Genome-wide analysis identifies critical DNA methylations within NTRKs genes in colorectal cancer

Zijian Chen (chenzj53@mail2.sysu.edu.cn)  
Sun Yat-sen University Sixth Affiliated Hospital  
https://orcid.org/0000-0002-2338-7258

Zenghong Huang  
Sun Yat-sen University Sixth Affiliated Hospital

Yanxin Luo  
Sun Yat-sen University Sixth Affiliated Hospital

Qi Zou  
Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, the Sixth Affiliated Hospital, Sun Yat-sen University

Liangliang Bai  
Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, the Sixth Affiliated Hospital, Sun Yat-sen University

Guannan Tang  
Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, the Sixth Affiliated Hospital, Sun Yat-sen University

Xiaolin Wang  
Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, the Sixth Affiliated Hospital, Sun Yat-sen University

Guangwen Cao  
Department of Epidemiology, Second Military Medical University

Meijin Huang  
Department of Colorectal Surgery, the Sixth Affiliated Hospital, Sun Yat-sen University

Jun Xiang  
Department of Gastrointestinal surgery, the Sixth Affiliated Hospital, Sun Yat-sen University

Huichuan Yu  
Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Disease, the Sixth Affiliated Hospital, Sun Yat-sen University, Guangzhou, China

Research

Keywords: NTRKs, NTRK3, Methylation, Colorectal Cancer, Prognosis

DOI: https://doi.org/10.21203/rs.3.rs-106027/v1

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

Background: Neurotrophic tropomyosin receptor kinases (NTRKs) are a gene family function as oncogene or tumor suppressor gene in distinct cancers. We aimed to investigate the methylation and expression profiles and prognostic value of NTRKs gene in colorectal cancer (CRC).

Methods: An analysis of DNA methylation and expression profiles in CRC patients was performed to explore the critical methylations within NTRKs genes. The methylation marker was validated in a retrospectively collected cohort of 229 CRC patients and tested in other tumor types from TCGA. DNA methylation status was determined by quantitative methylation-specific PCR (QMSP).

Results: The profiles in six CRC cohorts showed that NTRKs gene promoter was more frequently methylated in CRC compared to normal mucosa, which was associated with suppressed gene expression. We identified a specific methylated region within NTRK3 promoter targeted by cg27034819 and cg11525479 that best predicted survival outcome in CRC. NTRK3 promoter methylation showed independently predictive value for survival outcome in the validation cohort (P = 0.004, HR = 2.688, 95% CI = [1.355, 5.333]). Based on this, a nomogram predicting survival outcome was developed with a C-index of 0.705. Furthermore, the addition of NTRK3 promoter methylation improved the performance of currently-used prognostic model (AIC: 516.49 vs 513.91; LR: 39.06 vs 43.64, P = 0.032). Finally, NTRK3 promoter methylation also predicted survival in other tumors, including pancreatic cancer, glioblastoma and stomach adenocarcinoma.

Conclusions: This study highlights the essential value of NTRK3 methylation to evaluate prognosis risk, improve current prognostic models and the potential to guide treatment in CRC and other tumors.

Introduction

Colorectal cancer (CRC) is the second leading cause of cancer death over the past years (1, 2). Survival outcomes and optimal regimens vary in CRCs. At this point, some clinicopathological risk factors, such as TNM stage, tumor size, and tumor differentiation, have been used to stratify the risk of CRC death. Unfortunately, they fail to accurately distinguish patients with different outcomes (3), and several molecular biomarkers are being investigated and applied in current models to increase their prognostic values (4). CRC arises with the accumulation of gene mutations and epigenetic alterations (5, 6). Among them, the aberrant methylation in gene promoters is prevalent across multiple cancers, which can lead to the inactivation of tumor suppressor genes(7). Some of these aberrant methylations have been discovered and used to serve as prognostic biomarkers for CRC (8, 9).

Neurotropic tropomyosin receptor kinase (NTRK) gene family, including NTRK1, NTRK2 and NTRK3, encodes tropomyosin receptor kinases (TRK), which can induce cell proliferation, differentiation, apoptosis, and survival of neurons through the PI3K, RAS/MAPK/ERK and phospholipase C-gamma signalling transduction pathways (10, 11). The aberrations of NTRKs gene function were widely known to play an oncogenic role in multiple cancers. Among them, NTRKs gene fusion was the best-characterized aberration, which promotes tumorigenesis through the constitutive activation of downstream cell growth and proliferative pathways (11). The first TRK inhibitor, larotrectinib, has been approved by FDA for the treatment of advanced solid tumors with NTRKs gene fusion (12).

Similar to gene fusion, the aberrant expression of NTRKs gene is a critical event in cancers. NTRK1 promoted proliferation and metastasis of cancer cells and lead to poor prognosis in multiple cancers(13–17), while it suppressed cell proliferation in neuroblastoma (18). NTRK2 was shown to serve as an oncogene in multiple cancers(19–22), and its increased expression was associated with poor outcome(23, 24). Based on this, inhibition of NTRK2-encoded TRKB was shown to induce antitumor effects and cellular apoptosis (25, 26). Similar to NTRK1, NTRK3 has been demonstrated to
be an oncogene in breast cancer and gastric cancer (27, 28), but it acts as a tumor suppressor gene in CRC, neuroblastomas, and head and neck squamous cell carcinoma (10, 29, 30).

The expression of NTRKs gene can be modulated by promoter methylation. A hypermethylated NTRKs gene promoter is associated with suppressed expression in multiple cancers, such as CRC (10, 31), neuroblastoma (32), glioma (33), ovarian cancer (34) and prostate cancer (35). Given the critical role of NTRKs gene in multiple cancers, we therefore aimed to perform a comprehensive analysis of NTRKs gene on their methylation signature, expression profile and prognostic value in CRC using the methylation profile we previously established and the published dataset, and identify the optimal CpG site from NTRKs gene region as methylation biomarker that can be applied in the current clinical models of CRC to improve their prognostic values.

**Materials And Methods**

**Study cohorts**

In the discovery set, the clinical information and methylation profiles of CRC tissues and matched normal tissues we previously established using Illumina Infinium HumanMethylation450 (450K microarray) BeadChip (GSE48684 (31)) were combined with datasets of TCGA-COAD&READ (The Cancer Genome Atlas-colorectal carcinoma) cohort(36) and three Gene Expression Omnibus cohorts (GSE83889(37), GSE39582(38) and GSE87211(39)) to investigate methylation and expression profiles of NTRKs gene, and identify the critical CpG methylations within NTRKs genes in CRC. For the validation cohorts, we included 229 patients with histologically confirmed, stage I-IV CRC who underwent curative resection at the Sixth Affiliated Hospital of Sun Yat-sen University between 2009 and 2012. The patients were selected according to the exclusion criteria, including hereditary cancer, inflammatory bowel disease, and multiple primary cancers. Patients were treated and followed according to the NCCN (National Comprehensive Cancer Network) guideline-based institutional protocol as previously described. To avoid a potential effect of chemotherapy on genomic methylation status, patients received chemotherapy before curative resection when tissue sample was collected were excluded. The Institutional Review Board at the Sixth Affiliated Hospital of Sun Yat-sen University approved this study, and all the included patients have been given the written informed consent. To externally validate the findings in colorectal cancer and other cancers, we used the methylation array data and clinical information of 23 TCGA cohorts including a colon cancer cohort, a rectal cancer cohort and 21 cohorts of other tumor types, which was referred as external validation set. The cohort disposition for data analysis was illustrated in Fig. 1.

**Tissue Collection**

Formalin-fixed, paraffin-embedded (FFPE) CRC tissue specimens were available from the pathology registry. Hematoxylin and eosin (H&E) stained slides of cancer tissues were assessed by two experienced pathologists to confirm tumor histology based on the WHO classification of cancers and Fuhrman grade. Information of clinicopathological characteristics, such as age at diagnosis, gender, tumor size, tumor differentiation, and TNM stage, was extracted from the pathologic reports and medical records.

**DNA Isolation And Methylation Analysis**

Genomic DNA from FFPE tissue was isolated using the QIAamp DNA Mini Kit (Qiagen) and bisulfite-treated using according to the manufacturer's instructions. Gene methylation was analyzed by quantitative methylation-specific PCR (QMS) that was performed using an ABI Prism 7700 detection system (Biosystems) as previously described (10, 40). The oligo sequences of primers and probes were presented in Table S1. The primers and probe designed to target cg27034819 and cg11525479 were illustrated in Fig. 2E. Alu-C4 was used as a reference locus for normalization for input DNA following a previously described protocol. Each reaction mix was consisted of 0.45 μM primers, 0.15 μM probe, 1.5 mM MgCl2, 200 pM dNTPs, and 1.5 U HotStart Taq. The thermocycler conditions were as follows: 95 °C for 10
minutes followed by 45 cycles of 30 seconds at 94 °C, 30 seconds at 60 °C. All samples were run in duplicate in at least two independent reactions. The methylation status of each sample was determined using the Percentage of Methylated Reference (PMR) method(10). The PMR ≥ 50 and < 50 were defined as hypermethylation and hypomethylation respectively in QMSP assay.

Molecular Characterization Of CRC

The CpG Island Methylator Phenotype (CIMP) status and microsatellite instability (MSI) status of a subset of the colorectal neoplasms were assessed as described previously (41). The assessment of KRAS and BRAF mutation was performed in the Molecular Diagnostic Laboratory of the Sixth Affiliated Hospital of Sun Yat-sen University, as previously described (42).

Statistical analysis

All the statistical analyses were conducted using SPSS 20 or R 3.6.1. The NTRKs genes mRNA expression between CRC tissues and normal tissues were compared using student t-test. The Spearman correlation analysis was used to explore the relationship between promoter methylation and gene expression of NTRKs genes in TCGA cohort. The characteristics of CRC cases with and without NTRK3 promoter hypermethylation were compared using the Wilcoxon rank-sum tests or Chi-square tests. Kaplan-Meier curves and Log-rank tests were used to evaluate the prognostic factors of disease-free survival (DFS) among candidate variables. DFS was defined as the time from curative resection until local recurrence, distant metastasis, or last follow-up. In addition, Cox proportional hazards analyses were used to obtain HRs and corresponding 95% confidence intervals (CI) for the association between NTRK3 promoter hypermethylation and DFS. The variables that were considered clinically relevant or showed a significant difference (P < 0.05) in univariate Cox regression were entered into multivariate Cox proportional-hazards regression model using the backward stepwise selection method. Based on this model, nomograms subjected to internal validation set were generated for predicting 3-year and 5-year DFS outcomes, and the concordance index (C-index) was calculated to evaluate the predictive accuracy. The comparison between NTRK3 hypermethylation and known prognostic factors was assessed using likelihood ratio (LR) and Akaike information criterion (AIC) in competing models including or not including NTRK3 hypermethylation. In general, the model with a lower AIC and a higher LR is considered a better model. To validate these findings in colorectal and other cancers, the univariate Cox regression analysis of NTRK3 hypermethylation on survival outcome was re-performed in 23 TCGA cohorts. In TCGA dataset, the normalized β value of cg27034819 and cg11525479 (adjacent to cg27034819), targeting the downstream region of the promoter region in NTRK3, was extracted from 450K microarray and its association with survival outcome was analyzed. A P value < 0.05 was considered statistically significant with a two-tailed test.

Results

NTRKs gene was commonly suppressed by DNA methylation in CRC

Using the methylation profile we previously generated through 450K microarray, we found NTRKs gene, including NTRK1, NTRK2 and NTRK3, had more frequently methylated promoters in CRC samples when compared with matched normal mucosae (NTRK1, cancer = 0.444, normal = 0.397, P = 0.012; NTRK2, cancer = 0.251, normal = 0.167, P < 0.001; NTRK3, cancer = 0.395, normal = 0.144, P < 0.001; Fig. 2; Table S2). In support of methylation analysis results, we found the mRNA expression of NTRK2 and NTRK3 in CRC samples was commonly lower than that in normal mucosae using the expression profiles in five CRC cohorts (n = 1460). However, no significant difference was observed in NTRK1 mRNA expression (Fig. 2). Moreover, the lower expression of NTRKs gene is associated with their mean methylation of promoter (NTRK1, P = 0.049, R² = 0.011; NTRK2, P < 0.001, R² = 0.115; NTRK3, P < 0.001, R² = 0.091; Fig. 2).
A specific methylated region within *NTRK3* promoter best predicted CRC death

We next sought to identify the CpG site that could best predict CRC death and be feasibly used in a clinical assay. In the Cox proportional hazards analyses of 450K microarray probes targeting genomic loci within *NTRK*s gene, we found that the methylation of most CpG sites targeted by these probes was associated with poor survival outcomes in CRC (Fig. 2D, Table S3). Among them, the methylation of cg27034819 was top-ranked for predicting CRC death. Of note, we found the probe cg11525479 was very close to cg27034819 in their targeting loci (Fig. 2E), and the methylation of cg11525479 also had a predictive value for CRC death that was superior to most probes. These results suggested that the specific region within *NTRK3* promoter targeted by cg27034819 and cg11525479 could be used to stratify the death risk of CRC. Interestingly, the methylation of this specific region was shown to be associated with the loss of NT3-dependent tumor suppressor gene function of *NTRK3* in our previous in vitro and in vivo study (10). Therefore, a QMSP assay for determining the methylation of this specific region within *NTRK3* promoter was developed (Fig. 2E).

Cohort validation of *NTRK3* methylation for prognostic significance

We further validated the predictive value of this candidate region within *NTRK3* promoter in our institutional cohort. The baseline characteristics of this validation cohort were summarized in Table 1. *NTRK3* promoter hypermethylation was observed in 26 of 229 patients (11.35%), and it was more frequent in patients with MSI (*P* = 0.015; Table 1). Moreover, *NTRK3* promoter hypermethylation was associated with *KRAS* mutation (*P* = 0.001; Table 1). For other characteristics relevant to clinical outcomes of CRC, they did not show significant difference between patients with *NTRK3* promoter hypermethylation and hypomethylation, including age, sex, tumor size, tumor differentiation, lymphovascular invasion, perineural invasion, TNM stage, CIMP, MSI, *KRAS* mutation, *BRAF* mutation, Ki-67, CA19-9, and CEA.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>NTRK3 promotor methylation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>hypomethylation</td>
<td>hypermethylation</td>
</tr>
<tr>
<td><strong>Age (median = 62 years)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 62</td>
<td>111</td>
<td>100</td>
<td>11</td>
</tr>
<tr>
<td>≥ 62</td>
<td>118</td>
<td>103</td>
<td>15</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>100</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>Male</td>
<td>129</td>
<td>115</td>
<td>14</td>
</tr>
<tr>
<td><strong>Tumor localization</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>126</td>
<td>114</td>
<td>12</td>
</tr>
<tr>
<td>Rectum</td>
<td>103</td>
<td>89</td>
<td>14</td>
</tr>
<tr>
<td><strong>Tumor size (median = 4.5 cm)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 4.5</td>
<td>104</td>
<td>90</td>
<td>14</td>
</tr>
<tr>
<td>≥ 4.5</td>
<td>123</td>
<td>111</td>
<td>12</td>
</tr>
<tr>
<td>unknown</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Tumor differentiation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High or moderate</td>
<td>191</td>
<td>168</td>
<td>23</td>
</tr>
<tr>
<td>Poor</td>
<td>38</td>
<td>35</td>
<td>3</td>
</tr>
<tr>
<td><strong>Lymphovascular invasion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>210</td>
<td>188</td>
<td>22</td>
</tr>
<tr>
<td>Yes</td>
<td>17</td>
<td>13</td>
<td>4</td>
</tr>
<tr>
<td>unknown</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Perineural invasion</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>208</td>
<td>183</td>
<td>25</td>
</tr>
<tr>
<td>Yes</td>
<td>19</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>unknown</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TNM stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I-II</td>
<td>143</td>
<td>123</td>
<td>20</td>
</tr>
<tr>
<td>III-IV</td>
<td>85</td>
<td>79</td>
<td>6</td>
</tr>
<tr>
<td>unknown</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CIMP, CpG island methylator phenotype; KRAS, kirsten rat sarcoma viral oncogene; BRAF, B-Raf proto-oncogene, serine/threonine kinase; Ki-67, kiel67 antigen; CA19-9, carbohydrate antigen 19 − 9; CEA, carcinoembryonic antigen.
<table>
<thead>
<tr>
<th>Variables</th>
<th>Total</th>
<th>NTRK3 promotor methylation</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>hypomethylation</td>
<td>hypermethylation</td>
</tr>
<tr>
<td><strong>CIMP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>negative</td>
<td>222</td>
<td>198</td>
<td>24</td>
</tr>
<tr>
<td>Positive</td>
<td>7</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td><strong>Microsatellite status</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSS</td>
<td>125</td>
<td>113</td>
<td>12</td>
</tr>
<tr>
<td>MSI</td>
<td>51</td>
<td>39</td>
<td>12</td>
</tr>
<tr>
<td>unknown</td>
<td>53</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KRAS</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wild-type</td>
<td>119</td>
<td>110</td>
<td>9</td>
</tr>
<tr>
<td>mutation</td>
<td>62</td>
<td>46</td>
<td>16</td>
</tr>
<tr>
<td>unknown</td>
<td>48</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>BRAF</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>wild-type</td>
<td>174</td>
<td>151</td>
<td>23</td>
</tr>
<tr>
<td>mutation</td>
<td>8</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>unknown</td>
<td>47</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>KI-67</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 25%</td>
<td>100</td>
<td>89</td>
<td>11</td>
</tr>
<tr>
<td>&gt; 25%</td>
<td>84</td>
<td>72</td>
<td>12</td>
</tr>
<tr>
<td>unknown</td>
<td>45</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CA19-9</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 37</td>
<td>179</td>
<td>160</td>
<td>19</td>
</tr>
<tr>
<td>&gt; 37</td>
<td>32</td>
<td>29</td>
<td>3</td>
</tr>
<tr>
<td>unknown</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CEA</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤ 5</td>
<td>158</td>
<td>138</td>
<td>20</td>
</tr>
<tr>
<td>&gt; 5</td>
<td>57</td>
<td>54</td>
<td>3</td>
</tr>
<tr>
<td>unknown</td>
<td>14</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CIMP, CpG island methylator phenotype; KRAS, Kirsten rat sarcoma viral oncogene; BRAF, B-Raf proto-oncogene, serine/threonine kinase; Ki-67, kiel67 antigen; CA19-9, carbohydrate antigen 19–9; CEA, carcinoembryonic antigen.

In the Kaplan-Meier curve, significantly worse DFS outcomes were observed in patients with NTRK3 promotor hypermethylation compared to those with NTRK3 promotor hypomethylation (P = 0.012; Fig. 3A). The prognostic value of
NTRK3 promoter methylation status was further confirmed by univariate Cox proportional hazards ($P = 0.014$, HR = 2.194, 95% CI = [1.169, 4.117]; Table 2). Next, in the light of multivariate analysis, NTRK3 promoter hypermethylation was still a prognostic factor adjusted by age, TNM stage, and BRAF mutation ($P = 0.004$, HR = 2.688, 95% CI = [1.355, 5.333]; Table 2).
Table 2
Cox Proportional Hazard analyses on DFS in patients with CRC

<table>
<thead>
<tr>
<th>Variables</th>
<th>DFS in the patients with I-IV stage CRC</th>
<th>DFS in the patients with I-III stage CRC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Univariate</td>
<td>Multivariate</td>
</tr>
<tr>
<td></td>
<td>P value</td>
<td>HR[95%CI]</td>
</tr>
<tr>
<td>NTRK3 hypermethylation</td>
<td>0.014</td>
<td>2.194(1.169, 4.117)</td>
</tr>
<tr>
<td>Older age</td>
<td>0.011</td>
<td>1.952(1.168, 3.263)</td>
</tr>
<tr>
<td>Male (vs. female)</td>
<td>0.574</td>
<td>1.153(0.702, 1.894)</td>
</tr>
<tr>
<td>Rectal tumor (vs. colon tumor)</td>
<td>0.479</td>
<td>1.192(0.733, 1.939)</td>
</tr>
<tr>
<td>Tumor size ≥ 4.5</td>
<td>0.056</td>
<td>1.655(0.988, 2.772)</td>
</tr>
<tr>
<td>Poor differentiation</td>
<td>0.581</td>
<td>1.193(0.637, 2.233)</td>
</tr>
<tr>
<td>Lymphovascular invasion</td>
<td>0.052</td>
<td>2.087(0.995, 4.378)</td>
</tr>
<tr>
<td>Perineural invasion</td>
<td>0.350</td>
<td>1.455(0.663, 3.195)</td>
</tr>
<tr>
<td>Advanced TNM stage</td>
<td>0.003</td>
<td>2.125(1.300, 3.473)</td>
</tr>
<tr>
<td>CIMP positive status</td>
<td>0.265</td>
<td>1.935(0.607, 6.169)</td>
</tr>
<tr>
<td>MSI</td>
<td>0.316</td>
<td>1.350(0.751, 2.427)</td>
</tr>
<tr>
<td>KRAS mutation</td>
<td>0.113</td>
<td>1.566(0.899, 2.726)</td>
</tr>
<tr>
<td>BRAF mutation</td>
<td>0.010</td>
<td>3.376(1.338, 8.519)</td>
</tr>
<tr>
<td>Ki-67 &gt; 25%</td>
<td>0.458</td>
<td>0.806(0.456, 1.425)</td>
</tr>
<tr>
<td>CA19-9 &gt; 37</td>
<td>0.078</td>
<td>1.780(0.938, 3.377)</td>
</tr>
<tr>
<td>CEA &gt; 5</td>
<td>0.095</td>
<td>1.591(0.923, 2.741)</td>
</tr>
</tbody>
</table>

NTRK3, neurotropic tropomyosin receptor kinase 3; CIMP, CpG island methylator phenotype; MSI, microsatellite instability; KRAS, Kirsten rat sarcoma viral oncogene; BRAF, B-raf proto-oncogene, serine/threonine kinase; Ki-67, kie67 antigen; CA19-9, carbohydrate antigen 19–9; CEA, carcinoembryonic antigen.

In sensitivity analyses, NTRK3 promoter methylation was still independently associated with poor DFS outcome after the exclusion of patients with stage IV disease (P = 0.015, HR = 2.630, 95% CI = [1.206, 5.734]; Fig. 3B, Table 2). CIMP (P =
0.003, HR = 2.806, 95% CI = [1.432, 5.500]; Table S4), MSI (\(P = 0.008\), HR = 3.483, 95% CI = [1.391, 8.717]; Table S5), or BRAF mutation (\(P = 0.025\), HR = 2.603, 95% CI = [1.125, 6.022]; Table S6) in multivariate Cox analysis.

**A Nomogram For Predicting DFS In CRC Patients**

A nomogram for predicting 3-year and 5-year DFS outcome was generated using the variables from the multivariate Cox model, including \(NTRK3\) methylation, age at diagnosis, TNM stage, and \(BRAF\) mutation (Fig. 3C). The calibration curves for the nomogram were shown (Fig. 3D). The C-index of the nomogram for predicting DFS was 0.705.

**\(NTRK3\) methylation adds values to current prognostic panels**

The model 1 had a lower AIC and a higher LR compared with the model 2 (AIC: 597.73 vs 600.69; LR: 6.91 vs 5.95, \(P = 0.005\); Table 3), indicating that \(NTRK3\) hypermethylation alone is better in predicting prognosis than rough TNM staging alone. In the comparison between model 2 and 3, after \(NTRK3\) hypermethylation was added to TNM stage, a lower AIC and a higher LR were observed (AIC: 600.69 vs 592.41; LR: 5.95 vs 16.23, \(P = 0.002\); Table 3). These results suggest \(NTRK3\) hypermethylation could increase prognostic values of TNM staging.

**Table 3**

<table>
<thead>
<tr>
<th>Models</th>
<th>N</th>
<th>AIC</th>
<th>LR</th>
<th>(P) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model 1</td>
<td>219</td>
<td>597.73</td>
<td>6.91</td>
<td></td>
</tr>
<tr>
<td>Model 2</td>
<td>219</td>
<td>600.69</td>
<td>5.95</td>
<td>0.005(^a)</td>
</tr>
<tr>
<td>Model 3</td>
<td>219</td>
<td>592.41</td>
<td>16.23</td>
<td>0.002(^b)</td>
</tr>
<tr>
<td>Model 4</td>
<td>219</td>
<td>516.49</td>
<td>39.06</td>
<td></td>
</tr>
<tr>
<td>Model 5</td>
<td>219</td>
<td>513.91</td>
<td>43.64</td>
<td>0.032(^c)</td>
</tr>
</tbody>
</table>

N, patient counts in each model; AIC, Akaike information criterion value; LR, likelihood ratio.

\(^a\) \(P\) values for the LR test in model 1 compared with model 2; \(^b\) \(P\) values for the LR test in model 2 compared with model 3; \(^c\) \(P\) values for the LR test in model 4 compared with model 5.

To determine the values of \(NTRK3\) hypermethylation in commonly-used models using multiple clinicopathological variables, model 4 was built using the variables included in the model recommended by AJCC(43, 44). As expected, after \(NTRK3\) hypermethylation was included, model 5 had a lower AIC and a higher LR in comparison to model 4 (AIC: 516.49 vs 513.91; LR: 39.06 vs 43.64, \(P = 0.032\); Table 3). Thus, the model recommended by AJCC may get increased discriminatory ability in predicting prognosis with \(NTRK3\) hypermethylation.

**Prognostic significance of \(NTRK3\) methylation in multiple tumors**
Both the methylation of cg27034819 and cg11525479 were analyzed on their associations with survival outcome in 23 tumors using TCGA methylation profiles generated by 450K microarray. Overall, similar to the conflicting findings from in vitro and in vivo studies on NTRKs gene, the association of their methylations with survival outcome varied in different tumors. The methylation of cg27034819 was significantly associated with worse survival outcome in COAD ($P = 0.008$, $HR = 1.91$, 95% CI = [1.18, 3.09]), READ ($P = 0.006$, $HR = 3.73$, 95% CI = [1.37, 10.15]), KICH ($P = 0.005$, $HR = 7.36$, 95% CI = [1.83, 29.58]), and PAAD ($P = 0.001$, $HR = 1.96$, 95% CI = [1.30, 2.96]) cohorts. However, it was significantly associated with better survival outcome in GBM ($P = 0.037$, $HR = 0.63$, 95% CI = [0.41, 0.97]), SKCM ($P = 0.002$, $HR = 0.58$, 95% CI = [0.41, 0.81]) and STAD ($P = 0.013$, $HR = 0.57$, 95% CI = [0.37, 0.89]) cohorts (Fig. 3E). In cg11525479 methylation analysis, a similar predictive value for worse survival were found in COAD, READ, and PAAD, and a similar predictive value for better survival were found in GBM, SKCM and STAD (Fig. 3F). These results suggested a robust prognostic value of the methylation of the specific promoter region targeted by cg27034819 and cg11525479 in multiple tumors.

Discussions

Based on the merging of six cohorts, we found the NTRKs gene promoter was more frequently methylated in CRC compared to normal mucosa, which was associated with suppressed expression of NTRK2 and NTRK3. Through a screen of probes targeting NTRKs gene, we identified a specific methylated region within NTRK3 promoter targeted by cg27034819 and cg11525479 that was the most promising prognostic marker for CRC. We developed a QMSP assay to determine the methylation of this region that could be easily applied in clinical assay and validate its predictive value for survival outcome in a cohort of 229 CRC patients and 23 TCGA cohorts including a colon cancer cohort, a rectal cancer cohort and 21 cohorts of other tumor types. Using NTRK3 promoter methylation, age, TNM stage, and BRAF mutation, a novel nomogram predicting DFS outcome was developed and validated with a good prognostic performance. Also, we investigated the values of NTRK3 promoter methylation that added to current prognostic panels, in which we observed a meaningful performance improvement of AJCC model and TNM staging alone after the introduction of NTRK3 promoter methylation.

The conflicting findings in previous studies have revealed the complicated roles of NTRKs in different cancers. In our results, the mRNA expression of NTRK2 and NTRK3 was commonly lower in CRC samples in comparison to normal tissues, while the difference was not observed for NTRK1. In addition, this decreased expression of NTRKs was associated with promoter methylation. These results indicated that NTRK2 and NTRK3 may play a more important role of tumor suppressor in CRC, and methylation silencing of NTRK2 and NTRK3 would contribute more to CRC tumorigenesis. However, a decreased mRNA expression of NTRK1 was found and attributed to the methylated promoter in neuroblastoma and ovarian cancer (32, 34).

In the discovery set, we identified the CpG site targeted by cg27034819 as the most promising methylation biomarker for prognosis, and further assessed its prognostic value in other cohorts. To apply this high-throughput screen finding to a clinical assay in large cohorts of patients, we developed a PCR-based assay covering this genomic region that can be easily used with low cost. The robustness of this assay includes the finding that the probe cg11525479 that was very close to cg27034819 also had a prognostic value superior to most probes. The use of this assay in our previous in vitro study on NTRK3 also strengthen the reliability (10). Using this assay in our CRC cohort, we found that NTRK3 promoter hypermethylation was associated with worse DFS validation. This association maintained after the adjustment with all clinicopathological predictors. Furthermore, we revealed that NTRK3 promoter hypermethylation is highly associated with MSI and KRAS mutation that is known as response biomarkers for cancer treatment and have conflicting predictive value for survival (4). Therefore, we conducted sensitivity analyses excluding the patients with MSI or KRAS mutation. Consequently, the adverse impact of these molecular phenotypes on our methylation biomarker was minimal or not observed. Therefore, we are convinced that NTRK3 promoter methylation determined by the QMSP assay we developed was an independent prognostic factor in CRC.
We used AIC and likelihood-ratio test to compare the discriminatory ability of predictive models for survival outcome as previously described (45). First, \textit{NTRK3} methylation status is better than rough TNM stage (I, II, III) in predicting prognosis in CRC. Then, the addition of \textit{NTRK3} promoter methylation status in TNM stage and the AJCC models was shown to improve the predictive performance for DFS in CRC patients. Thus, \textit{NTRK3} promoter methylation is a valuable prognostic marker in CRC patients.

In previous studies, \textit{NTRK3} has been demonstrated to be an oncogene or a tumor suppressor gene in different cancer types (10, 27–30, 46). These conflicting findings from in vitro and in vivo studies is similar to the results of our cohort analyses on \textit{NTRK3} methylation. In our analyses, \textit{NTRK3} hypermethylation was associated with worse survival in some tumors, such as CRC, kidney chromophobe, and pancreatic adenocarcinoma, but it is related to a better outcome in other tumors, including glioblastoma multiforme, skin Cutaneous Melanoma, and stomach adenocarcinoma. \textit{NTRK3} promoter methylation could serve as a prognostic marker in multiple cancers, and its prognosis role is different in various cancers.

We have to admit that our study has some limitations. First, despite the multiple supports from discovery set, SYSU cohort and TCGA cohorts, external validation for \textit{NTRK3} promoter methylation and the new nomogram using the QMSP assay we developed would strengthen our findings. Next, although we excluded patients receiving chemo/radiotherapy before sample collection (neoadjuvant treatment) to avoid the confounding effect on genomic methylation and clonal selection, the retrospective nature of the SYSU cohort does not totally rule out other potential confounding factors.

**Conclusions**

We found \textit{NTRKs} gene was commonly suppressed by promoter methylation in CRC compared to normal mucosa. We identified the cg27034819-cg11525479 region within \textit{NTRK3} promoter as the most promising predictive marker for survival outcome, and it was validated in our CRC cohort and 23 TCGA cohorts including a colon cancer cohort, a rectal cancer cohort and 21 cohorts of other tumor types. A novel nomogram included \textit{NTRK3} promoter methylation and other independent predictors was developed. In addition, we observed a performance improvement of currently used prognostic models after the introduction of \textit{NTRK3} promoter methylation. These findings have essential implications for prognosis stratification in clinical decision-making for CRC management.

**Abbreviations**

ACC, Adrenocortical carcinoma

AIC, Akaike information criterion

BLCA, Bladder urothelial carcinoma

\textit{BRAF}, B-Raf proto-oncogene, serine/threonine kinase

BRCA, Breast invasive carcinoma

CA 19-9, Carbohydrate antigen 19-9

CEA, Carcinoembryonic antigen.

CI, Confidence intervals

CIMP, CpG island methylator phenotype

C-index, Concordance index
COAD, Colon adenocarcinoma
COADREAD, Colorectal adenocarcinoma
CRC, Colorectal cancer
DFS, Disease-free survival
ESCA, Esophageal carcinoma
FFPE, Formalin-fixed paraffin-embedded
FHCRC, Fred Hutchinson Cancer Research Center
GBM, Glioblastoma multiforme
HNSC, Head and Neck squamous cell carcinoma
HR, Hazard ratio
Ki-67, Kiel67 antigen
KICH, Kidney Chromophobe
KIRC, Kidney renal clear cell carcinoma
KRAS, Kirsten rat sarcoma viral oncogene
LAML, Acute Myeloid Leukemia
LGG, Brain Lower Grade Glioma
LIHC, Liver hepatocellular carcinoma
LR, Likelihood ratio
LUAD, Lung adenocarcinoma
LUSC, Lung squamous cell carcinoma
MESO, Mesothelioma
MSI, Microsatellite instability
NTRK1, Neurotropic tropomyosin receptor kinase 1
NTRK2, Neurotropic tropomyosin receptor kinase 2
NTRK3, Neurotropic tropomyosin receptor kinase 3
NTRKs, Neurotropic tropomyosin receptor kinases
PAAD, Pancreatic adenocarcinoma
PMR, Percentage of Methylated Reference
QMSP, Quantitative methylation-specific PCR
READ, Rectum adenocarcinoma
SARC, Sarcoma
SKCM, Skin Cutaneous Melanoma
STAD, Stomach adenocarcinoma
SYSU, Sun Yat-sen University
TCGA, The Cancer Genome Atlas
TRK, Tropomyosin receptor kinases
UCEC, Uterine Corpus Endometrial Carcinoma
UCS, Uterine Carcinosarcoma
UVM, Uveal Melanoma

Declarations

Ethics approval and consent to participate

Ethical approval was given by the Institutional Review Board at the Sixth Affiliated Hospital of Sun Yat-sen University. All the included patients have been given the written informed consent.

Consent for publication

Informed consent was obtained from all individual participants included in the study, giving their authorization to access their clinical information and tumor samples for research purpose.

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and publication.

Funding

This research was funded by the National Basic Research Program of China (No. 2015CB554001, JW), the National Natural Science Foundation of China (No. 81972245, YL; No. 81902877, HY), the Natural Science Fund for Distinguished Young Scholars of Guangdong Province (No. 2016A030306002, YL), the Natural Science Foundation of Guangdong Province (No. 2016A030310222, HY; No. 2018A0303130303, HY), Sun Yat-Sen University Clinical Research 5010 Program (No. 2018026, YL), the “Five Five” Talent Team Construction Project of the Sixth Affiliated Hospital Of Sun Yat-Sen University (No. P20150227202010244, JW; No. P20150227202010251, YL), the Program of Introducing Talents of Discipline to Universities, and National Key Clinical Discipline (2012).

Authors’ contributions
Collection of data, ZC, ZH, HY, QZ, LB, XW, GT, XF, MH; data analyses, ZC, HY; manuscript writing, ZC, ZH, HY, MH, JW, GC; Supervision, study designation and manuscript editing, HY, JX and YL.

Acknowledgements

We thank all patients for participation in this study.

References


Figures
Identification of critical methylated CpGs within NTRKs genes

DMP analyses to select colorectal cancer-specific CpGs within NTRKs genes:
FHCRC cohort (n=105)

Gene expression- and survival-associated CpGs selection:
5 CRC cohorts (n=1402)

Critical methylated CpGs within NTRKs genes:
cg27034819-cg11525479

Validation cohort:
SYSU cohort
(n = 229)

External cohort:
TCGA cohorts
(23 cancers)

Survival analysis

Pan-cancer analysis

Figure 1

The flow diagram of cohort analyses. The Fred Hutchinson Cancer Research Center (FHCRC) cohort was an in-house retrospective cohort with 64 colorectal cancer (CRC) samples and 41 normal colon samples. The DNA methylation profile of FHCRC cohort was used to identify the differentially methylated probes (DMPs) within NTRKs genes between tumor and normal samples. We further used the clinical information, DNA methylation and mRNA expression profiles of TCGA-COAD & READ, GSE83889, GSE39582 and GSE87211 cohorts that prospectively enrolled patients with CRC to screen the critical methylated CpG sites within NTRKs genes. The SYSU (Sun Yat-sen University) cohort was an in-house retrospective cohort of 229 patients diagnosed with CRC that was used to validate the identified critical methylated CpGs. The external cohort is a set of 23 tumor types from TCGA, which was used to demonstrate the significance of identified critical methylated CpGs in pan-cancer.
Figure 2

The methylation and expression profiles of NTRKs gene in CRC. A-C. The distribution and correlation of methylation and expression profiles of NTRK1, NTRK2 and NTRK3 gene in CRC and normal mucosa tissue. The heatmaps showed the results of probe-dimensional hierarchical clustering analysis of FHCRC cohort based on \( \beta \) values of all probes within NTRK1 (A, left), NTRK2 (B, left), and NTRK3 (C, left). Each row represented a probe and each column represented a CRC or normal mucosa sample in the heatmap. Probes targeting promoter region were annotated on heatmap left. The red lines annotated in the heatmap for NTRK3 indicated cg27034819 (top) and cg11525479 (bottom). NTRKs gene expression signature in each cohort showed suppressed mRNA expression of NTRK2 (B, middle) and NTRK3 (C, middle) in CRC tissues compared with normal mucosa tissue. Differences between CRC tissue and normal mucosa tissue were assessed with student t-test, and \( P \) values summarized with asterisks were shown in the upper space of the plot (***(P < 0.001)). The Spearman correlation analysis between mean promoter methylation of gene expression in each NTRK gene was shown in right panel. D. Volcano plot of statistical significance against hazard ratio for DFS among all CpG sites within NTRKs
gene targeted by 450K microarray probes. The top-ranked significantly probe was annotated according to the P values in the univariate Cox analyses. See supplementary tables for the full results of univariate Cox analysis relevant to this panel. E. A QMSP assay was developed using primers and probe targeting cg27034819-cg11525479 region.

Figure 3

Prognostic significance of NTRK3 promoter methylation. A-B. Kaplan-Meier curves of disease-free survival according to NTRK3 promoter methylation in CRC. Kaplan-Meier curves for the whole validation cohort (A) and stage I-III subgroup (B) were shown. The P value for each log-rank test was presented in the plots. Number at risk showed the quantity of CRC patients with NTRK3 hyper- or hypo- methylation and among them the quantity of survivors or dead patients respectively. C-D. A nomogram and calibration curve for predicting DFS in CRC. A nomogram to predict individual patient-level 3-year, and 5-year DFS based on clinicopathological risk factors and NTRK3 methylation (C). Calibration plots for the validation sample of the above nomogram (D). Actual DFS statue measured via Kaplan-Meier analysis is shown on the Y-axis, and the nomogram-predicted probability of DFS statue is shown on the X-axis. The average nomogram-predicted probability
of DFS was plotted against actually observed DFS estimated by Kaplan-Meier. 95% confidence intervals of the Kaplan-Meier estimates are indicated with vertical lines. Grayline indicates the reference line, showing where an ideal nomogram would lie. Instructions for users: Locate the status on each variable axis, and draw a straight line up to the Points axis to determine how many points toward risk the patient should receive from each variable. Sum the points and locate this number on the Total Points axis. Draw a straight line down from the total points to the 3-year or 5-year DFS Probability axis to ascertain the patient’s specific possibility of maintaining DFS until 3 or 5 years. E-F. Pan-cancer analysis of the prognostic significance of NTRK3 promoter methylation. The forest plots showed the values of the HR and CI for the prediction of the survival outcomes in univariate Cox analysis for the methylation of cg27034819 (E) and cg11525479 (F) in 23 TCGA cancer types. The x-axes presenting HRs were log2-scaled.

**Supplementary Files**

This is a list of supplementary files associated with this preprint. Click to download.

- AdditionalFile1SupplementaryTables.docx