Genes & Diseases (2023) 10, 657–659

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.keaipublishing.com/en/journals/genes-diseases

RAPID COMMUNICATION

Genomic analysis of schistosomiasis-associated colorectal cancer reveals a unique mutational landscape and therapeutic implications

The schistosomiasis-associated colorectal cancer (SA-CRC) is thought to be caused by the chronic inflammation as a result of schistosomal ova deposition in the submucosal layer of the intestine (Fig. S1). It is reported that SA-CRC differs from sporadic colorectal cancer (S-CRC) in terms of epidemiological features, histological patterns, and clinical outcomes,[1] which is also supported by our institutional data (Table S1 and Fig. S2). Here we investigate the genomic alternations and probable molecular determinants of pathogenesis in Chinese SA-CRC.

Although there is no uniform definition, most reports define SA-CRC as CRC patients with the history of schistosomiasis and the pathological evidence of schistosomal ova deposition. Thirty patients diagnosed with SA-CRC in Changzheng Hospital from 2014 to 2020 were included in this study (Table S2). The median age at the time of diagnosis was 69.3 years old. Among these patients, SA-CRC mainly occurred at the rectum (83.3%) and were moderately differentiated (86.7%). The pathological stages were significantly correlated with the relapse-free survival (RFS) (Fig. S3). The archival formalin-fixed paraffin-embedded (FFPE) tumor tissues and paired adjacent non-tumor tissues were subject to whole exome sequencing (WES). Detailed methods including experiment processing and data analysis were available in "Supplementary Materials and Methods".

Bioinformatic analysis revealed a total of 2476 non-synonymous mutations in 1978 genes, with a range of 6–164 mutations per sample (Fig. 1A), suggesting a relatively stable number of mutations across the cohort. The median tumor mutation burden (TMB) for SA-CRC was 1.61/MB, which was lower than that of S-CRC (2.03/MB). In addition, the microsatellite instability (MSI) score for each sample was estimated, ranging from 0.31% to 12.72%. Thus, SA-CRC was assumed to be in low microsatellite instability (MSI-L) or microsatellite stable (MSS) status, compared to the MSI scores calculated by use of the Cancer Genome Atlas (TCGA) database[2] (Fig. S4). It was consistent with few hints mismatch repair (MMR)-related gene mutations, and was further confirmed by immunohistochemistry (IHC) for MMR proteins (Fig. S5). Moreover, the number of neoepitopes was also predicted, ranging from 2 to 722, which was similar to the neoepitopes load of MSS CRC but much less than the neoepitopes load of high microsatellite instability (MSI-H) CRC.[3] In a word, these results showed that SA-CRC could be categorized into MSI-L CRC with low TMB, which would unlikely to benefit from the immunotherapy.

We identified 71 recurrently mutated genes across the 30 samples (Fig. 1B), most of which had been previously linked to the S-CRC, including the common CRC-related driver genes (*APC*, *TP53* and *KRAS*). Other non-CRC driver genes (*IRS4*, *SETBP1*, *TNC* and *CREBBP*) were also present. However, 4 *de-novo* driver genes were identified, of which *KRT76* and *PLEKHO1* were novel. *PLEKHO1* was reported to act as a colonic tumor suppressor. The loss or deletion of *PLEKHO1* would promote the tumorigenesis. It was also reported that loss of *KRT76* increased carcinogen-induced tumors in tongue and stomach. In this study, 2 frameshift deletions and 1 missense mutation were identified in *PLEKHO1* and 4 samples were found to be with in-frame deletions of *KRT76* (Fig. S6). These results might suggest a possible loss of function in both genes and their roles in the pathogenesis of SA-CRC. In addition, a total of 18,882 somatic copy number alteration (SCNA) events were identified with 628 events per sample (Fig. 1C). We can clearly see that copy number gain events constituted of the majority of all SCNA events. Gains of 7q, 8q, chromosome 12, and 20q occurred in more than half of the samples. Four amplification regions and 26 deletion

Peer review under responsibility of Chongqing Medical University.

https://doi.org/10.1016/j.gendis.2022.05.026

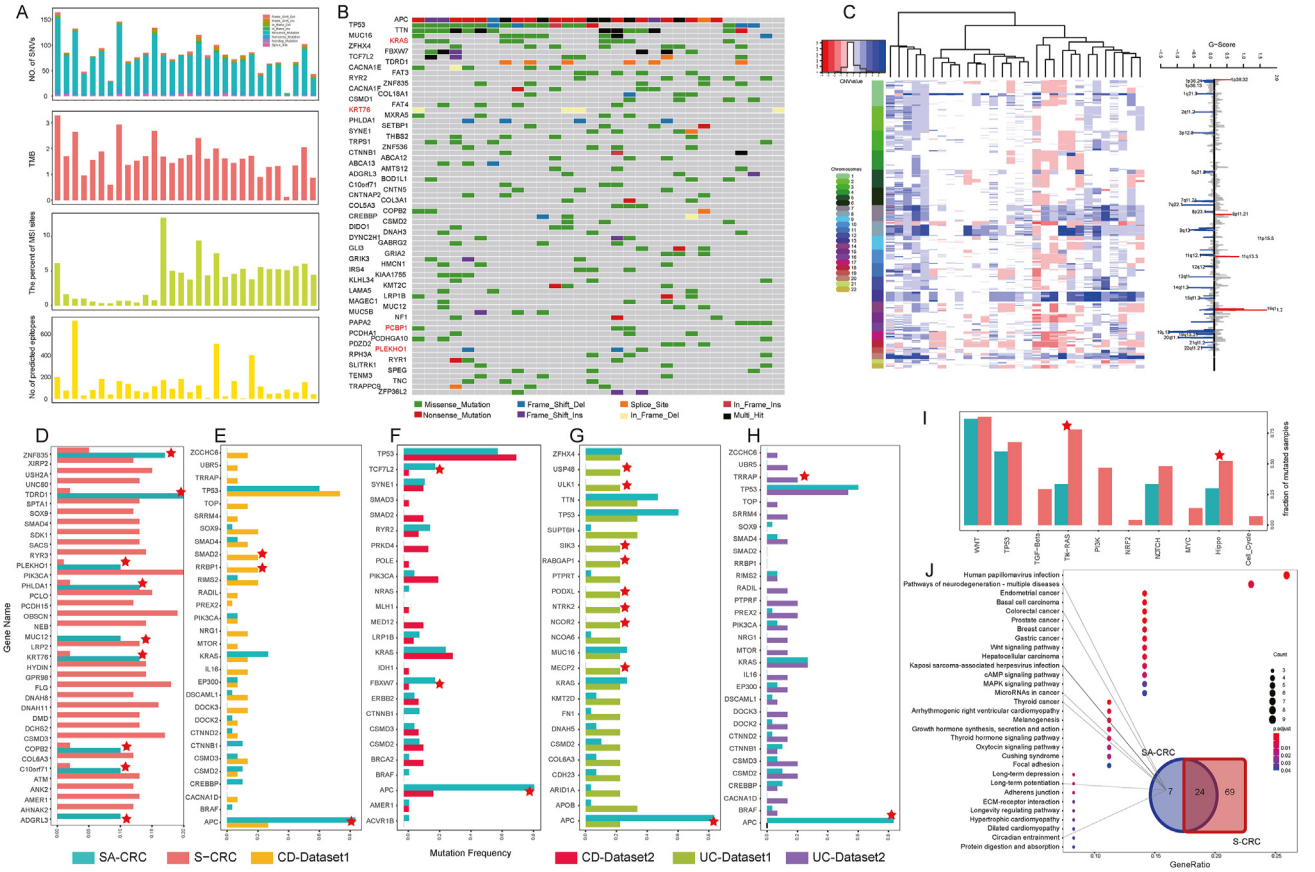

Figure 1 Genomic landscape of schistosomiasis-associated colorectal cancer. **(A)** Genetic characteristics of the SA-CRC. The number of mutations, the tumor mutation burden, the estimated MSI scores, and the predicted number of neo-epitopes of each sample were present. **(B)** Mutational landscape of recurrently mutated genes in SA-CRCs. Red arrow corresponds to the known driver gene, and the yellow arrow correspond to the novelly predicted driver genes. **(C)** Landscape of somatic copy number alternations (SCNAs) in SA-CRCs. Heatmap of SCNAs across all the samples (left) and the significant enriched SCNA regions (right) were showed. **(D)** Recurrently mutated genes with significant differences in frequency between SA-CRC and S-CRC. Gold star denotes the genes highly mutated in SA-CRC. **(E—H)** Comparison of recurrently mutated genes in frequency between SA-CRC and other types of CRC. Red star denotes the genes showing significant differences in frequency between two datasets. **(I)** Comparison of known oncogenic signaling pathways in frequency between SA-CRC and S-CRC. **(J)** Dot plot of the conventional enriched pathways in SA-CRC. The small Venn plot shows the overlap of enriched pathways between SA-CRC and S-CRC.

regions were found to be significantly enriched across the cohort. Several oncogenes were found in the enriched regions, including CCND1 and FGF3/4/19 in 11q13.3, MUC2/6/3A/12/17 in 11p15.5 and 7q22.1.

To better understand the process of carcinogenesis in SA-CRC, we compared the genomic profiles of SA-CRC with those of other types of CRC, including S-CRC and inflammatory bowel disease (IBD)-associated CRC. First, we found 9 recurrently mutated genes, including *KRT76* and *PLEKHO1*, significantly highly mutated in SA-CRC compared to S-CRC, and some common driver genes, such as *PIK3CA*, *ATM* and *SMAD4*, were rarely seen in SA-CRC (Fig. 1D). However, most of the top 20 mutated genes showed similar frequency distributions between SA-CRC and S-CRC. Second, APC was the only gene with significant difference in frequency compared to IBD-associated CRC (Crohn's disease-associated colorectal cancer (CD-CRC) and ulcerative colitis-associated colorectal cancer (UC-CRC), Fisher test, $P < 0.05$, Fig. 1E—H). Besides, there were some other genes showing significant difference in frequency in each dataset, such as the known driver genes TCF7L2 and FBXW7. Third, most of the SCNA regions enriched in the SA-CRC were not found in the S-CRC, except one focal amplification region (8p11.21) and two focal deletion regions (21q11.2, 14q11.2) (Fig. 1C). Therefore, SA-CRC has a distinct genomic alternation profile from other types of CRC.

Notably, we further identified two known oncogenic signaling pathways to be significantly different in frequency between SA-CRC and S-CRC (Hipp: $P = 0.02$, $OR = 0.39$; RTK-RAS: $P < 0.001$, $OR = 0.14$, Fisher test, Fig. 1I). TGF-Beta and PI3K pathways enriched in the S-CRC were not detected in SA-CRC because of the paucity of mutations in tumor suppressor gene SMAD4 and oncogene PIK3CA. As for the conventional Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis, 31 pathways were identified and 7 of them were only enriched in SA-CRC (Fig. 1J). Among 7 enriched pathways, Kaposi sarcoma-associated herpesvirus (KSHV) infection had been detected at a high frequency in patient-derived bladder cancer tissue and might be associated with the pathogenesis of bladder cancer.[4] Two central nervous system-related

pathways: pathways of neurodegeneration—multiple diseases and long-term potentiation, might be worth to note, since schistosoma infections may lead to central nervous system lesions.[5]

As it was clinically important and challenging to identify SA-CRC patients with aggressive features, we sought to identify genomic alternations which were associated with patient outcomes (7 deaths, 11 recurrences). All the frequently mutated genes, enriched pathways, and significant enriched SCNA events were considered to detect their relations with overall survival (OS) and relapse-free survival (RFS) in SA-CRC. After adjustment for age and sex, the alternation of 4 genes (*HMCN1*, *MAGEC1*, *PCDHGA10*, and *SLITRK1*) was found to be significantly associated with OS and RFS (Table S3). In addition, ABCA12-mut and KMT2C-mut were significantly related to worse RFS. DNAH3-mut and MUC12-mut were significantly related to worse OS. However, in the S-CRC dataset, these genes showed no significant correlation with OS and RFS. These results suggested that the 8 genes might play critical roles in the tumor progression and prognosis of SA-CRC, and be worthy of further research.

Nevertheless, this study has some limitations. First, only WES data were available for analysis. Integration of other omics-based dataset, such as transcriptome and proteome, will provide extensive screening and examination of key genes and pathways which participate in the pathogenesis of SA-CRC. Second, the sample size was relatively small. However, epidemiological survey showed that SA-CRC only took the proportion of no more than 0.4% in all CRC cases.

In the East Asia, considerable evidence supports an etiological link between schistosomiasis and colorectal cancer. However, up to now, the pathogenesis of SA-CRC is still unknown. For the first time, this study investigates the genomic landscape of SA-CRC and detects the genomic differences with other types of CRC. The findings demonstrate a unique molecular profile of SA-CRC which provide clues to the etiology.

Author contributions

Haiyang Zhou contributed to the conception of the study; Anqi Wang, Jing Zhang, and Xinxin Li performed the experiment; Yu Dong wrote the manuscript. Caifeng Jiang and Haiyang Zhou revised it critically.

Conflict of interests

The authors declare that there are no competing interests.

Funding

This work was supported by the National Key R&D Program of China (No. 2019YFA0110601).

Data availability

The datasets generated during the current study are available in The National Genomics Data Center (https://ngdc.cncb.ac.cn/) with accession number PRJCA007606.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.gendis.2022.05.026.

References

1. Hamid HKS. *Schistosoma japonicum*-associated colorectal cancer: a review. *Am J Trop Med Hyg*. 2019;100(3):501—505.
2. Jia P, Yang X, Guo L, et al. MSIsensor-pro: fast, accurate, and matched-normal-sample-free detection of microsatellite instability. *Dev Reprod Biol*. 2020;18(1):65—71.
3. Din S, Wong K, Mueller MF, et al. Mutational analysis identifies therapeutic biomarkers in inflammatory bowel disease-associated colorectal cancers. *Clin Cancer Res*. 2018;24(20):5133—5142.
4. Lee S, Jang J, Jeon H, et al. Latent Kaposi's sarcoma-associated herpesvirus infection in bladder cancer cells promotes drug resistance by reducing reactive oxygen species. *J Microbiol*. 2016;54(11):782—788.
5. Ferrari TC, Moreira PR. Neuroschistosomiasis: clinical symptoms and pathogenesis. *Lancet Neurol*. 2011;10(9):853—864.

Dong Yu [a,1], Anqi Wang [b,1], Jing Zhang [c,1], Xinxing Li [d,1], Caifeng Jiang [e,**], Haiyang Zhou [b,*]

[a] *Center for Translational Medicine, Naval Medical University, Shanghai 200433, China*
[b] *Department of Colorectal Surgery, Changzheng Hospital, Naval Medical University, Shanghai 200003, China*
[c] *Department of Pathology, Changzheng Hospital, Naval Medical University, Shanghai 200003, China*
[d] *Department of General Surgery, Shanghai Tongji Hospital, Tongji University, Shanghai 200062, China*
[e] *Department of Gastroenterology, Dushu Lake Hospital Affiliated to Soochow University, Suzhou, Jiangsu 215123, China*

*Corresponding author.

**Corresponding author.
E-mail addresses: cfjiang1999@163.com (C. Jiang), haiyang1985_1@aliyun.com (H. Zhou)

25 April 2022
Available online 7 June 2022

[1] These authors contributed equally to this work.