Overexpression of HMGA2 Promotes Metastasis and Impacts Survival of Colorectal Cancers

Xiaochen Wang1, 4*, Xiyong Liu4*, Angela Ying-Jian Li4, Lirong Chen2, Lily Lai5, Her Helen Lin4, Shuya Hu4, Lifang Yao2, Jiaping Peng3, Sofia Loera6, Lijun Xue4, Bingsen Zhou4, Lun Zhou3, Shu Zheng3, Peiguo Chu6, Suzhan Zhang††, David Kong Ann††, and Yun Yen4††

1Department of Surgical Oncology, 2 Department of Pathology, 3Cancer Institute, Second Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China 310009 4Department of Molecular Pharmacology, 5Department of Surgery, 6Department of Anatomic Pathology, Beckman Research Institute, City of Hope Comprehensive Cancer, Duarte, CA 91010

Running title: HMGA2 as a biomarker of poor prognosis in colorectal cancer

*Co-first Authors: Xiaochen Wang and Xiyong Liu

†Co-corresponding author: Suzhan Zhang and David Ann

††Corresponding author: Yun Yen

Suzhan Zhang, Department of Surgical Oncology, Second Affiliated Hospital, Zhejiang University School of Medicine, 88 Jiefang Road, Hangzhou, China 310009. E-mail: zuci@zju.edu.cn

David Kong Ann, Department of Molecular Pharmacology, Beckman Research Institute, City of Hope Comprehensive Cancer Center, 1500 East Duarte Road, Duarte, CA 91010-3000. Phone: 626-359-8111, ext. 64967; Fax: 626-471-3607; E-mail: dann@coh.org

Yun Yen, Department of Molecular Pharmacology, Beckman Research Institute, City of Hope Comprehensive Cancer Center, 1500 East Duarte Road, Duarte, CA 91010-3000. Phone: 626-359-8111, ext. 65707; Fax: 626-471-3607; E-mail: vyen@coh.org
Abbreviations:

HMGA2: high mobility group A2;
CRC: colorectal cancer;
FFPE: formalin-fixed, paraffin-embedded tissue block;
IHC: immunohistochemistry;
MTA: multiple tissue array;
MTB: multiple tissue board;
OR: odds ratio;
HR: hazard ratio;
95% CI: 95% confidence interval;
OS: overall survival;
PFS: progress-free survival;
DSB: DNA double strand breaks;
NHEJ: non-homologous end joining.
Statement of Translational Relevance

The high mobility group A2 (HMGA2) protein, a non-histone chromosomal architectural protein, regulates the expression of a cohort of genes by either enhancing or suppressing their transcription. A change in the HMGA2 mRNA or protein level correlates with the malignant phenotype. In this report, we investigated the hypothesis that the expression level of HMGA2 is related to cancer survivability and serves as a prognostic biomarker for the clinical outcome of colorectal cancer. HMGA2 overexpression led to suppression of non-homologous end joining (NHEJ) pathway in repairing DNA double strand breaks (DSBs). CRC patients with HMAG2-positive staining in primary tumors had augmented efficacy of adjuvant radiotherapy. Therefore, HMGA2 is a potential predictive biomarker for colorectal cancers’ therapeutic outcome.
Abstract

**Purpose:** This study aims to address hypothesis that the high mobility group A2 (HMGA2), an oncofetal protein, relates to survivability and serves as a prognostic biomarker for colorectal cancer (CRC).

**Experimental Design:** This is a retro-prospective multiple center study. The HMGA2 expression level was determined by performing immunohistochemistry (IHC) on surgical tissue samples of 89 CRCs from a training set and 191 CRCs from a validation set. The Kaplan-Meier analysis and COX proportional hazard model were employed to analyze survivability.

**Results:** Multivariate logistic analysis indicated that the expression of HMGA2 significantly correlates with distant metastasis in training set (odd ratio, OR=3.53, 95% CI 1.37-9.70) and validation set (OR=6.38, 95% CI 1.47-43.95). Survival analysis revealed that the overexpression of HMGA2 is significantly associated with poor survival of CRC patients ($p < 0.05$). The adjusted hazard ratios (HRs) for overall survival were 2.38 (95% CI 1.30-4.34) and 2.14 (95% CI 1.21-3.79) in training and validation sets, respectively. Further investigation revealed that HMGA2 delays the clearance of $\gamma$-H2AX in HCT-116 and SW480 cells post $\gamma$-irradiation, which supports our finding that CRC patients with HMAG2 positive staining in primary tumors had augmented efficacy of adjuvant radiotherapy (HR=0.18, 95% CI 0.04-0.63).

**Conclusion:** Overexpression of HMGA2 is associated with metastasis and unequivocally occurred in parallel with reduced survival rates of patients with CRC. Therefore, HMGA2 may potentially serve as a biomarker for predicting aggressive CRC with poor survivability and as an indicator for better response of radiotherapy.

**Keywords:** high mobility group A2 protein, colorectum, adenocarcinoma, survival, metastasis
Introduction

Colorectal cancer (CRC), one of the most prevalent cancers in the world, consists of a group of histologically heterogeneous diseases involving distinct tumorigenic pathways. It ranks as the second leading death from cancer in the United States (1) and the fourth in China (2). Mortality from CRC is mainly due to metastatic disease detected in at least half of the cases (3). Untreated patients with liver metastases have a 5-year survival rate of a mere 2% (4). Even though numerous genes have been implicated in colon tumorigenesis, only a few of them have been validated as biomarkers for predicting metastasis and treatment response.

High mobility group A (HMGA) gene family includes four chromatin-binding proteins: HMGA1a, HMGA1b, HMGA1c, and HMGA2. Each of them contains three “AT-hooks”, the functional motif characteristic of the HMGA family (5). HMGA2 encodes a small, non-histone chromatin-associated protein that has no intrinsic transcriptional activity but can modulate transcription by altering the chromatin architecture (6, 7). In humans, the HMGA2 gene is located at chromosome 12q14 and encodes a 109 amino-acid protein. HMGA2 is expressed during embryogenesis but is absent or present at low levels in terminally differentiated tissues. In embryonic tissues, HMGA2 reportedly plays a critical role in stimulating normal cardiogenesis (8), and mouse central and peripheral neural stem cell self-renewal (9), thereby regulating cell growth and differentiation. HMGA2 is overexpressed in many malignant neoplasms (5, 10, 11), and overexpression of HMGA2 is a poor prognostic factor for lung cancer (12), oral squamous cell carcinoma (13), ovarian cancer (14), and metastatic breast cancer (15). Expression of HMGA2 has been found to be related to Dukes stages and metastasis of CRCs in a cross-section study (16). However, it is not known whether HMGA2 overexpression is associated with the survival of CRCs nor if HMGA2 is related to response of therapies for any cancer type.
In this study, we investigated whether the HMGA2 level is associated with the survival of CRC patients. We optimized the conditions of immunohistochemistry for HMGA2 and conducted an outcome study on a training set with 89 CRC cases collected from City of Hope National Medical Center. In addition, we validated our finding using 191 cases collected from the 2nd affiliated hospital of Zhejiang University. Our findings suggest that HMGA2 associates with poor prognosis and may serve as a biomarker for advanced CRC.
Materials and Methods

Patients
The eligible CRCs were collected based on inclusion and exclusion criteria. Inclusion criteria: (a) CRC with pathological diagnosis; (b) informed consent obtained or waiver of consent; (c) follow-up information available. Exclusion criteria: (a) failed to get informed consent; (b) multiple cancers; (c) lack of histological diagnosis; (d) no follow-up information. A series of assessable 89 CRCs who received surgical operation in City of Hope National Medical Center during 1980-1985 were recruited as training set. The participants in the training set include 82 Caucasian, 2 African-American, 3 Asian and 1 unknown. After surgical therapy, 37 cases have had adjuvant chemotherapy, and 11 cases had radiotherapy. Meanwhile, 191 consecutive assessable colorectal cancer patients who received surgical treatment at the 2nd affiliated hospital, Zhejiang University School of Medicine between 1999 and 2004 were entered as the validation set. All CRCs in the validation set are Chinese (Asian). In the validation set, 66 of 191 cases had adjuvant chemotherapy after surgery. All patients were followed up until June 2007 and details of their demographic and survival data were updated (Table 1). All TNM stage data was obtained from clinical and pathological diagnosis. Primary tumor samples were obtained from surgical operation.

Study design
To avoid biases, standards and conditions of Immunohistochemistry (IHC) for HMGA2 expression determination were optimized on the training set (multiple tissue board, MTB) and validated on the validation set (multiple tissue array, MTA). Careful chart review was conducted and pathoclinical data was abstracted. Variables assessed included: birth date, gender, date of diagnosis, date of operation, type of chemotherapy, date of
chemotherapy, type of radiotherapy, date of radiotherapy, TNM stage, relapse/metastasis status, date of relapse/metastasis, date of last follow-up and vital status at last follow-up. The above information was coded and entered into a CRC database. Double data entry and logic checks were used for error reduction.

Sample size was calculated using parameter estimates obtained from a pilot study previously conducted at City of Hope. Using nQuery Advisor 6.01 software, it was determined that a sample size of 190 patients would be needed for about 80% power with a two-sided $\alpha$ of 0.05.

All patients were periodically followed for survival; patients with curative operations were also followed for recurrence-free survival. The follow-up period was calculated from the date of surgery until the date of last contact. The time of disease-free survival was defined as at the time of initial surgical therapy to tumor recurrence. Metastasis or local relapse was considered evidence of tumor recurrence. Only deaths from CRC were considered as endpoint of disease specific survival.

**Construction of multiple tissue array (MTA) and multiple tissue board (MTB)**

In the training set, all tissue samples were reassembled as multiple tissue board (MTB). Each MTB contained 8-12 pieces of sections, and each piece is approximately 1mm $\times$ 10mm. The tumor blocks also contained both tumor and normal colorectal tissue samples as positive and negative controls for each IHC staining.

To validate the IHC results yield from MTB, the multiple tissue array blocks were constructed using 2mm cores taken from archival, routinely processed, paraffin-embedded, colorectal cancer specimens in the validation set. The specimens had been accessioned between 1999 and 2004 and were selected solely based on the availability of colorectal cancer in the block. Pathologic diagnosis, grade and stage were previously determined for each case by attending pathologists from the department of pathology.
Both MTAs and MTBs were stored at room temperature. To determine whether the storage time affected the signal of IHC staining, we conducted a cross-tabulation between overall HMGA2 staining and year of diagnosis among the above samples. Both likelihood ratio ($p=0.235$) and Pearson ($p=0.354$) tests yielded results indicating that the quality was not reduced due to increased storage time of samples.

**Quantitative immunohistochemistry assays**

Immunohistochemistry (IHC) was used to investigate the HMGA2 protein expression. The accuracy of IHC was validated by qRT-PCR on two parallel samples. The details of deparaffinization and IHC were described previously (17). Briefly, after deparaffinization, the endogenous peroxidase activity was blocked with 3% H$_2$O$_2$. The array slides were incubated with normal goat serum for 20 minutes, and then applied with primary antibody for 20 minutes at room temperature. After 7 minutes of hydrogen peroxide treatment, the array slides were incubated with horseradish peroxidase–labeled polymer conjugated with corresponding antibodies for 30 minutes. Then, 3, 3-diaminobenzidine (0.05 g of 3, 3-diaminobenzidine and 100 mL of 30% H$_2$O$_2$ in 100 mL of PBS) was applied for 5 and 10 minutes, respectively. Each slide was counterstained with hematoxylin (DAKO, Carpinteria, CA). PBS was used as a negative control.

The rabbit antibody against HMGA2 was purchased from Alexis BioCheck Company (Foster City, CA) and applied for IHC staining (1:1000 dilution). Pre-selection of HMGA2 antibody was according to western blot (Supplementary Fig. 1), the optimized IHC condition was based on a serial dilution. To reduce the image reader bias, an automated imaging system was employed to obtain digital images of the stained sections for subsequent quantitative analyses. In addition, each sample was evaluated by two independent investigators in a double-blind manner. Investigators reviewed and assessed the subcellular localization (e.g., cytoplasm vs. nucleus), staining intensity (e.g.,
integrated optical density), and/or percentage of stained cells (e.g., total area or percentage of cells positive) for each image. Discrepancies in samples were resolved after joint review by the readers.

In generally, HMGA2 was predominantly nuclear staining in IHC. However, the HMGA2 was heterogeneously expressed between and within tumors. Some perinuclear granulation in cytoplasm was also observed. HMGA2 expression was quantified using a visual grading system based on the extent of staining. Only immunoreactivity in the nucleus was evaluated. HMGA2 was: negative (-) : <5% positive nuclear staining from CRC cells; positive (+): >=5% and <50% positive nuclear staining from tumor cells; or strong positive (++): more than 50% positive nuclear staining from tumor cells. The HMGA2 IHC standard of negative (-), weak positive (+) and strong positive (++) were displayed on Fig. 1A, 1B and 1C, respectively. HMGA2 expression was predominantly observed in the primary cancer cells, but not in the adjacent normal colorectal epithelium (Fig. 1D, 1E & 1F). Meanwhile, strong positive HMGA2 also could be seen in metastatic CRC to liver (Fig. 1G, 1H and 1I).

Cell culture, HMGA2 transduction, γ-irradiation and Western blot analysis
HCT-116 and SW480 colorectal cancer cell lines were cultured in Leibovitz’s L15 medium (Gibco) supplemented with 10% fetal bovine serum (Hyclone). Lentiviral vectors pRRLsin.hCMV-HMGA2, pΔ8.7, and pVSV-G were constructed and used for lentiviral production in HEK 293T cells as previously described (18). HCT116 and SW480 cells were transduced with lentiviruses encoding HMGA2 or vector alone. The transduced cells were irradiated (3 Gy/min, room temperature) using a 137Cs source (Mark II, gamma irradiator). Cells were then harvested using Laemmli sample buffer at respective time points post-irradiation followed by Western analysis. The antibodies used in this study were as follow: anti-HMGA2 (Biocheck), anti-actin (Chemicon), anti-γ-H2AX (Upstate), and anti-H2AX (Santa Cruz).
Statistical analysis

Data were analyzed using the JMP Statistical Discovery Software version 8.0 (SAS Institute, Cary, NC). The missing cases were labeled with missing code to avoid counting in statistical analysis. Group comparisons for continuous data were done with t test for independent means or one-way ANOVA. For categorical data, chi-square analysis, Fisher’s exact test, or binomial test of proportions was used. Multivariate logistic regression models were used to adjust for covariate effects on the odds ratio (OR). Kaplan-Meier analysis and Cox hazard proportional model were applied for overall survival (OS) and progress-free survival (PFS) analysis. Multivariate analysis and stratification were applied to reduce confounders’ impact on estimation of odds ratio (OR) and hazard ratio (HR). Statistical significance was set at $p < 0.05$. 
Results

Expression of HMGA2 is associated with distant organ metastasis of CRCs.

Based upon the IHC staining of HMGA2, 32 of 89 CRCs on the validation set and 70 of 191 CRCs on the training set were defined as HMGA2 nuclear positive staining (including weak positive “+” and strong positive “++”). The TNM stage of CRCs was based on clinical diagnosis. According to the univariate analysis results, the HMGA2 staining was positively and significantly associated with distant organ metastasis of CRC ($p<0.05$), but not with age, sex, tumor location, tumor invasion and lymph node involvement, in both the training and validation sets (Table 1).

To validate this finding, non-conditional Logistic analysis was employed for uni- and multivariate analyses. The HMGA2 positive and negative were stratified as unfavorable and favorable subsets, respectively. Tumor invasion, lymph node involvement, and distant organ metastasis were considered as the endpoint in Logistic analysis. The expression of HMGA2 significantly impacted the risk of distant metastasis but not tumor invasion and lymph node involvement according to either uni- or multivariate analysis. After adjusting for age and sex, the odds ratio (OR) of HMGA2-positive CRC for metastasis were 3.53 (95% CI 1.37-9.70) and 6.38 (95% CI 1.47-43.95) for the training set and validation set, respectively (Supplement Table 1). Therefore, our analyses revealed that HMGA2 significantly impacts distant metastasis of CRC, suggesting that HMGA2 may be used to prognosticate CRCs.

Positive HMGA2 correlates with poor prognosis in CRCs.

To examine the hypothesis that HMGA2 may impact the survival of CRC patients, the COX hazard proportional model and Kaplan-Meier analysis were employed to analyze survivability. In the training set, the longest follow-up time is 213 months; 60 out of 89 patients with CRC died from CRC-related disorders and 67 cases had recurrence during
the observation period. In the validation set, the longest follow-up time is 99.3 months; a total of 52 cases died from CRC and 65 CRC cases relapsed.

The Kaplan-Meier analysis result for OS is displayed in Fig. 2A and 2B (upper panels). The Log-rank test indicated that the positive HMGA2 is significantly related to OS in both the training and validation sets ($p < 0.05$). In addition, the HMGA2 expression in primary CRC inversely correlated with the median survival time in the training and validation sets. The “strong positive” HMGA2 (++) exhibited a significantly reduced survivability ($p < 0.05$). The recurrence analysis (Fig 2C and 2D) confirmed that the strong positive HMGA2 status significantly increased the risk for recurrence in both training and validation sets.

To avoid confounder effects, the multivariate COX analysis was conducted on both training and validation sets (Fig 2E & 2F, and Supplement Table 2). In the training set, factors including TNM stage, tumor location, gender and age were applied to adjust the hazard ratio (HR). To reduce the variance yield from insufficient sample size of the analysis, HMGA2 (+) and HMGA2 (++) were merged together as HMGA2-positive in COX proportional analysis. As illustrated in Fig. 2E & 2F, HMGA2 and TNM stages were significantly associated with poor OS of CRC for both training and validation sets with HRs for HMGA2 of 2.38 (95% CI 1.30-4.34) and 2.14 (95% CI 1.21-3.79), respectively. The HR of TNM stage also exhibited a comparable value in both sets, supporting the accuracy of our study. The tumor location significantly affected the prognosis in the training set, but it was not a significant contributing factor for the validation set (Supplement Table 2). The unexpected variance may be caused by small sample size of rectal cancer (only 7 cases) in the training set (Table 1). The gender and age did not correlate significantly with CRC survival in both sets. Overall, our analyses revealed that HMGA2 was significantly associated with the poor survival of CRC patients.

**HMGA2 impacts the survival of CRC patients with Stage III-IV.**
To confirm our finding, we analyzed the OS and PFS in different stages. As shown in Table 2, HMGA2 significantly impacted the survival of patients with CRC at stage III-IV, but not stage I-II in the training set. Although the tendency in the stage I-II could be detected, it was not statistically significant ($p > 0.05$). By contrast, the HMGA2 expression was associated with worse OS and PFS for stage III-IV CRC with a statistical significance ($p < 0.05$). The HRs of positive HMGA2 for OS and PFS were 2.52 (95% CI 1.37-4.58) and 2.95 (95% CI 1.57-5.50), respectively. This finding was confirmed with the validation set. Stratification analysis on validation set revealed that the adjusted HRs of HMGA2 for OS and PFS were 1.91 (95% CI 0.98-3.66) and 2.05 (95% CI 1.05-3.95) for stage III-IV CRC, respectively (Table 2). Moreover, the Kaplan-Meier analysis revealed that the expression level of HMGA2 negatively relates to the median of OS time in a dose-dependent manner for the training set (Fig. 3A). Similar results also could be seen in the validation set (Fig. 3B).

The pathoclinical features of cancer for rectum and colon are different, suggesting that HMGA2 may impact the survival of colon and rectal cancer patients differently. In order to address this possibility, all CRCs were stratified by tumor location for further analysis in the second panel of Table 2. In the training set, the positive HMGA2 was associated with worse survival of colon cancer, the HRs of OS and PSF were 1.44 (95% CI 0.81-2.50) and 1.33 (95% CI 0.76-2.24), respectively. Nevertheless, the Kaplan-Meier analysis revealed that the expression level of HMGA2 was inversely correlated with median OS of colon cancer in training set (log-rank $p = 0.026$) (Fig. 3C). However, it was not feasible to assess rectal cancer due to insufficient number of cases (only seven). In the validation set, positive HMGA2 was significantly associated with worse OS and PFS of rectal cancer patients ($p < 0.05$) than colon cancer. The HRs of HMGA2 for OS and PFS were 2.79 (95% CI 1.33-5.96) and 2.85 (95% CI 1.36-6.09), respectively. An additional Kaplan-Meier analysis revealed that the expression level of HMGA2 inversely correlated with the median OS of rectal cancer in the validation set (Fig. 3D). Based on the above
finding, we conclude that the HMGA2 overexpression represents an important molecular change specific for stage III-IV CRC, significantly impacting the survival of patients with cancers located in colon or rectum.

**Expression of HMGA2 delays the clearance of $\gamma$-H2AX post $\gamma$-irradiation in HCT-116 and SW480 cells, suggesting a mechanism for the HMGA2-augmented sensitivity to radiotherapy in CRC patients.**

Previous studies have shown that HMGA2 overexpression led to suppression of non-homologous end joining (NHEJ) pathway in repairing double strand breaks (DSBs) of DNA (19, 20). Based on our previous findings, the persistence of $\gamma$-H2AX could serve as an indirect indicator for DNA damage-response and the failure to repair DSBs. To investigate whether HMGA2 expression in colon cancer cells can lead to DNA damage accumulation, the clearance rates of $\gamma$-H2AX in HCT116 and SW480 cells with distinct HMGA2 contexts were assessed at various time points post $\gamma$-ray exposure. HCT116 and SW480 cells were transduced with vector- or HMGA2-harboring lentiviruses to generate vector control or HMGA2-expressing cells (Fig. 4A). Overexpression of HMGA2 rendered enhanced and prolonged phosphorylation of H2AX post irradiation compared to vector control in both HCT116 (Fig. 4B, right panel) and SW480 (left panel) cells, suggesting that expression of HMGA2 rendered cells an inability to repair DNA timely.

The above finding implied that HMGA2 levels might be affecting the outcome of radiotherapy in CRC patients. To address this hypothesis, a multivariate analysis was conducted and stratified by HMGA2 negative and positive (Fig. 4C & 4D, Supplement Table 3). Those potential confounders, such as chemotherapy, TNM stage, tumor location, pathological grade, age and gender that potentially relates to outcome of therapy were incorporated into multivariate analysis model. The results indicated that the radiotherapy did not affect the survival in HMGA2 negative CRCs (HR=1.14, 95% CI 0.38-2.83), but it significantly reduced the relative risk from death in HMGA2 positive CRCs (HR=0.18,
95% CI 0.04-0.63). This finding suggests that HMGA2 might serve as a biomarker for predicting the response to radiotherapy in CRC patients.
Discussion

In this report, we examined the expression level of HMGA2 in CRCs to assess its potential role as a prognostic or predictive marker. Our findings revealed that the positive expression of HMGA2 was significantly associated with the distant metastasis in both training and validation sets ($p < 0.05$). The adjusted odd ratios (ORs) of HMGA2 for the risk of distant metastasis were 3.53 (95% CI 1.37-9.70) and 6.38 (95% CI 1.47-43.95) on training and validation sets, respectively. Previous studies, which showed that the $HMGA2$ mRNA expression was significantly escalated in invasive breast cancer cell lines (21, 22) and metastatic CRCs (16), further support our conclusion based on clinical specimens. We also found that the positive expression of HMGA2 was an independent indicator of poor prognosis. The multivariate COX proportional hazard analysis revealed that the HR of HMGA2 for OS was 2.38 (95% CI 1.30-4.34) and 2.14 (95% CI 1.21-3.79) on training and validation sets, respectively. Further analysis revealed that the HMGA2 overexpression impacts the survival of CRC patients with stage III-IV more significantly than those of stage I-II. The HRs of HMGA2 for OS of stage III-IV were 2.95 (95% CI 1.57-5.50) and 2.05 (95% CI 1.05-3.95) for the training set and the validation set, respectively. Our results are concordant with the previous studies, which indicated that a poor prognosis is associated with high level expression of HMGA2 in lung cancer (12), gastric cancer (23), oral cancer (13), ovarian cancer (14) and metastatic breast cancer (15). Therefore, our findings suggest that the positive expression of HMGA2 is associated with the distant metastasis of CRCs and a reduced patient survival, especially for those with advanced-stage CRC. To our knowledge, this is the first demonstration of HMGA2 overexpression as a biomarker for advanced CRC with poor prognosis.

Although the results from previous studies have suggested that HMGA2 functions as an oncogene that is overexpressed in a variety of human cancers, it is not yet well understood how HMGA2 promotes cancer invasion and metastasis in these cancers. As
for HMGA2, translocations often affect the third intron leading to the expression of fusion transcripts containing three AT-hooks and ectopic sequences of different origin (24). Also, a mutation in the breast cancer susceptibility gene BRCA1 has been shown to de-repress HMGA2 in breast cancer cells (25). Transgenic mice expressing a truncated form of the HMGA2 protein comprising the first three binding domains exhibit a giant phenotype with high incidence of lipomas. These results indicate that primarily truncation and/or aberrant expression of HMGA2 rather than its fusion to other genes is responsible for neoplastic transformation (26, 27). The overexpression of HMGA2 occurs in various benign and malignant tumors. Disruption of the HMGA2 gene by rearrangements affecting the chromosome region 12q13-15 and the attendant overexpression of the HMGA2 protein results in several benign mesenchymal tumors such as lipoma and uterine leiomyoma (28-30). Several independent studies have identified the HMGA2 transcript as a target for the let-7 family of microRNAs (31, 32). Notably, a study investigating breast cancer-initiating cells suggested that silencing of HMGA2 enhances differentiation but not self-renewal (33). In addition, HMGA2 interacts with pRb and enhances the E2F1 activity by displacing histone deacetylase 1 (HDAC1) (34). Based on the results obtained using uterine leiomyomas, a balance between HMGA2 and the p19ARF-TP53-CDKN1A has been proposed to determine the tumor size (35). All these data support our notion that the degree of HMGA2 overexpression is one of the key elements determining the aggressiveness of CRC, affecting the survival of CRC patients.

It has been suggested that the aberrant expression of many genes correlates with the prognosis of CRC, but only a few of them have been validated and are used to predict the treatment response (36). For instance, high-frequency microsatellite instability (MSI-H) is associated with good survivability but poor therapeutic response for stage II CRC patients who received adjuvant chemotherapy (36, 37). Oncotype Dx kit has been developed to predict survival and treatment response for stage II and Stage III CRCs. Ki-67, a proliferative marker, is associated with a better response to 5-FU-based adjuvant
chemotherapy for stage III CRCs (38). The mutant KRAS/BRAF is implicated for resistance to EGFR inhibiting drugs for late stage CRC (36). Our study demonstrates that HMGA2 is relevant to metastasis and impacts the survival in advanced stage CRC. Moreover, all these findings had been validated with participants from different racial background. There are many outcome-related factors for CRCs with different racial background, such as risk factors, genomic variations, medical care and social economic classes (39). Therefore, the fact that identical results were yielded from two sets with different racial backgrounds further validates that HMGA2 is potentially an independent useful prognostic factor for predicting metastasis and survival for CRC patients.

The expression of HMGA2 was shown to increase the cytotoxic effect of DNA double strand breaks (DSB) induced by certain topoisomerase type II inhibitors, irradiation and the chemotherapeutic agent cisplatin but not the genotoxic agent hydroxyurea, MMS and low pH (19, 20, 40). Our data demonstrated that overexpression of HMGA2 delayed the clearance of $\gamma$-H2AX after 3Gy $\gamma$-ray irradiation on HCT-116 and SW480 cells (Fig. 4B). HMGA2 disrupts the DSBs repair by altering the binding affinity between Ku and DNA ends, and delaying the release of DNA-PKcs from DSBs (20). Therefore, the inhibition of NHEJ by HMGA2 facilitated the accumulation of chromosomal aberrations, which suggested HMGA2 is an oncogenic molecule promoting malignancy (Supplement Fig. 2). Under a therapeutic dose of $\gamma$-irradiation, disruption of DSBs repair by HMGA2 dramatically increased the cytotoxic effect of $\gamma$-rays and augmented efficacy of adjuvant radiotherapy in CRCs (Supplement Fig. 2). Studies using human samples also revealed that HMGA2 is associated with the CRC response to radiotherapy (Fig. 4C & 4D, Supplement Table 3). Radiotherapy significantly reduced the relative risk death in HMGA2 positive CRCs (HR=0.18, 95% CI 0.04-0.64), but not in HMGA2 negative CRCs. However, this finding could not be validated as radiotherapy was not applicable for CRCs from the validation set, representing a limitation with this study. Nonetheless,
all these biomarkers including HMGA2 are useful in selecting treatment strategies such as radiotherapy against CRC and deserve further validation.

In summary, our finding that HMGA2 overexpression is an informative biomarker, which is associated with poor prognosis of patients with advanced CRC has potential implications for CRC survival prediction, choice of treatment regimens and future development of treatment strategies.
Acknowledgements

This work was support in part by NIH research grants R01DE10742 and DE14183 (to D.K. Ann), CA72767 (to Y. Yen). We thank Mariko Lee for digital image capturing and Frank Un and Mansze Kong for editing.
Figure Legend

Figure 1. Immunohistochemistry (IHC) staining of HMGA2. In upper panels (A, B and C), the standard of IHC staining with negative (-), weak positive (+) and strong positive (++) are shown. In middle panels, (D) display the HMGA2 staining in colon tumor and adjacent normal colon tissues with low magnification (100X), (E) and (F) showed normal and cancer section with high magnification (200X) respectively. In lower panels, HMGA2 expression in metastatic colon cancer to liver is illustrated in (G) (100X), (H) and (I) displayed normal liver and metastatic colon cancer respectively. The magnification is indicated in each panel.

Figure 2. Positive HMGA2 impacts overall survival (OS) and recurrence of CRCs. The Kaplan-Meier analysis for overall survival is shown on upper panel, and the recurrence analysis is illustrated in middle panel, and the multivariate COX proportional hazard analysis is displayed on lower panel. The data from training set and validation set are shown on left and right columns respectively. A, Kaplan-Meier analysis for OS of CRCs from training set with different HMGA2 levels; B, Kaplan-Meier analysis for OS of CRCs from validation set with different HMGA2 levels; C, recurrence analysis for CRCs from training set; D, recurrence analysis for CRCs from validation set; E, multivariate COX analysis for OS of CRCs from training set; F, multivariate COX analysis for OS of CRC from validation set. The detail results of (E) and (F) were displayed on Supplement Table 2.

Figure 3. HMGA2 overexpression is associated with poor OS on subset of CRCs with stage III-IV and different tumor location. A, OS of CRCs with stage III-IV from training set with different HMGA2 levels; B, OS of CRCs with stage III-IV from validation set. C, OS of colon cancer from training set with different HMGA2 levels. D, OS of rectal cancer from validation set with different HMGA2 levels.

Figure 4. Expression of HMGA2 delayed clearance of γ-H2AX post γ-ray exposure and augmented the response to radiotherapy in CRC patients. A, Expression of
HMGA2 in vector- or HMGA2-transduced CHT-116 and SW480 colorectal cancer cells. 

B, Western blot analysis of γ-H2AX elimination in HCT-116/vector and HCT-116/HMGA2 cells (left panel), and SW480/vector and SW480/HMGA2 cells (right panel). Cells were either nonirradiated or irradiated with 3-Gy γ-rays; protein lysates were prepared at different time points after γ-ray exposure as indicated. The multivariate COX analysis for OS were stratified by HMGA2 negative (C) and HMGA2 positive (D) CRCs. Detail results of (C) and (D) were displayed in Supplement Table 3.

Supplement Figure 1. Pre-selecting HMGA2 antibodies for IHC staining by Western Blot. Three candidate antibodies, P1, P2 and P3, against HMGA2 for IHC were test by Western analysis. Cell lysate was loaded as described in Lane 1, 2, 3 and 4. The HMGA2_P1 was selected for IHC staining because there was less non-specific signal.

Supplement Figure 2. Mechanism of HMGA2 enhances effect of radiotherapy through disrupting NHEJ DSBs repair pathway. Therapeutic dose of γ-ray causes DNA double strain breaks. The heterodimer of Ku70/Ku80 could bind to break ends of DNA. This binding could be disrupted by HMGA2. Meanwhile, HMGA2 could induce hyper- phosphorylation of DNA-PKcs at Thr-2609 and Ser-2056. Then, HMGA2 delayed the release of DNA-PKcs from DSB sites, which interrupted the clearance of γ-H2AX complex after γ-irradiation. Under physiological condition, impairment of DSBs repair caused spontaneous chromosome aberrations, which promoted malignancy of CRC. Under therapeutic dose of γ-ray, it dramatically enhanced cell death and augmented efficacy of adjuvant radiotherapy.
References

24. Kazmierczak B, Bullerdiek J, Pham KH, Bartnitzke S, Wiesner H. Intron 3 of HMGIC is the most frequent target of chromosomal aberrations in human tumors and has been conserved basically for at least 30 million years. Cancer genetics and cytogenetics 1998;103(2):175-7.
29. Geurts JM, Schoenmakers EF, Van de Ven WJ. Molecular characterization of a complex chromosomal rearrangement in a pleomorphic salivary gland adenoma involving the 3'-UTR of HMGIC. Cancer genetics and cytogenetics 1997;95(2):198-205.
| Table 1. Pathoclinical Characteristics and HMGA2 distribution of eligible CRCs from COH and ZJU |
|-----------------------------------------------|-----------------------------------------------|
| Training set (COH)                           | Validation Set (ZJU)                          |
| No. of Cases                                 | No. of Cases                                 |
| **Age†**                                     | **No. of Positive HMGA2 * (%)**               |
| <40                                          | 4 (0.0)                                      |
| 40-49                                        | 5 (40.0)                                     |
| 50-59                                        | 7 (35.0)                                     |
| 60-69                                        | 13 (43.3)                                    |
| 70-79                                        | 6 (27.3)                                     |
| >80                                          | 3 (42.9)                                     |
| **P value**                                  | **No. of Positive HMGA2 * (%)**               |
|                                              | 15 (46.7)                                    |
|                                              | 10 (34.5)                                    |
|                                              | 20 (42.6)                                    |
|                                              | 14 (28.6)                                    |
|                                              | 15 (38.5)                                    |
|                                              | 4 (33.3)                                     |
| **Gender**                                   | **No. of Positive HMGA2 * (%)**               |
| Male                                         | 18 (42.9)                                    |
| Female                                       | 14 (29.8)                                    |
| **P value**                                  | **No. of Positive HMGA2 * (%)**               |
|                                              | 45 (41.3)                                    |
|                                              | 25 (30.5)                                    |
| **Location†**                                | **No. of Positive HMGA2 * (%)**               |
| Rectum                                       | 5 (71.4)                                     |
| Colon*                                       | 26 (32.1)                                    |
| - Proximal‡                                   | 14 (28.6)                                    |
| - distal §                                   | 11 (36.7)                                    |
| **P value**                                  | **No. of Positive HMGA2 * (%)**               |
|                                              | 32 (36.8)                                    |
|                                              | 38 (36.5)                                    |
|                                              | 17 (30.9)                                    |
|                                              | 20 (41.7)                                    |
| **Tumor invasion†**                          | **No. of Positive HMGA2 * (%)**               |
| Within Propria                               | 7 (36.8)                                     |
| Out Propria                                 | 24 (34.8)                                    |
| **P value**                                  | **No. of Positive HMGA2 * (%)**               |
|                                              | 15 (32.6)                                    |
|                                              | 55 (37.9)                                    |
| **Lymph Node**                               | **No. of Positive HMGA2 * (%)**               |
| Negative                                     | 15 (39.5)                                    |
| Positive                                     | 17 (33.3)                                    |
| **P value**                                  | **No. of Positive HMGA2 * (%)**               |
|                                              | 38 (38.0)                                    |
|                                              | 32 (35.2)                                    |
| **Distant Metastasis**                       | **No. of Positive HMGA2 * (%)**               |
| No                                           | 13 (26.0)                                    |
| Yes                                          | 19 (48.7)                                    |
| **P value**                                  | **No. of Positive HMGA2 * (%)**               |
|                                              | 61 (34.0)                                    |
|                                              | 9 (81.8)                                     |
| **Note:** All information about TNM stage (tumor invasion, lymph node involvement, and distant metastasis) was based on clinical and pathological diagnosis at the time of first surgical operation. |
| **Positive HMGA2 includes weak positive and strong positive of nuclear staining score.** |
| † In training set, there is 1 unspecified age, 1 unspecified colon cancer, and 1 missing case. In validation set, there are 2 unspecified colon cancer and 1 missing case in tumor invasion in validation sets. |
| ‡ Proximal colon includes: hepatic flexure, transverse, cecum, appendix, ascending and splenic flexure. |
| § Distal colon includes: sigmoid and descending colon. |
| **Statistical significance, $p<0.05.$**
Table 2. Stratification and multivariate analysis for HMGA2 and survival of CRCs

<table>
<thead>
<tr>
<th></th>
<th>Training set (COH)</th>
<th>Validation Set (ZJU)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No. of Cases</td>
<td>HR of OS (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HR of OS (95% CI)</td>
</tr>
<tr>
<td><strong>TNM Stage</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I &amp; II</td>
<td>29</td>
<td>0.72 (0.15-2.61)</td>
</tr>
<tr>
<td>III &amp; IV</td>
<td>59</td>
<td>2.52 (1.37-4.58) †</td>
</tr>
<tr>
<td><strong>Tumor location</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>81</td>
<td>1.44 (0.81-2.50)</td>
</tr>
<tr>
<td>Rectum</td>
<td>7</td>
<td>N/A</td>
</tr>
<tr>
<td></td>
<td>104</td>
<td>1.15 (0.46-2.73)</td>
</tr>
</tbody>
</table>

Note: COH indicates samples were collected from City of Hope; and ZJU means samples were from Zhejiang University. Multivariate Cox proportional hazard analysis was conducted to evaluate hazard ratio of HMGA2 positive vs negative. HR is Hazard ratio. The HRs were adjusted by sex and age at diagnosis. OS is overall survival, PFS is progress free survival. The TNM stage was based on tumor invasion, lymph node involvement and distance organ metastasis. The stage IV CRCs were excluded in PFS analysis. HRs of RRM2 and RRM2B were based on high-expression versus low-expression. N/A means not available. † Statistics significant on COX analysis, \( p < 0.05 \).
Figure 1.

A. Negative colon cancer (-)
B. Weak Positive colon cancer (+)
C. Strong positive colon cancer (++)
D. Colon Normal vs cancer
E. Normal colon
F. Colon cancer
G. Liver metastatic Colon cancer
H. Normal Liver
I. Metastatic colon cancer
Figure 2.

A. Training Set (COH)

B. Validation Set (ZJU)

C. Training Set (COH)

D. Validation Set (ZJU)

E. Training Set (COH)

F. Validation Set (ZJU)
Figure 3.

Training Set (COH)

A

Stage III-IV CRCs (n=60)
Log-rank p=0.205

Median OS (95% CI)

<table>
<thead>
<tr>
<th></th>
<th>HMGA2 NU (-)</th>
<th>15 (8-38)</th>
<th>HMGA2 NU (+)</th>
<th>14 (11-23)</th>
<th>HMGA2 NU (++)</th>
<th>7 (2-17)</th>
</tr>
</thead>
</table>

No. of patients at risk

- HMGA2 (-) 22
- HMGA2 (+) 29
- HMGA2 (++) 9

Validation Set (ZJU)

B

Stage III-IV CRCs (n=92)
Log-rank p<0.001

Median OS (95% CI)

<table>
<thead>
<tr>
<th></th>
<th>3.3 months</th>
<th>37.2 months</th>
<th>86.1 months</th>
</tr>
</thead>
</table>

No. of patients at risk

- HMGA2 (-) 59
- HMGA2 (+) 28
- HMGA2 (++) 5

Training Set (COH)

C

Colon Cancer (n=81)
Log-rank p=0.026

Median OS (95% CI)

<table>
<thead>
<tr>
<th></th>
<th>HMGA2 NU (-)</th>
<th>27 (11-112)</th>
<th>HMGA2 NU (+)</th>
<th>21 (14-211)</th>
<th>HMGA2 NU (++)</th>
<th>7 (2-16)</th>
</tr>
</thead>
</table>

No. of patients at risk

- HMGA2 (-) 30
- HMGA2 (+) 39
- HMGA2 (++) 8

Validation Set (ZJU)

D

Rectal Cancer (n=92)
Log-rank p<0.001

Median OS (95% CI)

<table>
<thead>
<tr>
<th></th>
<th>2.6 months</th>
<th>56.1 months</th>
</tr>
</thead>
</table>

No. of patients at risk

- HMGA2 (-) 55
- HMGA2 (+) 29
- HMGA2 (++) 3
Figure 4.

A

<table>
<thead>
<tr>
<th></th>
<th>HCT-116</th>
<th>SW480</th>
</tr>
</thead>
<tbody>
<tr>
<td>HMGA2 vector</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>+</td>
<td>-</td>
<td>+</td>
</tr>
</tbody>
</table>

HMG A2

Actin

B

<table>
<thead>
<tr>
<th></th>
<th>HCT-116/vector</th>
<th>HCT-116/HMGA2</th>
<th>SW480/vector</th>
<th>SW480/HMGA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ-ray (3 Gy)</td>
<td>2h</td>
<td>4h</td>
<td>6h</td>
<td>2h</td>
</tr>
<tr>
<td>Post-irradiation</td>
<td>2h</td>
<td>4h</td>
<td>6h</td>
<td>2h</td>
</tr>
</tbody>
</table>

γ-H2AX

H2AX

Actin

C

HMGA2 Negative (n=54)

Radiotherapy
Chemotherapy
Gender
Age
Tumor Location
Pathological Grade
TNM Stage

HR (Range: 95% CI(Log 2))

D

HMGA2 Positive (n=31)

Radiotherapy
Chemotherapy
Gender
Age
Tumor Location
Pathological Grade
TNM Stage

HR (Range: 95% CI(Log 2))
Clinical Cancer Research

Overexpression of HMGA2 Promotes Metastasis and Impacts Survival of Colorectal Cancers

Xiaochen Wang, Xiyong Liu, Angela Y Li, et al.

Clin Cancer Res  Published OnlineFirst January 20, 2011.

Updated version  Access the most recent version of this article at: doi:10.1158/1078-0432.CCR-10-2542
Supplementary Material  Access the most recent supplemental material at: http://clincancerres.aacrjournals.org/content/suppl/2011/04/15/1078-0432.CCR-10-2542.DC1
Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.
Permissions  To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.