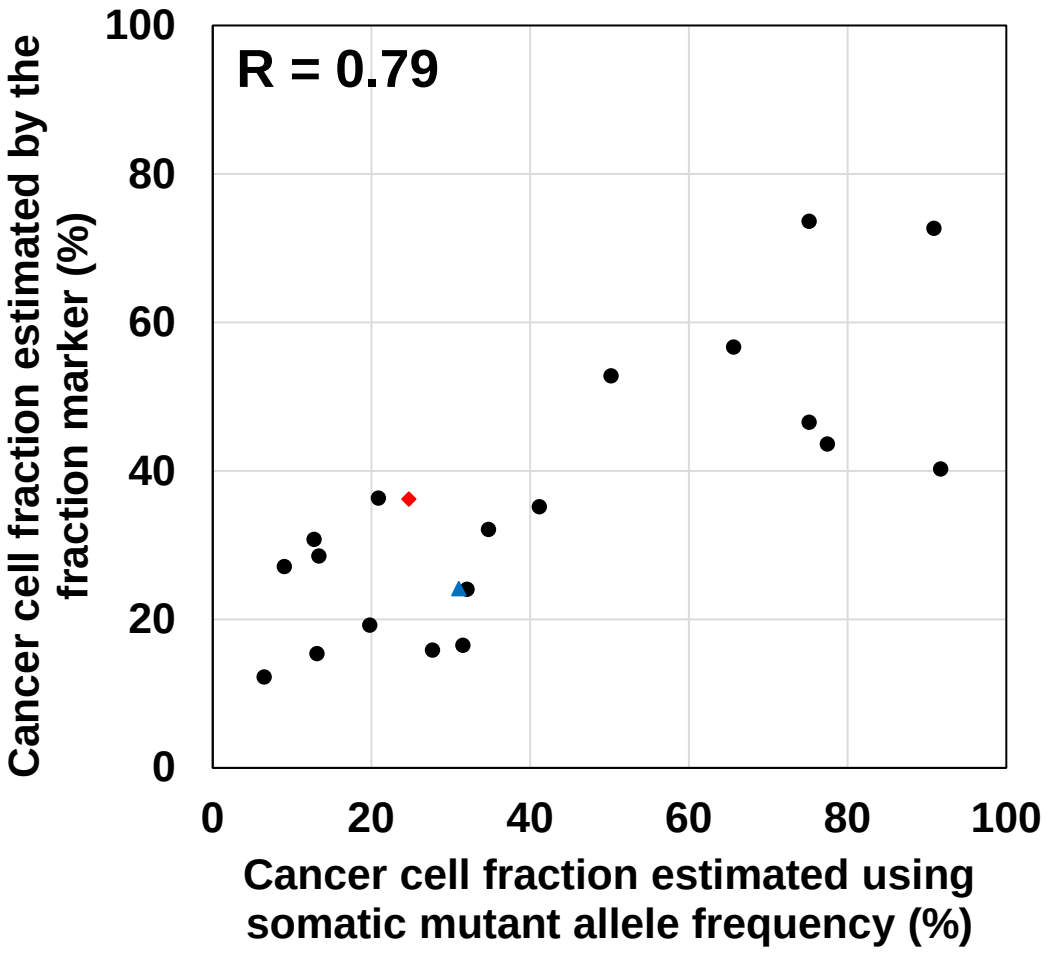
Figure 5



Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3

Table 1. Candidate genomic regions for fraction markers

No.	Gene symbol*	Chr*	Nt. number*	Probe ID	Relation to a CpG island	Position to a TSS*	No. of consecutive probes	Incidence of methylation in cell lines	Incidence of methylation in cancer tissues
1	*SIM1*	6	100465064	cg27252696	Island	-1386;-134;-174	3	6	15
2	*MIR129-2*	11	43581297	cg14416371	Island	24860;2407;1801;-84	3	7	17
3	*NR1I2*	3	119810109	cg00836482	Island	27398;29624;29624;28073;-968	2	7	18
4	*HOXB-AS4*	17	48633979	cg01452847	Island	4429;5303;-406;-1419	2	7	21

*According to human genome assembly hg38. Chr, chromosome; Nt. nucleotide; TSS, transcriptional start site

Supplementary Table 1. Patient background for pancreatic surgical specimens

Patient ID	Sex	Age	Pathological diagnosis	Pathological stage	Normal sample	Non-tumrous sample	Tumor sample
0	Male	33	Purchased normal pancreatic genomic DNA		1		
1	Female	75	Pancreatic cancer (adenocarcinoma)	I		1	1
2	Female	63	Pancreatic cancer (adenocarcinoma)	III		1	1
3	Female	73	Pancreatic cancer (adenocarcinoma)	III		1	1
4	Female	79	Pancreatic cancer (adenocarcinoma)	III		1	1
7	Male	78	IPMN			1	1
8	Male	38	IPMN			1	1
9	Male	75	IPMN			1	2*
10	Female	87	Carcinoma drived from IPMN	I			1
11	Female	72	Carcinoma drived from IPMN	I		1	
12	Female	72	Carcinoma drived from IPMN	I		1	1
13	Male	57	Chronic pancreatitis		1		
16	Male	80	Autoimmunepancreatitis		1		
19	Male	40	Chronic pancreatitis		1		
20	Male	75	Pancreatic cancer (adenocarcinoma)	III		1	1
21	Male	66	Pancreatic cancer (adenocarcinoma)	III		1	1
22	Female	69	Pancreatic cancer (adenocarcinoma)	III		1	1
23	Female	69	Pancreatic cancer (adenocarcinoma)	IV		1	1
26	Male	58	Pancreatic cancer (adenocarcinoma)	III		1	1
27	Male	56	Pancreatic cancer (adenocarcinoma)	III		1	1
28	Male	67	IPMN				2*
31	Male	74	IPMN			1	1
35	Female	64	Carcinoma drived from IPMN	I		1	3*
42	Male	66	Carcinoma drived from IPMN	I		1	1
48	Male	55	IPMN			1	1
52	Male	53	Carcinoma drived from IPMN	III		1	1
53	Male	78	Carcinoma drived from IPMN	I		1	1
60	Female	74	IPMN			1	1
62	Female	68	IPMN			1	1
63	Female	61	Pancreatic cancer (adenocarcinoma)	III		1	1
64	Female	67	Pancreatic cancer (adenocarcinoma)	IV		1	1
75	Male	46	Pancreatic cancer (adenocarcinoma)	IIb		1	1

*Two or three tumor samples were obtained from one patient.

Supplementary Table 2. Pancreatic tissue sample data extracted from the TCGA database

TCGA biospecimen ID	Sample	Sample ID	Sex	Age	Pathological diagnosis	Pathological Stage
TCGA-IB-AAUO	Cancerous tissue	PA01	Female	64	Adenocarcinoma Ductal Type	IIB
TCGA-H8-A6C1	Cancerous tissue	PA02	Male	53	Adenocarcinoma-Other Subtype	IIA
TCGA-XD-AAUG	Cancerous tissue	PA03	Female	66	Adenocarcinoma Ductal Type	IV
TCGA-LB-A7SX	Cancerous tissue	PA04	Female	74	Adenocarcinoma Ductal Type	IIB
TCGA-3A-A9IU	Cancerous tissue	PA05	Male	65	Adenocarcinoma Ductal Type	IIB
TCGA-3A-A9IB	Cancerous tissue	PA06	Female	69	Adenocarcinoma Ductal Type	IIB
TCGA-HZ-A8P0	Cancerous tissue	PA07	Male	76	Adenocarcinoma Ductal Type	IIB
TCGA-S4-A8RM	Cancerous tissue	PA08	Male	67	Adenocarcinoma Ductal Type	IIB
TCGA-S4-A8RP	Cancerous tissue	PA09	Female	77	Adenocarcinoma Ductal Type	IIB
TCGA-LB-A9Q5	Cancerous tissue	PA10	Female	63	Adenocarcinoma Ductal Type	IIB
TCGA-2J-AABP	Cancerous tissue	PA11	Female	58	Undifferentiated Carcinoma	IIB
TCGA-3A-A9IH	Cancerous tissue	PA12	Female	66	Adenocarcinoma Ductal Type	1A
TCGA-H6-A45N	Cancerous tissue	PA13	Female	88	Adenocarcinoma-Other Subtype	IIB
TCGA-HZ-A9TJ	Cancerous tissue	PA14	Male	70	Adenocarcinoma-Other Subtype	IV
TCGA-2J-AABH	Cancerous tissue	PA15	Male	61	Adenocarcinoma-Other Subtype	IIA
TCGA-FZ-5919	Cancerous tissue	PA16	Female	59	Adenocarcinoma Ductal Type	IIB
TCGA-FZ-5920	Cancerous tissue	PA17	Male	52	Adenocarcinoma Ductal Type	IIB
TCGA-FZ-5922	Cancerous tissue	PA18	Male	81	Adenocarcinoma Ductal Type	IIA
TCGA-FZ-5923	Cancerous tissue	PA19	Male	71	Adenocarcinoma Ductal Type	IV
TCGA-FZ-5924	Cancerous tissue	PA20	Male	83	Adenocarcinoma Ductal Type	IIA
TCGA-FZ-5926	Cancerous tissue	PA21	Female	73	Adenocarcinoma Ductal Type	III
TCGA-F2-6879	Cancerous tissue	PA22	Male	57	Adenocarcinoma-Other Subtype	IIB
TCGA-H6-A45N	Non-cancerous tissue		Female	88	Adenocarcinoma-Other Subtype	IIB
TCGA-H6-8124	Non-cancerous tissue		Female	56	Adenocarcinoma Ductal Type	IIB
TCGA-HV-A5A3	Non-cancerous tissue		Male	50	Adenocarcinoma Ductal Type	IIA
TCGA-FZ-5919	Non-cancerous tissue		Female	59	Adenocarcinoma Ductal Type	IIB
TCGA-FZ-5920	Non-cancerous tissue		Male	52	Adenocarcinoma Ductal Type	IIB
TCGA-FZ-5922	Non-cancerous tissue		Male	81	Adenocarcinoma Ductal Type	IIA
TCGA-FZ-5923	Non-cancerous tissue		Male	71	Adenocarcinoma Ductal Type	IV
TCGA-FZ-5924	Non-cancerous tissue		Male	83	Adenocarcinoma Ductal Type	IIA
TCGA-FZ-5926	Non-cancerous tissue		Female	73	Adenocarcinoma Ductal Type	III

Supplementary Table 3. Forty-two genes from 45 genomic regions specifically methylated in pancreatic cancer cells.

No.	Gene symbol*	Chr*	Nt. number*	Probe ID	Relation to a CpG island	Position to a TSS*	No. of consecutive probe	Incidence of methylation in cell lines	Incidence of methylation in cancer tissues	Copy number alternations**
1	*RNF220*	1	44407392	cg04023150	Island	2197;2103;2103;-1311	3	7	15	Deletion
			44407557	cg10224098	Island	2362;2268;2268;-1146	3	7	18	
			44407920	cg04541474	Island	2725;2631;2631;-783	3	7	14	
2	*KCNA3*	1	110674572	cg20302133	Island	462	3 and 2	7	20	Deletion
			110674784	cg26013553	Island	250	3 and 2	7	19	
			110674875	cg11595545	Island	159	3 and 2	7	20	
			110674953	cg01423964	Island	81	3 and 2	7	17	
			110675069	cg06750832	Island	-35	3	7	14	
3	*TIRM58*	1	247857134	cg20429172	Island	-66	3	6	19	Deletion
			247857330	cg20810478	Island	130	3	6	19	
			247857339	cg26157385	Island	139	3 and 2	7	20	
			247857390	cg23054189	Island	190	3 and 2	7	20	
			247857395	cg20146541	Island	195	3 and 2	7	20	
			247857510	cg07533148	Island	310	3 and 2	7	19	
			247857789	cg16021909	Island	589	3 and 2	7	19	
4	*CELF2*	10	11017751	cg17290701	Island	12454;-180	3	6	14	None
			11017755	cg26328510	Island	12458;-176	3	7	14	
			11017762	cg03813164	Island	12465;-169	3	7	15	
			11017764	cg12356890	Island	12467;-167	3	7	15	
5	*MIR129-2*	11	43581297	cg14416371	Island	24860;2407;1801;-84	3	7	17	None
			43581307	cg14944647	Island	24870;2417;1811;-74	3	7	14	
			43581329	cg01939477	Island	24892;2439;1833;-52	3	7	14	
			43581364	cg16407471	Island	24927;2474;1868;-17	3 and 2	7	16	
			43581370	cg05376374	Island	24933;2480;1874;-11	3 and 2	7	17	
6	*GABRG3*	15	26970850	cg03024760	Island	-595;-544;-433	3	6	14	None
			26970861	cg02281208	Island	-584;-533;-422	3	6	16	
			26970865	cg08182446	Island	-580;-529;-418	3	6	15	
7	*SALL1*	16	51150875	cg05151154	Island	401;-277;493;367;-277	3	6	15	None
			51150975	cg07498275	Island	301;-377;393;267;-377	3	7	14	
			51151090	cg08806408	Island	186;-492;278;152;-492	3	7	15	
8	*HOXB-AS3*	17	48577927	cg07438617	S_Shore	28296;12312;1431;12315;12309;4696;12346;26986;2185	3	6	14	None
			48578025	cg21546671	S_Shore	28394;12214;1333;12217;12211;4598;12248;26888;2087	3	6	15	
			48578032	cg14458834	S_Shore	28401;12207;1326;12210;12204;4591;12241;26881;2080	3	6	17	
			48578114	cg15565065	S_Shore	28483;12125;1244;12128;12122;4509;12159;26799;1998	3	6	16	
			48578199	cg08089301	S_Shore	28568;12040;1159;12043;12037;4424;12074;26714;1913	3	6	17	
			48578218	cg07015911	S_Shore	28587;12021;1140;12024;12018;4405;12055;26695;1894	3	6	16	
			48578226	cg09194159	S_Shore	28595;12013;1132;12016;12010;4397;12047;26687;1886	3	6	14	
9	GALR1	18	77250010	cg03659519	Island	-540	3	6	14	Amplification
			77250012	cg20872937	Island	-538	3	6	15	
			77250044	cg17911318	Island	-506	3	6	16	
10	*CHST8*	19	33621919	cg26565021	Island	-37;-53;-37;-66	3	6	16	None
			33622085	cg19594305	Island	129;113;129;100	3	6	17	
			33622104	cg16190732	Island	148;132;148;119	3	6	17	
11	*ZNF382*	19	36605246	cg25397945	Island	-74;428;-90;-68;-68;-83;12;-2;31;17;-2	3 and 2	7	22	None
			36605419	cg02587316	Island	99;601;83;105;105;90;-161;-175;-142;-156;-175	3 and 2	7	17	
			36605421	cg18630667	Island	101;603;85;107;107;92;-163;-177;-144;-158;-177	3 and 2	7	16	
			36605427	cg05020604	Island	107;609;91;113;113;98;-169;-183;-150;-164;-183	3	7	15	
			36605585	cg04332534	Island	265;767;249;271;271;256;-327;-341;-308;-322;-341	3	6	15	
12	*AC003006.7*	19	57708927	cg11294513	Island	26881;26881;268;242;26891	2	7	21	None
			57709002	cg05661282	Island	26956;26956;193;167;26966	2	7	20	
			57709148	cg27049766	Island	27102;27102;47;21;27112	3 and 2	7	16	
			57709289	cg03234186	S_Shore	27243;27243;-94;-120;27253	3 and 2	7	17	
			57709294	cg08668790	S_Shore	27248;27248;-99;-125;27258	3 and 2	7	16	
13	*AC093702.1*	2	45013869	cg25623768	Island	-200	3	6	14	None
			45013971	cg22882665	Island	-302	3	6	14	
			45014083	cg10476112	Island	-414	3	6	14	
14	*POU3F3*	2	104855502	cg01878345	Island	-10;2216	3	7	18	None
			104855507	cg17078686	Island	-5;2221	3	6	17	
			104855521	cg24472231	Island	9;2235	3	6	14	
			104855735	cg05506365	Island	223;2449	3	7	16	
15	*HOXD-AS2*	2	176122734	cg04739647	Island	14280;13;-30	3	7	15	None
			176122737	cg05167251	Island	14277;16;-27	3 and 2	7	18	
			176122877	cg02885007	Island	14137;156;113	3 and 2	7	18	
			176129634	cg17863912	Island	7380;-107;-61;-724;-389	2	7	16	
			176129637	cg19384289	Island	7377;-104;-58;-721;-386	2	7	19	
			176129937	cg15808943	Island	7077;196;242;-421;-86	3 and 2	7	16	
			176130036	cg24416513	Island	6978;295;341;-322;13	3 and 2	7	18	
			176130360	cg15520279	Island	6654;619;665;2;337	3 and 2	7	19	
16	*BHLHE23*	20	63007166	cg14060496	Island	-130;-130	3	6	15	None
			63007222	cg26492446	Island	-186;-186	3	7	15	
			63007236	cg27501878	Island	-200;-200	3	7	14	
17	*PEX5L*	3	180036815	cg02119363	Island	-44;112;116;145;221;239	3	7	15	None
			180036825	cg13473356	Island	-54;102;106;135;211;229	3 and 2	7	16	
			180036827	cg04894619	Island	-56;100;104;133;209;227	3 and 2	7	16	
18	*EVC2*	4	5708592	cg14654886	Island	957;-24;957;472	3 and 2	7	18	None
			5708610	cg27434509	Island	939;-42;939;454	3 and 2	7	19	
			5708645	cg17255450	Island	904;-77;904;419	3	6	18	
19	*FBXL7*	5	15500161	cg26134895	Island	-36;-1278	3	6	17	None
			15500605	cg06577205	Island	408;-834	3	7	18	
			15500724	cg14667871	Island	527;-715	3	7	15	
19a	-	5	73299798	cg03638905	Island	-	3	7	16	None
			73299858	cg12505170	Island	-	3	6	17	
			73299914	cg08893692	Island	-	3	7	17	
20	*SIM1*	6	100465030	cg21684012	Island	-1352;-100;-140	3	6	14	None
			100465064	cg27252696	Island	-1386;-134;-174	3	6	15	
			100465070	cg17380661	Island	-1392;-140;-180	3	6	15	
21	*HOXA10*	7	27164433	cg25188395	Island	15581;1098;1094;10748;6066;5305;5012	3	6	18	None
			27164730	cg03698009	Island	15284;801;797;10451;5769;5008;4715	3	6	14	
			27165044	cg20399871	Island	14970;487;483;10137;5455;4694;4401	3	7	15	
			27165109	cg26476852	Island	14905;422;418;10072;5390;4629;4336	3	7	17	
22	*VWC2*	7	49773437	cg04904331	Island	-225	3	6	15	None

			49773469	cg14045872	Island	-193	3 and 2	7	17	
			49773492	cg01893212	Island	-170	3 and 2	7	16	
			49773506	cg02467990	Island	-156	3 and 2	7	17	
			49773515	cg09493505	Island	-147	3 and 2	7	18	
			49773890	cg18206027	Island	228	3 and 2	7	18	
23	*VIPR2*	7	159145277	cg23572908	Island	-319;-417;-424	3	6	17	Deletion
			159145359	cg25189564	Island	-401;-499;-506	3	6	19	
			159145368	cg20673829	Island	-410;-508;-515	3	6	17	
24	*NKX2-6*	8	23706412	cg14428146	Island	187;-730	3	7	14	None
			23706457	cg15854847	Island	142;-685	3 and 2	7	17	
			23706495	cg05477514	Island	104;-647	3 and 2	7	17	
			23706518	cg13324546	Island	81;-624	3	7	16	
			23706643	cg03694713	Island	-44;-499	3	6	15	
			23706680	cg22747746	Island	-81;-462	3	7	16	
			23706781	cg10603004	Island	-182;-361	3	6	17	
25	*ADAM32*	8	39107507	cg22848598	Island	-143;-23;-65;-102;-450;-267;516	3	7	14	None
			39107603	cg26124318	Island	-47;73;31;-6;-354;-171;612	3	7	15	
			39107733	cg02637318	Island	83;203;161;124;-224;-41;742	3	7	15	
25a	-	8	56157129	cg02182795	Island	-	3	6	14	None
			56157348	cg11071231	Island	-	3	6	17	
			56157454	cg16504626	Island	-	3	6	14	
26	*BARHL2*	1	90718140	cg11823511	Island	-902	2	7	19	Deletion
			90718569	cg20311863	Island	-1331	2	7	17	
27	*CCNA1*	13	36431926	cg02478448	Island	-433;22;-598;405	2	7	18	None
			36431970	cg18348647	Island	-389;66;-554;449	2	7	17	
28	*NKX2-1*	14	36518096	cg23906738	Island	2133;1598;3054;2039;-1183;1426	2	7	17	None
			36518203	cg04347874	Island	2026;1491;2947;1932;-1076;1319	2	7	16	
29	*CLEC14A*	14	38255443	cg16404157	Island	927	2	7	16	None
			38255470	cg05057720	Island	900	2	7	17	
30	*SALRNA1*	14	60642089	cg19610529	N_Shore	2872	2	7	17	None
			60642489	cg23322933	Island	3272	2	7	16	
31	*LINC00925*	15	89379097	cg09969277	Island	17517;980;996;533;-28;965;16622;1047;9941;16609	2	7	17	Deletion
			89379114	cg09086835	Island	17534;997;1013;550;-11;982;16639;1064;9958;16626	2	7	20	
32	*HOXB-AS4*	17	48633979	cg01452847	Island	4429;5303;-406;-1419	2	7	21	None
			48634084	cg22851691	S_Shore	4534;5408;-511	2	7	20	
33	*AC010729.1*	2	5696049	cg26659079	Island	-172;3381	2	7	17	None
			5696099	cg18897632	Island	-122;3431	2	7	18	
34	*TLX2*	2	74515659	cg19656282	Island	1174;2195;2099	2	7	17	None
			74515882	cg07203423	Island	1397;2322	2	7	21	
35	*CPXM1*	20	2800595	cg22304612	Island	43	2	7	17	None
			2800616	cg07113642	Island	22	2	7	17	
35a	-	21	36693224	cg00495860	Island	-	2	7	16	None
			36693747	cg10445315	Island	-	2	7	16	
36	*NR1I2*	3	119810109	cg00836482	Island	27398;29624;29624;28073;-968	2	7	18	None
			119810301	cg25562664	Island	27590;29816;29816;28265;-776	2	7	16	
37	*ZIC1*	3	147419360	cg02232208	Island	25937	2	7	17	None
			147419393	cg15873149	Island	25970	2	7	18	
38	*UCHL1*	4	41256893	cg07068756	Island	59;-16;-54;123;-41;-40;-48;-19;-53;479;-165;-171	2	7	16	None
			41256918	cg16142306	Island	84;9;-29;148;-16;-15;-23;6;-28;504;-190;-196	2	7	16	
39	*BEND4*	4	42151678	cg02781618	Island	1201;1201;506	2	7	19	None
			42151691	cg24657817	Island	1188;1188;493	2	7	17	
40	OLIG3	6	137493591	cg01972751	Island	660	2	7	16	None
			137493823	cg12744820	Island	428	2	7	16	
41	*HOXA-AS3*	7	27152437	cg10739556	S_Shore	22459;5273;12072;4953;2384;145;125;162	2	7	17	None
			27152442	cg24398479	S_Shore	22464;5278;12077;4958;2389;140;120;157	2	7	18	
42	*NPTX2*	7	98616689	cg13314145	Island	-609	2	7	19	Amplification
			98616694	cg08315202	Island	-604	2	7	20	

*According to human genome assembly hg38.
**According to studies with reference number [27] and [28].

Supplementary Table 4. Conditions for PCR for bisulfite pyrosequencing

Gene symbol*	Primer	Primer sequence	Length (bp)	Annealing temprature (°C)	Sequencing primer	Sequencing primer sequence	Sequencing to analyze
SIM1	Forward	Biotine-GGTTTAGAGGGTAGTAAGATTTAGAGTT	334	54	Reverse	ACCAATAAAACTAAATAACA	CRAATCRACC CCRAACC
	Reverse	AACTACCCCCCCTAACTTCTTTATA					
MIR129-2	Forward	GGAGATAGAGGGATAGGATAGGTAG	274	54	Forward	AGGAGTGGTGAGATTGA	GTYGYGATGG AAYGYGTTGG GGAGATTTAG
	Reverse	ACCCTAAAACCAAACAAACTAAATC-Biotine					
NR1I2	Forward	TTTTTATTTTTTTATAGGAGGGTTATGA	492	54	Forward	TGTTTTTTTAGGAAGGG	YGTTTGYGAG ATTATT
	Reverse	CTACCCCCAAATATAATTTCAAACC-Biotine					
HOXB-AS4	Forward	GGAAAGATGTAAAAAATGGAGGTTAT	458	54	Forward	TGTAGGTGGAGGTTTTTA	GTTTTTGTYG GGYGYGGGTT GGGTTAG
	Reverse	AATAAAACTTCACCCTATTAATAAACTTCAA-Biotine					

*According to human genome assembly hg38.

Supplementary Table 5. *KRAS* mutations detected in 19 tumor samples and one non-tumurous sample

Tumor sample ID	Coverage	Variant frequencies (%)	Nucleotide change	Amino acid change
4T	13451	12.8	c.183A>T	p.Gln61His
9T1	7741	45.4	c.38G>A	p.Gly13Asp
9T2	16340	32.8	c.38G>A	p.Gly13Asp
12T	6869	37.6	c.35G>A	p.Gly12Asp
20T	4789	9.9	c.34G>C	p.Gly12Arg
26T	6360	6.7	c.34G>C	p.Gly12Arg
27T	7219	6.6	c.35G>A	p.Gly12Asp
28T1	6230	20.6	c.35G>A	p.Gly12Asp
28T2	7558	45.9	c.35G>A	p.Gly12Asp
35T1	6839	38.7	c.35G>A	p.Gly12Asp
35T2	6861	3.3	c.35G>A	p.Gly12Asp
35T3	5510	37.6	c.35G>A	p.Gly12Asp
48T	11683	16.0	c.35G>A	p.Gly12Asp
52T	9392	4.5	c.35G>T	p.Gly12Val
53T	7326	25.1	c.35G>A	p.Gly12Asp
60N	6460	4.2	c.35G>T	p.Gly12Val
60T	6249	13.8	c.35G>T	p.Gly12Val
62T	7330	15.8	c.35G>A	p.Gly12Asp
64T	1360	10.4	c.34G>C	p.Gly12Arg
75T	1600	17.4	c.35G>A	p.Gly12Asp

Supplementary Table 6. Somatic mutations detected in the two tumor samples without *KRAS* mutations but with substantial methylation levels

Tumor sample	Gene	Coverage	Variant frequencies (%)	Nucleotide change	Amino acid change
22T	*CDKN2A*	2558	15.5	c.45G>A	p.Trp15*
23T	*ALK*	3798	12.4	c.3707G>A	p.Gly1236Asp
	SMO	145	12.4	c.592G>T	p.Val198Phe
	CCND1	2239	11.3	c.674C>T	p.Ser225Phe
	CCND1	2279	11.2	c.677A>C	p.Tyr226Ser