miR-17-92a-1 cluster host gene (MIR17HG) evaluation and response to neoadjuvant chemoradiotherapy in rectal cancer

Abstract: Neoadjuvant chemoradiotherapy (NCRT) followed by surgery is the gold standard for the treatment of patients with locally advanced rectal cancer (LARC). However, response is variable, and no predictive markers have been validated. The amplification of 13q31–34 seemed to distinguish between nonresponders and responders to NCRT. The miR-17-92a-1 cluster host gene (MIR17HG), which is involved in the development, progression, and aggressiveness of colorectal cancer, and the ABCC4 gene, an ATP-binding cassette transporter, are located at this region. Moreover, the transcription factor c-Myc is closely related to MIR17HG. The aim of this study was to examine the role of MIR17HG, ABCC4, and CMYC gene copy numbers (CNs) in determining response to NCRT. We analyzed DNA CN of pretherapy biopsies from 108 LARC patients and the expression of microRNA (miR)-17, miR-18a, miR-19a, miR-19b-1, miR-20a, and miR-92a-1 in 34 biopsies. MIR17HG, CMYC, and ABCC4 gene CNs were frequently altered in pretreatment tumors, amplification being the most frequent alteration. With regard to response to therapy, 41% of responders showed MIR17HG deletion, while MIR17HG amplification was observed in 41% of nonresponders. With regard to pathological T stage (ypT), a higher percentage of ypT3–4 than ypT0–2 tumors showed MIR17HG amplification. Finally, a higher, albeit nonsignificant, variability in the expression of MIR17HG cluster members was detected in nonresponders compared to responders. No association was observed between clinical pathological parameters and ABCC4 or CMYC CN. Our data did not highlight a significant association between MIR17HG, CMYC, and ABCC4 gene CNs and response to NCRT in LARC. However, MIR17HG gene amplification would seem to be related to a lack of response. Evaluation of the expression of MIR17HG cluster members is warranted in a larger case series, together with functional studies, to evaluate the potential of this gene as a new predictive marker.

Keywords: rectal cancer, neoadjuvant chemoradiotherapy, gene copy number, MIR17HG, ABCC4, CMYC

Introduction

The treatment of choice for the majority of patients with locally advanced rectal cancer (LARC) is neoadjuvant chemoradiotherapy (NCRT). Although this treatment strategy is associated with an overall benefit for patients, it obtains a wide range of responses ranging from complete disease regression to no response. However, such responses cannot be predicted via conventional clinical pathological characteristics. Numerous studies have analyzed the correlation between expression levels of candidate genes or global genetic profiles and response to therapies, but the predictive role of such genes is still controversial, and needs to be validated prospectively in large independent case series.
In a previous paper on the characterization of genomic alterations in rectal pretherapy tumor specimens, we showed that a gain of several clones spanning the 13q31–34 region seems capable of distinguishing between responsive and unresponsive tumors. This region is often amplified in tumors, including colorectal cancer (CRC) and its preferential gain in nonresponders, and is of interest because of the importance of a number of genes located within this region, eg, ABCC4, CLDN10, WAVE3, XPO4, CDC16, GPC3, and GPC6. At the 13q31 chromosomal level, in addition to the GPC5 gene, there is another annotated gene (c13orf25) whose increased expression strongly correlates with the presence of the amplicon in malignant lymphoma and to a lesser extent in CRC. The c13orf25 transcript is the functional precursor of six microRNAs (miRs), ie, small, noncoding RNAs that regulate gene expression: miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1, and miR-92a-1. This polycistronic gene, called MIR17HG or Oncomir-1 because of its mainly oncogenic functions (ie, activities that enhance cell proliferation, inhibit apoptosis, and modulate angiogenesis), is related to the homologous MIR106a–363 cluster on chromosome X and the MIR106b–25 cluster on chromosome 7. Functional studies have confirmed the pivotal role of members of the MIR17HG in the development, progression, and aggressiveness of CRC. Furthermore, an association has been found between its increased expression, probably induced by c-Myc, and DNA copy number (CN) gain on 13q31, which ranges from TRG0-1 (absence of response) to TRG4 (complete response). The study was approved by the Area Vasta Romagna institutional review board, in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Written informed consent was obtained from all patients.

Sample size was determined using previous results, hypothesizing \(\alpha=0.05, \beta=0.20\), and an odds ratio equal to 4. Only patients who received the full therapy regimen were included in the study.

**Materials and methods**

**Patients and sample collection**

Formalin-fixed, paraffin-embedded (FFPE) pretherapy biopsies were retrospectively collected from a series of 120 patients with LARC (mid/low rectum, maximum of 12 cm from the anal verge) who were candidates for homogeneous NCRT. After computed tomography evaluation to determine tumor volume and stage, all patients were submitted to homogeneous treatment consisting of radiotherapy (50.4 Gy for 5 weeks with conventional fractionation), together with a daily dose of 225 mg/m² fluorouracil infused by central catheter during radiotherapy or 825 mg/m² capecitabine twice daily. Surgery was scheduled 6–8 weeks after completion of therapy. Assessment of pathological response was carried out using tumor-regression grade (TRG) according to Dworak et al’s classification, which ranges from TRG0-1 (absence of response) to TRG4 (complete response). The study was approved by the Area Vasta Romagna institutional review board, in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Written informed consent was obtained from all patients.

Sample size was determined using previous results, hypothesizing \(\alpha=0.05, \beta=0.20\), and an odds ratio equal to 4. Only patients who received the full therapy regimen were included in the study.

**DNA extraction and copy number analysis**

DNA was extracted from FFPE pretherapy biopsies using a QIAamp DNA FFPE tissue kit (Qiagen NV, Venlo, the Netherlands), according to the manufacturer’s instructions. Only biopsies containing >60% of tumor cells were macrodissected and processed. DNA quality was evaluated by spectrophotometric quantification (NanoDrop ND-1000; Thermo Fisher Scientific, Waltham, MA, USA) and agarose-gel electrophoresis, and samples with DNA fragments <200 bp were rejected. FFPE DNA (60 ng) were amplified using the GenomePlex® whole-genome amplification (WGA2) kit (Sigma-Aldrich, St Louis, MO, USA), according to the manufacturer’s instructions. Commercial genomic DNA (Thermo Fisher Scientific) was used as positive control for each reaction. WGA DNA was purified using the GenElute polymerase chain reaction (PCR) cleanup kit (Sigma-Aldrich). The WGA DNA quality was evaluated by loading DNA onto 1.5% agarose gel and considered suitable when mean size was about 400 bp (range 100–1,000 bp).

DNA CN was analyzed by real-time quantitative PCR (qPCR) using a TaqMan® CN assay (Thermo Fisher Scientific) for MIR17HG (Hs03861872), CMYC (Hs01764918),
and ABCC4 (Hs05289077). RNase P (4403326) was used as endogenous control. Multiplex qPCR reactions were performed in triplicate according to the manufacturer’s protocol and run on a 7500 real-time PCR system (Thermo Fisher Scientific). The maximum tolerated difference between triplicates to calculate the average cycle threshold (Ct) was 0.5. FFPE samples with RNase P Ct over 32 were qualified as noncontributory and not used for CN evaluation. A pool of genomic DNA extracted from colorectal FFPE tissues of healthy individuals and preamplified by WGA in the same manner as that of tumor samples was used for calibration. The relative CN, normalized to endogenous control, was calculated using $2^{-\Delta\Delta Ct}$, as described previously, using CopyCaller software version 2.0 (Thermo Fisher Scientific). Amplification was defined as CN >2.5, and deletion as CN ≤1.5.

miR extraction and real-time reverse-transcriptase quantitative PCR

Only samples with tumor content >80% were considered for miR expression analysis. Freshly cut tissue sections (20 μm) were first deparaffinized using xylene and ethanol washes. Total RNA was isolated using the Norgen FFPE RNA purification kit (Norgen Biotek, Thorold, ON, Canada) according to the manufacturer’s instructions and quantified using the NanoDrop ND-1000. The detection and quantification of mature miR was carried out by real-time reverse-transcriptase qPCR (RT-qPCR), according to the manufacturer’s protocol, starting with 10 ng of the RNA as input material and adopting small nuclear RNA (RNU6B) as endogenous reference. The following reagents were used: TaqMan miR reverse-transcription kit, TaqMan miR individual assays for hsa-miR-17 (assay ID 002308), hsa-miR-18a (assay ID 002422), hsa-miR-19a (assay ID 000395), hsa-miR-19b-1 (assay ID 000396), hsa-miR-20a (assay ID 000580), hsa-miR-92a-1 (assay ID 000431), and RNU6B (assay ID 001093), and TaqMan Universal PCR Master Mix without AmpErase® UNG (Thermo Fisher Scientific). RT-qPCR reactions for each target were performed in triplicate according to the manufacturer’s protocol and run on the 7500 real-time PCR system. The maximum tolerated difference between triplicates to calculate the average Ct was 0.5. The relative expression of each miR was calculated according to the $2^{-\Delta\Delta Ct}$ quantification method.

Statistical analysis

Categorical data are expressed as percentages, and continuous variables as medians and ranges. The $\chi^2$ or Fisher’s exact test was used to analyze the association between gene CN and clinical pathological characteristics. The Wilcoxon–Mann–Whitney and Kruskal–Wallis tests were used to analyze the potential correlation between miR expression and clinical pathological features or CN values. A box plot was used to visualize descriptive statistics of the different groups. $P$-values <0.05 were considered statistically significant. Statistical analyses were carried out with SAS statistical software (version 9.3; SAS Institute, Cary, NC, USA) and with Stata/MP 10.1 for Windows (StataCorp LP, College Station, TX, USA).

Results

Patient characteristics

The clinical pathological characteristics of patients included in this study are summarized in Table 1. Only 108 of 120 patients had pretherapy DNA of sufficient quality for the CN evaluation. With regard to response to NCRT, patients were grouped into two subsets according to TRG criteria: 59%

<table>
<thead>
<tr>
<th>Variable</th>
<th>n</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>33</td>
<td>31</td>
</tr>
<tr>
<td>Male</td>
<td>75</td>
<td>69</td>
</tr>
<tr>
<td>Median age, years (range)</td>
<td>67</td>
<td>(31–80)</td>
</tr>
<tr>
<td>Tumor invasion before NCRT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>uT2</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>uT3</td>
<td>95</td>
<td>88</td>
</tr>
<tr>
<td>uT4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Not available</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Lymph-node metastasis before NCRT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>uN0</td>
<td>53</td>
<td>49</td>
</tr>
<tr>
<td>uN1</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>uN2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>uNx</td>
<td>19</td>
<td>17</td>
</tr>
<tr>
<td>Not available</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Tumor invasion after NCRT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ypT0</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>ypT1</td>
<td>11</td>
<td>10</td>
</tr>
<tr>
<td>ypT2</td>
<td>31</td>
<td>29</td>
</tr>
<tr>
<td>ypT3</td>
<td>39</td>
<td>36</td>
</tr>
<tr>
<td>ypT4</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Not available</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Lymph-node metastasis after NCRT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ypN0</td>
<td>77</td>
<td>71</td>
</tr>
<tr>
<td>ypN1</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>ypN2</td>
<td>7</td>
<td>6</td>
</tr>
<tr>
<td>ypNx</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Not available</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>TRG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRG0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>TRG1</td>
<td>23</td>
<td>21</td>
</tr>
<tr>
<td>TRG2</td>
<td>38</td>
<td>35</td>
</tr>
<tr>
<td>TRG3</td>
<td>25</td>
<td>23</td>
</tr>
<tr>
<td>TRG4</td>
<td>19</td>
<td>18</td>
</tr>
</tbody>
</table>

Abbreviations: NCRT, neoadjuvant chemoradiotherapy; TRG, tumor-regression grade; u, ultrasound; yp, pathological, after a neoadjuvant therapy.

Table 1 Clinical pathological characteristics of patients in the study
with TRG0–2 were defined as nonresponders, while those with TRG3–4 (41%) were considered responders. Among these, 24% and 18% belonged to extreme response classes, i.e., absence of response (TRG0–1) and complete response (TRG4), respectively. Moreover, 45% of patients did not show any downstaging when the ultrasound stage (uT) and the pathological stage after therapy (ypT) were compared, while 55% showed a decrease in at least one T stage.

**Determination of MIR17HG, CMYC, and ABCC4 DNA copy number**

ABCC4 CN was available in 104 of 108 samples, whereas all cases were evaluable for MIR17HG and CMYC. 68%, 80%, and 73% of patients showed alterations in MIR17HG, CMYC, and ABCC4 CN, respectively, and 35%, 74%, and 62% of cases were amplified. We then examined whether the CN of selected genes was associated with clinical and pathological features. The distribution of deleted, normal, and amplified cases in response classes is reported for each gene in Figure 1 (not significant). Only MIR17HG showed a trend toward significance: although 32% of diploid patients were equally distributed among the response classes, 41% of responders showed MIR17HG deletion, while 41% of nonresponders showed MIR17HG amplification. The distribution of amplified samples in response classes was significant ($P=0.023$) for MIR17HG. Moreover, considering only extreme response classes, the difference in the percentage of amplified samples was even higher (42% in TRG0–1 versus 26% TRG4). Although no significant association was found between MIR17HG CN and age, sex, ypN, or ypT, it was interesting to see that within the latter group, 21%, 33%, and 37% of ypT0, ypT1–2, and ypT3–4 showed MIR17HG amplification, respectively. Moreover, a significant difference between ypT classes was observed when only amplified samples were considered ($P=0.026$). No relation with any clinical pathological features was observed for CMYC or ABCC4 CN.

**MIR17HG expression**

To study the potential role of MIR17HG in response to NCRT in LARC, expression levels of the six MIR17HG cluster members (miR-17, miR-18a, miR-19a, miR-19b-1, miR-20a, and miR-92a-1) were also quantified by RT-qPCR in 34 cases for which RNA was available. In particular, the expression of all six miRs was detected in 23 of 34 cases, miR-19b-1 and miR-92a-1 showed the highest expression level, while miR-18a showed the lowest. Although no significant association with response to treatment was found for any of the miRs analyzed, higher gene expression variability was detected in nonresponders than responders (Figure 2). In particular, considering only extreme response classes, miR-19a, miR-19b-1, and miR-92a-1 showed higher levels in TRG0–1 compared to TRG4, whereas miR-17, miR-18a, and miR-20a showed lower levels. The association between MIR17HG members expression and CN did not reach statistical significance.

Expression levels of each miR were compared with clinical and pathological characteristics of patients to evaluate whether changes in specific miRs belonging to the MIR17HG were associated with any of these parameters. However, no significant associations were observed between miR levels and any of the clinicopathological characteristics analyzed.

**Discussion**

In the present study, we investigated MIR17HG, CMYC, and ABCC4 gene CN alterations in LARC patients, focusing in particular on their relation to NCRT response. We found that all genes were frequently altered and a high number of amplifications detected, both common events in CRC, which is characterized by chromosomal instability, often showing 8q and 13q amplifications.26,27

Our results were suggestive of the involvement of MIR17HG gene amplification in the lack of response to NCRT, as 41% of responders showed deletion and 41% of nonresponders showed amplification of this locus. Moreover, patients who did not exhibit any regression (TRG0–1) showed an even higher percentage of MIR17HG gene amplification than completely responsive patients (TRG4). Moreover, a higher percentage of MIR17HG amplified tumors was observed in patients who did not show downstaging after NCRT (ypT3–4). These results, albeit not significant, are in...
Figure 2 Expression levels of each of the MIR17HG cluster members in different response classes. Relative expression is log₂ relative expression ($2^{-\Delta Ct}$). Box plots show median and 25th and 75th percentiles and range of expression levels. Black bars inside the box plots indicate the log₂ median value of microRNA (miR)-expression levels.

**Abbreviations:** NR, non responders; R, responders.
line with our previously published data showing that amplification of several clones spanning the 13q31–34 region helps to distinguish between completely unresponsive and responsive tumors, as the amplification occurs more frequently in TRG1 than in TRG4 patients.5

To determine the impact of CN dosage, we also analyzed expression levels of each MIR17HG cluster member in a small number of samples for which miR extraction was possible. Overall, miR-92a-1 and miR-19b-1 were transcribed at higher levels than the other four miRs, while miR-18a reached the lowest level, in agreement with results from other studies.10,14 A significant association between the expression of each miR and gene CN was not observed. However, although the amplification of the MIR17HG locus in CRC is associated with an increased expression of this miR cluster,14 the frequency and degree of overexpression is lower than those of other tumors harboring the amplification, eg, lymphomas.6 Our results, on the other hand, were suggestive of an upregulation of miR-19a, miR-19b-1, and miR-92a-1 in patients who did not show regression after therapy. There is some evidence that aberrant miR expression may play a crucial role in the response to anticancer drugs.28 Studies on RC cell lines have also concluded that the expression of some MIR17HG cluster members may be related to chemotheraphy response.29,30 Moreover, it has been demonstrated that miR-19b-1 overexpression is associated with fluorouracil resistance in cell lines.30 Finally, Chen et al showed that upregulated serum miR-19a could serve as a biomarker in predicting resistance to first-line FOLFOX (folic acid, fluorouracil, oxaliplatin) chemotherapy in CRC patients.31

Although less is known about the impact of miR on radiosensitivity or radioresistance,32 MIR17HG has been shown to play a role in response to radiation in tumors other than LARC. In their study on human mantle-cell lymphoma cells, Jiang et al were the first to show that this miR plays an important role in the radioresistance of tumor cells, as its upregulation is associated with a downregulation of PTEN and PHLPP2.17 With regard to LARC patients submitted to NCRT and miR analyses, the few studies published in this area to date have evaluated only small case series, used different response assessments, and found different predictive miR signatures.33–39 Among others, Hotchi et al reported a lower expression of miR-17, miR-20a, and miR-92a-1 in LARC responders than in nonresponders.40 Interestingly, a recent paper showed that lymph-node status after NCRT could be predicted on the basis of circulating miR-20a and miR-18b levels, their reduced expression being associated with postoperative lymph-node negativity.41 Our results indicated an upregulation of miR-19a, miR-19b-1, and miR-92a-1 in patients who did not have tumor regression after therapy, which is in line with their role in resistance to fluorouracil and radiation, respectively.

The oncogenic MIR17HG cluster has proved to be an important player in c-Myc signaling, being one of its main transcriptional targets.42 We found that 62% of pretherapy samples showed CMYC gene amplification, which however were equally distributed independently of the response to NCRT. Another critical gene mapped on 13q31–34 is ABCC4, an ABC transporter known to be involved in multidrug resistance. Yu et al demonstrated a significant association between sensitivity to NCRT and downregulation of ABCC4, as well as between high ABCC4 expression and shorter disease-free survival,20,21 indicating a possible role of this protein as a predictive and prognostic biomarker in LARC. Despite a high frequency of amplification, we did not find any association between this alteration and response to NCRT, suggesting that the different expression levels observed between responders and nonresponders in other studies may not be attributable to CN alterations.

To the best of our knowledge, ours is the first study to assess MIR17HG gene CN in relation to response to NCRT in LARC. The limitations of our work include the small number of samples for which miR extraction was possible, thus making all the results on MIR17HG gene expression of a preliminary nature. Another critical aspect may be the correct definition of response classes for which there is no general consensus. In fact, although the use of TRG and in particular the Dworak regression system is a widely accepted approach, other clinical pathological features may also need to be taken into account, ie, ypTNM classification.22,43 Significant variations in definitions of responders and nonresponders, in addition to the intrinsic subjectivity of these definitions, may be critical in this setting. For these reasons, we also analyzed data considering only extreme TRG response classes (complete response and no response), as we felt that these were less ambiguous and more easily recognizable response parameters and could thus better highlight markers that are indicative of response.

Although NCRT is one of the preferred treatment options in LARC, it can lead to important problems in nonresponders, who account for about a third of all patients treated.44 For these, there are no expected benefits and a poor postsurgery outcome. The availability of new molecular predictors of response would enable clinicians to rationalize therapy and choose the most effective therapeutic interventions for each patient, thus improving the cost–benefit
of treatment. Moreover, prior knowledge of resistance to conventional treatments would enable alternative therapies to be offered in both neoadjuvant and adjuvant settings. The relation between MIR17HG gene amplification and resistance to treatment, if confirmed, could be of value from a clinical point of view, as CN evaluation can also be performed by fluorescence in situ hybridization on diagnostic biopsies. Further analyses of MIR17HG expression in larger case series could also provide important information about the real role of these miR-cluster members in determining the response to NCRT in LARC. If miR-19a, miR-19b-1, and miR-92 overexpression in pretherapy biopsies were confirmed to be associated with the lack of response, the assessment of their levels in fluid biopsies could be proposed in prospective studies for a more rational and personalized management of LARC patients.

**Conclusion**

Of the three genes studied, only MIR17HG CN would seem to be related to response to NCRT in LARC. Further investigation into the potential application of MIR17HG evaluation as a novel predictive biomarker is warranted.

**Acknowledgment**

The authors wish to thank Ursula Elbling for editing the manuscript.

**Disclosure**

The authors report no conflicts of interest in this work.

**References**