

RESEARCH ARTICLE

# Association of SMAD4 mutation with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer

Amir Mehrvarz Sarshekeh, Shailesh Advani, Michael J. Overman, Ganiraju Manyam, Bryan K. Kee, David R. Fogelman, Arvind Dasari, Kanwal Raghav, Eduardo Vilar, Shanequa Manuel, Imad Shureiqi, Robert A. Wolff, Keyur P. Patel, Raja Luthra, Kenna Shaw, Cathy Eng, Dipen M. Maru, Mark J. Routbort, Funda Meric-Bernstam, Scott Kopetz\*

Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, United States of America

\* [skopetz@mdanderson.org](mailto:skopetz@mdanderson.org)



**OPEN ACCESS**

**Citation:** Sarshekeh AM, Advani S, Overman MJ, Manyam G, Kee BK, Fogelman DR, et al. (2017) Association of SMAD4 mutation with patient demographics, tumor characteristics, and clinical outcomes in colorectal cancer. PLoS ONE 12(3): e0173345. doi:10.1371/journal.pone.0173345

**Editor:** Surinder K. Batra, University of Nebraska Medical Center, UNITED STATES

**Received:** May 12, 2016

**Accepted:** February 20, 2017

**Published:** March 7, 2017

**Copyright:** © 2017 Sarshekeh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** A minimal dataset for all of our patients with Colorectal cancer has been uploaded as "Supporting information". This dataset was used for survival analysis based on mutational status. Per PLOS ONE guideline, age, gender and race were not included as the information would compromise patient confidentiality. Please note that per PLOS ONE guideline, dataset for the analysis of response to anti EGFR treatment was not provided as our patient population for this analysis was very small (N=44).

## Abstract

SMAD4 is an essential mediator in the transforming growth factor- $\beta$  pathway. Sporadic mutations of SMAD4 are present in 2.1–20.0% of colorectal cancers (CRCs) but data are limited. In this study, we aimed to evaluate clinicopathologic characteristics, prognosis, and clinical outcome associated with this mutation in CRC cases. Data for patients with metastatic or unresectable CRC who underwent genotyping for SMAD4 mutation and received treatment at The University of Texas MD Anderson Cancer Center from 2000 to 2014 were reviewed. Their tumors were sequenced using a hotspot panel predicted to cover 80% of the reported SMAD4 mutations, and further targeted resequencing that included full-length SMAD4 was performed on mutated tumors using a HiSeq sequencing system. Using The Cancer Genome Atlas data on CRC, the characteristics of SMAD4 and transforming growth factor- $\beta$  pathway mutations were evaluated according to different consensus molecular subtypes of CRC. Among 734 patients with CRC, 90 (12%) had SMAD4 mutations according to hotspot testing. SMAD4 mutation was associated with colon cancer more so than with rectal cancer (odds ratio 2.85;  $p < 0.001$ ), female sex (odds ratio 1.71;  $p = 0.02$ ), and shorter overall survival than in wild-type SMAD4 cases (median, 29 months versus 56 months; hazard ratio 2.08;  $p < 0.001$  [log-rank test]). SMAD4 mutation was not associated with age, stage at presentation, colonic location, distant metastasis, or tumor grade. A subset of patients with metastatic CRC ( $n = 44$ ) wild-type for KRAS, NRAS, and BRAF who received anti-epidermal growth factor receptor therapy with mutated SMAD4 ( $n = 13$ ) had shorter progression-free survival duration than did patients wild-type for SMAD4 ( $n = 31$ ) (median, 111 days versus 180 days;  $p = 0.003$  [log-rank test]). Full-length sequencing confirmed that missense mutations at R361 and P356 in the MH2 domain were the most common SMAD4 alterations. In The Cancer Genome Atlas data, SMAD4 mutation frequently occurred with KRAS, NRAS, and BRAF mutations and was more common in patients with the consensus molecular subtype 3 of CRC than in those with the other 3 subtypes. This is one of the largest retrospective studies to date characterizing SMAD4 mutations in CRC patients and demonstrates the

**Funding:** This study was supported in part by the Sheikh Khalifa Al Nahyan Ben Zayed Institute for Personalized Cancer Therapy and The University of Texas MD Anderson Cancer Center Support Grant No. P30CA016672 and National Institutes of Health Grant No. R01CA184843 (S.K.). There was no additional external funding received for this study.

**Competing interests:** The authors have declared that no competing interests exist.

prognostic role and lack of response of CRC to anti-epidermal growth factor receptor therapy. Further studies are required to validate these findings and the role of SMAD4 mutation in CRC.

## Introduction

Colorectal cancer (CRC) is the third most commonly diagnosed cancer in the United States, and researchers estimated that about 150,000 new cases of CRC would be diagnosed in 2016. Although screening practices have led to a decline in CRC mortality rates since the 1990s, researchers estimated that 49,190 CRC-related deaths would occur in the United States in 2016 [1]. The prognosis for CRC is widely variable, and about 20% of cases are metastatic at the time of presentation.

Over the past few decades, advances in molecular biology have helped identify and characterize genes and molecular pathways involved in carcinogenesis, disease progression, and resistance to treatment in CRC patients. Tumor genotyping and sequencing technology—now standard practice—assist clinicians in predicting disease behavior, prognosis, and treatment response, but more prognostic markers are required to further personalize treatment plans and differentiate among various subgroups of patients. Therefore, the identification of new markers remains essential to improving treatment outcomes and survival in CRC patients [2, 3].

The transforming growth factor (TGF)- $\beta$  signaling pathway has a critical role in CRC progression. This pathway is naturally involved in many biologic cellular processes such as cell proliferation, differentiation, apoptosis, and extracellular matrix production [4]. Alteration of TGF- $\beta$  has pivotal roles in carcinogenesis and cancer progression that are paradoxical. In the early stages of cancer development, activation of TGF- $\beta$  is associated with tumor suppression [5], but in the more advanced stages, TGF- $\beta$  is believed to promote metastasis, angiogenesis, and epithelial-to-mesenchymal transition [6, 7]. TGF- $\beta$  signaling is initiated by the binding of TGF- $\beta$  ligands to TGF- $\beta$  transmembrane protein kinase receptors. Upon activation of TGF- $\beta$  receptors (including TGF- $\beta$  receptors 1 and 2), phosphorylation of the receptor-activated SMADs and the SMAD protein family members (including SMAD2 and SMAD3) occurs, which enables SMAD2 and SMAD3 to bind to SMAD4 [8, 9]. The resulting complex relocates into the nucleus and regulates transcription of TGF- $\beta$ -related genes [10–13].

SMAD proteins are essential mediators of the TGF- $\beta$  signaling pathway. These proteins are highly homologous and harbor 2 conserved domains known as the N-terminal Mad homology domain-1 and C-terminal Mad homology domain-2 (MH2) [14]. The N-terminal Mad homology domain-1 is involved in negative regulation of MH2, nuclear import, and the binding of the DNA transcriptional co-regulator [14, 15]. The MH2 domain is involved in SMAD protein homo- and hetero-oligomerization, cytoplasmic anchoring, and transcription of target genes [10–12, 16, 17]. Among this protein family, SMAD4 (localized to band 18q21) has a central role as a common downstream regulator and tumor suppressor gene [18, 19]. In previous studies, the loss of SMAD4 function was an independent prognostic factor for decreased recurrence-free and overall survival (OS) in CRC patients, particularly those with advanced-stage disease [20–23]. On the other hand, higher levels of SMAD4 expression have been associated with better OS and disease-free survival in CRC patients [21]. In addition, inactivation of SMAD4 has rendered cells resistant to TGF- $\beta$ 1-induced growth inhibition [24]. Sporadic mutation of SMAD4 has been reported in 2.1–20.0% of CRC cases [25–30]. The occurrence of

SMAD4 mutation in CRC is considered a late event and usually happens in combination with other alterations [27, 31]. In a study performed on primary and metastatic tumor pairs, SMAD4 mutation was one of the frequent novel mutations in metastases; the finding that suggests SMAD4 is involved in clonal divergence [32]. Despite the essential role of this protein in TGF- $\beta$  signaling pathway, data regarding clinicopathologic features, treatment response, and outcome in patients with SMAD4-mutated tumors is limited. Therefore, the purpose of the present study was to evaluate characteristics, survival, and treatment response associated with SMAD4 mutation, and identify characteristics of SMAD4 and TGF- $\beta$  pathway mutations in CRC cases in The Cancer Genome Atlas (TCGA) database cases.

## Materials and methods

### Patients

Patients with metastatic or locally advanced unresectable CRC who received treatment at The University of Texas MD Anderson Cancer Center from 2000 to 2014 and whose tumors were genotyped for SMAD4 mutations, were included in this retrospective study. Tumors were sequenced using a hotspot panel (Ion Torrent, Life Technologies, Carlsbad, CA, USA) predicted to cover 80% of the reported mutations of SMAD4. Further targeted resequencing that included full-length SMAD4 was performed with a focus on mutated tumors using a HiSeq sequencing system (Illumina, San Diego, CA, USA) with full exome coverage to an average depth of 800. Patients were categorized based on their SMAD4 mutational status as SMAD4-mutated or SMAD4 wild-type.

Clinicopathologic information on the study patients was obtained from the CRC registry database at MD Anderson. The collected information included age at the time of presentation (<65 years versus  $\geq$ 65 years), sex, race, site of primary tumor (colon or rectum), tumor grade (well, moderately, or poorly differentiated), presence of distant metastasis at the time of presentation, date of diagnosis, date of last contact with MDACC, and vital status. Follow-up information on living patients was collected through December 2015. The average follow-up duration for all patients was 50 months. For a subset of patients who had available data on anti-epidermal growth factor receptor (EGFR) treatment, treatment response was evaluated according to SMAD4 mutational status. Patients with known SMAD4 mutational statuses who were wild-type for KRAS, NRAS, and BRAF genes and experienced progression of disease while receiving anti-EGFR treatment were included in treatment response analysis. For these patients, the dates of initiation and termination of anti-EGFR treatment and the anti-EGFR agent (cetuximab or panitumumab) used were retrospectively reviewed. All patients signed informed written consent and the study protocol was approved by the MD Anderson Institutional Review Board.

### Mutational status analysis

Consensus molecular subtypes (CMSs) of CRC were established by a large-scale analytical study interconnecting 6 independent CRC classification systems [33]. Each of the 4 molecular subtypes according to gene expression has very unique molecular characteristics representing the mesenchymal (CMS4), metabolic (CMS3), canonical (CMS2), and microsatellite instability/immune (CMS1) phenotypes. In the present study, CMS classifications of TCGA CRC samples were used to determine the frequency distribution for SMAD4 and TGF- $\beta$  signaling pathway mutations across the molecular subtypes [34]. The mutation information for the TCGA CRC-sequenced samples ( $n = 224$ ) within the members of the TGF- $\beta$  signaling pathway was obtained from the cBioPortal for Cancer Genomics database [25, 35, 36]. In this data set, only samples classified into one of the molecular subtypes ( $n = 171$ ) were used to assess

mutation frequency. The frequency of the mutations in the TGF- $\beta$  signaling pathway was quantified using the R-language ComplexHeatmap software package. The genetic alterations of KRAS, NRAS, and BRAF in CRC cells in TCGA CRC patient data ( $n = 220$ ) were reviewed, and the co-occurrence of SMAD4 mutation and these mutations was quantified.

## Statistical analysis

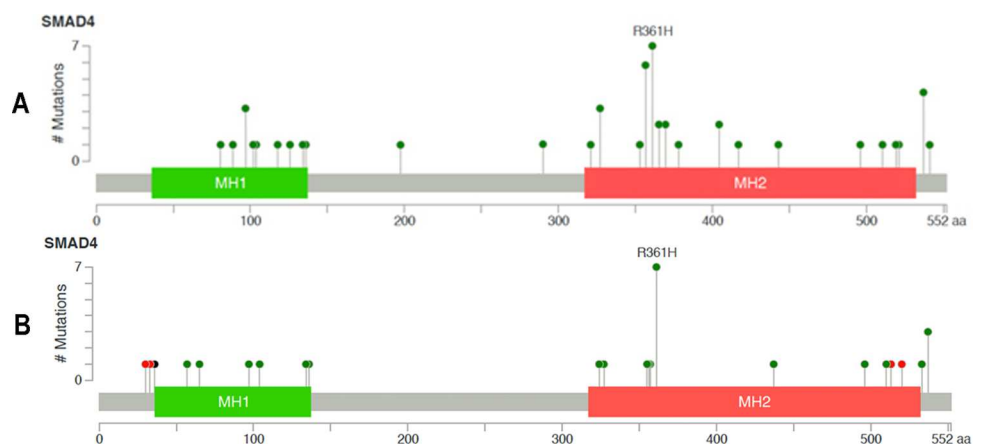
Patient demographic and clinical characteristics were compared according to SMAD4 mutational status. Comparisons across groups were done using chi-square tests and the Fisher's exact test. Logistic regression analysis was performed to calculate the odds ratio for development of SMAD4-positive tumors according to various demographic and clinical characteristics. The Fisher exact test was also used to determine the association of SMAD4 and TGF- $\beta$  protein mutations across the CMS classifications.

Also, survival analysis was performed to determine whether the SMAD4 mutation status plays a role in clinical outcomes of metastatic CRC. OS was defined as the time from diagnosis of CRC to death owing to any cause and analyzed using the Kaplan-Meier method. Log-rank testing was used to compare survival curves across mutational statuses. Univariate and multivariate Cox regression analyses were performed to determine the association of various factors with OS in metastatic CRC cases. Variables, which were associated with poor OS in the univariate analysis, were included in the multivariate Cox regression analysis. All  $p$  values were 2-sided, and statistical significance was set at  $p < 0.05$ . Statistical analyses were performed using the Stata (version 13.0) and SPSS Windows (version 16.0) software programs (SPSS Inc, Chicago, IL, USA).

## Results

### Mutation frequency and population characteristics

Of a total of 734 patients, 90 (12%) had SMAD4 mutations identified by hotspot testing. Full-length sequencing was performed in 49 mutated tumors in the study population. Missense mutations at R361 and P356 in the MH2 domain were the most common mutations of the SMAD4 gene (Fig 1).



**Fig 1. The prevalence and spectrum of SMAD4 mutations in the MD Anderson study patients and TCGA data.** (A) The prevalence and spectrum of SMAD4 mutations in the study patients who underwent full-length sequencing ( $n = 49$ ). (B) The prevalence and spectrum of SMAD4 mutations in TCGA data ( $n = 220$  patients).

doi:10.1371/journal.pone.0173345.g001

### Association of clinicopathologic features and OS with mutational status

The patients' demographic and tumor characteristics are listed in Table 1. The majority of the patients were male (n = 396 [54%]), and the median age at the time of diagnosis was 52 years. Overall, most patients had metastatic disease at the time of presentation (n = 600 [82%]).

We compared the demographic and clinical characteristics of our patients according to SMAD4 mutational status and found that SMAD4 mutation was associated with sex and tumor site. In univariate analysis, female patients were more likely to harbor SMAD4 mutations in their tumors than were male patients (odds ratio, 1.71, [95% confidence interval (CI) 1.07–2.74]; p = 0.02). Also, compared with white patients, other racial groups tended to have lower rates of SMAD4 mutation, though this difference was not significant (p = 0.062). Furthermore, SMAD4-mutant tumors were more likely to be located in the colon than in the rectum (odds ratio, 2.85 [95% CI 1.55–5.31]; p < 0.001).

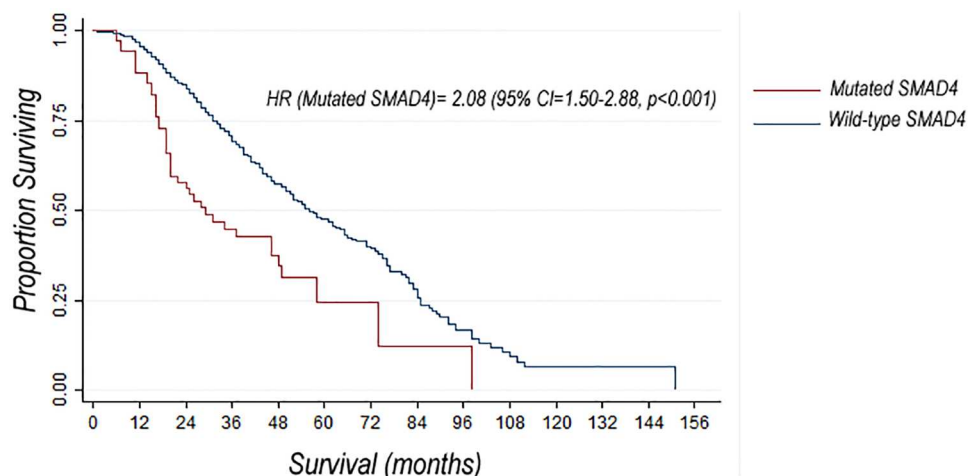
The median OS duration for all patients in the study was 62 months. We compared OS in the subset of patients with metastatic disease (n = 600). When compared with patients wild type for SMAD4, those with SMAD4 mutations had shorter OS duration from date of diagnosis (median, 29 months [95% CI, 20–48 months] versus 56 months [95% CI, 51–63 months]; p < 0.001) (Fig 2). We also performed univariate and multivariate Cox regression analyses of the association of SMAD4 mutation status and other factors with OS in patients with metastatic disease. In univariate analysis, SMAD4 mutation was associated with a poor OS (HR, 2.08 [95%

**Table 1. Demographic and tumor characteristics of CRC patients according to SMAD4 mutation status.**

	Wild-type SMAD4 (n = 644) (%)	Mutated SMAD4 (n = 90) (%)	P
<b>Sex</b>			0.020
• Female	286 (44)	52 (58)	
• Male	358 (56)	38 (42)	
<b>Race/Ethnicity</b>			0.062*
• White	478 (74)	71 (79)	
• Black	63 (10)	9 (10)	
• Hispanic	57 (9)	7 (8)	
• Other	46 (7)	3 (3)	
<b>Age at the time of presentation</b>			0.320
• <65 years	561 (87)	75 (83)	
• ≥65 years	83 (13)	15 (17)	
<b>Site of primary tumor</b>			0.001
• Colon	410 (64)	75 (83)	
• Rectum	234 (36)	15 (17)	
<b>Disease stage at the time of presentation (available data: n = 691)</b>			0.260*
• Stage II	14 (2)	0	
• Stage III	71 (11)	6 (7)	
• Stage IV	526 (82)	74 (82)	
<b>Presence of distant metastasis (available data: n = 691)</b>			0.110
• Yes	526 (82)	74 (82)	
• No	85 (13)	6 (7)	
<b>Tumor grade (available data: n = 534)</b>			0.828
• Well to moderately differentiated	357 (76)	50 (79)	
• Poorly differentiated	114 (24)	13 (20)	

\* F-test.

doi:10.1371/journal.pone.0173345.t001



**Fig 2. Comparison of survival curves in patients with metastatic CRC according to SMAD4 mutation status.** Median OS durations of 29 months (95% CI, 20–48 months) and 56 months (95% CI, 51–63 months) were observed in patients with mutated and wild-type SMAD4, respectively.

doi:10.1371/journal.pone.0173345.g002

CI, 1.50–2.88];  $p < 0.001$ ). In addition to SMAD4 mutation status, variables associated with OS in univariate Cox regression analysis were Hispanic ethnicity and poor tumor differentiation. In multivariate Cox regression analysis, SMAD4 mutation was associated with poor OS (HR, 1.39 [95% CI, 0.93–2.07];  $p = 0.10$ ) after adjusting for other characteristics (Table 2).

**Table 2. Univariate and multivariate Cox regression analyses of OS in metastatic CRC patients (n = 600).**

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	p	HR (95% CI)	p
<b>SMAD4 status</b>				
Wild-type	Ref		Ref	
Mutated	2.08 (1.50–2.88)	<0.001	1.39 (0.93–2.07)	0.100
<b>Sex</b>				
Female	Ref			
Male	0.98 (0.81–1.22)	0.920		
<b>Race/Ethnicity</b>				
White	Ref		Ref	
Black	1.22 (0.89–1.66)	0.220	1.3 (0.90–1.87)	0.160
Hispanic	1.45 (1.03–2.03)	0.030	1.31 (0.90–1.90)	0.170
Other	0.74 (0.50–1.16)	0.250	0.70 (0.38–1.28)	0.240
<b>Age at the time of presentation</b>				
<65 years	Ref			
≥65 years	1.08(0.81–1.46)	0.590		
<b>Site of primary tumor</b>				
Colon	Ref			
Rectum	0.88 (0.70–1.10)	0.260		
<b>Tumor grade (available data: n = 587)</b>				
Well to moderately differentiated	Ref		Ref	
Poorly differentiated	1.64 (1.25–2.14)	<0.001	1.64 (1.25–2.15)	<0.001

Ref: reference;

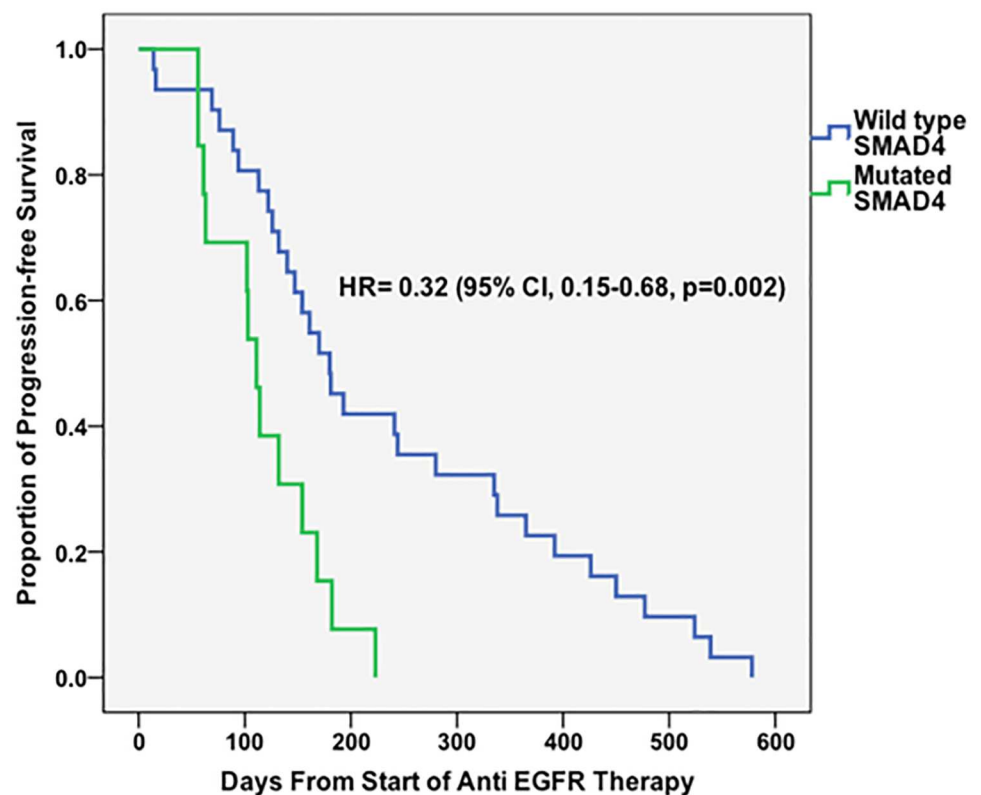
doi:10.1371/journal.pone.0173345.t002

## Response of CRC to anti-EGFR treatment

We included 44 patients (48% females [ $n = 21$ ]) wild-type for the KRAS, NRAS, and BRAF genes in our treatment response analysis, 13 of whom (30%) had SMAD4 mutations. The majority of these patients ( $n = 26$  [55%]) received line anti-EGFR treatment as third-line systemic therapy. Specifically, patients received cetuximab ( $n = 39$  [88%]), panitumumab ( $n = 4$  [9%]), or both ( $n = 1$  [2%]). Majority of patients received anti-EGFR treatment in combination with irinotecan (10 [77.0%] patients with mutated SMAD4 and 23 [74.2%] patients with wild-type SMAD4). More than two thirds of the patients ( $n = 31$  [71%]) had colonic primary tumors. Patients with SMAD4 mutations had significantly shorter progression-free survival durations than did those wild-type for SMAD4 (median, 111 days [95% CI, 96–125 days] versus 180 days [95% CI, 137–222 days]; HR, 0.32 [95% CI, 0.15–0.68];  $p = 0.003$ ). Progression-free survival curves are displayed in Fig 3.

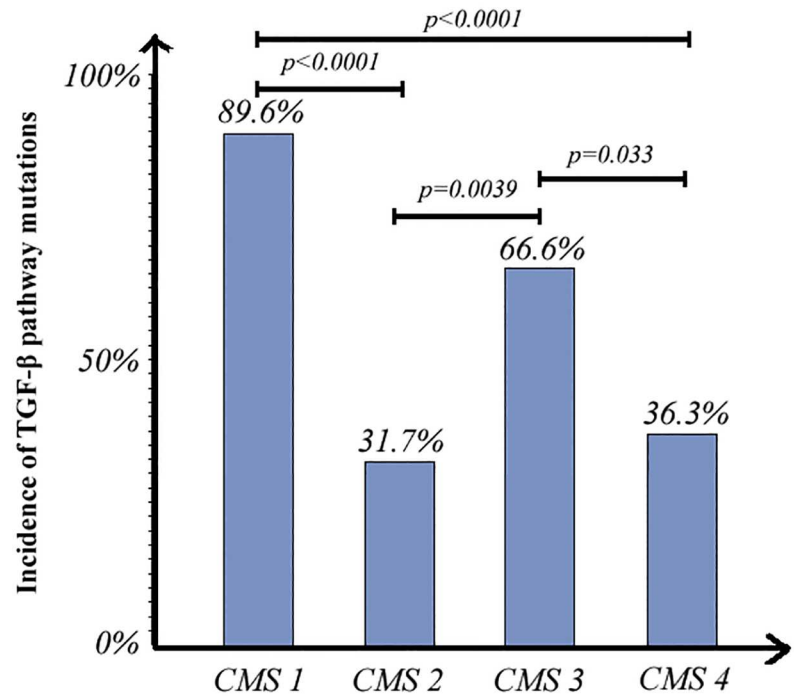
**Genetic analyses of TCGA data.** Figs 4 and 5 show the incidences of TGF- $\beta$  pathway and SMAD4 mutations, respectively, across different molecular subtypes. SMAD4 mutations were more common with the CMS3 (metabolic) (25% mutated) than with the other 3 molecular subtypes. Two thirds of the CMS3 samples (66.6%) and approximately one third of the CMS2 (31.8%) and CMS4 (36.4%) samples had at least one mutation in the TGF- $\beta$  signaling pathway (Fig 5).

An OncoPrint of the mutations in the TGF- $\beta$  signaling pathway across different consensus molecular subtypes in TCGA CRC cases is shown in Fig 6. In CMS3 and CMS4 samples, we observed a pattern of mutual exclusivity between SMAD4 mutations and other frequent



**Fig 3. Progression-free survival curves for CRC patients who received anti-EGFR treatment according to SMAD4 mutation status.** All patients were wild-type for KRAS, NRAS, and BRAF genes.

doi:10.1371/journal.pone.0173345.g003



Frequency of cases with TGF-β pathway mutations	26	27	16	12
Frequency of cases with no TGF-β pathway mutation	3	58	8	21

**Fig 4. Comparison of the incidences of TGF-β pathway mutations across different molecular subtypes.**

doi:10.1371/journal.pone.0173345.g004

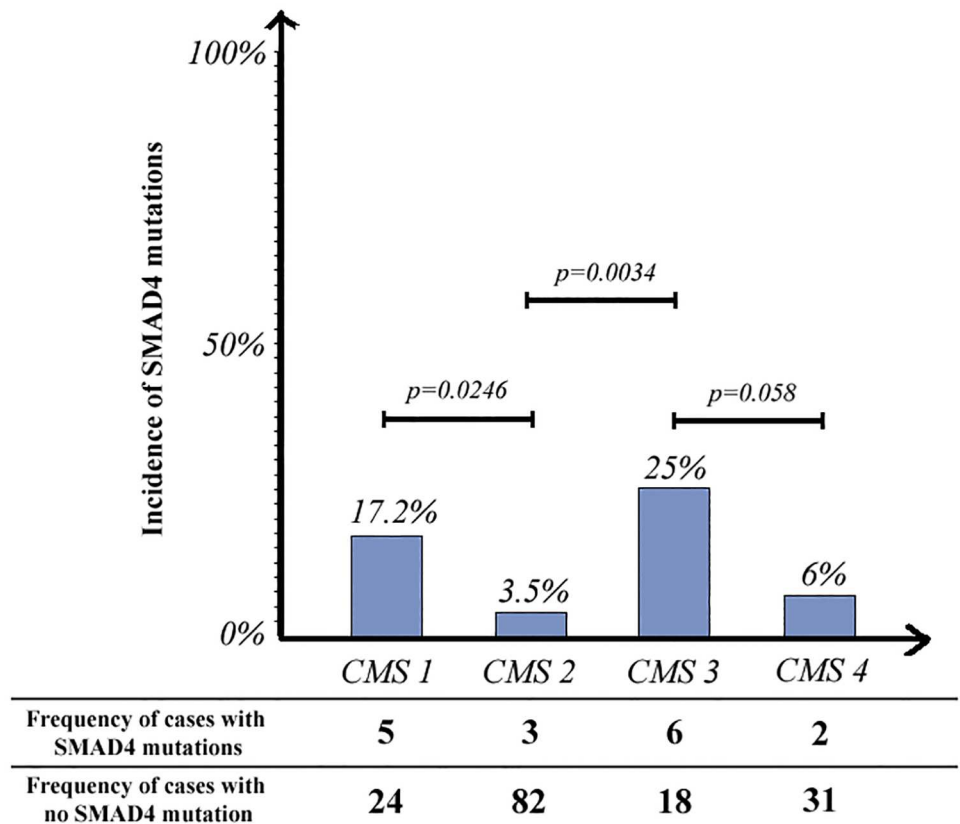
mutations in the TGF-β signaling pathway such as TGFBR2, CREBBP, ACVR2A, SMAD2, and SMAD3 mutations, which suggest functional relation of these proteins.

Also, an OncoPrint of TCGA CRC cases sequenced for SMAD4, KRAS, NRAS, and BRAF mutations is shown in Fig 7. We observed a statistically significant co-occurrence of SMAD4 mutations and these mutations in these cases (odds ratio, 3.14 [95% CI, 1.06–9.95]; p = 0.021 [chi-square test]).

### Discussion

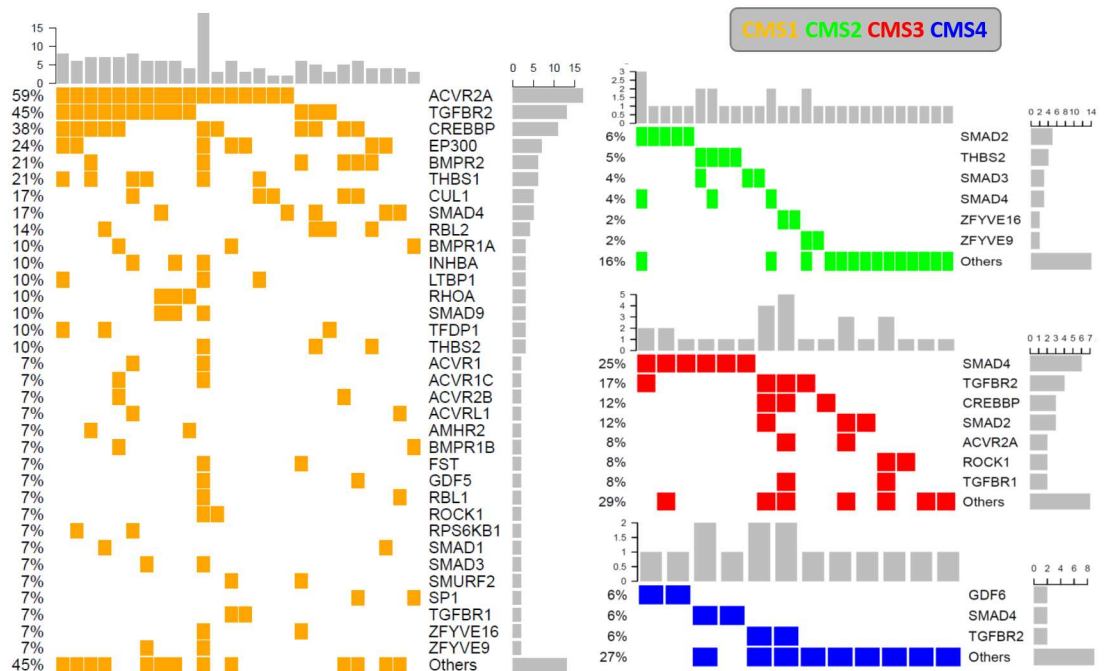
SMAD4 protein has a pivotal role in TGF-β signaling pathway by regulating gene expression following activation of TGF-β receptors [18, 37]. In prior studies, loss of SMAD4 and deletion of chromosome 18q have been extensively elaborated and shown to correlate with colorectal metastasis, resistance to 5-fluorouracil chemotherapy and poor outcome [22, 38–41]. Somatic mutations of SMAD4 are less common than loss of heterozygosity and are present in 2.1–20.0% of colorectal cancers [25–29]. To the best of our knowledge, this study is one of the largest retrospective studies to date characterizing SMAD4 mutations in CRC patients and demonstrates a prognostic role for this mutation in metastatic CRC cases. The prevalence and spectrum of SMAD4 mutations in our patients who underwent full-length sequencing (n = 49) were consistent with those in previous studies and data from TCGA [29, 31, 42]. Most of the mutations were missense mutations, and the hotspot regions affected by mutations were Arg361 and Pro356, which are located in the MH2 domain. Cancer associated SMAD4





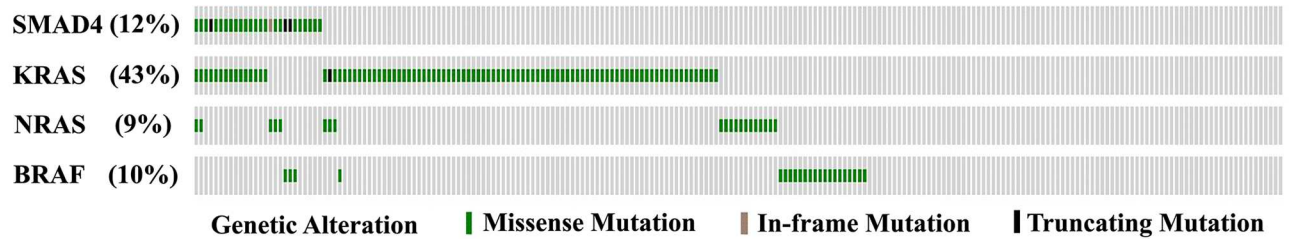
**Fig 5. Comparison of the incidences of SMAD4 mutations across different molecular subtypes.**

doi:10.1371/journal.pone.0173345.g005



**Fig 6. OncoPrint showing the distribution of TGF-β mutations across different molecular subtypes in TCGA CRC samples.**

doi:10.1371/journal.pone.0173345.g006



**Fig 7. OncoPrint of genomic alterations in TCGA CRC cases (n = 220) showing mutations of SMAD4, KRAS, NRAS, and BRAF.**

doi:10.1371/journal.pone.0173345.g007

mutations usually occur in this domain. Mutations in MH2 domain mainly affect residues close to protein interface involved in homo and hetero-oligomerization of SMAD4 with R-SMAD proteins which is required for activation [23, 29, 43] whereas, mutations in MH1 domain have been shown to alter protein stability, DNA binding and nuclear translocation [14, 15, 29, 44]. The majority of SMAD4 mutations were missense mutations, and their impact on SMAD4 protein expression remains to be determined.

We found that SMAD4 mutation was not related to age, race, tumor grade, or presence of distant metastasis. However, it was associated with colon tumors, female sex, and poor prognosis. Fleming et al, evaluated characteristics of 744 CRC patients and associations with SMAD4 mutation status. They demonstrated association of SMAD4 mutation with mucinous histology but they did not observe any further correlations with age, gender, or tumor stage [29]. In a more recent study, SMAD4 mutation was found more frequently in high-grade mucinous adenocarcinomas versus low-grade mucinous tumors [45]. Some studies have shown association of SMAD4 mutation with advanced-stage and aggressive phenotype of CRC [23, 32, 46]. Although our retrospective study did not demonstrate relationship between this mutation and CRC stage at presentation, it did demonstrate that SMAD4-mutated tumors exhibited more aggressive behavior than did nonmutated tumors. Inferior outcomes of CRC may also be related to poor treatment response in patients with SMAD4 mutations, as alteration of the SMAD4 protein rendered CRC resistant to 5-fluorouracil in prior studies [22, 41].

Data regarding response of CRC to anti-EGFR treatment in patients with SMAD4 mutations are limited. Although the sample size in our treatment response analysis was small, we observed a markedly shorter progression-free survival time in patients with SMAD4-mutated tumors than in those wild-type for SMAD4. Lupini et al [47] evaluated response of CRC to anti-EGFR antibody-based therapy in patients with different mutations (including KRAS, NRAS, BRAF, PI3KCA, and SMAD4 mutations) and found a higher number of nonresponders in the SMAD4-mutated arm, although the difference between the 2 treatment arms based on mutational status was not statistically significant (possibly owing to small numbers of patients: 4 versus 1).

In our analysis of TCGA CRC cases, the high rate of SMAD4 and TGF- $\beta$  pathway mutations in CMS1 samples is explained by microsatellite instability and hyper-mutation as CMS1 tumors harbor defective DNA mismatch repair. In addition to CMS1 samples, SMAD4 mutation was significantly more common in CMS3 cases than in CMS2 and CMS4 cases ( $p = 0.0034$  and  $p = 0.058$ , respectively). This mutation also exhibited significant co-occurrence with KRAS, NRAS, and BRAF mutations in TCGA in cases. These findings are consistent with those of previous studies describing overrepresentation of KRAS mutations and metabolic dysregulation in CMS3 CRC cases and are common features with gastric and pancreatic cancers [33, 48–50].

Our study has some limitations that are largely attributable to its retrospective nature. Our finding regarding resistance of CRC to anti-EGFR treatment must be validated in larger studies. Also, only a subset of our patients underwent full-length tumor sequencing. Since, the hot-spot panel did not fully cover all reported SMAD4 mutations, a more comprehensive genetic analysis would increase the accuracy of our findings. Furthermore, our sequencing method did not identify loss of heterozygosity status of SMAD4. Most of the SMAD4 mutations were missense mutations, and their impact on SMAD4 protein function requires further investigation.

In conclusion, we observed association of SMAD4 mutation with patient sex, colonic tumor location, and poor prognosis in CRC cases. Further studies are required to validate our findings and evaluate the implications of this mutation and TGF- $\beta$  pathway dysregulation in response of CRC to therapy.

## Supporting information

**S1 Data. Data on tumor characteristics, SMAD4 mutational status and survival of patients with CRC who were treated at MD Anderson Cancer Center.** The data was used for Tables 1 and 2.  
(XLS)

## Acknowledgments

The results presented here in part are based upon data generated by TCGA Research Network (<http://cancergenome.nih.gov/>). This work was presented in part at the 2016 American Society of Clinical Oncology Gastrointestinal Cancers Symposium in San Francisco, CA.

This study is supported in part by the Sheikh Khalifa Al Nahyan Ben Zayed Institute for Personalized Cancer Therapy and The University of Texas MD Anderson Cancer Center Support Grant No. P30CA016672 and National Institutes of Health Grant No. R01CA184843 (S. K.). There was no additional external funding received for this study.

## Author Contributions

**Conceptualization:** SK AMS.

**Data curation:** SM.

**Formal analysis:** SA AMS.

**Funding acquisition:** SK AMS SA GM.

**Investigation:** SK AMS SA GM.

**Methodology:** SK AMS SA GM.

**Project administration:** AMS SK.

**Resources:** MJO BKK DRF AD KR EV SM IS RAW KPP RL KS CE DMM MJR FMB SK.

**Software:** GM.

**Supervision:** SK.

**Validation:** SA AMS.

**Visualization:** AMS.

**Writing – original draft:** AMS.

**Writing – review & editing:** AMS MJO BKK DRF AD KR EV SM IS RAW KPP RL KS CE  
DMM MJR FMB SK.

## References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016; 66(1):7–30. Epub 2016/01/09. doi: [10.3322/caac.21332](https://doi.org/10.3322/caac.21332) PMID: [26742998](https://pubmed.ncbi.nlm.nih.gov/26742998/)
2. Duffy MJ, Crown J. A personalized approach to cancer treatment: how biomarkers can help. *Clin Chem.* 2008; 54(11):1770–9. Epub 2008/09/20. doi: [10.1373/clinchem.2008.110056](https://doi.org/10.1373/clinchem.2008.110056) PMID: [18801934](https://pubmed.ncbi.nlm.nih.gov/18801934/)
3. Deschoolmeester V, Baay M, Specenier P, Lardon F, Vermorken JB. A review of the most promising biomarkers in colorectal cancer: one step closer to targeted therapy. *Oncologist.* 2010; 15(7):699–731. Epub 2010/06/30. doi: [10.1634/theoncologist.2010-0025](https://doi.org/10.1634/theoncologist.2010-0025) PMID: [20584808](https://pubmed.ncbi.nlm.nih.gov/20584808/)
4. Massague J. TGFbeta in Cancer. *Cell.* 2008; 134(2):215–30. Epub 2008/07/30. doi: [10.1016/j.cell.2008.07.001](https://doi.org/10.1016/j.cell.2008.07.001) PMID: [18662538](https://pubmed.ncbi.nlm.nih.gov/18662538/)
5. Ikushima H, Miyazono K. TGFbeta signalling: a complex web in cancer progression. *Nat Rev Cancer.* 2010; 10(6):415–24. Epub 2010/05/25. doi: [10.1038/nrc2853](https://doi.org/10.1038/nrc2853) PMID: [20495575](https://pubmed.ncbi.nlm.nih.gov/20495575/)
6. Pickup M, Novitskiy S, Moses HL. The roles of TGFbeta in the tumour microenvironment. *Nat Rev Cancer.* 2013; 13(11):788–99. Epub 2013/10/18. doi: [10.1038/nrc3603](https://doi.org/10.1038/nrc3603) PMID: [24132110](https://pubmed.ncbi.nlm.nih.gov/24132110/)
7. Drabsch Y, ten Dijke P. TGF-beta signalling and its role in cancer progression and metastasis. *Cancer Metastasis Rev.* 2012; 31(3–4):553–68. Epub 2012/06/21. doi: [10.1007/s10555-012-9375-7](https://doi.org/10.1007/s10555-012-9375-7) PMID: [22714591](https://pubmed.ncbi.nlm.nih.gov/22714591/)
8. Abdollah S, Macias-Silva M, Tsukazaki T, Hayashi H, Attisano L, Wrana JL. TbetaRI phosphorylation of Smad2 on Ser465 and Ser467 is required for Smad2-Smad4 complex formation and signaling. *J Biol Chem.* 1997; 272(44):27678–85. Epub 1997/11/05. PMID: [9346908](https://pubmed.ncbi.nlm.nih.gov/9346908/)
9. Souchelnytskyi S, Tamaki K, Engstrom U, Wernstedt C, ten Dijke P, Heldin CH. Phosphorylation of Ser465 and Ser467 in the C terminus of Smad2 mediates interaction with Smad4 and is required for transforming growth factor-beta signaling. *J Biol Chem.* 1997; 272(44):28107–15. Epub 1997/11/05. PMID: [9346966](https://pubmed.ncbi.nlm.nih.gov/9346966/)
10. Shi Y, Hata A, Lo RS, Massague J, Pavletich NP. A structural basis for mutational inactivation of the tumour suppressor Smad4. *Nature.* 1997; 388(6637):87–93. Epub 1997/07/03. doi: [10.1038/40431](https://doi.org/10.1038/40431) PMID: [9214508](https://pubmed.ncbi.nlm.nih.gov/9214508/)
11. Chacko BM, Qin B, Correia JJ, Lam SS, de Caestecker MP, Lin K. The L3 loop and C-terminal phosphorylation jointly define Smad protein trimerization. *Nat Struct Biol.* 2001; 8(3):248–53. Epub 2001/02/27. doi: [10.1038/84995](https://doi.org/10.1038/84995) PMID: [11224571](https://pubmed.ncbi.nlm.nih.gov/11224571/)
12. Chacko BM, Qin BY, Tiwari A, Shi G, Lam S, Hayward LJ, et al. Structural basis of heteromeric smad protein assembly in TGF-beta signaling. *Mol Cell.* 2004; 15(5):813–23. Epub 2004/09/08. doi: [10.1016/j.molcel.2004.07.016](https://doi.org/10.1016/j.molcel.2004.07.016) PMID: [15350224](https://pubmed.ncbi.nlm.nih.gov/15350224/)
13. Xu Y, Pasche B. TGF-beta signaling alterations and susceptibility to colorectal cancer. *Hum Mol Genet.* 2007; 16 Spec No 1:R14–20. Epub 2007/09/26.
14. Brown KA, Pietenpol JA, Moses HL. A tale of two proteins: differential roles and regulation of Smad2 and Smad3 in TGF-beta signaling. *J Cell Biochem.* 2007; 101(1):9–33. Epub 2007/03/07. doi: [10.1002/jcb.21255](https://doi.org/10.1002/jcb.21255) PMID: [17340614](https://pubmed.ncbi.nlm.nih.gov/17340614/)
15. Hata A, Lo RS, Wotton D, Lagna G, Massague J. Mutations increasing autoinhibition inactivate tumour suppressors Smad2 and Smad4. *Nature.* 1997; 388(6637):82–7. Epub 1997/07/03. doi: [10.1038/40424](https://doi.org/10.1038/40424) PMID: [9214507](https://pubmed.ncbi.nlm.nih.gov/9214507/)
16. Kuang C, Chen Y. Tumor-derived C-terminal mutations of Smad4 with decreased DNA binding activity and enhanced intramolecular interaction. *Oncogene.* 2004; 23(5):1021–9. Epub 2003/12/03. doi: [10.1038/sj.onc.1207219](https://doi.org/10.1038/sj.onc.1207219) PMID: [14647410](https://pubmed.ncbi.nlm.nih.gov/14647410/)
17. Wu G, Chen YG, Ozdamar B, Gyuricza CA, Chong PA, Wrana JL, et al. Structural basis of Smad2 recognition by the Smad anchor for receptor activation. *Science.* 2000; 287(5450):92–7. Epub 1999/12/30. PMID: [10615055](https://pubmed.ncbi.nlm.nih.gov/10615055/)
18. Massouh J, Hata A. TGF-beta signalling through the Smad pathway. *Trends Cell Biol.* 1997; 7(5):187–92. Epub 1997/05/01. doi: [10.1016/S0962-8924\(97\)01036-2](https://doi.org/10.1016/S0962-8924(97)01036-2) PMID: [17708943](https://pubmed.ncbi.nlm.nih.gov/17708943/)
19. Derynck R, Zhang Y, Feng XH. Smads: transcriptional activators of TGF-beta responses. *Cell.* 1998; 95(6):737–40. Epub 1998/12/29. PMID: [9865691](https://pubmed.ncbi.nlm.nih.gov/9865691/)

20. Roth AD, Delorenzi M, Tejpar S, Yan P, Klingbiel D, Fiocca R, et al. Integrated analysis of molecular and clinical prognostic factors in stage II/III colon cancer. *J Natl Cancer Inst.* 2012; 104(21):1635–46. Epub 2012/10/30. doi: [10.1093/jnci/djs427](https://doi.org/10.1093/jnci/djs427) PMID: [23104212](https://pubmed.ncbi.nlm.nih.gov/23104212/)
21. Alazzouzi H, Alhopuro P, Salovaara R, Sammalkorpi H, Jarvinen H, Mecklin JP, et al. SMAD4 as a prognostic marker in colorectal cancer. *Clin Cancer Res.* 2005; 11(7):2606–11. Epub 2005/04/09. doi: [10.1158/1078-0432.CCR-04-1458](https://doi.org/10.1158/1078-0432.CCR-04-1458) PMID: [15814640](https://pubmed.ncbi.nlm.nih.gov/15814640/)
22. Alhopuro P, Alazzouzi H, Sammalkorpi H, Davalos V, Salovaara R, Hemminki A, et al. SMAD4 levels and response to 5-fluorouracil in colorectal cancer. *Clin Cancer Res.* 2005; 11(17):6311–6. Epub 2005/09/08. doi: [10.1158/1078-0432.CCR-05-0244](https://doi.org/10.1158/1078-0432.CCR-05-0244) PMID: [16144935](https://pubmed.ncbi.nlm.nih.gov/16144935/)
23. Miyaki M, Iijima T, Konishi M, Sakai K, Ishii A, Yasuno M, et al. Higher frequency of Smad4 gene mutation in human colorectal cancer with distant metastasis. *Oncogene.* 1999; 18(20):3098–103. Epub 1999/05/26. doi: [10.1038/sj.onc.1202642](https://doi.org/10.1038/sj.onc.1202642) PMID: [10340381](https://pubmed.ncbi.nlm.nih.gov/10340381/)
24. Papageorgis P, Cheng K, Ozturk S, Gong Y, Lambert AW, Abdolmaleky HM, et al. Smad4 inactivation promotes malignancy and drug resistance of colon cancer. *Cancer Res.* 2011; 71(3):998–1008. Epub 2011/01/20. doi: [10.1158/0008-5472.CAN-09-3269](https://doi.org/10.1158/0008-5472.CAN-09-3269) PMID: [21245094](https://pubmed.ncbi.nlm.nih.gov/21245094/)
25. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012; 487(7407):330–7. Epub 2012/07/20. doi: [10.1038/nature11252](https://doi.org/10.1038/nature11252) PMID: [22810696](https://pubmed.ncbi.nlm.nih.gov/22810696/)
26. Lee SY, Haq F, Kim D, Jun C, Jo HJ, Ahn SM, et al. Comparative genomic analysis of primary and synchronous metastatic colorectal cancers. *PLoS One.* 2014; 9(3):e90459. Epub 2014/03/07. doi: [10.1371/journal.pone.0090459](https://doi.org/10.1371/journal.pone.0090459) PMID: [24599305](https://pubmed.ncbi.nlm.nih.gov/24599305/)
27. Malapelle U, Pisapia P, Sgariglia R, Vigliar E, Biglietto M, Carlomagno C, et al. Less frequently mutated genes in colorectal cancer: evidences from next-generation sequencing of 653 routine cases. *J Clin Pathol.* 2016; 69(9):767–71. Epub 2016/01/23. doi: [10.1136/jclinpath-2015-203403](https://doi.org/10.1136/jclinpath-2015-203403) PMID: [26797410](https://pubmed.ncbi.nlm.nih.gov/26797410/)
28. Han SW, Kim HP, Shin JY, Jeong EG, Lee WC, Lee KH, et al. Targeted sequencing of cancer-related genes in colorectal cancer using next-generation sequencing. *PLoS One.* 2013; 8(5):e64271. Epub 2013/05/24. doi: [10.1371/journal.pone.0064271](https://doi.org/10.1371/journal.pone.0064271) PMID: [23700467](https://pubmed.ncbi.nlm.nih.gov/23700467/)
29. Fleming NI, Jorissen RN, Mouradov D, Christie M, Sakthianandeswaren A, Palmieri M, et al. SMAD2, SMAD3 and SMAD4 mutations in colorectal cancer. *Cancer Res.* 2013; 73(2):725–35. Epub 2012/11/10. doi: [10.1158/0008-5472.CAN-12-2706](https://doi.org/10.1158/0008-5472.CAN-12-2706) PMID: [23139211](https://pubmed.ncbi.nlm.nih.gov/23139211/)
30. Yu J, Wu WK, Li X, He J, Li XX, Ng SS, et al. Novel recurrently mutated genes and a prognostic mutation signature in colorectal cancer. *Gut.* 2015; 64(4):636–45. Epub 2014/06/22. doi: [10.1136/gutjnl-2013-306620](https://doi.org/10.1136/gutjnl-2013-306620) PMID: [24951259](https://pubmed.ncbi.nlm.nih.gov/24951259/)
31. Jauhri M, Bhatnagar A, Gupta S, Shokeen Y, Minhas S, Aggarwal S. Targeted molecular profiling of rare genetic alterations in colorectal cancer using next-generation sequencing. *Med Oncol.* 2016; 33(10):106. Epub 2016/08/29. doi: [10.1007/s12032-016-0820-2](https://doi.org/10.1007/s12032-016-0820-2) PMID: [27568332](https://pubmed.ncbi.nlm.nih.gov/27568332/)
32. Goswami RS, Patel KP, Singh RR, Meric-Bernstam F, Kopetz ES, Subbiah V, et al. Hotspot mutation panel testing reveals clonal evolution in a study of 265 paired primary and metastatic tumors. *Clin Cancer Res.* 2015; 21(11):2644–51. Epub 2015/02/20. doi: [10.1158/1078-0432.CCR-14-2391](https://doi.org/10.1158/1078-0432.CCR-14-2391) PMID: [25695693](https://pubmed.ncbi.nlm.nih.gov/25695693/)
33. Guinney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, et al. The consensus molecular subtypes of colorectal cancer. *Nat Med.* 2015; 21(11):1350–6. Epub 2015/10/13. doi: [10.1038/nm.3967](https://doi.org/10.1038/nm.3967) PMID: [26457759](https://pubmed.ncbi.nlm.nih.gov/26457759/)
34. KEGG PATHWAY Database. <http://www.genome.jp/kegg/pathway.html>.
35. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013; 6(269):pl1. Epub 2013/04/04. doi: [10.1126/scisignal.2004088](https://doi.org/10.1126/scisignal.2004088) PMID: [23550210](https://pubmed.ncbi.nlm.nih.gov/23550210/)
36. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012; 2(5):401–4. Epub 2012/05/17. doi: [10.1158/2159-8290.CD-12-0095](https://doi.org/10.1158/2159-8290.CD-12-0095) PMID: [22588877](https://pubmed.ncbi.nlm.nih.gov/22588877/)
37. Schmierer B, Hill CS. TGFbeta-SMAD signal transduction: molecular specificity and functional flexibility. *Nat Rev Mol Cell Biol.* 2007; 8(12):970–82. Epub 2007/11/15. doi: [10.1038/nrm2297](https://doi.org/10.1038/nrm2297) PMID: [18000526](https://pubmed.ncbi.nlm.nih.gov/18000526/)
38. Losi L, Bouzourene H, Benhattar J. Loss of Smad4 expression predicts liver metastasis in human colorectal cancer. *Oncol Rep.* 2007; 17(5):1095–9. Epub 2007/03/29. PMID: [17390050](https://pubmed.ncbi.nlm.nih.gov/17390050/)
39. Xiao DS, Wen JF, Li JH, Wang KS, Hu ZL, Zhou JH, et al. Effect of DPC4 gene on invasion and metastasis of colorectal carcinoma cells. *Acta Biochim Biophys Sin (Shanghai).* 2006; 38(12):883–92. Epub 2006/12/08.
40. Li X, Liu B, Xiao J, Yuan Y, Ma J, Zhang Y. Roles of VEGF-C and Smad4 in the lymphangiogenesis, lymphatic metastasis, and prognosis in colon cancer. *J Gastrointest Surg.* 2011; 15(11):2001–10. Epub 2011/07/26. doi: [10.1007/s11605-011-1627-2](https://doi.org/10.1007/s11605-011-1627-2) PMID: [21786062](https://pubmed.ncbi.nlm.nih.gov/21786062/)

41. Zhang B, Zhang B, Chen X, Bae S, Singh K, Washington MK, et al. Loss of Smad4 in colorectal cancer induces resistance to 5-fluorouracil through activating Akt pathway. *Br J Cancer*. 2014; 110(4):946–57. Epub 2014/01/05. doi: [10.1038/bjc.2013.789](https://doi.org/10.1038/bjc.2013.789) PMID: [24384683](https://pubmed.ncbi.nlm.nih.gov/24384683/)
42. Dallol A, Buhmeida A, Al-Ahwal MS, Al-Maghrabi J, Bajouh O, Al-Khayyat S, et al. Clinical significance of frequent somatic mutations detected by high-throughput targeted sequencing in archived colorectal cancer samples. *J Transl Med*. 2016; 14(1):118. Epub 2016/05/06. doi: [10.1186/s12967-016-0878-9](https://doi.org/10.1186/s12967-016-0878-9) PMID: [27146902](https://pubmed.ncbi.nlm.nih.gov/27146902/)
43. Iacobuzio-Donahue CA, Song J, Parmigiani G, Yeo CJ, Hruban RH, Kern SE. Missense mutations of MADH4: characterization of the mutational hot spot and functional consequences in human tumors. *Clin Cancer Res*. 2004; 10(5):1597–604. Epub 2004/03/12. PMID: [15014009](https://pubmed.ncbi.nlm.nih.gov/15014009/)
44. Moren A, Itoh S, Moustakas A, Dijke P, Heldin CH. Functional consequences of tumorigenic missense mutations in the amino-terminal domain of Smad4. *Oncogene*. 2000; 19(38):4396–404. Epub 2000/09/12. doi: [10.1038/sj.onc.1203798](https://doi.org/10.1038/sj.onc.1203798) PMID: [10980615](https://pubmed.ncbi.nlm.nih.gov/10980615/)
45. Yoshioka Y, Togashi Y, Chikugo T, Kogita A, Taguri M, Terashima M, et al. Clinicopathological and genetic differences between low-grade and high-grade colorectal mucinous adenocarcinomas. *Cancer*. 2015; 121(24):4359–68. Epub 2015/10/22. doi: [10.1002/cncr.29676](https://doi.org/10.1002/cncr.29676) PMID: [26488212](https://pubmed.ncbi.nlm.nih.gov/26488212/)
46. Mamot C, Mild G, Reuter J, Laffer U, Metzger U, Terracciano L, et al. Infrequent mutation of the tumour-suppressor gene Smad4 in early-stage colorectal cancer. *Br J Cancer*. 2003; 88(3):420–3. Epub 2003/02/06. doi: [10.1038/sj.bjc.6600733](https://doi.org/10.1038/sj.bjc.6600733) PMID: [12569386](https://pubmed.ncbi.nlm.nih.gov/12569386/)
47. Lupini L, Bassi C, Mlcochova J, Musa G, Russo M, Vychytilova-Faltejskova P, et al. Prediction of response to anti-EGFR antibody-based therapies by multigene sequencing in colorectal cancer patients. *BMC Cancer*. 2015; 15:808. Epub 2015/10/29. doi: [10.1186/s12885-015-1752-5](https://doi.org/10.1186/s12885-015-1752-5) PMID: [26508446](https://pubmed.ncbi.nlm.nih.gov/26508446/)
48. Lei Z, Tan IB, Das K, Deng N, Zouridis H, Pattison S, et al. Identification of molecular subtypes of gastric cancer with different responses to PI3-kinase inhibitors and 5-fluorouracil. *Gastroenterology*. 2013; 145(3):554–65. Epub 2013/05/21. doi: [10.1053/j.gastro.2013.05.010](https://doi.org/10.1053/j.gastro.2013.05.010) PMID: [23684942](https://pubmed.ncbi.nlm.nih.gov/23684942/)
49. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature*. 2014; 513(7517):202–9. Epub 2014/08/01. doi: [10.1038/nature13480](https://doi.org/10.1038/nature13480) PMID: [25079317](https://pubmed.ncbi.nlm.nih.gov/25079317/)
50. Vincent A, Herman J, Schulick R, Hruban RH, Goggins M. Pancreatic cancer. *Lancet*. 2011; 378(9791):607–20. Epub 2011/05/31. doi: [10.1016/S0140-6736\(10\)62307-0](https://doi.org/10.1016/S0140-6736(10)62307-0) PMID: [21620466](https://pubmed.ncbi.nlm.nih.gov/21620466/)