

Article

Validation of Gene Expression-Based Predictive Biomarkers for Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer

Tomoyuki Momma¹, Hirokazu Okayama^{1,*} , Yasuyuki Kanke¹, Satoshi Fukai¹, Hisashi Onozawa¹, Shotaro Fujita¹, Wataru Sakamoto¹, Motonobu Saito¹, Shinji Ohki^{1,2} and Koji Kono¹

¹ Department of Gastrointestinal Tract Surgery, Fukushima Medical University School of Medicine, 1 Hikarigaoka, Fukushima 960-1295, Japan; tmomma@fmu.ac.jp (T.M.); kanke33@fmu.ac.jp (Y.K.); sfukai@fmu.ac.jp (S.F.); hisa444@fmu.ac.jp (H.O.); newyork@fmu.ac.jp (S.F.); ws1024@fmu.ac.jp (W.S.); moto@fmu.ac.jp (M.S.); ohki@fmu.ac.jp (S.O.); kojikono@fmu.ac.jp (K.K.)

² Hospital Director, Shirakawa Kosei General Hospital, 2-1 Kamiyajiro, Shirakawa, Fukushima 961-0005, Japan

* Correspondence: okayama@fmu.ac.jp; Tel.: +81-24-547-1259

Simple Summary: There is a clinical need for predictive biomarkers that can identify patients with rectal cancer who do not respond to preoperative neoadjuvant chemoradiotherapy. In this study, we assembled multiple independent microarray datasets of biopsy specimens obtained from patients with rectal cancer before neoadjuvant treatment, including 237 non-responders and 152 responders. These datasets were utilized as the discovery cohorts or the validation cohorts, to develop and validate gene expression signatures predictive of treatment response. Using an in silico meta-analysis approach, here we tested not only our 4-gene signature built in this study but also nine different single-gene and multi-gene predictive signatures that were previously reported in the literature. Nevertheless, in the validation cohorts, none of the tested signatures were consistently differentially expressed between tumor specimens from non-responders and responders, and the meta-analyses revealed that those signatures had limited predictive values in clinical practice.

Abstract: Background: Neoadjuvant chemoradiotherapy (nCRT) followed by surgery is widely used for patients with locally advanced rectal cancer. However, response to nCRT varies substantially among patients, highlighting the need for predictive biomarkers that can distinguish non-responsive from responsive patients before nCRT. This study aimed to build novel multi-gene assays for predicting nCRT response, and to validate our signature and previously-reported signatures in multiple independent cohorts. Methods: Three microarray datasets of pre-therapeutic biopsies containing a total of 61 non-responders and 53 responders were used as the discovery cohorts to screen for genes that were consistently associated with nCRT response. The predictive values of signatures were tested in a meta-analysis using six independent datasets as the validation cohorts, consisted of a total of 176 non-responders and 99 responders. Results: We identified four genes, including *BRCA1*, *GPR110*, *TNFK*, and *WDR4* in the discovery cohorts. Although our 4-gene signature and nine published signatures were evaluated, they were unable to predict nCRT response in the validation cohorts. Conclusions: Although this is one of the largest studies addressing the validity of gene expression-based classifiers using pre-treatment biopsies from patients with rectal cancer, our findings do not support their clinically meaningful values to be predictive of nCRT response.

Keywords: locally advanced rectal cancer; neoadjuvant chemoradiotherapy; predictive biomarker; pre-treatment biopsy; gene expression signature; microarray; meta-analysis



Citation: Momma, T.; Okayama, H.; Kanke, Y.; Fukai, S.; Onozawa, H.; Fujita, S.; Sakamoto, W.; Saito, M.; Ohki, S.; Kono, K. Validation of Gene Expression-Based Predictive Biomarkers for Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer.

Cancers **2021**, *13*, 4642.

<https://doi.org/10.3390/cancers13184642>

Academic Editors: Luis Bujanda, Ajay Goel and Ane Etxart

Received: 2 August 2021

Accepted: 15 September 2021

Published: 16 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Colorectal cancer remains one of the leading causes of cancer death worldwide, with rectal cancer accounting for one-third of these cases [1]. During the past decade,

advances in treatment strategies with the use of standardized surgical technique, combined with preoperative (neoadjuvant) local and systemic therapies, have provided a dramatic reduction in local recurrence rate and improved survival outcomes for patients with locally advanced rectal cancer (LARC) [2,3]. The most commonly used multidisciplinary approach for LARC patients is an intravenous or oral 5-fluorouracil (5FU)-based neoadjuvant chemoradiotherapy (nCRT) followed by a standardized surgical technique (total mesorectal excision) and postoperative adjuvant chemotherapy, which is the standard of care in Western countries [3–5]. This could lead to reduced local recurrence rate and improved disease-free survival, with approximately 20% of patients achieving a pathologic complete response (pCR: ypT0M0) at the time of surgery, whereas a considerable proportion of patients exhibit resistance to nCRT, thereby resulting in only minimal to no regression or disease progression, even during nCRT [6,7]. Such heterogeneous responses to nCRT among patients with LARC can finally impact long-term oncological outcomes [6–9]. It is also important to balance the risk of local and metastatic recurrence, avoiding over-treatment, preserving organ function and patients' quality of life [3]. Accordingly, to determine the optimal treatment plan, there is a critical need of predictive biomarkers that can discriminate between non-responsive patients with LARC from those of responsive. Identifying potential non-responders and responders before neoadjuvant treatment may help clinicians to consider more personalized multidisciplinary strategies that include intensive preoperative treatment, such as total neoadjuvant therapy (TNT), upfront surgery to avoid unnecessary treatment-related toxicities, as well as non-operative “watch and wait” management [2,3].

Tumor tissue-based molecular predictors of response to nCRT in patients with LARC have been extensively studied using pre-treatment biopsy specimens. Particularly, many researchers reported gene expression signatures as predictive biomarkers for nCRT response based on high-throughput technologies, such as microarrays. However, in 2011, Brettingham-Moore et al. evaluated several gene signatures reported in earlier studies [10–12] (published between 2005 and 2008), and revealed that these published signatures had a limited accuracy in independent samples of LARC [13]. They thus concluded that gene expression-based signatures based on microarray analyses could not reliably predict nCRT response [13]. More recently, several promising classifiers based on gene expression have been reported [14–22]. These signatures were typically identified and developed in discovery or training cohorts and their predictive performance was confirmed in internal or external validation cohorts. It is important to note that despite the high prediction accuracy of those reported signatures, they had very few overlap in the identified genes among studies and also lacked prospective validation in clinical trials. Accordingly, none of the gene signatures are currently available for predicting nCRT response in patients with LARC in clinical practice.

The aim of this study was to build transcriptional multi-gene assays for predicting nCRT response in LARC, based on comprehensive screening in multiple datasets of pre-treatment biopsies. We also aimed at validating the predictive performance of various signatures, including ours and several previously-reported multi-gene and single-gene predictors reported since 2011. In order to test those signatures, we used an *in silico* meta-analysis approach based on six microarray datasets, which have been publicly available since 2016, of pre-therapeutic biopsies from patients with LARC that were independent of those used in developmental studies.

2. Materials and Methods

2.1. Microarray-Based Gene Expression Data Analysis

All expression datasets used in this study are publicly accessible in the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo>, accessed on 12 June 2021). We used the preprocessed values obtained from each microarray dataset. If a gene is represented by multiple probes, only the probe with the highest mean expression was used.

To calculate each gene signature, a representative score was determined as the sum of expression values for genes that are upregulated in non-responders minus the sum of expression values for genes that are downregulated in non-responders.

2.2. Selection of Microarray Datasets for Comprehensive Screening (the Discovery Cohorts)

For comprehensive screening of differentially expressed genes between non-responders and responders to nCRT, we obtained three microarray datasets of pre-therapeutic biopsies from LARC, including GSE35452, GSE45404 and GSE53781. They were each originally used to develop and validate multi-gene signatures predictive of nCRT response in three different studies, reported between 2014 and 2015.

2.3. Selection of Microarray Datasets for Validation (the Validation Cohorts)

We searched the GEO database in May 2021 with the search terms, (rectal cancer or rectal adenocarcinoma or rectal carcinoma) and (chemoradiotherapy or radiochemotherapy or chemoradiation or biopsy), filtered by series type (Expression Profiling by Array). Thirty-four GEO series identified by the initial GEO search were screened on the basis of eligibility criteria that required each dataset to be based on genome-wide expression data for pre-treatment biopsies from LARC patients who underwent nCRT and to include response information for more than 30 patients. We only included datasets that were published in 2012 or later. The resulting six independent datasets, including GSE46862, GSE68204, GSE94104, GSE119409, GSE133057, and GSE150082, were used as the validation cohorts.

2.4. Statistical Analysis

Unpaired *t*-test was used to determine differences between two groups. The discriminatory ability of different gene signatures was assessed in each dataset by performing the receiver operating characteristic (ROC) analysis based on continuous values for each signature, and areas under the ROC curve (AUCs) and their 95% confidence intervals (95% CI) were calculated as a measure of overall prediction accuracy. The pooled AUC was calculated using the fixed-effects model. Statistical heterogeneity between studies was assessed using I^2 statistic. Statistical analyses were performed using GraphPad Prism v6.04 (Graphpad Software, San Diego, CA, USA), SPSS Statistics version 26 (IBM Corporation, Armonk, NY, USA), or EZR (Saitama Medical Center, Jichi Medical University, Saitama, Japan), a graphical user interface for R 2.6.1 (The R Foundation for Statistical Computing, Vienna, Austria). All statistical tests were two-sided, and *p* values less than 0.05 were considered statistically significant.

3. Results

3.1. Identification of Four Common Genes Differentially Expressed between Non-Responders and Responders in the Discovery Datasets

The overall workflow of this study is illustrated in Figure 1. In 2011, Brettingham-Moore et al. critically reported that previously-published gene signatures had limited accuracy for predicting nCRT response [13]. Subsequently, in 2014, through 2015, multi-gene predictors for nCRT were reported, and microarray datasets of pre-treatment biopsies obtained from LARC patients, including GSE35452 (*n* = 46), GSE45404 (*n* = 42), and GSE53781 (*n* = 26) were deposited in the GEO database [14,15,18]. The present study utilized these three datasets as the discovery cohorts, consisting of a total of 61 non-responders and 53 responders, analyzed on Affymetrix or CodeLink platforms (Table 1). The detailed characteristics of the datasets are also demonstrated in Table S1.

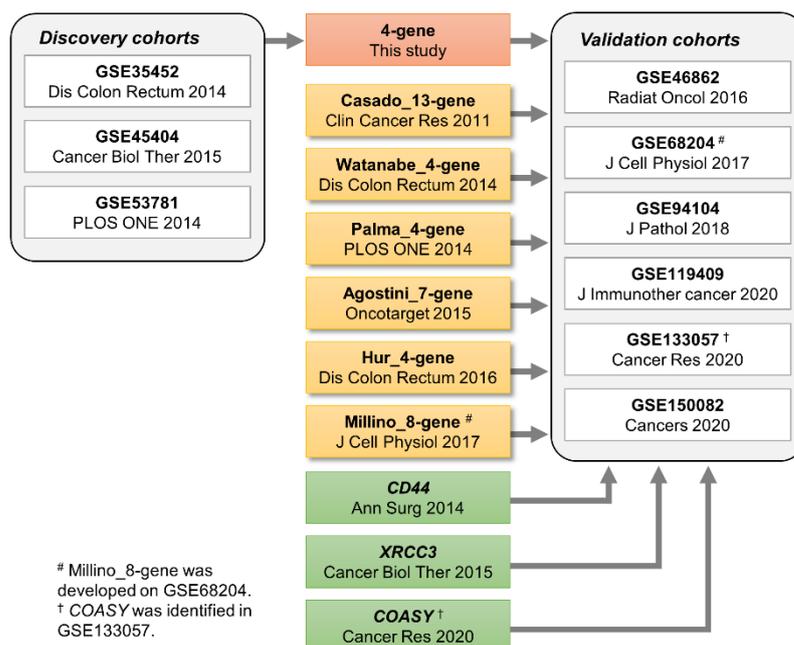


Figure 1. Overall study design. Three microarray datasets, published between 2014 and 2015, were used as the discovery cohorts to screen genes for building the 4-gene signature. The systematic search of the GEO database identified six microarray datasets as the validation cohorts to test the predictive performance of our signature and previously-published multi-gene or single-gene signatures.

Table 1. GEO datasets of pre-treatment biopsies from patients with LARC who underwent nCRT.

Purpose	GEO Accession	Platform	Total	Non-Responder	Responder	References
Discovery	GSE35452	Affymetrix Human Genome U133 Plus 2.0 Array	46	22	24	[14]
	GSE45404	Affymetrix Human Genome U133 Plus 2.0 Array	42	23	19	[18]
	GSE53781	CodeLink Human Whole Genome Array	26	16	10	[15]
Validation	GSE46862	Affymetrix Human Gene 1.0 ST Array	69	41	28	[23]
	GSE68204	Agilent-014850 Whole Human Genome Microarray 4x44K G4112F	38	22	16	[19]
	GSE94104	Illumina HumanHT-12 WG-DASL V4.0 R2 expression beadchip	40	29	11	[24]
	GSE119409	Affymetrix Human Genome U133 Plus 2.0 Array	56	41	15	[25]
	GSE133057	Illumina human-6 v2.0 expression beadchip	33	20	13	[22]
	GSE150082	Agilent-026652 Whole Human Genome Microarray 4x44K v2	39	23	16	[26]

With the use of the three discovery datasets, we first conducted comprehensive screening of genes that were differentially expressed between pre-therapeutic biopsies obtained from non-responders and responders. As demonstrated in Figure S1A,B four common genes, including G-protein-coupled receptor 110 (*GPR110*), TRAF2 and NCK

interacting kinase (*TNIK*), WD repeat domain 4 (*WDR4*) and BRCA1 DNA repair associated (*BRCA1*), in those three different datasets were identified. In tumor biopsies from non-responders, *GPR110* and *TNIK* were significantly upregulated, while *WDR4* and *BRCA1* were significantly downregulated, as compared to those of responders ($p < 0.05$). We then generated an expression signature on the basis of those four genes calculated as $((GPR110 + TNIK) - (WDR4 + BRCA1))$, showing higher levels of the 4-gene signature in biopsies from non-responders than those of responders with AUCs ranging from 0.77 to 0.93 in each of the discovery cohort (Figures 2A and S1C).

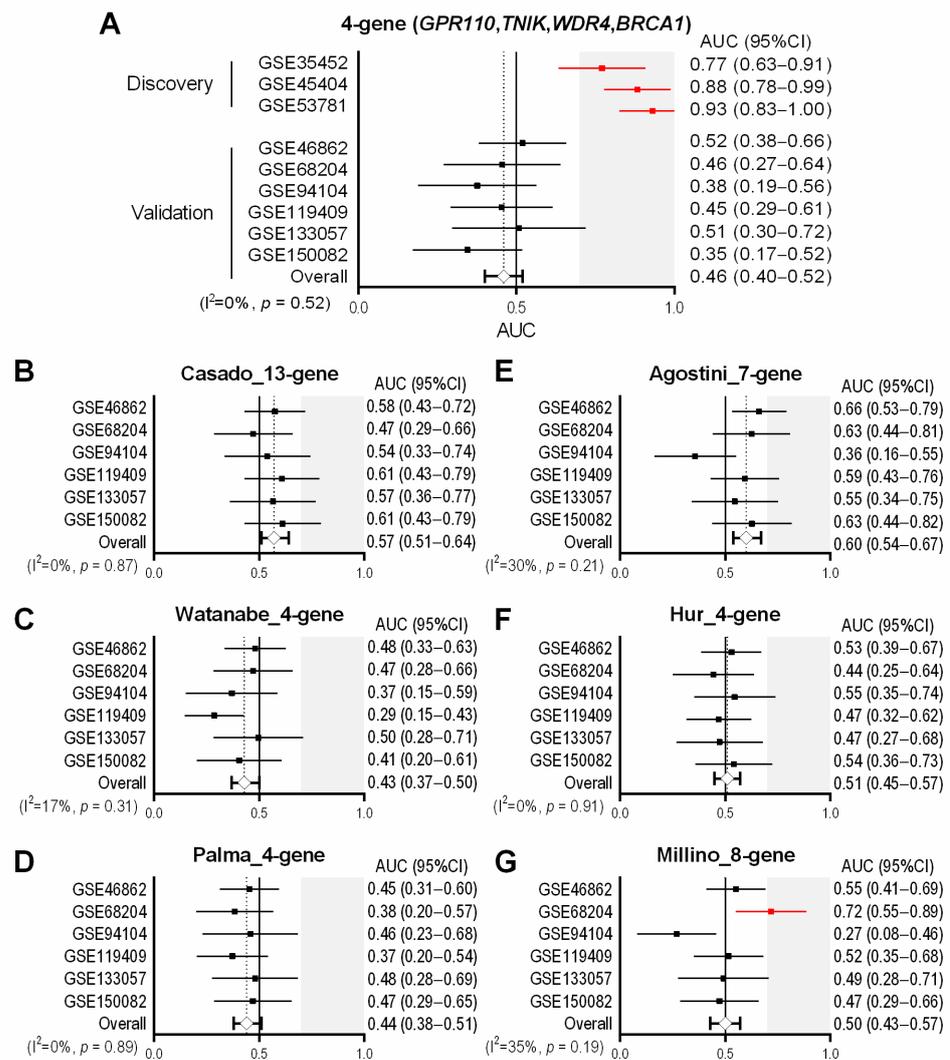


Figure 2. Assessment of the predictive performance of multi-gene signatures in the validation cohorts. (A–G) Forest plots of the area under the receiver operating characteristics curve (AUC) values and 95% confidence intervals (95% CI) for tested signatures, including the 4-gene (A), the Casado_13-gene (B), the Watanabe_4-gene (C), the Palma_4-gene (D), the Agostini_7-gene (E), the Hur_4-gene (F), and the Millino_8-gene (G). Datasets that were used to for the development of signatures (red squares and lines) were excluded from the meta-analyses.

3.2. Assessment of the Predictive Performance of the 4-Gene Signature in the Validation Cohorts

We then sought to determine the reproducibility of the predictive performance of our 4-gene signature in multiple independent datasets. For this purpose, the systematic search of the GEO database was conducted, and the search terms and the eligibility criteria were described in Materials and Methods, and in Figure S2. We identified six datasets of pre-therapeutic biopsies from LARC patients who underwent nCRT followed by surgical

resection, including GSE46862 ($n = 69$), GSE68204 ($n = 38$), GSE94104 ($n = 40$), GSE119409 ($n = 56$), GSE133057 ($n = 33$), and GSE150082 ($n = 39$), each analyzed on different microarray platforms, including Affymetrix, Illumina, and Agilent (Figures 1 and S2, Tables 1 and S1). These six GEO datasets were used as the validation cohorts that contained a total of 176 non-responders and 99 responders [19,22–26]. However, the levels of the 4-gene signature did not significantly differ between non-responders and responders in any of the six cohorts analyzed ($p > 0.05$, Figure S3), showing no prediction accuracy to differentiate non-responders from responders with AUCs around 0.50 in each cohort (Figures S3 and 2A). The pooled AUC was 0.46 (95% CI 0.40–0.52) in the meta-analysis for the six validation cohorts (Figure 2A). Although we further analyzed each of the four genes, *GPR110*, *TNFK*, *WDR4*, and *BRCA1*, at the single-gene level, it did not consistently differ between non-responders and responders in the validation cohorts (Figure S4).

3.3. Previously-Published Multi-Gene Signatures in the Validation Cohorts

Although our attempt to develop a gene signature predictive of nCRT response was unsuccessful, we next evaluated the predictive performance of six previously-reported multi-gene signatures that were identified from the literature published since 2011, as listed in Table 2 and illustrated in Figure 1. Signature scores were calculated as the sum of expression values for upregulated genes in non-responders minus the sum of expression values for downregulated genes in non-responders. For each of the six validation datasets, the levels of the six candidate signatures were examined in non-responders and responders (Figures S5A–F and S6) and then AUC values and 95% CIs were plotted (Figure 2B–G). Since the Millino_8-gene was built based on GSE68204 [19], we confirmed reasonably higher levels of this signature score in non-responders compared to responders (Figure S5F) with an AUC of 0.72 (95% CI 0.55–0.89) in GSE68204 (Figures S6B and 2G). Additionally, the Agostini_7-gene were significantly higher in non-responders in GSE46862 ($p < 0.05$, Figure S5D), showing an AUC of 0.66 (95% CI 0.53–0.79) (Figures S6A and 2C). The Watanabe_4-gene in GSE119409 and the Millino_8-gene in GSE94104 showed statistically significant difference between non-responders and responders ($p < 0.05$, Figure S5B,F), but in the opposite direction (AUCs < 0.5 , Figures S6C,D and 2E,G). Collectively, except for the two analyses mentioned above (the Millino_8-gene in GSE68204 and the Agostini_7-gene in GSE46862), none of the six signatures showed significant higher levels in non-responders compared to responders in any of the validation cohorts ($p > 0.05$, Figure S5A–F), giving limited or no discriminative power with diverse AUCs ranging from 0.27 to 0.63 in each dataset (Figure 2B–G). Likewise, the meta-analyses demonstrated that those six signatures had limited predictive values with the pooled AUCs of less than 0.6, namely, 0.57 (95% CI 0.51–0.64) for the Casado_13-gene, 0.43 (95% CI 0.37–0.50) for the Watanabe_4-gene, 0.44 (95% CI 0.38–0.51) for the Palma_13-gene, 0.60 (95% CI 0.54–0.67) for the Agostini_7-gene, 0.51 (95% CI 0.45–0.57) for the Hur_4-gene, and 0.50 (95% CI 0.43–0.57) for the Millino_8-gene (Figure 2B–G).

Table 2. Multi-gene and single-gene expression signatures tested in meta-analyses.

Signature	Genes Upregulated in Non-Responders vs. Responders	Genes Downregulated in Non-Responders vs. Responders	References
4-gene	<i>GPR110 (ADGRF1), TNFK</i>	<i>WDR4, BRCA1</i>	This study
Casado_13-gene	<i>BAK1, MLH1, TYMS, CKB, GPX2, HIG2 (HILPDA), PH4 (P4HTM)</i>	<i>ALDH1A1, CDKN1A, FOS, RELB, STAT3, TFF3</i>	[17]
Watanabe_4-gene	<i>FRMD3, SAMD5, TMC7</i>	<i>LRRIQ3 (LRRC44)</i>	[14]
Palma_4-gene	-	<i>GNG4, MYC, POLA1, RRM1</i>	[15]

Table 2. Cont.

Signature	Genes Upregulated in Non-Responders vs. Responders	Genes Downregulated in Non-Responders vs. Responders	References
Agostini_7-gene	<i>AKR1C3</i>	<i>CXCL11, CXCL10, IDO1, CXCL9, MMP12, HLA-DRA</i>	[16]
Hur_4-gene	<i>TP53</i>	<i>MKI67, CDKN1A, CD133 (PROM1)</i>	[20]
Millino_8-gene	<i>ITGA2, NRG1, KLF7</i>	<i>TMEM188 (CNEP1R1), TRAM1, BCL2L13, MYO1B, GTSE1</i>	[19]
<i>CD44</i>	<i>CD44</i>	-	[21]
<i>XRCC3</i>	-	<i>XRCC3</i>	[18]
<i>COASY</i>	<i>COASY</i>	-	[22]

3.4. Previously-Published Single-Gene Signatures in the Validation Cohorts

As demonstrated in Figures S7A–C, S8A–F, and 3A–C, we further tested three different single-gene signatures, including the expression of *CD44*, *XRCC3* and *COASY* (Table 2), using the six validation cohorts (Table 1). Since *COASY* was originally identified and reported using GSE133057 [22], we confirmed that the expression of *COASY* was significantly increased in non-responders compared to responders ($p < 0.05$, Figure S7C), showing an AUC of 0.71 (95% CI 0.53–0.89) in GSE133057 (Figures S8E and 3C). However, none of the other analyses showed significant difference of each single-gene levels between non-responders and responders ($p > 0.05$, Figure S7A–C) and AUCs ranged from 0.44 to 0.67 (Figure S8A–F). As demonstrated in Figure 3A–C, the meta-analyses revealed that those three genes had limited predictive values with the pooled AUCs: 0.57 (95% CI 0.51–0.64) for *CD44*, 0.50 (95% CI 0.44–0.57) for *XRCC3*, and 0.53 (95% CI 0.47–0.60) for *COASY*.

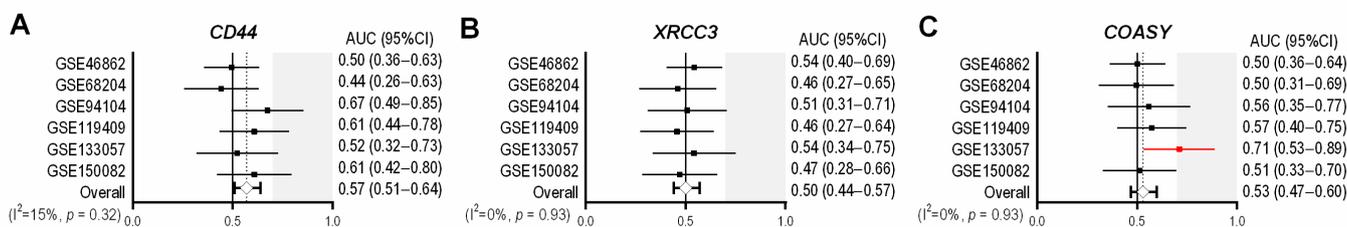


Figure 3. Assessment of the predictive performance of single-gene signatures in the validation cohorts. (A–C) Forest plots of the area under the receiver operating characteristics curve (AUC) values and 95% confidence intervals (95% CI) for tested genes, including *CD44* (A), *XRCC3* (B), and *COASY* (C). Datasets that were used to identify genes (red square and line) were excluded from the meta-analyses.

4. Discussion

Here, we analyzed a total of nine microarray cohorts of pre-therapeutic biopsy specimens obtained from 389 patients with LARC who received nCRT followed by surgery. We initially built our own signature to be predictive of nCRT response based on comprehensive analyses in the three discovery datasets. Thereafter, using the six independent validation cohorts of pre-treatment biopsies, we evaluated the performance of the 4-gene signature we developed, as well as several transcriptional signatures that were previously-reported in the literature. Even though these signatures showed clear discrimination between non-responder and responders at least in their developmental cohorts, we found that none of the tested signature scores were consistently different between tumor specimens from non-responders and responders. This resulted in a weak or no discriminative ability of each signature in the pooled analyses. Overall, our findings do not support

clinically-meaningful predictive values of those signatures on the basis of microarray data for responsiveness to nCRT in patients with LARC.

The present study identified a set of genes, including *GPR110*, *TNIK*, *WDR4*, and *BRCA1*, to be each significantly associated with nCRT response in the discovery cohorts of pre-treatment biopsies from LARC. It is worth noting that *BRCA1* plays a pivotal role in DNA repair and cell cycle regulation in response to DNA damage, potentially resulting from chemotherapy and radiotherapy. Indeed, several studies reported that *BRCA1* mRNA expression had predictive impact on responses to chemotherapy, as well as to chemoradiotherapy in many types of cancer, such as breast cancer, lung cancer, and esophageal cancer [27–31], suggesting *BRCA1* as an attractive candidate for predicting nCRT response in LARC. It has been reported that *GPR110* (*ADGRF1*) can induce cell cycle arrest and chemoresistance in breast cancer [32]. *TNIK* appeared to be an essential factor for WNT signaling and stemness in colorectal cancer [33] and might be responsible for chemoresistance in osteosarcoma [34]. However, the 4-gene signature, consisting of *GPR110*, *TNIK*, *WDR4*, and *BRCA1*, showed no association with nCRT response in any of the validation cohorts of LARC biopsies. Likewise, despite the promising correlations of the previously-reported signatures in the original publications [14–17,19,20], we found little evidence that they have sufficient clinical utility to guide decision-making in the current practice. This discrepancy could be due to small sized patient cohorts, technical differences between laboratories, disparities between various microarray probes and platforms, and insufficient analytical validation. It is also important to note that tumor biopsy-based transcriptomic profiles are likely prone to sampling bias due to intra-tumor heterogeneity. Moreover, there are differences and controversies in nCRT regimens, including chemotherapeutic drugs and radiation doses [2,3]. Optimal timing of surgical resection after nCRT also remains debatable, as longer intervals from nCRT to surgery appeared to lead to a greater potential of achieving a pCR in a meta-analysis [35], but delaying surgery was significantly associated with increased morbidity in a randomized trial [36]. Histological tumor regression grade is assessed using various grading systems with subjective categorization to reflect therapeutic response, leading to interobserver variability [37,38]. Such considerable variabilities among previously-published studies and datasets, including nCRT regimens, intervals from CRT to surgery, grading systems for tumor regression, and definitions of non-responder might also be a potential bias. We suggest that future gene signature studies of LARC using microarray or other expression platforms would be required to address those limitations.

In 2011, Brettingham-Moore et al. demonstrated that the microarray-based gene signatures, which were reported between 2005 and 2008, did not retain their predictive power in an independent cohort. Hence, they suggested that alternative approaches for predictive studies in LARC should be considered [13]. Although the present study again demonstrated that the gene expression signatures derived from microarray technologies were not capable of discriminating non-responders from responders, most recent studies have extensively searched for promising candidates to predict nCRT response in LARC using more robust expression assays and platforms. For instance, multi-gene expression assays developed on the NanoString nCounter system may provide more accurate prediction to nCRT response in LARC [39,40]. DNA-based molecular markers, including mutations in *KRAS* or *TP53* and microsatellite instability (MSI), may also be associated with responses to neoadjuvant treatment in LARC [41–43]. Other promising tissue-based predictive biomarkers, including the expression of microRNAs [38], differentially methylated CpGs [44], immune profiles [45,46], multi-protein expression assays by immunohistochemistry [47], are necessary to be validated on large, retrospective, and prospective cohorts of pre-treatment LARC biopsies, although the translation of biomarkers into clinical practice remains challenging. In addition, integrative approaches that include not only tissue-based biomarkers but also liquid-based molecular assays and imaging modalities may contribute to more sensitive stratification of patients with LARC receiving nCRT [2,38].

5. Conclusions

In this study, we developed a novel gene expression-based classifier for predicting nCRT response based on microarray cohorts of pre-treatment biopsies from patients with LARC. Although our signature and previously published signatures were tested in multiple independent datasets for validation purposes, none of them were capable of classifying patients with LARC into responders and non-responders to nCRT. We suggest that current gene expression-based signatures using microarray platforms are not robust enough to predict nCRT response or guide clinical decision-making in patients with LARC.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/cancers13184642/s1>, Figure S1: Identification of 4 common genes that were significantly differentially expressed between biopsy specimens from non-responders and responders in three discovery cohorts. Figure S2: Dataset selection flowchart for the validation cohorts. Figure S3: The performance of the 4-gene signature in six validation datasets. Figure S4: Differential expression analyses of four single genes in the 4-gene signature between non-responders and responders in nine cohorts. Figure S5: The levels of six multi-gene signatures in non-responders and responders in six validation cohorts. Figure S6: The performance of six multi-gene signatures in six validation cohorts. Figure S7: The levels of six multi-gene signatures in non-responders and responders in six validation cohorts. Figure S8: The performance of three single-gene signatures in six validation cohorts. Table S1: Characteristics of GEO datasets.

Author Contributions: Conceptualization, T.M. and H.O. (Hirokazu Okayama); methodology, H.O. (Hirokazu Okayama) and Y.K.; software, H.O. (Hirokazu Okayama) and Y.K.; validation, T.M. and H.O. (Hirokazu Okayama); formal analysis, H.O. (Hirokazu Okayama) and Y.K.; investigation, T.M. and H.O. (Hirokazu Okayama); resources, T.M., H.O. (Hirokazu Okayama), S.F. (Satoshi Fukai), H.O. (Hisashi Onozawa), S.F. (Shotaro Fujita) and W.S.; data curation, H.O. (Hirokazu Okayama); writing—original draft preparation, T.M. and H.O. (Hirokazu Okayama); writing—review and editing, T.M., H.O. (Hirokazu Okayama), W.S., and M.S.; visualization, H.O. (Hirokazu Okayama); supervision, S.O. and K.K.; project administration, H.O. (Hirokazu Okayama) and K.K.; funding acquisition, T.M., H.O. (Hirokazu Okayama), M.S. and S.O. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Japan Society for the Promotion of Science KAKENHI, grant number 16K10545.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used in this study are openly accessible through the Gene Expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo>, accessed on 12 June 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. Keller, D.S.; Berho, M.; Perez, R.O.; Wexner, S.D.; Chand, M. The multidisciplinary management of rectal cancer. *Nat. Rev. Gastroenterol. Hepatol.* **2020**, *17*, 414–429. [[CrossRef](#)]
3. Smith, J.J.; Garcia-Aguilar, J. Advances and challenges in treatment of locally advanced rectal cancer. *J. Clin. Oncol.* **2015**, *33*, 1797–1808. [[CrossRef](#)] [[PubMed](#)]
4. Benson, A.B.; Venook, A.P.; Al-Hawary, M.M.; Cederquist, L.; Chen, Y.J.; Ciombor, K.K.; Cohen, S.; Cooper, H.S.; Deming, D.; Engstrom, P.F.; et al. Rectal Cancer, Version 2.2018, NCCN Clinical Practice Guidelines in Oncology. *J. Natl. Compr. Canc. Netw.* **2018**, *16*, 874–901. [[CrossRef](#)] [[PubMed](#)]
5. Glynne-Jones, R.; Wyrwicz, L.; Tiret, E.; Brown, G.; Rodel, C.; Cervantes, A.; Arnold, D.; Committee, E.G. Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* **2017**, *28*, iv22–iv40. [[CrossRef](#)]
6. Maas, M.; Nelemans, P.J.; Valentini, V.; Das, P.; Rodel, C.; Kuo, L.J.; Calvo, F.A.; Garcia-Aguilar, J.; Glynne-Jones, R.; Haustermans, K.; et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: A pooled analysis of individual patient data. *Lancet Oncol.* **2010**, *11*, 835–844. [[CrossRef](#)]

7. Park, I.J.; You, Y.N.; Agarwal, A.; Skibber, J.M.; Rodriguez-Bigas, M.A.; Eng, C.; Feig, B.W.; Das, P.; Krishnan, S.; Crane, C.H.; et al. Neoadjuvant treatment response as an early response indicator for patients with rectal cancer. *J. Clin. Oncol.* **2012**, *30*, 1770–1776. [[CrossRef](#)] [[PubMed](#)]
8. Fokas, E.; Liersch, T.; Fietkau, R.; Hohenberger, W.; Beissbarth, T.; Hess, C.; Becker, H.; Ghadimi, M.; Mrak, K.; Merkel, S.; et al. Tumor regression grading after preoperative chemoradiotherapy for locally advanced rectal carcinoma revisited: Updated results of the CAO/ARO/AIO-94 trial. *J. Clin. Oncol.* **2014**, *32*, 1554–1562. [[CrossRef](#)]
9. Fokas, E.; Strobel, P.; Fietkau, R.; Ghadimi, M.; Liersch, T.; Grabenbauer, G.G.; Hartmann, A.; Kaufmann, M.; Sauer, R.; Graeven, U.; et al. Tumor Regression Grading After Preoperative Chemoradiotherapy as a Prognostic Factor and Individual-Level Surrogate for Disease-Free Survival in Rectal Cancer. *J. Natl. Cancer Inst.* **2017**, *109*, dx095. [[CrossRef](#)]
10. Ghadimi, B.M.; Grade, M.; Difilippantonio, M.J.; Varma, S.; Simon, R.; Montagna, C.; Fuzesi, L.; Langer, C.; Becker, H.; Liersch, T.; et al. Effectiveness of gene expression profiling for response prediction of rectal adenocarcinomas to preoperative chemoradiotherapy. *J. Clin. Oncol.* **2005**, *23*, 1826–1838. [[CrossRef](#)]
11. Kim, I.J.; Lim, S.B.; Kang, H.C.; Chang, H.J.; Ahn, S.A.; Park, H.W.; Jang, S.G.; Park, J.H.; Kim, D.Y.; Jung, K.H.; et al. Microarray gene expression profiling for predicting complete response to preoperative chemoradiotherapy in patients with advanced rectal cancer. *Dis. Colon. Rectum.* **2007**, *50*, 1342–1353. [[CrossRef](#)]
12. Rimkus, C.; Friederichs, J.; Boulesteix, A.L.; Theisen, J.; Mages, J.; Becker, K.; Nekarda, H.; Rosenberg, R.; Janssen, K.P.; Siewert, J.R. Microarray-based prediction of tumor response to neoadjuvant radiochemotherapy of patients with locally advanced rectal cancer. *Clin. Gastroenterol. Hepatol.* **2008**, *6*, 53–61. [[CrossRef](#)]
13. Brettingham-Moore, K.H.; Duong, C.P.; Greenawalt, D.M.; Heriot, A.G.; Ellul, J.; Dow, C.A.; Murray, W.K.; Hicks, R.J.; Tjandra, J.; Chao, M.; et al. Pretreatment transcriptional profiling for predicting response to neoadjuvant chemoradiotherapy in rectal adenocarcinoma. *Clin. Cancer Res.* **2011**, *17*, 3039–3047. [[CrossRef](#)]
14. Watanabe, T.; Kobunai, T.; Akiyoshi, T.; Matsuda, K.; Ishihara, S.; Nozawa, K. Prediction of response to preoperative chemoradiotherapy in rectal cancer by using reverse transcriptase polymerase chain reaction analysis of four genes. *Dis. Colon. Rectum.* **2014**, *57*, 23–31. [[CrossRef](#)] [[PubMed](#)]
15. Palma, P.; Cano, C.; Conde-Muino, R.; Comino, A.; Bueno, P.; Ferron, J.A.; Cuadros, M. Expression profiling of rectal tumors defines response to neoadjuvant treatment related genes. *PLoS ONE* **2014**, *9*, e112189. [[CrossRef](#)] [[PubMed](#)]
16. Agostini, M.; Janssen, K.P.; Kim, I.J.; D’Angelo, E.; Pizzini, S.; Zangrando, A.; Zanon, C.; Pastrello, C.; Maretto, I.; Digito, M.; et al. An integrative approach for the identification of prognostic and predictive biomarkers in rectal cancer. *Oncotarget* **2015**, *6*, 32561–32574. [[CrossRef](#)]
17. Casado, E.; Garcia, V.M.; Sanchez, J.J.; Blanco, M.; Maurel, J.; Feliu, J.; Fernandez-Martos, C.; de Castro, J.; Castelo, B.; Belda-Iniesta, C.; et al. A combined strategy of SAGE and quantitative PCR Provides a 13-gene signature that predicts preoperative chemoradiotherapy response and outcome in rectal cancer. *Clin. Cancer Res.* **2011**, *17*, 4145–4154. [[CrossRef](#)] [[PubMed](#)]
18. Agostini, M.; Zangrando, A.; Pastrello, C.; D’Angelo, E.; Romano, G.; Giovannoni, R.; Giordan, M.; Maretto, I.; Bedin, C.; Zanon, C.; et al. A functional biological network centered on XRCC3: A new possible marker of chemoradiotherapy resistance in rectal cancer patients. *Cancer Biol. Ther.* **2015**, *16*, 1160–1171. [[CrossRef](#)] [[PubMed](#)]
19. Millino, C.; Maretto, I.; Pacchioni, B.; Digito, M.; De Paoli, A.; Canzonieri, V.; D’Angelo, E.; Agostini, M.; Rizzolio, F.; Giordano, A.; et al. Gene and MicroRNA Expression Are Predictive of Tumor Response in Rectal Adenocarcinoma Patients Treated With Preoperative Chemoradiotherapy. *J. Cell. Physiol.* **2017**, *232*, 426–435. [[CrossRef](#)]
20. Hur, H.; Tulina, I.; Cho, M.S.; Min, B.S.; Koom, W.S.; Lim, J.S.; Ahn, J.B.; Kim, N.K. Biomarker-Based Scoring System for Prediction of Tumor Response After Preoperative Chemoradiotherapy in Rectal Cancer by Reverse Transcriptase Polymerase Chain Reaction Analysis. *Dis. Colon. Rectum.* **2016**, *59*, 1174–1182. [[CrossRef](#)]
21. Huh, J.W.; Lee, J.H.; Kim, H.R. Pretreatment expression of 13 molecular markers as a predictor of tumor responses after neoadjuvant chemoradiation in rectal cancer. *Ann. Surg.* **2014**, *259*, 508–515. [[CrossRef](#)]
22. Ferrandon, S.; DeVecchio, J.; Duraes, L.; Chouhan, H.; Karagkounis, G.; Davenport, J.; Orloff, M.; Liska, D.; Kalady, M.F. CoA Synthase (COASY) Mediates Radiation Resistance via PI3K Signaling in Rectal Cancer. *Cancer Res.* **2020**, *80*, 334–346. [[CrossRef](#)]
23. Gim, J.; Cho, Y.B.; Hong, H.K.; Kim, H.C.; Yun, S.H.; Wu, H.G.; Jeong, S.Y.; Joung, J.G.; Park, T.; Park, W.Y.; et al. Predicting multi-class responses to preoperative chemoradiotherapy in rectal cancer patients. *Radiat. Oncol.* **2016**, *11*, 50. [[CrossRef](#)]
24. Alderdice, M.; Richman, S.D.; Gollins, S.; Stewart, J.P.; Hurt, C.; Adams, R.; McCorry, A.M.; Roddy, A.C.; Vimalachandran, D.; Isella, C.; et al. Prospective patient stratification into robust cancer-cell intrinsic subtypes from colorectal cancer biopsies. *J. Pathol.* **2018**, *245*, 19–28. [[CrossRef](#)]
25. Ji, D.; Song, C.; Li, Y.; Xia, J.; Wu, Y.; Jia, J.; Cui, X.; Yu, S.; Gu, J. Combination of radiotherapy and suppression of Tregs enhances abscopal antitumor effect and inhibits metastasis in rectal cancer. *J. Immunother. Cancer* **2020**, *8*, e000826. [[CrossRef](#)]
26. Sendoya, J.M.; Iseas, S.; Coraglio, M.; Golubicki, M.; Robbio, J.; Salanova, R.; Kujaruk, M.; Mikolaitis, V.; Rizzolo, M.; Ruiz, G.; et al. Pre-Existing Tumoral B Cell Infiltration and Impaired Genome Maintenance Correlate with Response to Chemoradiotherapy in Locally Advanced Rectal Cancer. *Cancers* **2020**, *12*, 2227. [[CrossRef](#)] [[PubMed](#)]
27. Kennedy, R.D.; Quinn, J.E.; Mullan, P.B.; Johnston, P.G.; Harkin, D.P. The role of BRCA1 in the cellular response to chemotherapy. *J. Natl. Cancer Inst.* **2004**, *96*, 1659–1668. [[CrossRef](#)] [[PubMed](#)]
28. Chalasani, P.; Livingston, R. Differential chemotherapeutic sensitivity for breast tumors with “BRCAness”: A review. *Oncologist* **2013**, *18*, 909–916. [[CrossRef](#)] [[PubMed](#)]

29. Gao, Y.; Zhu, J.; Zhang, X.; Wu, Q.; Jiang, S.; Liu, Y.; Hu, Z.; Liu, B.; Chen, X. BRCA1 mRNA expression as a predictive and prognostic marker in advanced esophageal squamous cell carcinoma treated with cisplatin- or docetaxel-based chemotherapy/chemoradiotherapy. *PLoS ONE* **2013**, *8*, e52589. [[CrossRef](#)]
30. Xu, S.; Yu, Y.; Rong, J.; Hu, D.; Zhang, L.; Fu, S.; Yang, H.; Fan, J.; Yang, L.; Wu, J. Expression of BRCA1 and ERCC1 as predictive clinical outcome after radiochemotherapy in patients with locoregionally moderate-advanced nasopharyngeal carcinoma. *Oncotarget* **2017**, *8*, 31355–31367. [[CrossRef](#)]
31. He, Q.; Zhang, M.; Zhang, J.; Zhong, S.; Liu, Y.; Shen, J.; He, J.; Jiang, L.; Yang, C.; Zeng, Y.; et al. Predictive value of BRCA1 expression on the efficacy of chemotherapy based on anti-microtubule agents: A pooled analysis across different malignancies and agents. *Ann. Transl. Med.* **2016**, *4*, 110. [[CrossRef](#)] [[PubMed](#)]
32. Abdulkareem, N.M.; Bhat, R.; Qin, L.; Vasaikar, S.; Gopinathan, A.; Mitchell, T.; Shea, M.J.; Nanda, S.; Thangavel, H.; Zhang, B.; et al. A novel role of ADGRF1 (GPR110) in promoting cellular quiescence and chemoresistance in human epidermal growth factor receptor 2-positive breast cancer. *FASEB J.* **2021**, *35*, e21719. [[CrossRef](#)] [[PubMed](#)]
33. Masuda, M.; Uno, Y.; Ohbayashi, N.; Ohata, H.; Mimata, A.; Kukimoto-Niino, M.; Moriyama, H.; Kashimoto, S.; Inoue, T.; Goto, N.; et al. TNIK inhibition abrogates colorectal cancer stemness. *Nat. Commun.* **2016**, *7*, 12586. [[CrossRef](#)]
34. Hirozane, T.; Masuda, M.; Sugano, T.; Sekita, T.; Goto, N.; Aoyama, T.; Sakagami, T.; Uno, Y.; Moriyama, H.; Sawa, M.; et al. Direct conversion of osteosarcoma to adipocytes by targeting TNIK. *JCI Insight.* **2021**, *6*. [[CrossRef](#)] [[PubMed](#)]
35. Petrelli, F.; Sgroi, G.; Sarti, E.; Barni, S. Increasing the Interval Between Neoadjuvant Chemoradiotherapy and Surgery in Rectal Cancer: A Meta-analysis of Published Studies. *Ann. Surg.* **2016**, *263*, 458–464. [[CrossRef](#)]
36. Lefevre, J.H.; Mineur, L.; Kotti, S.; Rullier, E.; Rouanet, P.; de Chaisemartin, C.; Meunier, B.; Mehrdad, J.; Cotte, E.; Desrame, J.; et al. Effect of Interval (7 or 11 weeks) Between Neoadjuvant Radiochemotherapy and Surgery on Complete Pathologic Response in Rectal Cancer: A Multicenter, Randomized, Controlled Trial (GRECCAR-6). *J. Clin. Oncol.* **2016**, *34*, 3773–3780. [[CrossRef](#)] [[PubMed](#)]
37. Kim, S.H.; Chang, H.J.; Kim, D.Y.; Park, J.W.; Baek, J.Y.; Kim, S.Y.; Park, S.C.; Oh, J.H.; Yu, A.; Nam, B.H. What Is the Ideal Tumor Regression Grading System in Rectal Cancer Patients after Preoperative Chemoradiotherapy? *Cancer Res. Treat.* **2016**, *48*, 998–1009. [[CrossRef](#)]
38. De Palma, F.D.E.; Luglio, G.; Tropeano, F.P.; Pagano, G.; D’Armiento, M.; Kroemer, G.; Maiuri, M.C.; De Palma, G.D. The Role of Micro-RNAs and Circulating Tumor Markers as Predictors of Response to Neoadjuvant Therapy in Locally Advanced Rectal Cancer. *Int. J. Mol. Sci.* **2020**, *21*, 7040. [[CrossRef](#)]
39. Cho, E.; Park, I.J.; Yeom, S.S.; Hong, S.M.; Lee, J.B.; Kim, Y.W.; Kim, M.J.; Lim, H.M.; Lim, S.B.; Yu, C.S.; et al. A Multigene Model for Predicting Tumor Responsiveness After Preoperative Chemoradiotherapy for Rectal Cancer. *Int. J. Radiat. Oncol. Biol. Phys.* **2019**, *105*, 834–842. [[CrossRef](#)]
40. Park, I.J.; Yu, Y.S.; Mustafa, B.; Park, J.Y.; Seo, Y.B.; Kim, G.D.; Kim, J.; Kim, C.M.; Noh, H.D.; Hong, S.M.; et al. A Nine-Gene Signature for Predicting the Response to Preoperative Chemoradiotherapy in Patients with Locally Advanced Rectal Cancer. *Cancers* **2020**, *12*, 800. [[CrossRef](#)]
41. Kamran, S.C.; Lennerz, J.K.; Margolis, C.A.; Liu, D.; Reardon, B.; Wankowicz, S.A.; Van Seventer, E.E.; Tracy, A.; Wo, J.Y.; Carter, S.L.; et al. Integrative Molecular Characterization of Resistance to Neoadjuvant Chemoradiation in Rectal Cancer. *Clin. Cancer Res.* **2019**, *25*, 5561–5571. [[CrossRef](#)]
42. Garcia-Aguilar, J.; Chen, Z.; Smith, D.D.; Li, W.; Madoff, R.D.; Cataldo, P.; Marcet, J.; Pastor, C. Identification of a biomarker profile associated with resistance to neoadjuvant chemoradiation therapy in rectal cancer. *Ann. Surg.* **2011**, *254*, 486–492, discussion 492–483. [[CrossRef](#)] [[PubMed](#)]
43. Toomey, S.; Gunther, J.; Carr, A.; Weksberg, D.C.; Thomas, V.; Salvucci, M.; Bacon, O.; Sherif, E.M.; Fay, J.; Kay, E.W.; et al. Genomic and Transcriptomic Characterisation of Response to Neoadjuvant Chemoradiotherapy in Locally Advanced Rectal Cancer. *Cancers* **2020**, *12*, 1808. [[CrossRef](#)]
44. do Canto, L.M.; Barros-Filho, M.C.; Rainho, C.A.; Marinho, D.; Kupper, B.E.C.; Begnami, M.; Scapulatempo-Neto, C.; Havelund, B.M.; Lindebjerg, J.; Marchi, F.A.; et al. Comprehensive Analysis of DNA Methylation and Prediction of Response to Neoadjuvant Therapy in Locally Advanced Rectal Cancer. *Cancers* **2020**, *12*, 3079. [[CrossRef](#)] [[PubMed](#)]
45. Akiyoshi, T.; Tanaka, N.; Kiyotani, K.; Gotoh, O.; Yamamoto, N.; Oba, K.; Fukunaga, Y.; Ueno, M.; Mori, S. Immunogenomic profiles associated with response to neoadjuvant chemoradiotherapy in patients with rectal cancer. *Br. J. Surg.* **2019**, *106*, 1381–1392. [[CrossRef](#)]
46. El Sissy, C.; Kirilovsky, A.; Van den Eynde, M.; Musina, A.M.; Anitei, M.G.; Romero, A.; Marliot, F.; Junca, A.; Doyen, J.; Mlecnik, B.; et al. A Diagnostic Biopsy-Adapted Immunoscore Predicts Response to Neoadjuvant Treatment and Selects Patients with Rectal Cancer Eligible for a Watch-and-Wait Strategy. *Clin. Cancer Res.* **2020**, *26*, 5198–5207. [[CrossRef](#)]
47. Goncalves-Ribeiro, S.; Sanz-Pamplona, R.; Vidal, A.; Sanjuan, X.; Guillen Diaz-Maroto, N.; Soriano, A.; Guardiola, J.; Albert, N.; Martinez-Villacampa, M.; Lopez, I.; et al. Prediction of pathological response to neoadjuvant treatment in rectal cancer with a two-protein immunohistochemical score derived from stromal gene-profiling. *Ann. Oncol.* **2017**, *28*, 2160–2168. [[CrossRef](#)]