Introduction

Next-generation sequencing technologies have significantly transformed cancer genomics research through provision of reliable information about individual tumors, paving ways for precision therapy (1). WGS enhances the sequencing of the entire cancer patient genome and can be employed to detect all germline and somatic mutations, including copy number alternations, gene fusions and chromosomal rearrangements (2). Colorectal cancer (CRC) is one of the leading causes of cancer–associated mortality worldwide with about 1.8 million new cases and 860,000 deaths reported in 2018 (3). CRC is less common in young individuals and often times, young patients with CRC present with late-stage disease (stage III or IV) (4). Stage
III patients constitute a considerable proportion of most CRCs, and account for about one-third of all reported cases (5). More than two-thirds of all colorectal patients undergo curative surgery, with a considerable percentage of stage III patients experiencing tumor relapse manifesting as metastasis to distant organs or metachronous colorectal lesions within 5 years of follow up (6). Despite increasing scientific evidences that CRC is a heterogeneous disease, and genetic characteristics of the tumors influence patient prognosis and response to targeted therapies (7), the exploration of individual mutation profiles of same stage, advanced CRC patients has not received significant attention. We here explored the genomic landscape of a young long-term surviving stage III CRC patient and compared our data with those of conventional stage III patients reported in The Cancer Genome Atlas Network (TCGA) database. Mutations in APC, TP53, KRAS, SMAD4, FBXW7 and PIK3CA observed in TCGA patients were not recorded in our study. However, mutations in MUC4, MUC16, ARID1B, BAZ1A, BRCA2, CTNNND1 and NCOA2 rarely reported in TCGA patients were predominant in our patient. Additionally, we observed loss of heterozygosity (LOH) in POLE, RET, BMPR1A, NCOA4 and 30 other genes in contrast to deletion and amplification events recorded in TCGA patients. Put together, we offered substantial insights into the genomic features of the patient and provided a valuable resource for further study into the mutations that characterize advanced CRC which may be useful to design clinical therapy for personalized medicine. We present the following case in accordance with the CARE Guideline.

**Case presentation**

A 40-year old female patient presented CRC related symptoms and was admitted to the First Affiliated Anhui Medical University on October 23, 2008. She is a primary school teacher; has harmonious social relations and is married with two kids. The onset of symptoms was 4-month prior to her admission and this included abdominal discomfort, yellow watery stool and severe back pain. She had previously taken some self-prescribed intestinal drugs which made the symptoms subside. Much later, she developed diarrhea, and her tenesmus got worse. Although her body temperature and urine were normal, she experienced abdominal distension and pain. She also became anorexic and began to lose weight. She however denied any known family history of cancer. Diagnosis through colonoscopy revealed she had rectal invasive moderately differentiated adenocarcinoma. Her tumor size was 3 cm × 3 cm, and the carcinoma had invaded the serosa layer, covering 3/4 of the intestine tube. Upon examination of the mesenteric, two of her four lymph nodes showed metastases, although no obviously enlarged lymph nodes were found in the mesenteric roots. The patient was diagnosed with a stage III (T3N1M0, IIB) CRC. Examination of the rectum left lateral position showed there was a smooth mass 4 cm from the verge of the anus. The boundary was clear and the mass had a size of 2 cm × 2 cm. Pelvic CT scan revealed that the bowel was markedly narrowed about 10 cm from the anus and the site was seen as an irregular tissue shadow, with a length of 4 cm. From the colonoscopy, infiltrating lesion with four walls was observed at 12 cm, with an ulcerated surface. After laparotomy, the mass was found on the anterior wall of the rectum and the peritoneal reflex, 8 cm from the anal verge. Radical resection (Dixon operation) of the tumor was performed under anesthesia with intraoperative implantation of 600 mg of human flufenac into her abdominal cavity. Oxaliplatin in combination with Huaier granules, a traditional Chinese medicine, was given as adjuvant therapy. Huaier granules was applied in this case due to its affordability and demonstrated anti-tumor effects in various cancer including CRC. The clinical validity of Huaier granules combined with chemotherapy as adjuvant therapy after curative surgery in CRC patients has been demonstrated in several studies (8-10). Follow-up examination showed that the biochemical index was normal at the 8th day after surgery; the Neutrophils were slightly higher, while the lymphocytes were a little lower. The patient received relevant treatments under the guidance of a Doctor in compliance with the treatment plan and there were no adverse reactions to the treatment regimen. The patient had an overall survival of over 8-year and was lost to follow-up in July 2018 due to a change in her contact information (see Figure 1 for medical history timeline). To obtain the mutation profile of the patient, genomic DNA was isolated from the Formalin-Fixed, Paraffin-Embedded (FFPE) tumor and matched adjacent normal tissues using genomic DNA isolation kit (QIAGEN, Hilden, Germany). Extracted DNA was quantified using a NanoDrop ND-1000 spectrophotometer and the integrity was assessed with agarose gel electrophoresis. WGS of the prepared libraries was performed using the Illumina X10 (llumina Inc., San Diego, CA, USA) with 150-bp paired-end reads. Germline variants were called with GATK HaplotypeCaller joint
v3.8 (11). Variants which passed VQSR module, coverage \( \geq 3 \) were retained while the variants showing genotype “./.” in samples were filtered. Three callers: Lancet, Mutect v1.1.4, and SomaticIndelDetector 2.3-9 (12) were employed for somatic variants calling. Variants from the three callers were merged and all variants that flag as “PASS” were kept. Further annotations for both germline and somatic variants were added to each mutation using ANNOVAR (13) and VEP (14), respectively. Several publicly available databases such as 1,000 Genome Project, the Exome Aggregation Consortium, the Genome Aggregation Database and the compiled scores prediction system dbnsfp33a were also employed. Somatic structural variations were identified using FACETS and DELLY (15), thereafter, duplication, deletion and hemizygous segments were kept and annotated with ENCODE gene symbol in R v3.5.1 (16). Functional

**Figure 1** Medical history timeline of patient.
effects of the germline mutations were assessed using Mutation Assessor, Mutation Taster, M-CAP, Polyphen2, PROVEAN and SIFT (17), and that of somatic mutations with VEP. Germline mutations predicted to be deleterious by at least two algorithms were considered potential driver mutations while genes predicted by VEP as either possibly or probably damaging were considered as potential driver genes. Germline mutation analysis revealed a total of 4,532,691 SNPs/Indels. Among these, 19,102 SNPs/Indels were detected in the exonic regions, of which 11,090 were predicted to be protein altering, including 8,413 non-synonymous SNVs, 951 frameshift deletions, 310 frameshift insertions, 542 non-frameshift deletions, 163 non-frameshift insertions, 224 stop-gain, and 9 stop-loss. We filtered the variants using a set of publicly available datasets to exclude >1% MAF variants reported in the 1,000 Genomes Project, the Exome Aggregation Consortium and the Genome Aggregation Database and got 194 germline variants including 34 duplications, 12 loss of function and 69 hemizygous. Subsequently, we kept genes which were only altered in coding sequences and got 34 genes with LOH hemizygous. We queried the relevant data sets to select the most informative somatic variants. FACETS identified 34 genes with LOH in our patient and other stage III CRC cases, we queried TCGA database with the mutated genes. As we had an array of mutations, we focused our query on somatic mutations including copy number alterations. TCGA [colorectal adenocarcinoma (COAD)] comprises 8 cohort studies; DCFI Cell reports 2012, Genentech Nature 2012, TCGA Firehose Legacy, TCGA Nature 2012, TCGA PanCancer Atlas, MSKCC Genome Biology, 2014, MSKCC Cancer Cell, 2018, MSK Nature Medicine 2019 and contains mutation data of 3814 patients which cut across stages I to IV. For the database query, we excluded studies that did not provide information on patients’ age, tumor staging and survival outcomes. We also filtered out germline mutations as well as somatic mutations with unknown significance. However, we included somatic driver mutations annotated in OncoKB, cBioPortal and COSMIC. Of all the 4 databases queried, MUC16 predicted with potential driver mutations in our study, was modestly mutated in stage III patients reported in the DCFI Cell reports 2012 and TCGA Firehose Legacy studies. No profiles were returned for the MSKCC Cancer Cell 2018 project while higher mutation frequencies were recorded in TCGA PanCancer Atlas stage III patients. Furthermore, we straightly compared the exact different mutations in our patient with those of stage III CRC patients in TCGA database. Mutations in APC, TP53, KRAS, SMAD4, FBXW7 and PIK3CA predicted as drivers in TCGA stage III patients were not recorded in our study. However, mutations in MUC4, MUC16, ARID1B, BRCA2, CTNND1 and NCOA2 rarely reported in TCGA patients were predominant in our patient. Additionally, we observed there were no clear-cut correlations between patient age and survival outcomes in TCGA stage III CRC patients as no consistent trend was observed for a given age group or survival time. We got 1,996 structural events, including 1,367 translocations, 143 inversions, 306 tandem duplications and 180 deletions with DELLY; however, only in-frame events and fusions with distance over 50 Kb between two genes were kept. Ultimately, 53 significant structural events including 31 genes were reported. FACETS identified 107 large segments with copy number alternations, including 34 duplications, 12 loss of function and 69 hemizygous. Subsequently, we kept genes which were only alternated in coding sequences and got 34 genes with LOH and 1 with duplication. To investigate the differences between the mutations recorded in our patient and other stage III CRC cases, we queried LOH in POLE, RET, BAP1, NCOA4 and 30 other genes in contrast to deletion and amplification events recorded in TCGA stage III patients. We observed LOH in BRE, RAS, BAP1, NCOA4 and 30 other genes in contrast to deletion and amplification events recorded in TCGA stage III patients. We observed LOH in POLE, RET, BAP1, NCOA4 and 30 other genes in contrast to deletion and amplification events recorded in TCGA stage III patients.
were not captured in TCGA database, our literature search revealed that mutations in the gene is a rare occurrence in CRC patients.

**Discussion**

Germline mutations analysis identified 20 SNPs/Indels reported in the CGC and CPG databases. The germline mutations ranged from recurrent cancer susceptibility genes with quantifiable risks to rare genes not conventionally associated with CRC (18,19). Recently, Gong et al. (20), profiled 618 multi-stage Chinese CRC patients including 226 stage III and reported pathogenic germline mutation in 1 of every 3 patients younger than 50, stressing the importance of genetic testing in all Chinese patients younger than 50. In our somatic SNVs and Indels analysis, mutations in \textit{MUC4}, \textit{MUC16}, \textit{ARID1B}, \textit{BAZ1A}, \textit{BRCA2}, \textit{CTNND1} and \textit{NCOA2} were the most informative. However, functional annotations of the genes predicted \textit{MUC16} as a potential driver gene in our patient with others being either moderately tolerated or benign. Comparison with TCGA data showed that \textit{MUC16} had modest mutation frequency in other stage III patients regardless of their age. Although \textit{MUC16} mutations are often associated with ovarian cancer (21), emerging studies have linked mutations in the gene with other malignant conditions, including colon cancer (22,23). Further studies are however required to gain full insight into its oncogenic roles. A lack of correlation between \textit{MUC16} mutation and survival outcomes in different age groups of TCGA stage III patients may indicate these mutations have no significant influence on the patients’ prognosis. Also, we found no correlation between the age of the patients and their overall survival. Reports on the influence of age on the prognosis of metastatic CRC has not been consistent. Lieu et al. (4) asserted that age was a significant predictor of overall survival in metastatic CRC (stages III and IV) with the younger and older patients showing worse survival than patients of middle age while Schellerer et al. (24) presented opposing results. Inclusion of greater number of patients of particular age group, evaluation of age as a continuous variable, use of older databases among others were cited as probable reasons for the disparity (4). We developed a broad overview of copy number variations in our patient. Totally, we identified 35 copy number alternations including 34 LOH and 1 duplication. TCGA database query of these genes in conventional stage III CRC patients showed they had deep deletions and amplification events. LOH of one gene or another is thought to be relatively common in cancer of all types,

<table>
<thead>
<tr>
<th>Gene</th>
<th>Database</th>
<th>Chromosome</th>
<th>Variant classification</th>
<th>SIFT</th>
<th>PolyPhen</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARID1B</td>
<td>CGC723</td>
<td>chr6:157099313-157099315:CAC:-</td>
<td>In frame deletion</td>
<td>NSR</td>
<td>NSR</td>
</tr>
<tr>
<td>BAZ1A</td>
<td>CGC723</td>
<td>chr14:35234140-35234140:C:T</td>
<td>Splice site</td>
<td>NSR</td>
<td>NSR</td>
</tr>
<tr>
<td>BRCA2</td>
<td>CGC723_CPG114</td>
<td>chr13:32912123-32912123:G:A</td>
<td>Missense Mutation</td>
<td>Tolerated (0.87)</td>
<td>Benign (0.011)</td>
</tr>
<tr>
<td>CTNND1</td>
<td>CGC723</td>
<td>chr11:57556564-57556720:TTTTGATCTAGACTGGGCTTTCTC</td>
<td>Splice site</td>
<td>NSR</td>
<td>NSR</td>
</tr>
<tr>
<td>MUC16</td>
<td>CGC723</td>
<td>chr19:9069747-9069747:G:A</td>
<td>Missense mutation</td>
<td>NSR</td>
<td>Probably damaging (0.981)</td>
</tr>
<tr>
<td>MUC4</td>
<td>CGC723</td>
<td>chr3:195509573-195509573:A:G</td>
<td>Missense mutation</td>
<td>Tolerated low confidence (0.28)</td>
<td>Benign (0.316)</td>
</tr>
<tr>
<td>NCOA2</td>
<td>CGC723</td>
<td>chr8:71069450-71069450:G:A</td>
<td>Missense mutation</td>
<td>Tolerated (0.16)</td>
<td>Benign (0.001)</td>
</tr>
</tbody>
</table>

CRC, colorectal cancer; NSR, no score returned.
however, it is particularly significant in individuals who have inherited a predisposition for cancer suggesting the high possibility of genetic predisposition in our patient (25). Structural variation analysis identified mutations in BAGE2, a B Melanoma Antigen Family Member 2, which to the best of our knowledge has only been reported in CRC patients from the Han Chinese nationality (26). It remains to be determined whether this gene could be a biomarker or a precision therapy target in the population. The disparity observed between our mutation data and that of TCGA demonstrate the highly genetically heterogeneous nature of same stage CRC and this could be due to the differences in the anatomical pathology of CRC as the disease affects different regions of the digestive tract (27). Moreover, genes that drive tumor progression in different regions may be dissimilar, and most CRC studies do not pay specific attention to separating these regions (28). Notably, the use of WGS technique and multiple variant calling tools are the strengths of our study. WGS approach provides a complete coverage of the coding and noncoding regions, enhancing a comprehensive assessment of the patient genome. It also offers a more robust determination of copy number variations, rearrangements and other structural variations due to the longer reads length (29). The use of multiple tools for variant calling facilitates a more accurate and consistent variant identification, providing clinical-grade variant information for genomic medicine (30). However, our study is not without its limitations. Although WGS gave a comprehensive overview of the various alterations in the cancer patient genome, the differential expression profiles of the mutated genes could not be evaluated as we had no transcriptomics data. In conclusion, whole-genome profiling and comparison of our data with TCGA database produced a genomic mutation profile of the patient and offered a valuable resource for further study into the mutations that characterize long-term surviving stage III CRC which may be useful to design clinical therapy for personalized medicine.

Acknowledgments

None.

Footnote

Conflicts of Interest: 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. This study was approved by the Ethics Committee of the First Affiliated Hospital of Anhui Medical University (No. PJ-20190110) and is in accordance with the Helsinki Declaration as revised in 2013. Written informed consent was obtained from the patient for publication of this Case report and any accompanying images.

References


