Original Article

Analysis of transcript-wide profile regulated by microsatellite instability of colorectal cancer

Ying Xu^, Xiaofeng Wang, Yimin Chu, Ji Li, Weiyi Wang, Xiangyu Hu^, Fengli Zhou, Haiqin Zhang, Lu Zhou, Rong Kuai, Yunfei Jin, Daming Yang, Haixia Peng

Digestive Endoscopy Center, Tongren Hospital, Shanghai Jiaotong University School of Medicine, Shanghai, China

Contributions: (I) Conception and design: Y Xu, X Wang, D Yang; (II) Administrative support: Y Xu, H Peng, D Yang; (III) Provision of study materials or patients: Y Xu, X Wang, Y Chu, J Li, H Zhang, L Zhou; (IV) Collection and assembly of data: Y Xu, X Wang, W Wang, X Hu, F Zhou; (V) Data analysis and interpretation: Y Xu, X Wang, X Hu, R Kuai, Y Jin; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Daming Yang; Haixia Peng. Digestive Endoscopy Center, Tongren Hospital, Shanghai Jiaotong University School of Medicine, No. 1111, Xianxia Road, Shanghai 200336, China. Email: YDM1100@shtrhospital.com; PHX1101@shtrhospital.com.

Background: Microsatellite instability-high (MSI-H) is a form of genomic instability present in 15% of colorectal cancer (CRC) cases. Several differential gene analyses have been conducted on CRC; however, none have specifically explored the differentially expressed genes in MSI-H CRC. Research on the different gene expressions between MSI-H CRC and microsatellite stable (MSS) CRC, and their different patterns of metastasis will provide invaluable insights for diagnosis, prognosis, and treatment.

Methods: In this study, the differential expression of 46,602 genes were analyzed across 613 different tissue samples from The Cancer Genome Atlas (TCGA)-colon adenocarcinoma (COAD) and TCGA-rectum adenocarcinoma (READ) as part of a gene association analysis. R package TCGAbiolinks (version 2.18.0) was used to download the data set, and DESeq2 (version 1.30.1) was used for the differential gene analysis. The resulting genes were then analyzed for shared pathways with R package clusterProfiler (version 3.0.4).

Results: A total of 237 significantly differentially expressed genes (P_{adj}<0.05) were found between MSI-H and MSS CRC. Differentially expressed genes include insulin like growth factor 2 (*IGF2*) and fibroblast growth factor 3 (*FGF3*), and the enriched pathways mostly involve hearing, digestive regulation, and neurogenesis.463 differentially expressed genes were found between metastatic and non-metastatic CRC. Notably differentially expressed genes in metastatic CRC include DEAD-box helicase 53 (*DDX53*) and adiponectin, C1Q and collagen domain containing (*ADIPOQ*), and enriched pathways include the immune system, cell adhesion, and cell signaling. For MSI-H CRC, a total of 34 genes were significantly differently expressed between metastatic and non-metastatic CRC. These include notum, palmitoleoyl-protein carboxylesterase (*NOTUM*), serpin family B member 2 (*SERPINB2*), and several keratin (*KRT*) genes, and the pathway analysis showed the major enrichment of the hormonal and secretion and regulation pathways. Of the differentially expressed genes in metastatic CRC, 25 were immunity related and include fatty acid binding protein 4 (*FABP4*), and the pathway analysis showed the enrichment of humoral immunity and lymphocyte regulation.

Conclusions: Of the biologically plausible differentially expressed genes, the most notable were *NOTUM*, *KRT6A*, *KRT14*, *SERPINB2*, and serum amyloid A1 (*SAA1*). *NOTUM*, *KRT6A*, and *KRT14* are active in the Wnt pathway. All five are also involved in various inflammation pathways.

Keywords: Colorectal cancer (CRC); microsatellite instability-high (MSI-H); transcriptome analysis; differential expressed genes; functional enrichment

^ ORCID: Ying Xu, 0000-0002-7356-1218; Xiangyu Hu, 0000-0003-2782-3384.

Submitted Oct 15, 2021. Accepted for publication Feb 14, 2022.
doi: 10.21037/atm-21-6126
View this article at: https://dx.doi.org/10.21037/atm-21-6126

Introduction

Microsatellites are a common type of DNA repeats with very low base pair numbers in their DNA motif (1–6 bp). During DNA synthesis, this can result in mistakes during DNA annealing, which in turn can result in a mismatch in frame and loops of single-strand DNA. These errors are normally repaired by a mechanism called mismatch repair (MMR), whereby the mismatched areas are digested, and the DNA polymerase is allowed to attempt replication again. In cases in which MMR is disabled, mutation rates throughout the genome increase dramatically, often resulting in cancer (1). This results in a feature called microsatellite instability (MSI), which is characterized by a high quantity of mutations in microsatellite locations. Thus, the presence of MSI implies the disabling of MMR mechanisms (2). This results in major mutational differences for MSI-high (MSI-H) CRC, most notably a higher chance for mutations in Kirsten rat sarcoma virus (*KRAS*), neuroblastoma RAS viral oncogene homolog (*NRAS*), B-Raf proto-oncogene, serine/threonine kinase (*BRAF*), and receptor tyrosine-protein kinase erbB-2 (*HER2*) when compared to MSS CRC (3). This is likely due to the unique mutagenic effects of MSI, which tend to cause large numbers of frame shift mutations. MSI is present in many types of cancers, but is most commonly associated with colorectal cancer (CRC); 15% of CRC patients have MSI (4). The most common cause of MMR dysfunction is the epigenetic or genetic disabling of MMR-related genes, most commonly MutL homolog 1 (*MLH1*), which is responsible for roughly 3/4 of cases of MSI. The other 1/4 of cases is caused by Lynch syndrome, a hereditary genetic mutation that disables 1 of the 4 key MMR genes [i.e., *MLH1*, *MLH2*, *MLH6*, or *PMS1* homolog 2, mismatch repair system component (*PMS2*)] (5). Since CRC is the 2nd most lethal type of cancer after lung cancer in most developed countries, finding the most effective methods for combating MSI CRC is vital to the future of cancer prevention and treatment (6). Compared to MSS CRC, MSI CRC have higher rates of somatic mutations, especially in receptor genes, resulting in greater immune system activation. Most studies have shown that the outcomes for MSI-H CRC are better than those of MSS CRC (7,8). Indeed, MSI-H patients have lower rates of metastasis, higher survival, and slower progression (7,8). However, another study has found that the opposite is true in cases of germ cell tumors (9). In terms of treatment, programmed death-ligand 1 (PD-L1) inhibitors like Atezolizumab were shown to have greater effectiveness in MSI-H CRC when compared to MSS CRC (10).

Metastasis occurs in roughly 40% of CRC cases in The Cancer Genome Atlas (TCGA)-colon adenocarcinoma (COAD) and TCGA-rectum adenocarcinoma (READ) data sets. Driving mutations behind metastasis in CRC include *APC*, *KRAS*, *BRAF*, phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*PIK3CA*), mothers against decapentaplegic homolog 4 (*SMAD4*), and *p53* (11). The differential gene expression associated with CRC metastasis and MSI-H CRC are well known; however, no specific analysis appears to have been conducted on which expression patterns are correlated with metastasis in MSI-H. An expression analysis based on TCGA data showed that MSI-H CRC forms a cluster separate from all other CRC types (12). A main feature of MSI-H is its high mutation rate, especially that of frame shift mutations, the expression of which are not selectively repressed, resulting in large numbers of neoantigens on tumor cells (13). This results in a higher immune response compared to MSS CRC, which has a lower mutational load and fewer neoantigens (12,14).

A previous study has examined the expression differences between MSI and MSS CRC (15). The focus and novelty of this study is the comparison of gene expression in MSI-H and MSS and how they relate to metastasis. While some genetic analyses have been conducted regarding the TCGA gene expression data for MSI, none have focused specifically on its relationship with metastasis.

To identify patterns in the genetic expression associated with MSI-H CRC in general and metastatic MSI-H CRC and related immunology pathways in particular, we compared gene expression data across 613 CRC samples from TCGA-COAD and TCGA-READ. The resulting genes were compared to a list of immunity-related genes to establish biological feasibility and analyzed for interactions. Search Tool for the Retrieval of Interacting Genes (STRING) and Metascape analyses were also conducted for the list of differentially expressed genes obtained from each

analysis. We hope to use this information to improve our understanding of MSI-H gene expression and better predict disease progression.

We present the following article in accordance with the STREGA reporting checklist (available at https://atm.amegroups.com/article/view/10.21037/atm-21-6126/rc).

Methods

Patients and data sets

R package TCGA Biolinks (version 2.22.2) was used to retrieve messenger RNA (mRNA) expression information for 613 CRC samples from TCGA-COAD and TCGA-READ after removing duplicates. This selection comprised the total of CRC samples in TCGA. Patients' clinical data were also obtained from TCGA on May 20, 2021. All the data retrieval processes used R version 4.0.5. Patients' demographic data are set out online available: https://cdn.amegroups.cn/static/public/atm-21-6126-1.xlsx. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Sample classification

Samples were classified as metastatic if either "tumor, node, metastasis (TNM) grading" or "lymph infiltration" categories in TCGA clinical data were positive for metastasis and otherwise as non-metastatic. The MSI classification of TCGA data points was obtained from another paper, where MSI status was determined via MSI-Mono-Dinucleotide assay (16).

Identification of immunity-related genes

A list of immunity-related keywords was used to find immunity-related genes (e.g., immunity, cytokine, virus, and immunoglobin) from the Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.kegg.jp/blastkoala/) and MSigDB (https://www.gsea-msigdb.org/gsea/msigdb/index.jsp) databases. After removing the overlapping genes, 731 immunity-related genes remained.

Differential gene analysis

R (version 4.0.5) package DESeq2 (version 1.30.1) was used to conduct the analysis of differential gene expression across the following five conditions: (I) metastasis *vs.* non-metastasis; (II) MSI-H *vs.* non-MSI-H; (III) MSI-H metastasis *vs.* MSI-H non-metastasis; (IV) metastasis *vs.* non-metastasis for immunological genes; and (V) immunological genes for MSI-H metastasis *vs.* MSI-H non-metastasis. $P_{adj}<0.05$ and $|\log_2[\text{fold change (FC)}]|>0$ were used as the cutoffs for significant differential expression. A sample was defined as metastatic if there was either lymphatic invasion, or its TNM staging had a metastasis value ≥1. A MSI status is based on TCGA clinical information category "molecular subtype"; every patient that had the MSI-H subtype was considered MSS. DESeq2 automatically corrects for any read biases when converting from raw reads to fragments per kilobase million (FPKM).

Interaction networks and biological function analysis

The STRING (https://cn.string-db.org/) database was used to build an interaction network for each set of differentially expressed genes, and the involved genes were downloaded and the duplicates were removed. R package clusterProfiler (version 3.0.4) was then used to perform the Gene Ontology (GO) and KEGG analyses on these gene lists, using a false discovery rate (FDR) <0.05 as the cutoff for significant enrichment.

Statistical analysis

Statistical analysis was conducted using R (version 4.0.5) and R packages TCGA Biolinks (version 2.22.2), DESeq2 (version 1.30.1), and clusterProfiler (version 3.0.4).

Results

CRC metastasis

Of the 613 samples, 252 were classified as metastatic, and 361 as non-metastatic. A total of 463 differentially expressed genes were found in the metastasis condition, of which 273 were upregulated and 191 were downregulated. These included the under expression of various rRNA pseudogenes, such as *RNA5SP141* and *RNA5SP145*. Up-expressed genes, such as DEAD-box helicase 53 (*DDX53*), a known oncogene, and adiponectin, C1Q and collagen domain containing (*ADIPOQ*), have a positive correlation with metastasis and cancer (see *Figure 1A,1B*) (17). *DDX53* promotes stem cell-like activity and taxol resistance in cancer cells (18,19). ADIPOQ upregulation has been previously shown to positively correlate with metastasis in

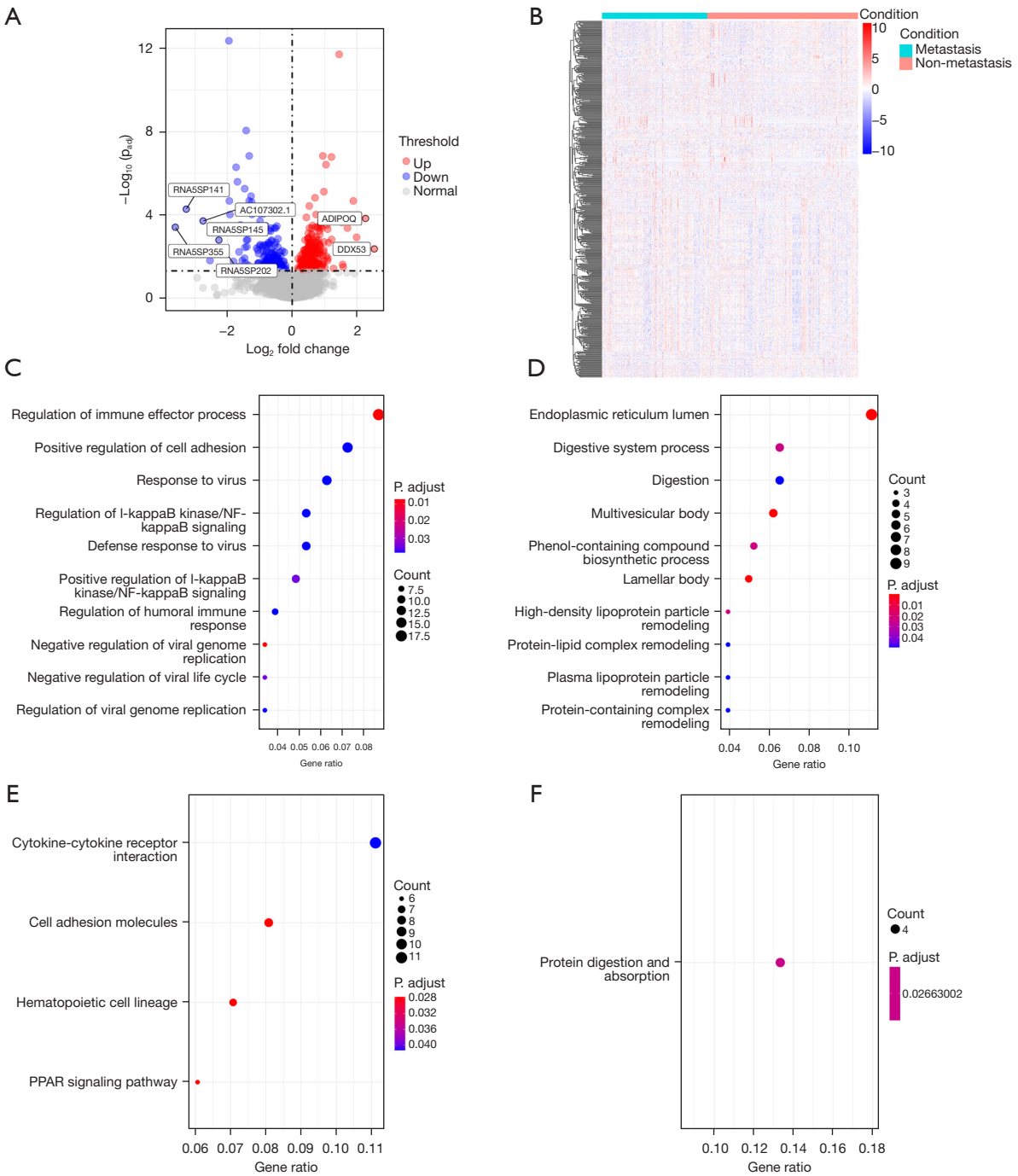

Figure 1 Differentially expressed genes in metastatic CRC. (A) Volcano plot of differential gene expression. Labeled genes have P_{adj}<0.05 and $|\log_2 FC| > 2$. (B) Heat map of differentially expressed genes between metastatic *vs.* non-metastatic CRC. Vertical axis indicates genes, and horizontal axis indicates patients. Color of cells indicates relative $\log_2$fold gene expression. (C) Upregulated genes in the GO analysis. (D) Downregulated genes in the GO analysis. (E) Upregulated genes in the KEGG analysis. (F) Downregulated genes in the KEGG analysis. CRC, colorectal cancer; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

CRC (20).

The STRING gene list expansion resulted in a total of 474 genes of interest. The GO analysis identified mostly immune-related, cell-adhesion, and cellular signaling genes, while the KEGG analysis identified genes for digestion and cellular processes (see *Figure 1C-1*F). The Metascape analysis showed that insulin-like growth factor regulation and defective colony stimulating factor 2 receptor subunit beta (*CSF2RB*) were enriched in the gene list (see Figure S1A). *CSF2RB* is a part of apoptosis pathways, and was previously recognized in the analysis of another gene expression dataset as being down-regulated in CRC (21).

MSI-H

Sixty-three samples were classified as MSI-H and 550 were classified as MSS. In the MSI condition, there were a total of 409 significantly differentially expressed genes, of which 107 were upregulated and 302 were downregulated. Only a single gene, regenerating family member 3 gamma (*REG3G*), had a $\log_2$FC >2. *REG3G* is known to play a role in pancreatic oncogenesis. Multiple known oncogenes were downregulated, such as insulin like growth factor 2 (*IGF2*) and fibroblast growth factor 3 (*FGF3*), as were various other genes, like keratin 5 (KRT5, a keratin coding gene) and claudin 18 (CLDN18) (a tight junction protein) (see *Figure 2A,2B*). IGF2 is a well-known oncogene, *FGF3* over-expression is associated with early cancer (22), and CLDN18 is over-expressed in gastric cancer (23).

The STRING gene list expansion resulted in a total of 416 genes of interest. The GO analysis mostly showed genes related to hearing and ear development (see *Figure 2C*), while the KEGG analysis showed neuro-interaction and digestion-related genes (see *Figure 2D*). The Metascape analysis showed that chemical synaptic transmission and sensory organ development were enriched in the gene list (see Figure S1B).

MSI-H metastasis

Of the 63 MSI-H samples, 21 were classified as metastatic while 42 were classified non-metastatic. Between these two groups, there were 34 differentially expressed genes, of which 14 were upregulated and 20 were downregulated. The downregulated genes included notum, palmitoleoyl-protein carboxylesterase (NOTUM), which is known to affect the Wnt signaling pathway, and 2 other keratin-related genes (i.e., *KRT14* and *KRT6A*), which are known cancer biomarkers (24), and serpin family B member 2 (*SERPINB2*), a gene related to autophagy and senescence in cancer (25). The upregulated genes included serum amyloid A1 (*SAA1*), a known tumor biomarker (see *Figure 3A,3B*).

The STRING gene list expansion included a total of 44 genes of interest. The GO analysis showed mostly hormone homeostasis genes, while the KEGG analysis showed neuro ligand receptor and insulin secretion genes (see *Figure 3C,3D*). The Metascape analysis showed that hormone level regulation and protein secretion regulation were enriched in the gene list (see Figure S1C).

Immunity-related genes in metastatic CRC MSI-H

Of the 463 differentially expressed genes in metastatic CRC, 25 genes were also on the list of immunity-related genes. Of these, 17 were upregulated and 8 were downregulated (see *Figure 4A,4B*). The upregulated genes included fatty acid binding protein 4 (*FABP4*). The downregulated genes included CAMP responsive element binding protein 3 like 3 (*CREB3L3*), which is a cyclic adenosine monophosphate (cAMP) regulating gene that is involved in inflammation and hepatocellular carcinoma (see *Figure 4C,4D*). FABP4 is a major predictor of metastasis in ovarian cancer (26).

The STRING gene list expansion produced a total of 35 genes of interest. The GO analysis showed mostly humoral and adaptive immune system related genes, while the KEGG analysis showed adenosine monophosphate-activated protein kinas signaling, cytokine interaction, and complement and coagulation cascades. The Metascape analysis showed that adaptive immune response and lymphocyte-mediated immunity were enriched in the gene list (see Figure S1D).

Discussion

The prognostic value of MSI status in CRC is well recognized, and MSI status is used to recommend treatment protocols. MSI-H CRC is generally associated with better patient outcomes and different patterns in metastasis (7,8). In this study, we focused on describing the genetic expression profiles of metastasizing and non-metastasizing MSI-H CRC, and their relationship to immunity. For metastasis in CRC, the most notable differentially expressed genes were *DDX53* and *ADIPOQ*. *DDX53* regulates stem-cell-like behavior in cancer cells and interacts with SRY-box transcription factor 2 (*SOX2*), a well-known oncogene (19). The GO and KEGG results mostly involved the immune

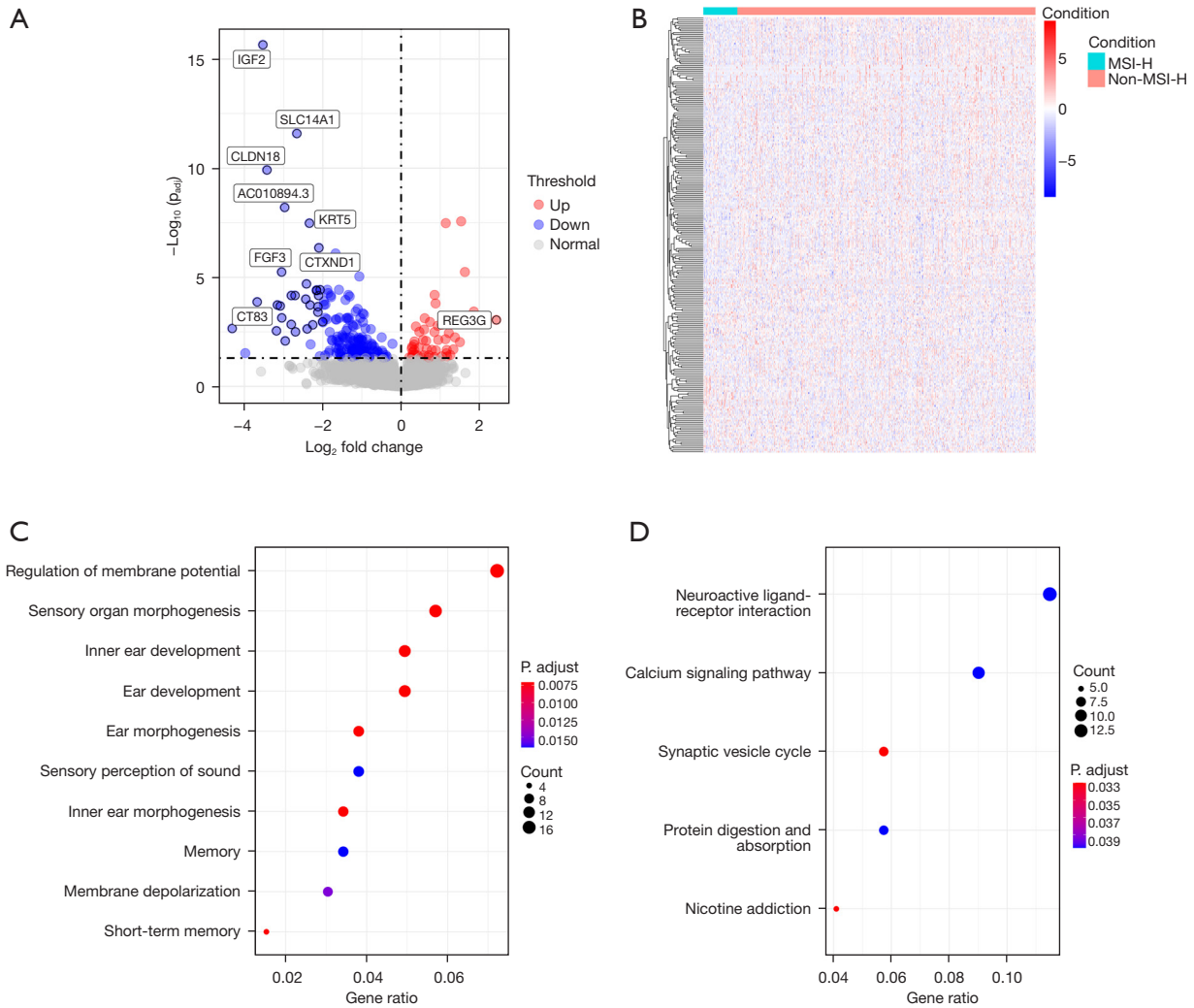

Figure 2 Differentially expressed genes in MSI-H CRC as compared to MSS CRC. (A) Volcano plot of differential gene expression. Labeled genes have P_{adj}<0.05 and |log_2FC| >2. (B) Heat map of differential gene expression in MSI-H *vs.* MSS CRC. Vertical axis indicates genes, and horizontal axis indicates patients. Color of cells indicates relative log_2fold gene expression. (C) Differentially regulated genes in the GO analysis. (D) Differentially regulated genes in the KEGG analysis. MSI-H, microsatellite instability-high; CRC, colorectal cancer; MSS, microsatellite stable; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

system, cell adhesion, and cell signaling, all factors that contribute to oncogenesis.

For MSI-H CRC, most of the differentially expressed genes, which included well-known oncogenes like *IGF2* and *FGF3*, were downregulated. The GO and KEGG analyses showed mostly hearing, ear development, digestive regulation, and neurogenesis. This is likely because the *SOX2*, transcription factor and known oncogene (27) are the hub of a cluster of genes, including forkhead box G1 (*FOXG1*) and EYA transcriptional coactivator and phosphatase 1 (*EYA1*), both of which are involved in embryonic hearing development and oncogenesis/cancer progression (28) (see Figure S2).

Our analysis of differentially expressed genes in metastatic MSI-H showed that high levels of *SERPINB2* expression is associated with increased survival and decreased changes in metastasis in pancreatic cancer (25). Our analysis of metastasis in MSI-H CRC showed that *SERPINB2* also had a lower expression in metastatic tumors. *SERPINB2* also plays an important role in immunity,

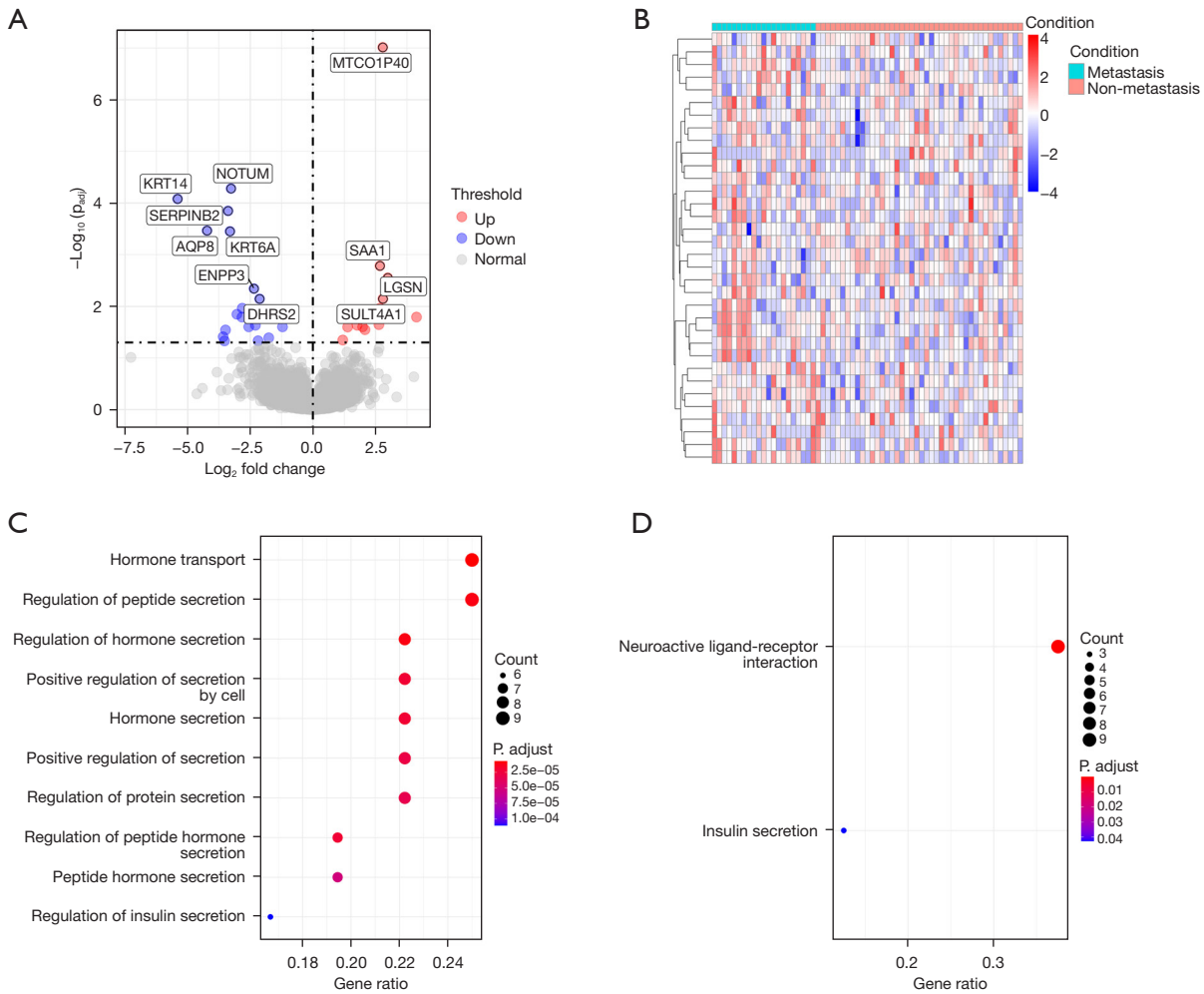

Figure 3 Differentially expressed genes in metastasis of MSI-H CRC as compared to non-metastatic MSI-H CRC. (A) Volcano plot of differential gene expression. Labeled genes have $P_{adj}<0.05$ and $|\log_2 FC| >2$. (B) Heat map of gene expression in metastatic vs. non-metastatic MSI-H CRC. Vertical axis indicates genes, and horizontal axis indicates patients. Color of cells indicates relative $\log_2$fold gene expression. (C) Differentially regulated genes in the GO analysis. (D) Differentially regulated genes in the KEGG analysis. MSI-H, microsatellite instability-high; CRC, colorectal cancer; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

specifically by helping to regulate the inflammation response (29).

Several genes affecting the Wnt pathway are also correlated to metastasis in MSI-H CRC, notably including NOTUM and several KRT genes. The activation of the Wnt pathway is known to contribute to oncogenesis and cancer progression, and *NOTUM* negatively regulates the pathway, which fits with its decreased expression in metastatic MSI-H CRC (30). The Wnt pathway has also been shown to play a positive role in gut inflammation (31).

The disruption of heightened *KRT* expression has been shown to inhibit prostate cancer cells *in vitro* by regulating the Wnt pathway (32). KRT is also known to be involved in the immune response, especially inflammation (33), and to be positively correlated with various conditions, such as inflammatory bowel disease and cancer (34). In addition, *KRT* plays a role in the immune system detection of tumor cells (35).

SAA1 expression is upregulated in metastatic MSI-H CRC, and its upregulation is correlated with cancer progression (36). In addition, it also plays a vital role in the regulation of inflammation, as it is significantly upregulated

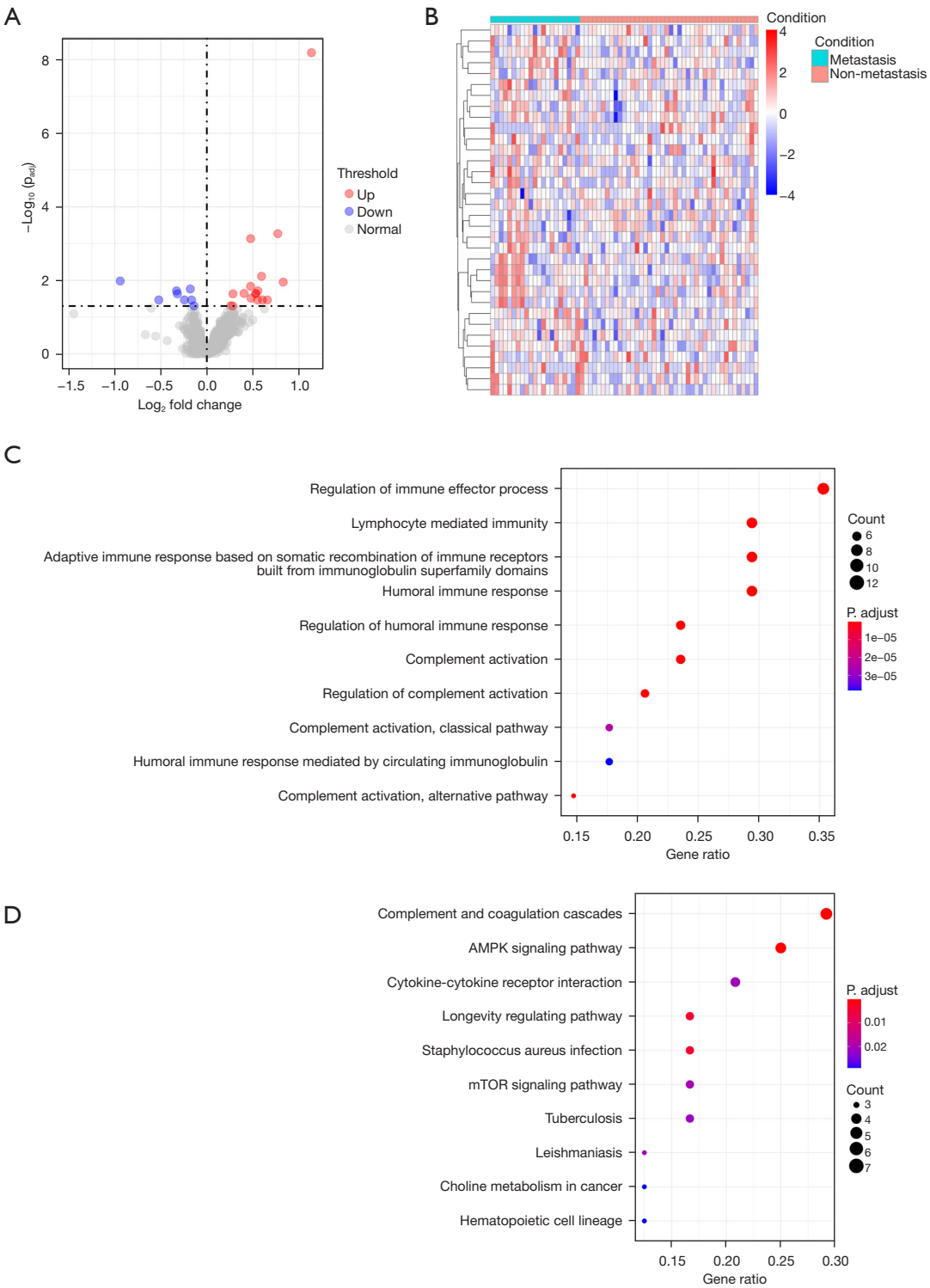

Figure 4 Differentially expressed immune system genes in metastatic CRC. (A) Volcano plot of differential gene expression. Labeled genes have P_{adj}<0.05 and |log_2FC| >2. (B) Heat map of immune system genes in metastatic *vs.* non-metastatic CRC. Vertical axis indicates genes, and horizontal axis indicates patients. Color of cells indicates relative log_2fold gene expression. (C) Differentially regulated genes in the GO analysis. (D) Differentially regulated genes in the KEGG analysis. CRC, colorectal cancer; FC, fold change; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

in irritable bowel disease, and correlated with CRC (37).

The GO analysis of differential gene expression in metastatic MSI-H showed a major enrichment of the hormonal and secretion and regulation pathways. The Metascape analysis of the expanded gene set had similar results, but also included G alpha signaling pathways, which help activate cAMP, and negatively regulate cell population proliferation.

Our analysis of immunity-related genes showed relatively few differentially expressed genes with both high absolute $\log_2$FC and low P values. The main exception was FABP4, whose expression levels is positively correlated with cancer metastasis in ovarian cancer (26). As expected, in our analysis, it was also positively correlated with metastasis.

The Metascape, GO, and KEGG analyses showed that lymphocytic and adaptive immunity were the most enriched pathways.

In addition, despite the known effectiveness of *PD-L1* inhibitors like Atezolizumab, our study did not find a statistically significant difference in expression of the coding gene, *CD247*, in any of our analyses. The reason for this is unknown.

This study was limited by the relatively low number of samples, which could have led to false positives in our list of differentially expressed genes. Future works are needed to confirm gene expression correlations found in our study.

Conclusions

In summary, the relative gene expressions of metastatic and MSI-H CRC in TCGA-COAD and TCGA-READ were analyzed, and several genes of interest were identified. Of these, several of the most statistically significant, such as *NOTUM*, *SERPINB2*, *SAA1*, *KRT14*, and *KRT6A*, were found to be associated with inflammation. Given the correlation between chronic gut inflammation and CRC, this may indicate that inflammation-induced CRC is likely to be MSI-H. Such findings can be used to better improve prognosis and treatment plans for CRC patients with chronic gut inflammation, such as those with Crohn's disease.

Acknowledgments

Funding: This work is supported by Science and Technology Commission of Changning District of Shanghai (No. CNKW2018Y02); the Research Fund of Key Laboratory for Translational Research and Innovative Therapeutics of Gastrointestinal Oncology (No. ZDSYS-2021-01); the Interdisciplinary Program of Shanghai Jiao Tong University (No. ZH2018QNB24); the Scientific Research Project of Medical Group of Shanghai Sixth People's Hospital (No. ly202003); the Research Fund of Tongren Hospital, Shanghai Jiaotong University School of Medicine [No. 2020TRYJ (LB) 01].

Footnote

Reporting Checklist: The authors have completed the STREGA reporting checklist. Available at https://atm.amegroups.com/article/view/10.21037/atm-21-6126/rc

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://atm.amegroups.com/article/view/10.21037/atm-21-6126/coif). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

References

1. Jiricny J. The multifaceted mismatch-repair system. Nat Rev Mol Cell Biol 2006;7:335-46.
2. de la Chapelle A. Microsatellite instability. N Engl J Med 2003;349:209-10.
3. Timmermann B, Kerick M, Roehr C, et al. Somatic mutation profiles of MSI and MSS colorectal cancer identified by whole exome next generation sequencing and bioinformatics analysis. PLoS One 2010;5:e15661.
4. Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology 2010;138:2073-2087.e3.

5. Lynch HT, Snyder CL, Shaw TG, et al. Milestones of Lynch syndrome: 1895-2015. Nat Rev Cancer 2015;15:181-94.
6. American Cancer Society. Cancer Facts & Figures. 2020. Available online: https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2020.htmlv
7. Gryfe R, Kim H, Hsieh ET, et al. Tumor microsatellite instability and clinical outcome in young patients with colorectal cancer. N Engl J Med 2000;342:69-77.
8. Popat S, Hubner R, Houlston RS. Systematic review of microsatellite instability and colorectal cancer prognosis. J Clin Oncol 2005;23:609-18.
9. Mayer F, Gillis AJ, Dinjens W, et al. Microsatellite instability of germ cell tumors is associated with resistance to systemic treatment. Cancer Res 2002;62:2758-60.
10. Li K, Luo H, Huang L, et al. Microsatellite instability: a review of what the oncologist should know. Cancer Cell Int 2020;20:16.
11. Huang D, Sun W, Zhou Y, et al. Mutations of key driver genes in colorectal cancer progression and metastasis. Cancer Metastasis Rev 2018;37:173-87.
12. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature 2014;513:202-9.
13. Roudko V, Bozkus CC, Orfanelli T, et al. Shared Immunogenic Poly-Epitope Frameshift Mutations in Microsatellite Unstable Tumors. Cell 2020;183:1634-49.e17.
14. Ratti M, Lampis A, Hahne JC, et al. Microsatellite instability in gastric cancer: molecular bases, clinical perspectives, and new treatment approaches. Cell Mol Life Sci 2018;75:4151-62.
15. Xiao J, Li W, Huang Y, et al. A next-generation sequencing-based strategy combining microsatellite instability and tumor mutation burden for comprehensive molecular diagnosis of advanced colorectal cancer. BMC Cancer 2021;21:282.
16. Liu Y, Sethi NS, Hinoue T, et al. Comparative Molecular Analysis of Gastrointestinal Adenocarcinomas. Cancer Cell 2018;33:721-35.e8.
17. Xu Y, He B, Pan Y, et al. The roles of ADIPOQ genetic variations in cancer risk: evidence from published studies. Mol Biol Rep 2013;40:1135-44.
18. Kim H, Kim Y, Jeoung D. DDX53 Promotes Cancer Stem Cell-Like Properties and Autophagy. Mol Cells 2017;40:54-65.
19. Kim Y, Yeon M, Jeoung D. DDX53 Regulates Cancer Stem Cell-Like Properties by Binding to SOX-2. Mol Cells 2017;40:322-30.
20. Tang HQ, Meng YL, Lu QL, et al. Decreased long noncoding RNA ADIPOQ promoted cell proliferation and metastasis via miR-219c-3p/TP53 pathway in colorectal carcinoma. Eur Rev Med Pharmacol Sci 2020;24:7645-54.
21. Slattery ML, Mullany LE, Sakoda LC, et al. Dysregulated genes and miRNAs in the apoptosis pathway in colorectal cancer patients. Apoptosis 2018;23:237-50.
22. Sun YW, Chen KM, Imamura Kawasawa Y, et al. Hypomethylated Fgf3 is a potential biomarker for early detection of oral cancer in mice treated with the tobacco carcinogen dibenzodef,pchrysene. PLoS One 2017;12:e0186873.
23. Nakayama I, Shinozaki E, Sakata S, et al. Enrichment of CLDN18-ARHGAP fusion gene in gastric cancers in young adults. Cancer Sci 2019;110:1352-63.
24. Karantza V. Keratins in health and cancer: more than mere epithelial cell markers. Oncogene 2011;30:127-38.
25. Harris NLE, Vennin C, Conway JRW, et al. SerpinB2 regulates stromal remodelling and local invasion in pancreatic cancer. Oncogene 2017;36:4288-98.
26. Gharpure KM, Pradeep S, Sans M, et al. FABP4 as a key determinant of metastatic potential of ovarian cancer. Nat Commun 2018;9:2923.
27. Novak D, Hüser L, Elton JJ, et al. SOX2 in development and cancer biology. Semin Cancer Biol 2020;67:74-82.
28. Cai S, Cheng X, Liu Y, et al. EYA1 promotes tumor angiogenesis by activating the PI3K pathway in colorectal cancer. Exp Cell Res 2018;367:37-46.
29. Schroder WA, Le TT, Major L, et al. A physiological function of inflammation-associated SerpinB2 is regulation of adaptive immunity. J Immunol 2010;184:2663-70.
30. Traister A, Shi W, Filmus J. Mammalian Notum induces the release of glypicans and other GPI-anchored proteins from the cell surface. Biochem J 2008;410:503-11.
31. Moparthi L, Koch S. Wnt signaling in intestinal inflammation. Differentiation 2019;108:24-32.
32. Du H, Wang X, Dong R, et al. miR-601 inhibits proliferation, migration and invasion of prostate cancer stem cells by targeting KRT5 to inactivate the Wnt signaling pathway. Int J Clin Exp Pathol 2019;12:4361-79.
33. Hobbs RP, Lessard JC, Coulombe PA. Keratin intermediate filament proteins - novel regulators of inflammation and immunity in skin. J Cell Sci 2012;125:5257-8.
34. Owens DW, Lane EB. Keratin mutations and intestinal pathology. J Pathol 2004;204:377-85.
35. Sequeira I, Watt FM. The role of keratins in modulating

carcinogenesis via communication with cells of the immune system. Cell Stress 2019;3:136-8.
36. Mattarollo SR, Smyth MJ. A novel axis of innate immunity in cancer. Nat Immunol 2010;11:981-2.
37. Tang MS, Bowcutt R, Leung JM, et al. Integrated Analysis of Biopsies from Inflammatory Bowel Disease Patients Identifies SAA1 as a Link Between Mucosal Microbes with TH17 and TH22 Cells. Inflamm Bowel Dis 2017;23:1544-54.

(English Language Editor: L. Huleatt)

Cite this article as: Xu Y, Wang X, Chu Y, Li J, Wang W, Hu X, Zhou F, Zhang H, Zhou L, Kuai R, Jin Y, Yang D, Peng H. Analysis of transcript-wide profile regulated by microsatellite instability of colorectal cancer. Ann Transl Med 2022;10(4):169. doi: 10.21037/atm-21-6126

