HHS Public Access

Author manuscript

Gastroenterology. Author manuscript; available in PMC 2020 June 01.

Published in final edited form as:

Gastroenterology. 2019 June ; 156(8): 2338–2341.e3. doi:10.1053/j.gastro.2019.02.027.

Gene Expression Signature in Surgical Tissues and Endoscopic Biopsies Identifies High-risk T1 Colorectal Cancers

Raju Kandimalla[1,*], **Tsuyoshi Ozawa**[2,*], **Feng Gao**[3,*], **Xin Wang**[3,#], **Ajay Goel**[1,#], and **T1 Colorectal Cancer Study Group**

[1]Center for Gastrointestinal Research; Center for Translational Genomics and Oncology, Baylor Scott & White Research Institute and Charles A Sammons Cancer Center, Baylor University Medical Center, Dallas, Texas, USA

[2]Department of Surgical Oncology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan

[3]Department of Biomedical Sciences, City University of Hong Kong, Hong Kong, China

Keywords

T1 colorectal cancer; submucosal colorectal cancer; lymph node metastasis; risk stratification

INTRODUCTION

Recent evidence suggests that endoscopic submucosal dissection (ESD) alone is sufficient for treating T1 colorectal cancer (CRC) patients who are at low risk for developing lymph-node metastasis (LNM); while the more extensive radical surgery (RS) is only needed for high-risk patients[1, 2]. Unfortunately the current risk-stratification criteria based on the post-endoscopic pathological examination, which includes, positive surgical margins, poor tumor-differentiation, presence of vascular or lymphatic invasion, depth of submucosal invasion (>1000 um) and high-grade tumor budding; tend to overestimate the degree of risk, and inadvertently categorize >70–80% of T1 CRCs into the high-risk category. However, in reality, based upon post-surgical pathology results, only 8–16% of all T1 CRC patients are truly LNM–positive, while a majority of them unnecessarily undergo RS procedures[1, 3]. This

Corresponding authors: Ajay Goel, PhD; Professor and Director, Center for Gastrointestinal Research; Center for Translational Genomics and Oncology, Baylor Scott & White Research Institute, Baylor University Medical Center, Dallas, Texas, 3410 Worth Street, Suite 610, Dallas, TX 75246, USA. Phone: 214-820-2603; FAX 214-818-9292; Ajay.Goel@BSWHealth.org; and Xin Wang, PhD; Department of Biomedical Sciences, City University of Hong Kong, 1B-102, 1/F, Block 1, 31 To Yuen Street, Kowloon Tong, Hong Kong. Phone: +852 3442 2367; Fax: +852 3442 0549; Xin.Wang@cityu.edu.hk.

*Author names in bold designate shared co-first authorship;

Author contributions: RK, TO and FG are involved in study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript. T1 study group helped in critical revision of the manuscript for important intellectual content and material support. AG and XW are involved in study concept and design, drafting of the manuscript, critical revision of the manuscript for important intellectual content, statistical analysis, obtained funding, material support and study supervision.

#Authors contributed equally

Conflicts of interest: None

is a significant clinical challenge, as these surgeries are expensive and associate with various complications including higher mortality rates[4, 5]. We previously reported a miRNA signature that allowed robust detection of high-risk T1 CRC patients[6]. While, recent evidence has highlighted the importance of gene expression in stratifying CRC patients into distinct consensus molecular subtypes (CMS)[7], no studies have yet undertaken a comprehensive effort to identify gene expression signatures for the detection of LNM in T1 CRC patients.

METHODS

In this retrospective study, two independent genome-wide mRNA-expression datasets were analyzed for mRNA-biomarker discovery (N=125, TCGA) and validation (N=56, GSE39582), followed by analytical training and validation in two independent T1-CRC cohorts (N=136 and N=67) using RT–PCR assays. The performance of our gene panel was evaluated in matched biopsy specimens (N=90) collected during colonoscopy. The presence of LNM was measure as the primary outcome. Detailed methodological description is provided in Supplementary Figure 1A and Supplementary Methods.

RESULTS

During the biomarker discovery in 16 LNM–positive and 109 LNM–negative T1 and T2 CRC patients, we identified 31 candidate genes (Supplementary Figure 1B). Subsequent bioinformatic and statistical analysis yielded a reduced panel of 16 genes, which robustly discriminated LNM positive vs. negative patients (Supplementary Figure 1C), using a trained Random Forest (RF) analysis in the T1 and T2 patients within the TCGA–dataset. While individual genes resulted in area under the curve (AUC) values ranging from 0.65 to 0.76, the combined 16–gene classifier achieved an AUC of 1.00 for detecting LNM patients in T1/T2 patients, as well as an impressive AUC of 1.00 for just the T1 CRC patients. Importantly, subsequent validation using the trained RF classifier in an independent validation cohort (GSE39582) resulted in an AUC of 0.85 (CI:0.74–0.96) and 0.94 (CI:0.79–1.00) in identifying LNM-positive T1/T2 and T1 CRC patients, respectively (Supplementary Figure 1D).

Next, we validated the performance of this signature using qRT–PCR assays for all 16 genes in our clinical training cohort (n=136; 19 LNM–positive and 117 LNM–negative), followed by a binary logistic regression analysis which resulted in a final model comprising of 8 genes. The risk scores for LNM in each patient were calculated as follows: Logit (P) = (3.700×*AMT*) + (−3.106×*MMP9*) + (−2.812×*FOXA1*) + (1.746×*LYZ*) + (1.718×*MMP1*) + (−0.473×*C2CD4A*) + (−0.901×*PIGR*) + (−1.836×*RCC1*) + (−6.468). The patients were dichotomized into low and high-risk groups based on the mRNA risk-scores obtained from Youden๖s index-derived cut-off thresholds. The resulting 8-gene risk-assessment model demonstrated an excellent performance for detecting LNM with an AUC value of 0.88 (CI: 0.79–0.97, p=0.0000002, power >0.99; Figure 1A), with a specificity of 0.86 and a sensitivity of 0.79. Subsequently, we applied the same statistical model and coefficients derived from the training cohort to an independent validation cohort of 67 T1 CRC patients (8 LNM-positive and 59 LNM-negative), which once again confirmed the robustness of our

risk-assessment model in identifying LNM-positive patients with an excellent AUC value of 0.84 (CI: 0.71–0.97, p=0.002, power = 0.96; Figure 1A). In addition, our mRNA-risk model provided extra net cost benefit compared to current strategies that are based upon considering either no patients or all patients for intervention across a range of thresholds, in the training and validation cohorts (Figure 1A and B).

To better appreciate the clinical significance of our 8-gene classifier, we analyzed its performance in identifying LNM in the context of other clinical variables, in our training cohort. In the univariate analysis, both the mRNA classifier and lymphovascular invasion were significant in detecting LNM with an odds ratio (OR) of 23.67 (95% CI: 6.97–80.38) and 3.5 (95% CI: 1.18–10.35), respectively (Figure 1C, upper panel). However, in multivariate analysis, only the mRNA classifier emerged as an independent feature for detecting LNM in T1 CRC patients (OR = 18.83 [95% CI: 4.34–81.54]; Figure 1C, lower panel).

When we utilized pathological criteria currently used in the clinic in our training cohort patients, it resulted in stratifying 84% patients (114 of the 136) into a high-risk category and the remaining 16% (22 of 136) into a low-risk group. However, examination of post-surgical tissues revealed that while such a risk-stratification approach was adequate for the patients in the low-risk group (because there were no LNM-positive cases), only 17% (19 of 114) of patients were actually high-risk; underscoring that 70% (95 of 136) of patients were erroneously categorized as high-risk and underwent unnecessary RS (Figure 1D, upper panel). In contrast, when we analyzed these same patients using our mRNA classifier, it stratified only 18% patients into the high-risk group (31 of the 136), while the remaining 82% (105 of the 136) patients were deemed as low-risk. Of the 31 patients who were classified as high-risk, 15 were associated with LNM (48%), indicating that only 11.7% (16 of 136) of the all T1 CRC patients were over-treated, which is significantly superior compared to currently used pathological features (70% vs. 11.7%; Figure 1D, lower panel). In addition to the availability of surgically resected specimens in our training cohort, we also had access to 90 matched, biopsy specimens (11 LNM-positive and 79 LNM-negative). Interestingly, we observed a significant correlation for 5 of the 8 genes in matched biopsy specimens (*AMT, FOXA1, C2CD4A, PIGR* and *LYZ*; Figure 1E, upper panel); with a corresponding AUC of 0.72 (95% CI: 0.52–0.91, $p = 0.02$) for the 8-gene signature for detecting LNM, even in pre-surgical biopsy specimens (Figure 1E, lower panel).

DISCUSSION

We for the first time identify and validate a novel, mRNA-based signature for the identification of LNM in submucosal T1 CRC patients. Although we are enthused by the performance of our biomarkers even in pre-surgical biopsies, we would like to acknowledge that the overall performance of our assay in these biopsy tissues was somewhat lower that resected specimens (AUC of 0.72 vs. 0.88). Furthermore, it is encouraging to notice that the robustness of our current mRNA assay was quite comparable to the miRNA biomarkers that we reported previously[6]. Currently we are planning prospective studies to evaluate the performance of both mRNA and miRNA panels – individually, as well as in combination - to determine whether a combined classifier may offer superior accuracy in identifying LNM in

T1 CRCs, especially in pre-surgical biopsies. If successful, pre-surgical use of this signature might lead to reduction in ESD–associated complications (perforation or bleeding), as well as in reducing the overall healthcare economic burden from such expensive surgical procedures[8]. Ours was a retrospective study; hence, future prospective studies must independently validate the robustness of these biomarkers prior to their translation into clinical practice.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We would like to acknowledge the valuable participation and contribution of the 'T1 Colorectal Cancer Study Group' collaborators for their critical revision of the manuscript, intellectual feedback and material support; without which, this study would not have been possible. The study group consisted of Hiroaki Nozawa, Keisuke Hata, Hiroshi Nagata, Satoshi Okada and Toshiaki Watanabe from the Department of Surgical Oncology, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan; Daisuke Izumi and Hideo Baba from the Department of Gastroenterological Surgery, Graduate School of Medical Science, Kumamoto University, Kumamoto, Japan; James Fleshman from the Department of Colorectal Surgery, Baylor University Medical Center, Dallas, TX, USA; Maria Pellise and Francesc Balaguer from the Gastroenterology Department, Hospital Clínic de Barcelona, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain; Luis Bujanda from the Department of Gastroenterology, Instituto Biodonostia, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Universidad del País Vasco (UPV/EHU), San Sebastián, Spain; Miriam Cuatrecasas from the Pathology Department, Hospital Clínic de Barcelona, Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBERehd), Institut d'Investigacions Biomediques August Pi i Sunyer (IDIBAPS), Universitat de Barcelona, Barcelona, Spain.

Funding: This work was supported by CA72851, CA181572, CA184792, CA202797 and CA187956 grants from the National Cancer Institute, National Institutes of Health; RP140784 from the Cancer Prevention Research Institute of Texas; grants from the Sammons Cancer Center and Baylor Foundation, as well as funds from the Baylor Scott & White Research Institute, Dallas, TX, USA awarded to Ajay Goel. In addition, this work was also supported by a start-up grant for new faculty (7200455) and a VPRT grant (9610337) from the City University of Hong Kong, a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CityU 21101115), as well as a grant from The Science Technology and Innovation Committee of Shenzhen Municipality (JCYJ20170307091256048) awarded to Xin Wang.

References

1. Ikematsu H, Yoda Y, Matsuda T, et al. Long-term outcomes after resection for submucosal invasive colorectal cancers. Gastroenterology 2013;144:551–9; quiz e14. [PubMed: 23232297]
2. Yoda Y, Ikematsu H, Matsuda T, et al. A large-scale multicenter study of long-term outcomes after endoscopic resection for submucosal invasive colorectal cancer. Endoscopy 2013;45:718–24. [PubMed: 23918621]
3. Tamaru Y, Oka S, Tanaka S, et al. Long-term outcomes after treatment for T1 colorectal carcinoma: a multicenter retrospective cohort study of Hiroshima GI Endoscopy Research Group. J Gastroenterol 2017.
4. Bokey EL, Chapuis PH, Fung C, et al. Postoperative morbidity and mortality following resection of the colon and rectum for cancer. Dis Colon Rectum 1995;38:480–6; discussion 486–7. [PubMed: 7736878]
5. Paun BC, Cassie S, MacLean AR, et al. Postoperative complications following surgery for rectal cancer. Ann Surg 2010;251:807–18. [PubMed: 20395841]
6. Ozawa T, Kandimalla R, Gao F, et al. A MicroRNA Signature Associated With Metastasis of T1 Colorectal Cancers to Lymph Nodes. Gastroenterology 2018;154:844–848 e7. [PubMed: 29199088]
7. Guinney J, Dienstmann R, Wang X, et al. The consensus molecular subtypes of colorectal cancer. Nat Med 2015;21:1350–6. [PubMed: 26457759]

8. Wada Y, Kudo SE, Tanaka S, et al. Predictive factors for complications in endoscopic resection of large colorectal lesions: a multicenter prospective study. Surg Endosc 2015;29:1216–22. [PubMed: 25159643]

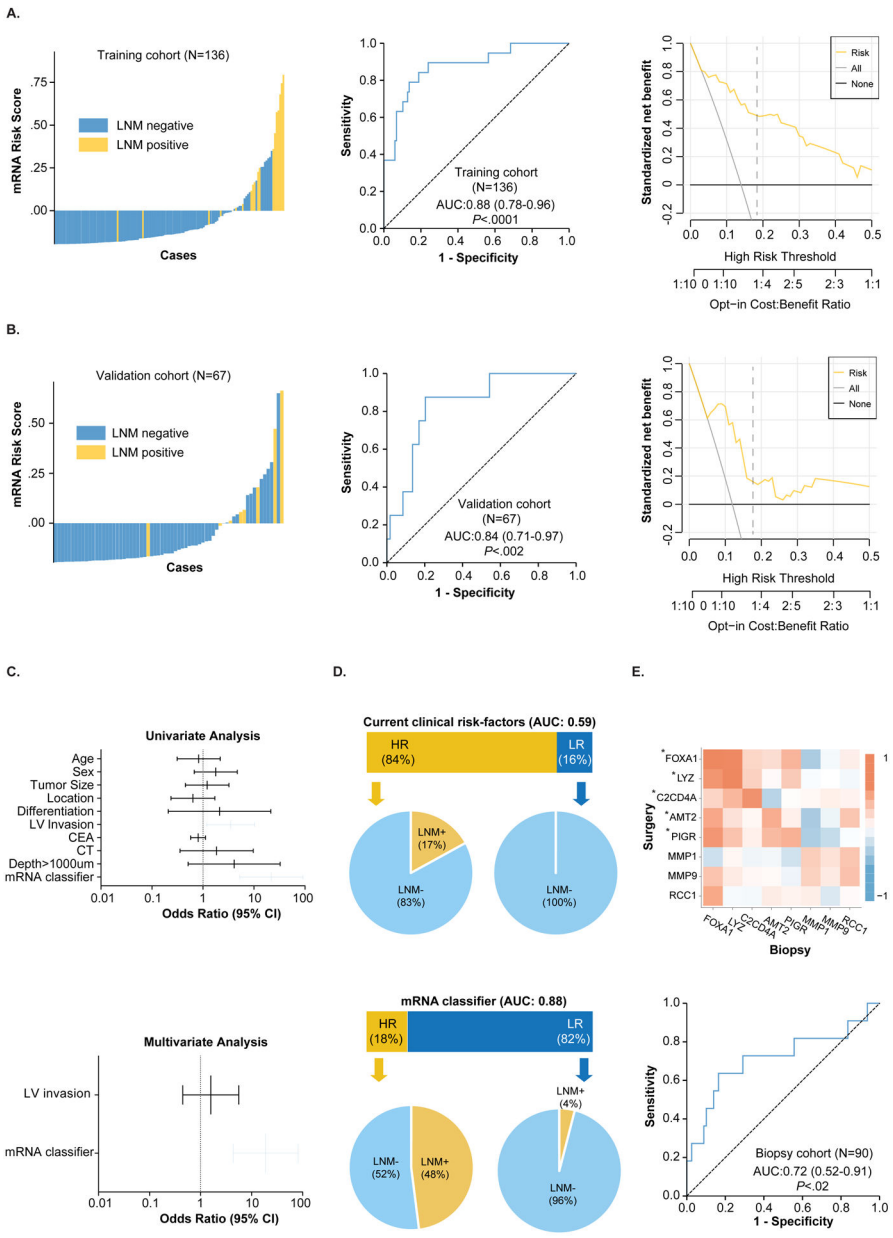

Figure 1:

Training and validation of the mRNA classifier in detecting lymph node metastasis in T1 CRC patients. Comparison of the mRNA classifier with the clinicopathological risk factors for lymph node metastasis detection and its performance in pre-surgical biopsy specimens. **A)** A waterfall plot depicting the 8–gene mRNA classifier risk-scores between LNM–positive and negative patients (left), ROC plot for the mRNA classifier (middle), and cost:benefit decision curves (right), in the patients from our training cohort; illustrating that our model provided extra cost-benefit compared to currently used strategies that are based on considering either no patients or all patients for intervention across a range of thresholds. **B)** A waterfall plot depicting the 8-gene mRNA classifier risk-scores between LNM-positive and negative patients in patients (left), ROC plot for the mRNA classifier (middle), and

cost:benefit decision curves (right), in the patients from our validation cohort. **C)** Univariate analysis illustrating odds ratios (OR) for the mRNA classifier, as well as various clinicopathological risk factors for the presence of lymph node metastasis (LNM) in T1 CRC patients (upper panel); Multivariate analysis illustrating odds ratios for the variables that demonstrated statistical significance in univariate analysis (mRNA classifier and lymphovascular invasion) for detecting LNM in T1 CRC patients (lower panel). **D)** Risk-stratification based on the currently used clinicopathological factors led to the overtreatment of 70% T1 CRC patients (upper panel), while mRNA classifier led to the overtreatment of only 11.7% T1 CRC patients (lower panel). **E)** Correlation matrix of the 8-gene classifier between pre-surgical biopsy specimens and surgically resected tissues (*indicates significant correlation; upper panel), and the ROC plot for the mRNA classifier in pre-surgical biopsy specimens (lower panel).