Landscape of alternative splicing characterizes the epithelial-mesenchymal transition subtype of gastric cancer

Yukyung Jun[1,2,3,§], Yun-Suhk Suh[1,4,5,6,§], SungHee Park[1], Jong-Il Kim[7], Sanghyuk Lee[2,8], Wan-Ping Lee[1,9,10,*], Olga Anczuków[1*], Han-Kwang Yang[4,5,11,*], Charles Lee[1,2,8,9,*]

[1]The Jackson Laboratory for Genomic Medicine, Farmington, CT, USA; [2]Ewha-JAX Cancer Immunotherapy Research Center, Ewha Womans University, Seoul, Korea; [3]Center for Supercomputing Applications, Division of National Supercomputing, Korea Institute of Science and Technology Information, Daejeon, Korea; [4]Department of Surgery, Seoul National University College of Medicine, Seoul, Korea; [5]Department of Surgery, Seoul National University Hospital, Seoul, Korea; [6]Department of Surgery, Seoul National University Bundang Hospital, Seongnam, Korea; [7]Department of Biomedical Sciences, Seoul National University College of Medicine, Seoul, Korea; [8]Department of Life Science, Ewha Womans University, Seoul, Korea; [9]The First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, China; [10]School of Cyber Science and Engineering, Xi'an Jiaotong University, Xi'an, China; [11]Cancer Research Institute, Seoul National University College of Medicine, Seoul, Korea

ABSTRACT

Alternative splicing (AS) has been known to play roles in the tumorigenesis, progression, and metastasis of gastric cancer. However, the landscape pattern of AS and its molecular characteristics in gastric cancer has not been comprehensively investigated yet. Using RNA-Seq data for paired gastric tumor and corresponding normal gastric mucosa of 83 gastric cancer patients, we observed 8 robust AS events that were differentially regulated between tumor and normal tissues and showed variable amount of differential AS across patients. Based on the differential exon usage of these 8 differential AS events, we then developed a patient classification scheme that divided patients into three groups. Molecular characterization using gene set enrichment analysis revealed roles of epithelial-mesenchymal transition (EMT) process in patient classification. Next, we investigated splicing regulators that had binding sites statistically enriched and were differentially expressed between epithelial and mesenchymal subgroups, which identified three RNA-binding proteins. Our patient stratification scheme was validated in independent cohorts of gastric cancer patients. The mesenchymal subtype was associated with the EMT subtype of the ACRG cohort and with the genomically stable (GS) subtype of the TCGA cohort, and the expression of three RBPs was significantly different with other subtypes in those data sets as well. Our study provides a novel patient classification scheme based on AS and highlights possible roles of AS on the EMT process in gastric cancer.

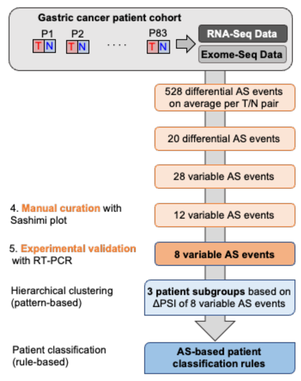
INTRODUCTION

Alternative splicing (AS) is a major mechanism for increasing transcript diversity in eukaryotes, impacting ~95% of multi-exon genes that express multiple spliced isoforms in humans. Several transcriptomic studies have reported AS of genes implicated in gastric cancer. Yet the role of AS in gastric cancer pathogenesis, and its potential clinical utility, have not been fully explored due to the lack of transcription-wide analysis on matched normal-tumor pair samples. A systematic analysis of AS events in matched normal-tumor gastric samples would allow identification of cancer-specific events, in addition to inter-tumor AS variation, and is currently lacking. Further, deeper insights into AS events and their regulation are greatly needed to better understand the fundamental molecular features of, and potential causative mechanisms leading to, gastric cancer.

PATIENT SUBTYPE CLASSIFICATION BASED ON ALTERNATIVE SPLICING EVENTS

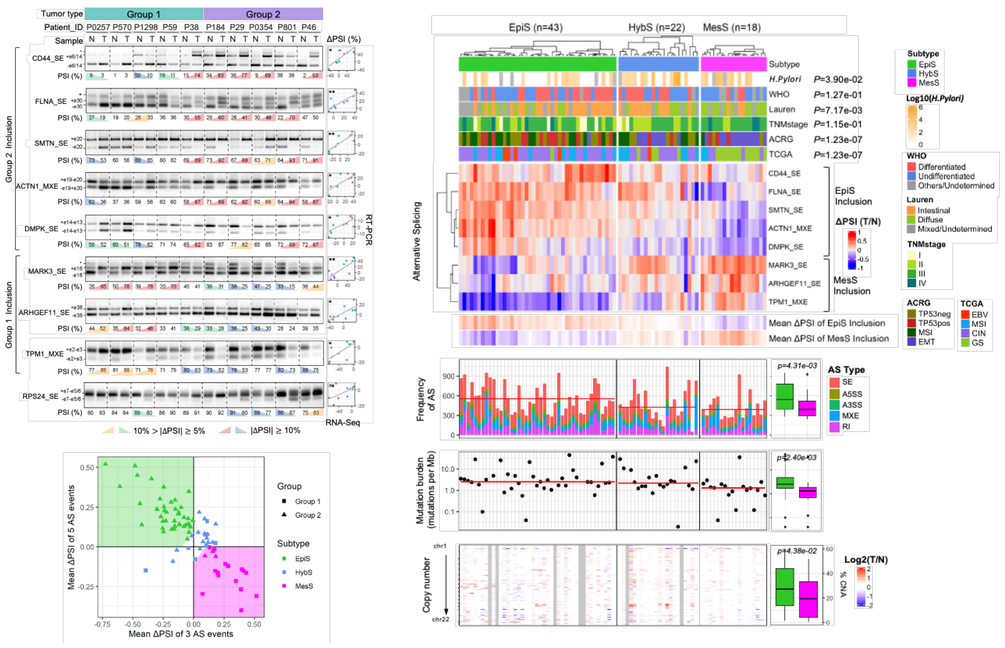
- **Summary of AS-based patient classification scheme**

We performed RNA-seq of paired tumor and adjacent normal mucosa tissues from 83 gastric cancer patients. We first looked for *differential* AS events between tumor and normal tissue and identified 20 AS events in 18 genes. We next sought to identify AS events that could be used for patient classification. Five such variable AS events were identified upon hierarchical clustering and 23 additional variable AS events were identified to facilitate the further dichotomization of patients. All events were manually inspected, and merged to produce a group of 12 non-redundant variable AS events in nine genes and eight AS events were validated experimentally.



- **Associations between AS-based classification and clinical and molecular subtypes**

The AS-based classification divided 83 patients into 43 EpiS, 22 HybS, and 18 MesS subtypes. The frequency of AS events, mutation rate, and copy number alterations were significantly higher in the EpiS subtype compared to the MesS subtype.

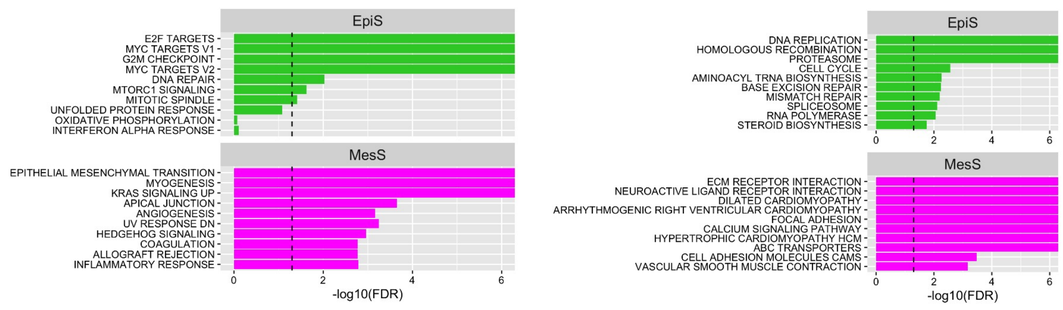


We compared our AS-based patient classification scheme with established clinical and molecular subtype classifications. MesS tumors exhibit higher *H. pylori* copy numbers compared to EpiS tumors (*P*=3.9e-02), in line with previous findings suggesting that *H. pylori* promotes EMT in gastric cancer[1]. We observed a significant association between our MesS subtype and the diffuse subtype, whereas patients classified as EpiS were enriched in the intestinal subtype (*P*=7.17e-03). Compared to the TCGA and ACRG molecular classifications, our MesS subtype, which rarely contains copy number alterations, was significantly associated with the TCGA GS subtype (*P*=1.23e-07) as well as with the ACRG EMT subtype (*P*=1.23e-07).

MOLECULAR CHARACTERIZATION OF AS EVENT-BASED GASTRIC CANCER SUBTYPES

- **Gene set enrichment analysis (GSEA)[2] between the EpiS and MesS subtypes**

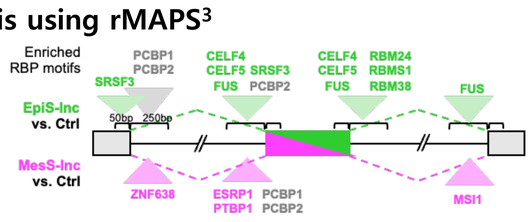
The EpiS subtype showed an enrichment of terms related to E2F targets, MYC targets, and cell cycle regulation, which are typical proliferation signals for epithelial cancer cells, as well as DNA damage response pathways. The MesS subtype showed various terms related to EMT, as well as ECM-receptor interaction, focal adhesion, and cell adhesion molecules.



SPLICING FACTORS RBM24 AND ESRP1 REGULATE AS EVENTS IN GASTRIC CANCER SUBTYPES

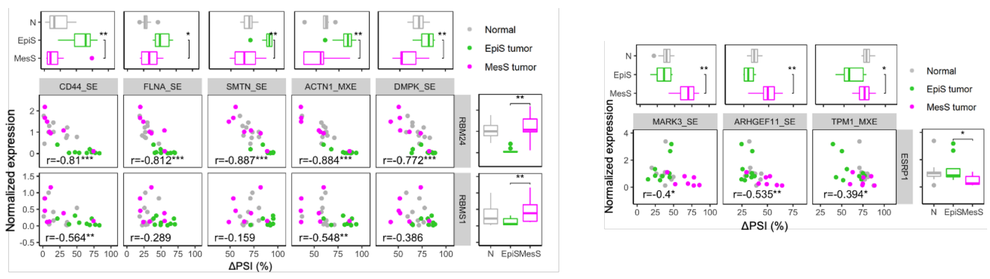
- **RBP binding motif enrichment analysis using rMAPS[3]**

For the five AS events included in EpiS tumors, we identified binding sites for nine RBPs specifically enriched in exonic or intronic regions near the spliced exons. Similarly, we found binding motifs for four RBPs enriched in the three AS events included in MesS tumors.



- **Correlation between the RBP expression and AS events**

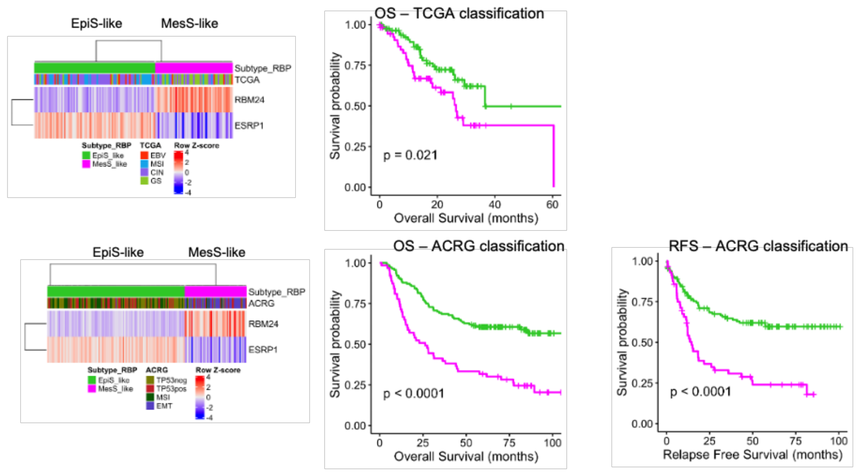
The correlation between the expression of regulatory RBPs and exon inclusion of all eight MesS and EpiS AS events was further validated experimentally. Gene expression of *RBM24*, and *ESRP1* were significantly correlated with the level of exon inclusion for most AS events.



VALIDATION OF THE RBP EVENT SIGNATURE

- **Survival analysis in the TCGA and ACRG cohorts**

We discovered that TCGA patients with the MesS-like subtype had a significantly worse prognosis than those with the EpiS-like subtype (log-rank *P*=2.10e-02). Similarly, ACRG patients with the MesS-like subtype, which was highly associated with the ACRG EMT subtype, showed worse survival than the EpiS-like subtype, both for overall and relapse-free survival (log-rank *P*<0.0001).



CONCLUSION

We systematically investigated the landscape and roles of AS in gastric cancer using RNA-Seq data from matched tumor and normal samples. Association of AS with the EMT program was firmly established and suggested AS-based patient stratification schemes, which highlighted the potential of AS analysis as a tool for precision medicine. To our knowledge, this study presents the most comprehensive analysis to date of AS in the context of patient classification, molecular mechanisms, and prognosis in gastric cancer.

REFERENCES

1> Baud J, et al. Helicobacter pylori initiates a mesenchymal transition through ZEB1 in gastric epithelial cells. PLoS One 2013.

2> Subramanian A et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005.

3> Park JW et al. rMAPS: RNA map analysis and plotting server for alternative exon regulation. Nucleic Acids Res 2016.