Transcriptomic expression profiling identifies ITGBL1, an epithelial to mesenchymal transition (EMT)-associated gene, is a promising recurrence prediction biomarker in colorectal cancer

Takatoshi Matsuyama 1,4, Toshiaki Ishikawa 2, Naoki Takahashi 3,5, Yasuhide Yamada 3, Masamichi Yasuno 4, Tatsuyuki Kawano 4, Hiroyuki Uetake 2 and Ajay Goel 1*

Abstract

The current histopathological risk-stratification criteria in colorectal cancer (CRC) patients following a curative surgery remain inadequate. In this study, we undertook a systematic, genomewide, biomarker discovery approach to identify and validate key EMT-associated genes that may facilitate recurrence prediction in CRC. Genomewide RNA expression profiling results from two datasets (GSE17538; N = 173 and GSE41258; N = 307) were used for biomarker discovery. These results were independently validated in two, large, clinical cohorts (testing cohort; N = 201 and validation cohort; N = 468). We performed Gene Set Enrichment Analysis (GSEA) for understanding the function of the candidate markers, and evaluated their correlation with the mesenchymal CMS4 subtype. We identified integrin subunit beta like 1 (ITGBL1) as a promising candidate biomarker, and its high expression associated with poor overall survival (OS) in stage I-IV patients and relapse-free survival (RFS) in stage I-III patients. Subgroup validation in multiple independent patient cohorts confirmed these findings, and demonstrated that high ITGBL1 expression correlated with shorter RFS in stage II patients. We developed a RFS prediction model which robustly predicted RFS (the area under the receiver operating curve (AUROC): 0.74; hazard ratio (HR): 2.72) in CRC patients. ITGBL1 is a promising EMT-associated biomarker for recurrence prediction in CRC patients, which may contribute to improved risk-stratification in CRC.

Keywords: ITGBL1, Prognostic marker, Epithelial mesenchymal transition, Colorectal cancer

Colorectal cancer (CRC) remains one of the primary causes of cancer-related deaths worldwide [1]. Although surgery remains the best treatment choice, a significant majority of stage II and III CRC patients develop disease recurrence following a curative resection; highlighting the inadequacy of currently used TNM classification for patient prognostication. Due to the high recurrence rates, patients with stage III disease routinely receive adjuvant chemotherapy [2]. Even though a clear benefit of adjuvant treatment in stage II CRC patients remains debatable, adjuvant chemotherapy is thought to be a reasonable treatment modality for the subgroup of high-risk stage II patients [3]. Nonetheless, given the relatively poor therapeutic response and high cancer recurrence rates, the current histopathological risk-stratification criteria remain inadequate. To address this concern, researchers have attempted to develop various biomarkers for patient stratification [4]; however, due a variety of biological and technical reasons, most of these biomarkers fail independent validations and are hence still not adopted in the clinical settings.

Epithelial-to-mesenchymal transition (EMT) is considered an essential regulatory process that mediates invasion and metastasis in cancer [5]. Recently, four consensus
Fig. 1 (See legend on next page.)
molecular subtypes (CMS) were identified in CRC patients following a comprehensive gene expression profiling [6]. Among these subgroups, the CMS4 subtype, characterized by the upregulation of EMT-associated genes, unequivocally emerged as a distinct subtype with worse overall survival (OS) and relapse-free survival (RFS). Although CMS classification holds promise in future, at this time, its clinical application for risk-stratification in CRC patients remains unclear. Nonetheless, given the strong association of CMS4 subgroups with an EMT phenotype, there is an emerging interest to develop EMT-associated biomarkers, which may serve as surrogates for the CMS4 subtype, and may allow for more improved patient stratification.

Recently, our group has shown that biomarkers highly expressed in liver metastasis are involved in distant metastasis and the EMT process [7, 8]. In this study, using a genomewide transcriptomic profiling of matched primary CRC and corresponding liver metastasis tissues, followed by their comparison in patients with and without disease recurrence, we identified a novel, EMT-related biomarker that robustly stratified low and high-risk CRC patients. Gene Set Enrichment Analysis (GSEA) revealed that high expression of integrin subunit beta like 1 (ITGBL1) strongly correlated with an EMT-phenotype, and significantly discriminated CRC patients with the CMS4 vs. the others subtypes. Subsequent clinical validation efforts revealed that high expression of ITGBL1 associated with poor OS and RFS in multiple, large, independent CRC patient cohorts, which allowed us to conclude that ITGBL1 is an attractive and promising prognostic biomarker in CRC.

Results and discussion

Overexpression of metastatic-recurrence-related genes in CRC

We first used a systematic biomarker discovery step to identify metastatic recurrence-specific genes for CRC from the publicly available GSE17538 and GSE41258 datasets. We identified two genes, ITGBL1 and SPP1 (osteopontin), which were differentially expressed between the primary CRC vs. metastatic tissues, recurrence vs. non-recurrence groups and normal vs. cancers (>2 fold change, and adjusted \( P < 0.05 \); Fig. 1a–c). Since, SPP1 has been extensively studied in CRC [9], while the clinical significance of ITGBL1 remains poorly but gaining a lot of attention in the field of cancer research [10], we selected ITGBL1 for further evaluation. The detailed methods are provided in the Additional file 1. The flow chart for the study design is illustrated in Additional file 2.

**ITGBL1 expression strongly correlates with an epithelial mesenchymal transition in CRC**

To gain further insight into the molecular function of ITGBL1 in CRC, we performed GSEA using genes that had a positive correlation with ITGBL1 expression. Based on the normalized enrichment score (NES), the EMT gene set emerged to be most strongly correlated with ITGBL1 expression (NES 2.099, \( P < 0.001 \), False discovery rate 0.016; Fig. 1d). Interestingly, several additional EMT-associated genes were also significantly correlated with the ITGBL1 expression (Fig. 1d); suggesting that ITGBL1 expression may serve as an important indicator of an EMT phenotype in CRC. Recent evidence indicates that an EMT phenotype is associated with the dissociation of the primary tumor cells from the primary site, followed by intravasation into blood and/or lymphatic vessels, establishing metastasis [5]. Through such an EMT process, CRCs with High ITGBL1 expression may lead to advanced disease, and present a higher risk for metastasis, which becomes the basis for developing recurrence prediction biomarkers.

**ITGBL1 serves as a surrogate for predicting the CMS4 subtype in CRC**

We next evaluated the expression of ITGBL1 in the context of CMS status in two public datasets (GSE39582 and GSE33113). We found that ITGBL1 expression was specifically higher in the CMS4 subtype vs. other subtypes in both patient cohorts. The AUROC for distinguishing CMS4 vs. CMS1–3 subtypes in CRC were 0.84 in GSE39582 and 0.91 in GSE33113 (Fig. 1e and f).

**ITGBL1 expression associates with poor RFS in CRC patients**

Furthermore, to investigate the clinical significance of ITGBL1 expression for risk-stratification of disease recurrence in stage II CRC patients, the group in which adjuvant chemotherapy decision-making is most desirable, we analyzed RFS in patients from the GSE39582 and GSE33113 datasets (Fig. 1g and i, respectively). In line with our earlier findings, we observed that high ITGBL1 expression group consistently demonstrated shorter RFS in stage II patients; yet again confirming the
The ITGBL1 protein expression is specifically higher in metastatic tissues from CRC patients

For a better understanding of the expression pattern of ITGBL1, we performed immunohistochemical (IHC) analysis. We found that ITGBL1 expression in normal colonic mucosa was quite weak (Additional file 3: Figure...
Fig. 2 (See legend on next page.)
S2D). However, ITGBL1 expression gradually increased from the luminal region to the invasive front in primary CRC, indicating that elevation of ITGBL1 expression might facilitate higher metastatic potential at the invasive front in primary CRC (Additional file 3: Figure S2A, B, and C). Likewise, liver metastasis revealed extremely high expression of ITGBL1 compared to adjacent hepatocytes (Additional file 3: Figure S2E).

**High ITGBL1 expression correlated with advanced stage, and presence of lymphovascular and distant metastasis in CRC patients**

We next investigated the level of ITGBL1 expression in relationship with various clinicopathological variables in two independent clinical testing and validation cohorts of 669 CRC patients (Additional file 4: Table S1). High ITGBL1 expression significantly correlated with increased tumor size, higher T stage, lymphovascular invasion, and the presence of distant metastasis in both cohorts (Table 1). Furthermore, when all CRC patients were segregated based upon the TNM stage, a gradual increase in ITGBL1 expression levels was observed from the low to high stages in both cohorts (Fig. 2a and d).

**Overexpression of ITGBL1 correlated with poor survival in CRC patients**

Next, we examined ITGBL1 expression with regard to its prognostic significance in the testing (n = 201), and validation cohorts (n = 468). In both cohorts, we noted that high ITGBL1 expression level correlated with shorter RFS in stage I-III patients (Fig. 2b and e), as well as a shorter OS in stage I-IV patients (Fig. 2c and f).

Cox’s univariate and multivariate analyses for RFS showed that high ITGBL1 expression was an independent prognostic factor for RFS in stage II CRC patients in the validation cohort (Additional file 5: Table S2). This observation was supported by univariate and multivariate analyses for OS, with high ITGBL1 expression showing an HR of 3.16 (Fig. 2j). Taken together, these findings indicate that high ITGBL1 expression has important clinical significance and could potentially serve as an important biomarker for predicting recurrence in CRC patients.

We finally constructed a RFS prediction model with various combinations of parameters including ITGBL1 expression using the Cox’s proportional hazard model in stage II CRC patients. AUROC at five years of this prediction model including Rectum, T4, MSS and ITGBL1 expression further improved from 0.61 to 0.74 (Fig. 2k); highlighting the recurrence predictive potential of ITGBL1 in CRC.

**Conclusion**

In conclusion, high ITGBL1 expression in primary tumors was associated with recurrence in CRC patients following curative surgery. Our study identified ITGBL1 as a novel, promising EMT-associated gene that could help in risk stratification and recurrence prediction in CRC patients.

**Additional files**

- Additional file 1: Detailed materials and methods. (DOCX 40 kb)
- Additional file 2: The study design. (DOCX 32 kb)
- Additional file 3: IHC staining for ITGBL1. (DOCX 2489 kb)
- Additional file 4: Table S1. The clinicopathological features of patients in this study. (DOCX 21 kb)
- Additional file 5: Table S2. Univariate and multivariate analysis of RFS in stage II patients of validation cohort. (DOCX 23 kb)

**Abbreviations**

AUROC: Area under the receiver operating curve; CMS: Consensus molecular subtypes; CRC: Colorectal cancer; EMT: Epithelial to mesenchymal transition; GSEA: Gene Set Enrichment Analysis; HR: Hazard ratio; ITGBL1: Integrin subunit beta like 1; MSS: Microsatellite stable; NES: Normalized enrichment score; HCC: Hepatocellular carcinoma; OS: Overall survival; RFS: Relapse-free survival

**Acknowledgements**

We thank Yoko Takagi and Junko Inoue for preparing the samples. We also thank Dr. Carson Harrod for proofreading the manuscript.

**Funding**

The present work was supported by the grants CA187956, 187956 from the National Cancer Institute, National Institute of Health, a grant (RP140784) from the Cancer Prevention Research Institute of Texas (CPRIT), pilot grants from the Baylor Sammons Cancer Center and Foundation, as well as funds from the Baylor Research Institute.
Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' contributions
TM was involved in study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript. TI, NT, YY, MY, TK and HU were involved in critical revision of the manuscript for important intellectual content and material support. AG was 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. All authors read and approved the final manuscript.

Ethics approval and consent to participate
All participants provided informed written consent, and the study protocol was approved by the Institutional Review Board of Tokyo Medical and Dental University and National Cancer Center Hospital.

Consent for publication
All subjects have written informed consent.

Competing interests
The authors declare that they have no competing interests.

Publisher’s Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details
1Center 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, 3410 Worth Street, Suite 610, Dallas, TX 75246, USA. 2Department of Specialized Surgery, Tokyo Medical and Dental University Graduate School of Medicine, Tokyo, Japan. 3Department of Gastroenterology, National Cancer Center Hospital, Tokyo, Japan. 4Department of Gastrointestinal Surgery, Tokyo Medical and Dental University Graduate School of Medicine, Tokyo, Japan. 5Department of Gastroenterology, Saitama Cancer Center Hospital, Saitama, Japan.

Received: 27 November 2018 Accepted: 18 January 2019
Published online: 04 February 2019

References
1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D, Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer. 2013;136: E359–86.
2. Andre T, Boni C, Moumoudji-Boudiaf L, Navarro M, Taberner J, Hickish T, Topham C, Zaninelli M, Clingan P, Bridgewater J, et al. Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. N Engl J Med. 2004;350:2343–51.
3. Benson AB 3rd, Schrag D, Somerfield MR, Cohen AM, Figueredo AT, Flynn PJ, Krozanowski MK, Maroun J, McAllister P, Van Cutsem E, et al. American Society of Clinical Oncology recommendations on adjuvant chemotherapy for stage II colon cancer. J Clin Oncol. 2004;22:3408–19.
4. Lopez NE, Weiss AC, Robles J, Fanta P, Ramamoorthy SL. A systematic review of clinically available gene expression profiling assays for stage II colorectal cancer: initial steps toward genetic staging. Ann J Surg. 2016;212:700–14.
5. Heerboth S, Housman G, Leary M, Longacre M, Byler S, Lapinska K, Willbanks A, Sarkar S. EMT and tumor metastasis. Clin Transl Med. 2015;4:6.
6. Guiney J, Dienstmann R, Wang X, de Reynies A, Schlicker A, Soneson C, Maria L, Roepman P, Nyamundanda G, Angelino P, et al. The consensus molecular subtypes of colorectal cancer. Nat Med. 2015;21:1350–6.
7. Hur K, Toiyama Y, Takahashi M, Balague F, Nagasaka T, Koike J, Hemmi H, Koi M, Boland CR, Goel A. MicroRNA-200c modulates epithelial-to-mesenchymal transition (EMT) in human colorectal cancer metastasis. Gut. 2013;62:1315–26.
8. Hur K, Toiyama Y, Ogawa Y, Ide S, Imakka H, Boland CR, Goel A: Circulating microRNA-203 predicts prognosis and metastasis in human colorectal cancer. Gut. 2017;66:654–65.
9. Zhao M, Liang F, Zhang B, Yan W, Zhang J. The impact of osteopontin on prognosis and clinicopathology of colorectal cancer patients: a systematic meta-analysis. Sci Rep. 2015;5:12713.
10. Li XQ, Xu X, Li DM, Kong PZ, Sun Y, Liu PF, Wang QS, Feng YM. ITGB8L1 is a Runx2 transcriptional target and promotes breast cancer bone metastasis by activating the TGFbeta signaling pathway. Cancer Res. 2015;75:3502–13.

Ready to submit your research? Choose BMC and benefit from:
- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.
Learn more biomedcentral.com/submissions