ETV5 expression positively correlates with promoter methylation and predicts response for 5-FU-based adjuvant therapy response in proximal colon cancer

Anil K Giri,*

1 Institute for Molecular Medicine Finland (FIMM), University of Helsinki, FI-00290 Helsinki, Finland

*Corresponding author
Abstract

Discovery of markers predictive for 5-Fluorouracil (5-FU)-based adjuvant chemotherapy (adjCTX) response in patients with locally advanced stage II and III colorectal cancer (CRC) is necessary for early identification of potential responders as only 20-65% of CRC patients benefit from the treatment. PEA3 subfamily of ETS transcription factors (ETV1, ETV4, and ETV5) are upregulated in multiple cancers including colon cancers. However, the underlying epigenetic mechanism regulating their overexpression and their role in predicting therapy response in colon cancer is largely unexplored. In this study, using gene expression and methylation data from The Cancer Genome Atlas (TCGA) project, we showed that promoter DNA methylation negatively correlates with ETV4 expression ($\rho = -0.17$, $p=5.6x10^{-3}$) and positively correlates with ETV5 expression ($\rho = 0.22$, $p=1.43x10^{-4}$) in colon cancer tissue. Further, our analysis in 662 colon cancer patients treated with 5-FU-based-adjCTX revealed that higher ETV5 expression associated with shorter relapse-free survival (RFS) of treated patients with proximal tumors (Hazard ratio = 3.30 - 6.22, $p=0.005-0.02$). We also observed higher expression of signaling molecules involved in cellular proliferation (e.g. GNB5, DUSP4, FYN) in patients with high ETV5 level, suggesting that the increased cellular proliferation due to overexpression of these genes could drive the therapy resistance. The present study suggests ETV5 expression as a strong predictive biomarker for 5-FU-based adjCTX response in stage II/III CRC patients with proximal tumors.

Keywords: ETV5, ETV4, 5-Fluorouracil, colon cancer, GNB5, DNA methylation, PEA3
Introduction

Colon rectal adenocarcinoma (CRC) is the fourth most commonly diagnosed cancer (2 million cases in 2018) in the world and kills nearly 1 million people annually mostly due to the spread of tumor cells to other secondary organs (e.g. liver) in the later stage (stage III and IV) of the disease. Therefore, successful treatment of the locally advanced early-stage (stage I, II, and III) cancer is necessary in order to prevent disease progression and improve the overall survival of the patients. Usually, the early-stage patients are cured by surgical removal of the tumor only without the use of chemotherapy, however, systematic use of 5-Fluorouracil (5-FU)-based-adjCTX is recommended for stage II and III cases with high risk (e.g. with perineural invasion and poor histological differentiation) of reoccurrence. However, the use of adjCTX after surgery cures only 20% of additional stage III patients over surgery alone (cures 50% of cases) and improves the chance of 10-year overall survival only by 10%–20% in stage II patients. Further, it incurs considerable toxicity (e.g. myelosuppression, diarrhea) and economic costs to the patients. The higher toxicity and low efficacy of 5-FU-based adjCTX demand novel and reliable molecular markers that can predict the treatment response in early-stage (II and III) patients and help to stratify patients with different response.

Attempts to predict response for adjuvant chemotherapy have identified molecular alterations (e.g. microsatellites status, TP53 mutations, genetic polymorphism in MTHFR and DPYD) as a predictive marker for 5-FU-based adjCTX response. Further, recent studies exploring gene expression signatures as predictive markers for treatment response in colon cancer have identified ESR1, and CD8 expression as predictors for 5-FU-based adjCTX response in CRC. However, none of the identified markers can successfully segregate the responders from nonresponders before initiating the therapy indicating the need to search for additional novel markers predictive for adjCTX response.

Members of the polyoma enhancer activator 3 (PEA3) subfamily of E26 transformation-specific (ETS) domain-containing transcription factors which include ETV1, ETV4, and ETV5 promote cancer cell proliferation and survival in solid tumors including gastric, ovarian cancer, and colon cancers and are being targeted for therapy. However, the role of epigenetic mechanism especially DNA methylation in regulating their expression in colon cancer is largely unexplored. Further, the available evidence suggests that PEA3 subfamily members could be a potential biomarker for therapy response against CRC, and their role in predicting adjCTX response has not been explored in any cancers including CRC. The current study explores the role of DNA methylation in regulating PEA3 members gene expression using The Cancer Genome Atlas (TCGA) colon adenocarcinoma cohort. Further, we explored the PEA3 subfamily ETS transcription factors as
a predictive biomarker for 5-FU based adjCTX response in stage II and III CRC patients. Our analysis identified and validated ETV5 as a predictive marker for 5-FU-based adjCTX response in 662 stage II and II patients.

**Material and Method**

**Processing of gene expression and DNA methylation data from the TCGA cohort:** The expectation-maximization genes normalized RNA-Seq data for 328 colon adenocarcinoma patients samples (41 normal and 287 cancerous tissue) profiled in the TCGA project were downloaded using the Firehose tool (http://gdac.broadinstitute.org/). The data were further normalized using voom function in the limma package, and Z-transformed before the differential and correlation analyses. We also downloaded methylation data for 482,481 CpGs using Infinium HumanMethylation450 Beadchip for 38 normal and 297 cancerous tissue samples. Methylation status at a CpG site was measured as beta value (β), which is the ratio of the methylated probe intensity and the overall intensity (sum of methylated and unmethylated probe intensities designed for a particular CpG in 450K beadchip). β ranges from 0 to 1, indicating no methylation (β = 0) to complete methylation of the CpGs (β = 1). In all the analyses, we performed appropriate quality control of the published data before their downstream analysis as described previously. Briefly, we removed all the CpGs with missing values and CpGs assessed by probes that have a tendency of cross-hybridization, as specified in the supplementary file of Chen et al.23 We used BMIQ normalization to remove any possible bias due to design differences in the type of probes (the type I and type II probes) present in the Illumina 450K platform before averaging of probes in the promoter region. The average methylation of probes between 1,500 bases upstream of the transcription start site (TSS) was defined as promoter methylation level for a gene.

**Discovery cohort:** We downloaded the clinical and normalized gene expression profile of 477 stage II and III colon cancer patients collected under The French national Cartes d’Identité des Tumeurs (CIT) program from the Gene Expression Omnibus (GEO) platform (GSE39582). These primary patients with tumors have been treated with 5-Fluorouracil based adjuvant chemotherapy after surgery and monitored for relapse (distant and/or locoregional recurrence; median follow-up of 51.5 months). The recurrence-free survival (RFS) for the patients has been calculated as the time from surgery to the first recurrence. Patients have been staged according to the American Joint Committee on cancer tumor node metastasis (TNM) staging system. The location of the tumor has been noted as distal and tumor based on the anatomical location. The gene expression data have been generated on Affymetrix U133 Plus 2.0 chips and normalized using the robust multi-array average method.
implemented in the R package affy. Gene expression was summarized as the average expression levels of all the probes of the genes and was used for differential and survival analysis.

**Validation cohort:** For validation of the 5-FU-based adjCTX response prediction ability of ETV5, we downloaded the clinical and normalized gene expression data (GSE14333) for 185 CRC patients collected and published by Jorissen et al.\textsuperscript{27} from the GEO platform. The gene expression data have been collected from specimens derived from primary carcinomas and snap-frozen in liquid nitrogen immediately after surgery for storage at $-80^\circ$C. The patients have received standard adjuvant chemotherapy (either single-agent 5-fluorouracil/capecitabine or 5-fluorouracil and oxaliplatin) or postoperative concurrent chemoradiotherapy (50.4 Gy in 28 fractions with concurrent 5-fluorouracil) according to hospital protocols\textsuperscript{27}. The median follow-up for the patients was 37.2 months (range 0.9 to 118.6 months). Disease-free survival (DFS) was calculated as the duration from surgical operation to cancer recurrence, second cancer, or death from any cause. The grading for tumor stage has been determined using AJCC cancer staging manual\textsuperscript{20} and the position of the tumor has been noted as left, right, colon, and rectum. We categorized the cancers in the rectum and left colon as distal and in the right colon as proximal in origin.

**Differential expression analysis:**
5-FU-based-adjCTX treated patients with ETV5 expression $> 3$rd quartile was classified as ETV5-high group and with expression and $< 1$st quartile was classified as ETV5-low group. Differential expression analysis was performed using the Wilcoxon rank-sum test.

**Survival analysis:** The effect of average ETV5 over RFS or DFS was determined using the Cox regression analysis. The hazard ratio has been calculated as the exponential of the regression coefficient obtained from the fitted regression model. The significance of the model was tested using the log-rank test.

**Enrichment analysis:** The biological pathway enrichment for 22 genes against the human genome as the background was performed using Genecodis 4.0\textsuperscript{28} and FDR adjusted hypergeometric test p-values were used to identified enriched pathways.

**Statistical analysis:** All the statistical analysis has been performed using R version 3.5.3. The Non-parametric Wilcoxon rank test has been used to compare the gene expression between two groups. The survival analysis has been performed using Cox-proportional regression as implemented in the “survival” package and the survival plots have been drawn using “ggplot” and “GGally” package in R.

**Results**
Cancer tissue has higher expression of ETV4 and ETV5 genes that correlate with promoter methylation in CRC patients. First, we compared the expression level of ETV1, ETV4, and ETV5 between the normal and cancerous tissue in TCGA data and observed higher expression of only ETV4 (Wilcoxon test $P=4.90\times10^{-25}$) and ETV5 (Wilcoxon test $P=5.17\times10^{-9}$) in tumor suggesting their possible role in tumor biology (Figure 1A, and B). Further, we checked the role of promoter DNA methylation over expression of ETV4 and ETV5 in the cancer tissue as promoter hypermethylation can shutdown gene expression. Our analysis revealed that promoter DNA methylation nominally differs between cancer and normal tissue for both ETV4 and ETV5 (Figure 1C, and D). Further analysis revealed that promoter methylation negatively correlates with ETV4 expression ($\rho=-0.17$, $p=5.6\times10^{-3}$) whereas positively correlates with ETV5 expression ($\rho=0.22$, $p=1.43\times10^{-4}$) in cancer tissue suggesting that DNA methylation play a strong role in regulating ETV5 and ETV4 expression in colon cancer tissue (Figure 1E, and F).
Figure 1: DNA methylation regulates ETV4 and ETV5 expression in colon cancer. ETV4 (A) and ETV5 (B) are overexpressed in cancerous tissue compared to healthy tissue in TCGA samples. Scatter plot showing promoter methylation difference between cancer and normal tissue for ETV4 (C) and ETV5 (D) genes. The p-value has been calculated using the Wilcoxon nonparametric test. Scatterplot showing spearman’s correlation between promoter methylation and expression levels of ETV4 (C) and ETV5 (D) genes.
and ETV5 (D) in the cancerous tissue of CRC patients from TCGA data. The Spearman correlation coefficient and respective p-values have been shown in the figure.

**ETV5 expression differs in CRC patients with proximal and distal tumors**

After observing that ETV4 and ETV5 are overexpressed in colon cancerous tissue from TCGA cohort, we studied their role in predicting adjCTX response in CRC using two publicly available datasets. Since the location of the tumor can affect the response to chemotherapy, first we checked the expression level of ETV4 and ETV5 in stage II/III CRC patients (both treated and untreated) with proximal and distal tumors. We observed a significant difference in ETV4 expression between distal and proximal tumors only in discovery cohort samples (Figure 2A) but not in validation cohort (Figure 2B) whereas ETV5 expression differed both in discovery (Wilcoxon test P=8.05x10^-5) and validation cohorts (Wilcoxon test P=0.005) (Figure 1B and C). The result suggests a tumor-side specific role of ETV5 in CRC. Further, in order to find the effect of ETV5 over disease progression, we compared the ETV5 expression level between adjCTX-treated patients with stage II and III and found no consistent significant difference across disease stages (Suppl. figure 1) in the discovery and validation cohorts (Suppl. Table 1).
Figure 2: Comparison of ETV4 (A and B) and ETV5 (C and D) expression in distal and proximal tumors of combined stage II and III CRC patients from discovery (left side), and validation (right side) cohort. The P-value has been calculated using the Wilcoxon test.

**ETV5 correlates with disease free survival of stage II patients treated with adjuvant**

In order to assess the effect of ETV4 and ETV5 expression over chemotherapy response, we used Cox-regression with RFS as the treatment outcome and observed that higher ETV5 expression is a strong and selective predictor for poor RFS (HR= 2.29, P=0.00178) in treated patients in the discovery cohort (Figure 3A). Further, stratified analysis revealed that ETV5 expression significantly predicts for RFS of patients with both distal (HR=2.21, P=0.01) and proximal tumors (HR= 3.30, p=0.02) in the adjCTX treated patients in the discovery cohort (Figure 3B and C). Furthermore, survival analysis in 85 stages II/III adjCTX treated patients from the validation cohort revealed that ETV5 expression significantly associates with poor RFS in patients with proximal tumors (HR=6.22, P=5.77x10^-3) but not with distal tumor (HR=1.21, P=0.57) (Figure 3D, E, and F). We also checked the effect of tumor location (proximal or distal) over RFS of adjCTX-treated patients and did not observe a significant effect in both the discovery (HR=1.03, P=0.92) and validation (HR=0.68, P=0.06) datasets (Suppl. Fig. 2). The results suggest that the association of ETV5 with poor RFS in 5-FU based adjCTX treated patients with proximal tumors is not confounded by the survival difference due to the location of the tumor. We did not observe a significant association of ETV4 with RFS of patients either with distal or proximal tumors.
Figure 3: Effect of ETV5 expression over RFS in 5-FU-based adjCTX treated CRC patients. Kaplan–Meier survival curves for the probability of RFS of treated-CRC patients with low and high ETV5 has been shown for GSE39582- combined (A), distal, (B) proximal (C) tumors. Similarly, the probability for RFS of CRC patients in GSE14333 cohort has been shown for combined (D), distal (E), and proximal (F) tumors. The hazard ratio and p-value have been calculated using cox-regression and has been shown in the figure. For visualization purpose, patients with the ETV5 expression > 3rd quartile has been categorized as high (red) and with expression < 1st quartile has been categorized as low (low). The log-rank analysis was used to test for significance between the survival curves.

Higher ETV5 level associated with genes involved in cell differentiation

After observing that higher ETV5 expression correlates with shorter RFS in 5-FU-based adjCTX treated patients with proximal tumors. We explored the possible mechanism by which ETV5 may induce poor response for therapy. A differential analysis between adjCTX treated patients with higher (the 4th quartile) and lower (the 1st quartile) ETV5 expression revealed a strong difference (P < 10^-s) in the expression of 22 genes enriched in the cell differentiation process (Suppl. Table 2) in both
cohorts. The genes included signaling molecules (GNB5, DUSP4, FYN), and drug transporter (ABCA3) and 5-fluorouracil target genes (e.g. ACE2) that have been already implicated in the resistance of 5-fluorouracil and other chemotherapies (Figure 4). The result suggests that ETV5 may induce drug resistance by activating signaling molecules and transcription factors involved in cell proliferation.

Figure 4: Higher ETV5 expression may induce therapy-resistance by higher expression of cell differentiation-associated genes. GNB5 (A), DUSP4 (B), FYN (C) and ABCA3 (D) expression level in patients with high (>3rd quartile) and low (<1rd quartile) ETV5 expression in GSE39582 (left) and GSE14333 (right). P-value has been calculated using Wilcoxon nonparametric test.

Discussion

5-FU-based adjCTX after surgery is the primary choice of treatment in early-stage colon cancer due to survival advantage over surgery alone. However, higher toxicity and inability to segregate
responders from nonresponders using available markers is a notable challenge to the clinical success
of therapy. Therefore, there is an urgent need to identify additional predictive markers for the 5FU-
based-adjCTX response in CRC. To address the need, we investigated the expression regulation of
members of the PEA3 subfamily of ETS transcription factors by DNA methylation and the potential
of their expression as a predictive marker for chemotherapy response in stage II and III colorectal
cancer patients using 2 publicly available independent colon cancer datasets.
Our analysis revealed higher expression ETV4 and ETV5 genes in colon cancer tissue compared to
normal tissue in TCGA samples (Figure 1A and B). Further, ETV4 expression showed a negative
correlation with promoter methylation (Figure 1E) suggesting that DNA methylation mediates
expression by hindering transcription factor binding36. Additionally, we observed a paradoxical
positive correlation between ETV5 promoter hypermethylation and expression in cancer tissue
suggesting that hypermethylation may facilitate gene expression either by the opening of chromatin37
or mechanical inhibition of transcriptional repressor binding38 or allowing transcription from an
alternative promoter39. To our knowledge, there has been no study exploring the methylation-
expression relation of PEA3 member proteins in colon cancers, and the results need to be validated
in independent CRC cohorts. Further, a detailed mechanistic study in higher experimental model
systems is needed on how the promoter hypermethylation increases the gene expression.
Additionally, we explored the role of ETV4 and ETV5 in 5-FU-based-adjCTX response prediction
utilizing two publicly available CRC patient cohort. We observed higher expression of ETV5 in the
proximal CRC tumor as compared to the distal tumor (Figure 2C and D) in both cohorts which is in
accordance with more aggressive and high-grade histology of proximal tumors compared to
distal tumors40 as higher ETV5 expression has been associated with faster cell proliferation and
aggressive phenotypes40. Further, ETV5 overexpression stimulates CRC angiogenesis through
activation of VGFR by PDGFR-β/Src/STAT3 signaling18 and increases bevacizumab resistance
through AKT, ERK, and p38 signaling decreasing overall survival of CRC patients18. Survival
analysis revealed that higher ETV5 expression significantly associated with shorter RFS in proximal
colon cancer in stage II and III patients (Figure 3). Further, patients with higher ETV5 levels also had
higher GNB5 expression (Figure 4A) which is a downstream Gbg subunit of G-protein coupled
receptor (GPCR)30. GNB5 transduces the signal to various downstream effectors including RAS-
MEK-ERK and PI3K-AKT pathways leading to cellular proliferation and viability30. GNB5 has
emerged as the major target to overcome cetuximab resistance in colorectal cancer30. Furthermore,
we observed an expression difference of DUSP4 in patients with a high level of ETV5. DUSP4 is a
MAPK phosphatase and its role has been contradictory as downregulation of DUSP4 enhances cell
proliferation and invasiveness in colorectal carcinomas41 but inhibits growth in human colorectal
cancer cell lines. These results suggest that higher ETV5 expression may induce drug resistance by upregulation of genes involved in colon cancer proliferation and growth. However, studies in larger human cohort and animal model systems can fully explain the detailed mechanism for the ETV5 role in 5-FU-based adjCTX resistance.

The current study identified ETV5 as a biomarker of 5-FU-based adjCTX response in CRC patients with evidence II level as defined by Simon et al., and revealed that higher ETV5 is associated with poor response in proximal CRC tumors. These results suggest that ETV5 could be useful for the identification of responders before administration 5-FU-based adjCTX when included along with other already established clinicopathological markers.

**Funding**

None

**Author contribution**

AKG conceptualized the study, analyzed the data, and wrote the manuscript.

**Acknowledgment**

The author is very much thankful to Prof. Tero Aittokallio and Aleksandr Ianevski for their valuable suggestions during the analysis and editing of the manuscript.

**Reference**

1. Arnold, M. *et al.* Global patterns and trends in colorectal cancer incidence and mortality. *Gut* 66, 683-691 (2017).
2. Yu, Y. *et al.* The long-term survival characteristics of a cohort of colorectal cancer patients and baseline variables associated with survival outcomes with or without time-varying effects. *BMC Med* 17, 150 (2019).
3. Argiles, G. *et al.* Localised Colon Cancer: ESMO Clinical Practice Guidelines for Diagnosis, Treatment and Follow-up. *Annals of Oncology* (2020).
4. Casadaban, L. *et al.* Adjuvant chemotherapy is associated with improved survival in patients with stage II colon cancer. *Cancer* 122, 3277-3287 (2016).
5. Auclin, E. *et al.* Subgroups and prognostication in stage III colon cancer: future perspectives for adjuvant therapy. *Ann Oncol* 28, 958-968 (2017).
6. Breugom, A.J. *et al.* Adjuvant chemotherapy after preoperative (chemo)radiotherapy and surgery for patients with rectal cancer: a systematic review and meta-analysis of individual patient data. *The Lancet Oncology* 16, 200-207 (2015).
7. Giri, A.K. & Aittokallio, T. DNMT Inhibitors Increase Methylation in the Cancer Genome. *Frontiers in Pharmacology* 10(2019).
8. Giri, A.K., Ianevski, A. & Aittokallio, T. Genome-wide off-targets of drugs: risks and opportunities. *Cell Biol Toxicol* **35**, 485-487 (2019).

9. Ribic, C.M., *et al.* Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N Engl J Med* **349**, 247-257 (2003).

10. Kandioler, D. *et al.* TP53 Mutational Status and Prediction of Benefit from Adjuvant 5-Fluorouracil in Stage III Colon Cancer Patients. *EBioMedicine* **2**, 825-30 (2015).

11. Nahid, N.A. *et al.* DPYD*2A and MTHFR C677T predict toxicity and efficacy, respectively, in patients on chemotherapy with 5-fluorouracil for colorectal cancer. *Cancer Chemother Pharmacol* **81**, 119-129 (2018).

12. Hariprakash, J.M. *et al.* Pharmacogenetic landscape of DPYD variants in south Asian populations by integration of genome-scale data. *Pharmacogenomics* **19**, 227-241 (2018).

13. Henricks, L.M., Opdam, F.L., Beijnen, J.H., Cats, A. & Schellens, J.H.M. DPYD genotype-guided dose individualization to improve patient safety of fluoropyrimidine therapy: call for a drug label update. *Ann Oncol* **28**, 2915-2922 (2017).

14. Ye, S.B. *et al.* The Predictive Value of Estrogen Receptor 1 on Adjuvant Chemotherapy in Locally Advanced Colorectal Cancer: A Retrospective Analysis With Independent Validation and Its Potential Mechanism. *Front Oncol* **10**, 214 (2020).

15. Allen, W.L. *et al.* Transcriptional subtyping and CD8 immunohistochemistry identify poor prognosis stage II/III colorectal cancer patients who benefit from adjuvant chemotherapy. *JCO Precis Oncol* **2018** (2018).

16. Oh, H.H. & Joo, Y.E. Novel biomarkers for the diagnosis and prognosis of colorectal cancer. *Intest Res* **18**, 168-183 (2020).

17. Keld, R. *et al.* PEA3/ETV4-related transcription factors coupled with active ERK signalling are associated with poor prognosis in gastric adenocarcinoma. *Br J Cancer* **105**, 124-30 (2011).

18. Llaurado, M. *et al.* ETV5 transcription factor is overexpressed in ovarian cancer and regulates cell adhesion in ovarian cancer cells. *Int J Cancer* **130**, 1532-43 (2012).

19. Cheng, X. *et al.* ETS variant 5 promotes colorectal cancer angiogenesis by targeting platelet-derived growth factor BB. *Int J Cancer* **145**, 179-191 (2019).

20. Hsing, M., Wang, Y., Rennie, P.S., Cox, M.E. & Cherkasov, A. ETS transcription factors as emerging drug targets in cancer. *Med Res Rev* **40**, 413-430 (2020).

21. Ritchie, M.E. *et al.* limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, e47 (2015).
22. Giri, A.K. et al. DNA methylation profiling reveals the presence of population-specific
signatures correlating with phenotypic characteristics. *Molecular Genetics and Genomics* **292**, 655-662 (2017).

23. Chen, Y.A. et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina
Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203-9 (2013).

24. Teschendorff, A.E. et al. A beta-mixture quantile normalization method for correcting probe
design bias in Illumina Infinium 450 k DNA methylation data. *Bioinformatics* **29**, 189-96 (2013).

25. Marisa, L. et al. Gene expression classification of colon cancer into molecular subtypes:
characterization, validation, and prognostic value. *PLoS Med* **10**, e1001453 (2013).

26. American Joint Committee on Cancer (1997) AJCC cancer staging manual, 5th edition.
Philadelphia: Lippincott-Raven.

27. Jorissen, R.N. et al. Metastasis-Associated Gene Expression Changes Predict Poor Outcomes
in Patients with Dukes Stage B and C Colorectal Cancer. *Clin Cancer Res* **15**, 7642-7651 (2009).

28. Tabas-Madrid, D., Nogales-Cadenas, R. & Pascual-Montano, A. GeneCodis3: a non-
redundant and modular enrichment analysis tool for functional genomics. *Nucleic Acids Res* **40**, W478-83 (2012).

29. Stintzing S, Tejpar S, Gibbs P, Thiebach L, Lenz HJ. Understanding the role of primary tumour
localisation in colorectal cancer treatment and outcomes. *Eur J Cancer* **84**, 69-80 (2017).

30. Park, S.M. et al. Systems analysis identifies potential target genes to overcome cetuximab
resistance in colorectal cancer cells. *FEBS J* **286**, 1305-1318 (2019).

31. Du, F. et al. miR-137 alleviates doxorubicin resistance in breast cancer through inhibition of
epithelial-mesenchymal transition by targeting DUSP4. *Cell Death Dis* **10**, 922 (2019).

32. Chen, J., Elfiky, A., Han, M., Chen, C. & Saif, M.W. The role of Src in colon cancer and its
therapeutic implications. *Clin Colorectal Cancer* **13**, 5-13 (2014).

33. Sribenja, S., et al. Thymosin β10 as a predictive biomarker of response to 5-fluorouracil
chemotherapy in cholangiocarcinoma. *Ann Hepatol.* **15**, 577-585(2016).

34. Xie, B. et al. The anti-cancer drug 5-fluorouracil affects cell cycle regulators and potential
regulatory long non-coding RNAs in yeast. *RNA Biol* **16**, 727-741 (2019).

35. Wilkinson, N.W. et al. Long-term survival results of surgery alone versus surgery plus 5-
fluorouracil and leucovorin for stage II and stage III colon cancer: pooled analysis of NSABP
C-01 through C-05. A baseline from which to compare modern adjuvant trials. *Ann Surg Oncol* **17**, 959-66 (2010).
36. Esteller, M. Epigenetic gene silencing in cancer: the DNA hypermethylome. *Hum Mol Genet* **16** Spec No 1, R50-9 (2007)

37. Smith, J., Sen, S., Weeks, R.J., Eccles, M.R. & Chatterjee, A. Promoter DNA Hypermethylation and Paradoxical Gene Activation. *Trends Cancer* **6**, 392-406 (2020).

38. Nabilsi, N.H., Broaddus, R.R. & Loose, D.S. DNA methylation inhibits p53-mediated survivin repression. *Oncogene* **28**, 2046-50 (2009).

39. Renaud, S. *et al.* Expression of the CTCF-paralogous cancer-testis gene, brother of the regulator of imprinted sites (BORIS), is regulated by three alternative promoters modulated by CpG methylation and by CTCF and p53 transcription factors. *Nucleic Acids Res* **35**, 7372-88 (2007).

40. Bazzocco, S. *et al.* Highly Expressed Genes in Rapidly Proliferating Tumor Cells as New Targets for Colorectal Cancer Treatment. *Clin Cancer Res* **21**, 3695-704 (2015).

41. Ichimanda, M. *et al.* Downregulation of dual-specificity phosphatase 4 enhances cell proliferation and invasiveness in colorectal carcinomas. *Cancer Sci* **109**, 250-258 (2018).

42. Ratsada, P. *et al.* DUSP4 is involved in the enhanced proliferation and survival of DUSP4-overexpressing cancer cells. *Biochemical and Biophysical Research Communications* **528**, 586-593 (2020).

43. Simon, R.M., Paik, S. & Hayes, D.F. Use of archived specimens in evaluation of prognostic and predictive biomarkers. *J Natl Cancer Inst* **101**, 1446-52 (2009).
Suppl. Table 1: Clinical characteristic of studied patients

| Traits                          | Discovery (GSE39582) | Validation (GSE14333) |
|--------------------------------|----------------------|------------------------|
|                                | Treated (210)        | Untreated (267)        | Treated (85) | Untreated (100) |
| Age (<60/>=60)                 | 109/101              | 64/203                 | 38/47        | 19/81          |
| Sex (M/F)                      | 116/94               | 152/115                | 45/40        | 53/47          |
| Location (distal/proximal)     | 137/73               | 144/123                | 46/39        | 54/46          |
| Stage (II/III)                 | 58/152               | 210/57                 | 22/63        | 72/28          |

Suppl. Table 2: Biological process enrichment analysis for 22 common genes differentially expressed in ETV5 high vs low patients in both discovery and validation cohort.

| SN | Process                          | Hyp_pval_adj | Genes                                                                 |
|----|----------------------------------|--------------|----------------------------------------------------------------------|
| 1  | Cell differentiation             | 2.18x10-4    | CDX2, SOX13, FYN, MECOM, ETV5, ETV1, DPYSL2                          |
| 2  | Multicellular organism development | 2.11x10-3   | CEBPA, CDX2, FYN, MECOM, DPYSL2, ANO1                                |
| 3  | Negative regulation of transcription | 1.57x10-4 | CEBPA, CDX2, RUNX1                                                   |
| 4  | Response to drug                 | 6.02x10-4    | FYN, ABCA3, DPYSL2                                                   |

FDR corrected p-value for hypogeometric test for enrichment has been shown in the table.
Suppl. Fig. 1: Comparison of ETV5 expression across different stages of relapse (coded as 1) and non-relapse (coded as 0) patients in adjCTX treated patients from GSE39582(A) and GSE1433 cohort (B). P-values have been calculated using Wilcoxon nonparametric test.
Suppl. Fig. 2: Effect of tumor location on the survival of the patients treated with 5-FU-based adjCTX in GSE39582 (A) and GSE14333 (B) cohort. The hazard ratio and p-value from the cox-regression has been shown in the figure.