Cancer Genetics 231–232 (2019) 54–61

Cancer Genetics

ORIGINAL ARTICLE

Genomic landscape of synchronous tubulovillous adenoma and multiple non-familial colon cancers from a single patient

Kyung Kim[a,1], Su-Hye Choi[a,1], Jeeyun Lee[a], Won-Suk Lee[b,*]

[a] *Division of Hematology–Oncology, Department of Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea;* [b] *Department of Surgery, Gil Medical Center, Gachon University, Incheon, Republic of Korea*

Abstract

Colorectal cancer (CRC) is one of the leading causes of cancer-related death. We analyzed genomic of non-familial tubulovillous adenoma (TVA) and two synchronous malignant colorectal adenocarcinomas from a single patient. The number of somatic mutations was higher in the tumor sample (especially, AV 50 cm adenocarcinoma sample) than TVA sample, and also the allele frequency of mutation was higher on colon adenocarcinoma samples than TVA. Although they were very low frequency of sharing same genomic alterations between the three lesions, *APC* gene mutation was present in all three lesions, which confirm that *APC* gene mutation was an early event in this patient. The genetic alterations of *APC, KRAS* and *TP53*, which play an important role in the development of carcinoma from TVA, were shown through whole exome sequencing data of each sample. Of note, of the two synchronous adenocarcinoma samples, one lesion was *KRAS* mutant while the other one was *KRAS* wild-type. Our findings implicate that *KRAS* mutation may need to be taken from every primary cancer in a patient with multiple primary sites since *KRAS* status may differ amongst synchronous cancer lesions.

Keywords Colorectal cancer, Whole exome sequencing, Tubulovillous adenoma.

Introduction

Colorectal cancer (CRC) is the third most common cancer in men and the second most common cancer in women globally. The adenoma-carcinoma sequence has been described in colon tumorigenesis in the past decades [1,2]. Briefly, the tumorigenesis begins with the transformation of normal epithelium to an adenoma, which then progresses *in situ* into a carcinoma, and ultimately into an invasive and metastatic tumor [1–3].

The alteration of the *APC, KRAS* and *TP53* genes in the genetic pathway, where normal colonic epithelia proceed to malignant carcinoma, is associated with progression to malignancy [4]. Mutations in the V-Ki-ras2 Kirsten rat sarcoma viral oncogene homolog (*KRAS*), most frequently detected in codons 12, 13, and 61, occur in approximately 40% of metastatic CRC patients [5,6]. *KRAS* mutations are well established negative selection factor for Cetuximab indication [7,8]; for instance, only *KRAS* wildtype CRC patients demonstrated significantly improved survival outcome following Cetuximab-based treatment.

In addition, mutational load in CRC has been shown to be associated with microsatellite instability (MSI-H) tumor which is a strong predictor for immune-targeted therapy [9]. Tumor mutational burden (TMB) is an important indicator for immunetherapy [10]. In case of multiple cancers, they may show different TMB which underscores the importance of genomic profiling of each tumor when the patient has multiple cancers occurring simultaneously.

In this study, we have sequenced tubulovillous adenoma (TVA) and 2 malignant adenocarcinomas which occurred simultaneously in a single patient. Our aims of this study were to: (1) genomically characterize each lesion from a single

Received October 5, 2018; received in revised form January 8, 2019; accepted January 8, 2019

* Corresponding author.

E-mail address: lws@gilhospital.com

[1] Both authors contributed equally to this work.

2210-7762/$ - see front matter © 2019 Published by Elsevier Inc.
https://doi.org/10.1016/j.cancergen.2019.01.004

patient) (TVA and two sited cancers) (2) demonstrate genomic differences amongst different lesions occurring in a single patient in terms of somatic mutation and mutational load.

Methods and materials

Tumor specimens and ethical statement

Two synchronous sigmoid colon cancer fresh frozen tumors and an ascending colon TVA with adjoining non-tumor colon mucosal tissues include no necrosis or hemorrhage were collected at the Gil Medical Center, Incheon, Korea. This study was approved by the Institutional Review Board of Gil Medical Center, Incheon, Korea.

Exome-enriched library preparation experimental process

DNA fragmentation

First, the purified whole genome DNA was randomly fragmented by using a Covaris and then the DNA was analyzed by using the Agilent Bioanalyzer 2100. Assuming successful fragmentation, a peak between 200–300 bp was expected. Following successful fragmentation the resulting overhangs need to be converted to blunt ends, which was performed using the Sure Select Library Prep Kit, followed by a clean-up step (DNA isolation) using the AMPure XP Beads. The DNA was then stored at for −20 °C up to 7 days.

Adopter ligation & 3' ends A-tailing

The success of the fragmentation between fragmented DNA and index adaptors was increased, and the 3' ends were adenylation to reduce the self-ligation of the blunt fragment. Immediately after adenylation, the adapter is coupled to the newly adenylation genome DNA, and then purified by using AMPure XP beads. The sample was stored at −20 °C for up to 7 days.

Size selection

The samples were extracted from 2% of Agarose gel and extract an extract of 300–400 bp. Due to the adapter ligation (∼60 bp/adapter) the size of the target band was shifted from the 200–300 bp range to the 300–400 bp range. Following gel extraction and column purification of the DNA, successfully ligated DNA fragments which including adapter sequences are enhanced via PCR using adapter specific primers. DNA was once again isolated using the AMPure XP Beads and analyzed using the Agilent Bioanlyzer 2100. A sharp peak between the 300–400 bp ranges was expected.

Exome capture

First DNA was denatured and hybridized overnight to the Sure Select RNA probes. The Sure Select Library Probes are 120 bp biotinylated-RNA probes which target all human exons. Following hybridization, the biotinylated probes (and hybridized sample DNA) are captured via streptavidin beads, washed, and then DNA is eluted from the probes. The RNA probes were then digested and successfully captured, and DNA fragments that included the ligated adapter sequences were first purified and then enhanced via PCR using adapter specific primers. DNA was then purified again using the AMPure XP Beads and analyzed using the Agilent Bioanalyzer 2100. Once again a peak in the 300–400 bp range was expected. Following this, samples were quantified using QPCR. During the enhancement PCR step, unique index sequences were incorporated into each molecule, which allows for sample pooling following QPCR quantification. After samples were pooled they were then prepped for cluster generation.

High- throughput sequencing

The enriched exome library was loaded on the flow cells of Illumina cBot for cluster generation. The flow cells with clusters of the improved exome libraries were transferred to an Illumina HiSeq2500. High-throughput sequencing was then performed for each captured library to insure that each sample met the desired sequencing depth of 100X.

Identification of the sample origin

In order to verify that each sample was derived from the same patient, we analyzed the germline variants with bioinformatics method [11]. We sequenced three samples from tumor region, and one blood sample for normal. After whole-exome sequencing of these samples, we used 'Burrow-Wheeler Aligner (BWA)-MEM' for alignment to reference genome [12]. And then, we did PCR duplicate removal, recalibration of base and local realignment around known indels with GATK for further analysis [13]. We used 'Haplotype-caller' to detect germline mutations that each sample had. We defined the list of germline mutations of each sample and filtered these mutations with the following criteria: mutations were in exonic region; not included in 'Catalogue of somatic mutation in cancer (COSMIC)' database [14].

Data analysis

Alignment to the reference genomes (hg19 for human) was performed using Burrows-Wheeler Aligner (BWA) [12]. After pre-processing of exome sequencing data using GATK [13] (local reorganization, duplicate marking and base recalibration) Samtools [15] and Picard, somatic variants were identified for point mutations and indels severally using MuTect [16] and SomaticIndelDector [17]. The ANNOVAR [18] was used to select somatic variants located in exon sequences and predict their functional consequences.

The data was then filtered due to good quality variants expected to have potentially harmful effects consequence ('non-synonymous SNV', 'splicing', 'frameshift substitution', 'stopgain SNV', 'stoploss SNV') using NCBI, UCSC [19] database, SIFT [20] and PolyPhen [21]. Exclusion mutations with additional minor allele frequencies (MAF)$>$ 0.001

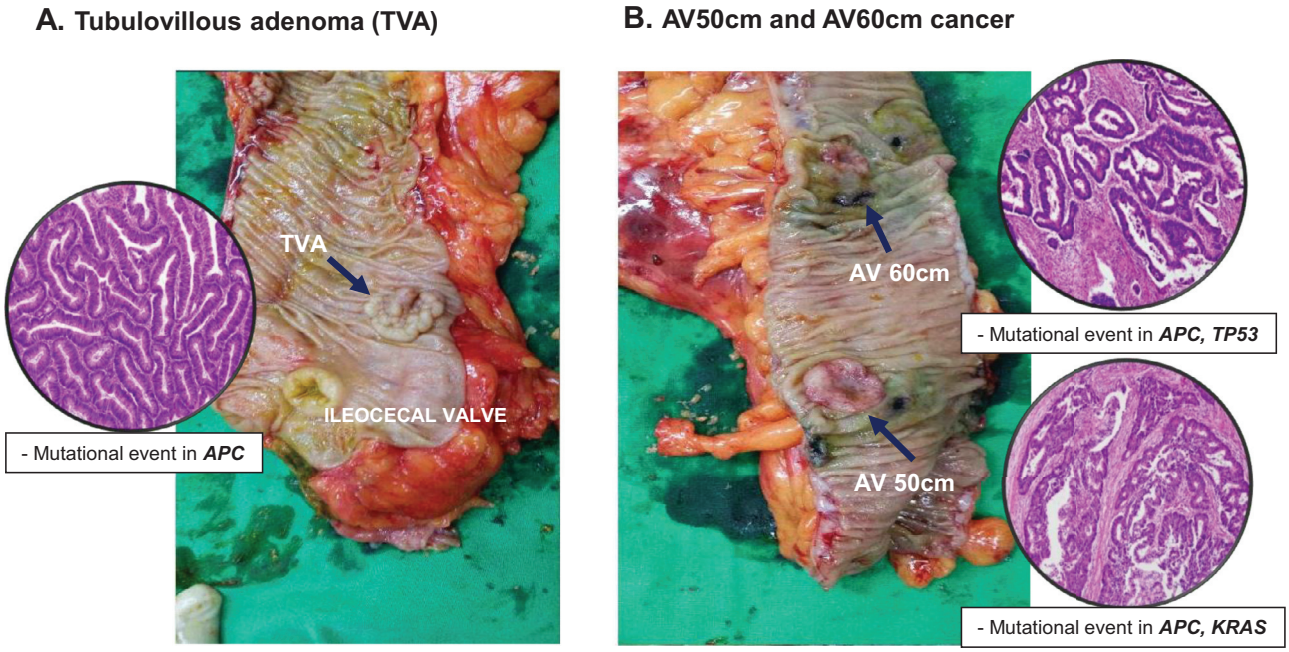

Fig. 1 Gross specimen and pathologic pictures (hematoxylin and eosin stain x200).

(Liberal threshold selected based on population frequency of both diseases) observed in NCBI dbSNP [22] (Release 147), 1000Genomes [23] and ExAC [24] were further filtered.

Results

A 74-year-old man presented with ongoing lower abdominal pain. The patient had two distal sigmoid colon cancers in close proximity with an ascending colon TVA without any metastasis. The patient's initial carcinoembryonic antigen level was 0.74 ng/mL. The patient had no prior history of any cancer or comorbid diseases. The patient underwent subtotal colectomy for cecal TVA and two synchronous sigmoid colon cancer (Fig. 1). The pathologic features were as follows: microsatellite stable, *KRAS* wild type, both AV 50 cm and AV 60 cm cancers indicated T2 lesions, no lymphovascular invasion, no perineural invasion, and 0 out of 27 lymph node metastases with negative resection margins. Currently, the patient is in a disease-free state without any evidence of recurrence.

Genomic concordance between synchronous TA, TVA and colon cancer from a single patient

Using the paired-end sequencing strategy, we simultaneously generated sequence data from normal colon tissues, primary colon tumors at AV 50 cm and AV 60 cm, and a right colon TVA from the single patient described above, with the corresponding depth of 33X–43X. The generated reads were mapped 99% on average (normal, 99.22%; TVA, 99.27%; AV 50 cm cancer, 99.35%; AV 60 cm cancer 99.24%). Coverage of depth in the target area was 308X on average (normal, 293X; TVA, 300X; AV 50 cm cancer, 302%; AV 60 cm cancer, 339%). The widespread somatic changes were noted throughout the genome, from somatic single nucleotide variants (SNVs) to small insertions or deletions (indels) (Fig. 2A).

Concordance of germline mutations in several samples derived from one patient

We analyzed that how many germline mutations are overlapped between samples using the final list of filtered germline mutations of each sample. These samples shared most of same mutations of germline. We found that the germline variants in each sample overlap with a high probability (Fig. 4). (The rate of concordance of germline mutations between blood and TVA, between blood and AV 50 cm cancer, between blood and AV 60 cm cancer $\geq$99.1%; between TVA and AV 50 cm cancer, TVA and AV 60 cm cancer $\geq$98.2%; AV 50 cm cancer and AV 60 cm cancer $\geq$96.7%)

Mutation analysis

Alignments to the reference genome with opened gaps were performed using MuTect [16], and the SomaticIndelDetector [17] was used to map out SNVs and indels from the cancer and matched adjacent normal tissues. In total, 79–258 exonic point mutations and indels per sample were detected in primary tumors.

Common gene alterations

In order to identify significant mutations to be malignant in colon cancer, we compared somatic mutations of each

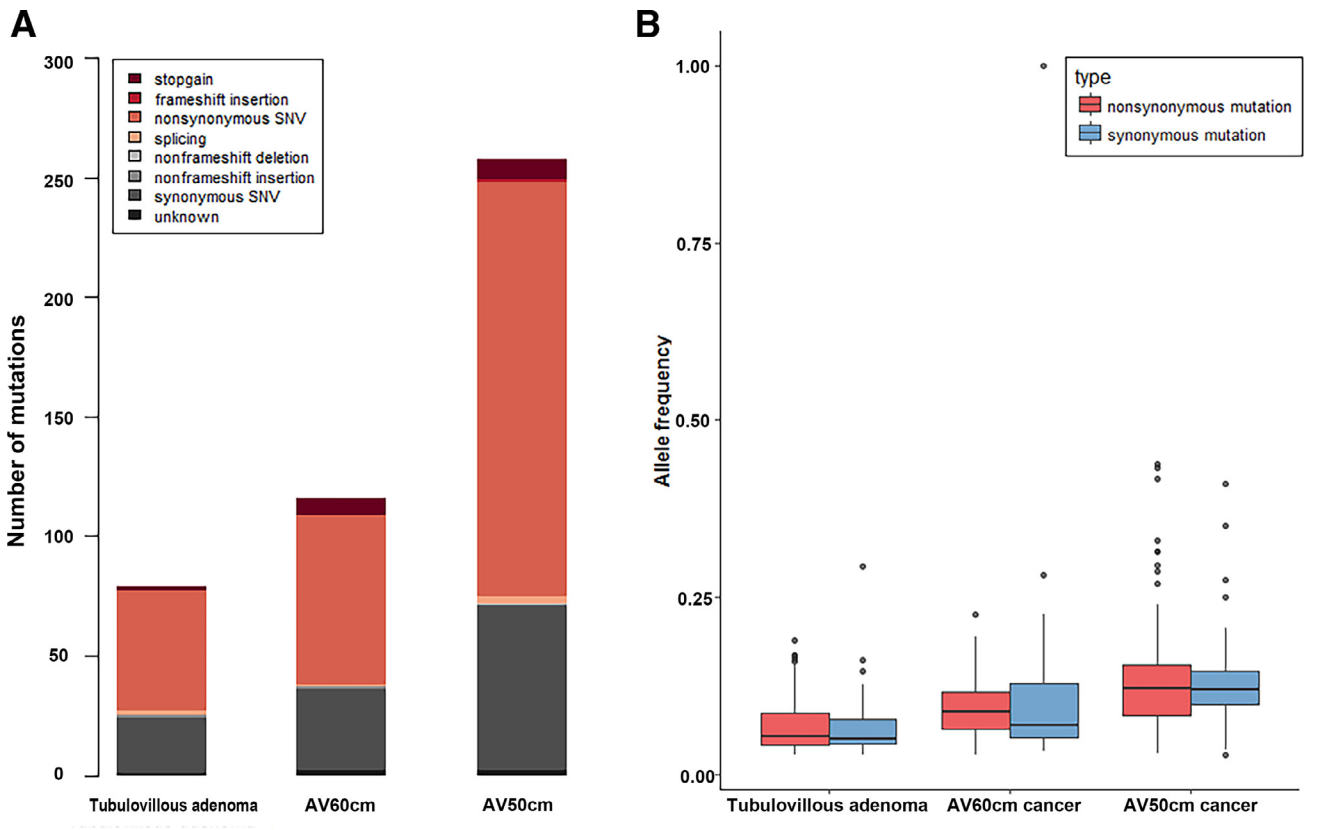

Fig. 2 Abundance of somatic mutations profiles in adenoma and 2 colon cancer genomes. (A) The number of somatic mutations is shown with the 8 functional categories. (B) Mutant allele frequencies (*y*-axis) of point mutations are shown for adenoma and 2 colon cancer genomes ($^{**}p < 0.01$).

sample. There are no common somatic mutations in TVA and 2 colon cancers. Although there is no exact point mutation in these samples, common gene alterations were found at the gene level (Fig. 3A). Interestingly, TVA and two colon cancers have mutations in the *APC* gene in common.

Mutant allele frequency

To extract possible genetic changes during TVA to adenocarcinoma and synchronous adenocarcinoma to adenocarcinoma, we performed an in-depth investigation of the mutations of matched TVA and primary cancer samples. A huge part of SNVs were shared by both stages, no significant differences were observed between the allele frequency of nonsynonymous and splice site SNVs in cancer-related genes (i.e., TVA, $p = 0.89$; AV60 cm cancer, $p = 0.40$; AV50 cm cancer, $p = 0.81$;). However, the AV 50 cm cancer had higher mutant allele frequencies than other the samples ($p < 0.01$) (Fig. 2B). Non-silent mutations were observed in *APC*, which was the only altered gene that was shared between all three samples. The AV 50 cm and AV 60 cm cancer samples shared only 2 gene mutations (VEGFC and COL6A3), and the TVA with AV 50 cm and AV 60 cm cancer samples shared one mutation each, respectively (i.e., CTR9 and MYH8) (Fig. 3B, Supplement_table1).

Discussion

In this study, we identified three different sets of genetic characters (TVA, AV 50 cm, AV 60 cm) from a patient. There were more somatic mutations in AV 60 cm than TVA, the highest number of somatic mutations in AV 50 cm. In addition, the mutational allele frequency was the highest in the AV 50 cm (Fig. 2, Supplement Table 2). Interestingly, we found that TVA, AV 60 cm and AV 50 cm have *APC* gene alteration, commonly (Fig. 3). In genetic pathway to be malignant carcinoma, *APC* alteration is an early event. In the development of colon cancer, a sequential change of *KRAS* and *TP53* is necessary after *APC* alteration [4,25].

AV 50 cm cancer had mutations in *KRAS* gene in contrast with AV 60 cm cancer. And AV 60 cm cancer had mutations in *TP53* gene in contrast with AV 50 cm cancer (Fig. 4). The reason why the mutation of *TP53* and the *KRAS* gene is not found in AV 50 cm and AV 60 cm is thought to be the tumor heterogeneity. Hence, multiple primary cancer lesions should be considered for *KRAS* mutational testing in every lesion based on our finding that the two lesions had discordance in *KRAS* mutational status. Previous studies have shown that *KRAS* mutation is a mutation that imposes an impact on carcinogenesis, especially on colon cancer [4]. In light of this, it is known that the tissue examination of colorectal TVA is a very important event for CRC early diagnosis.

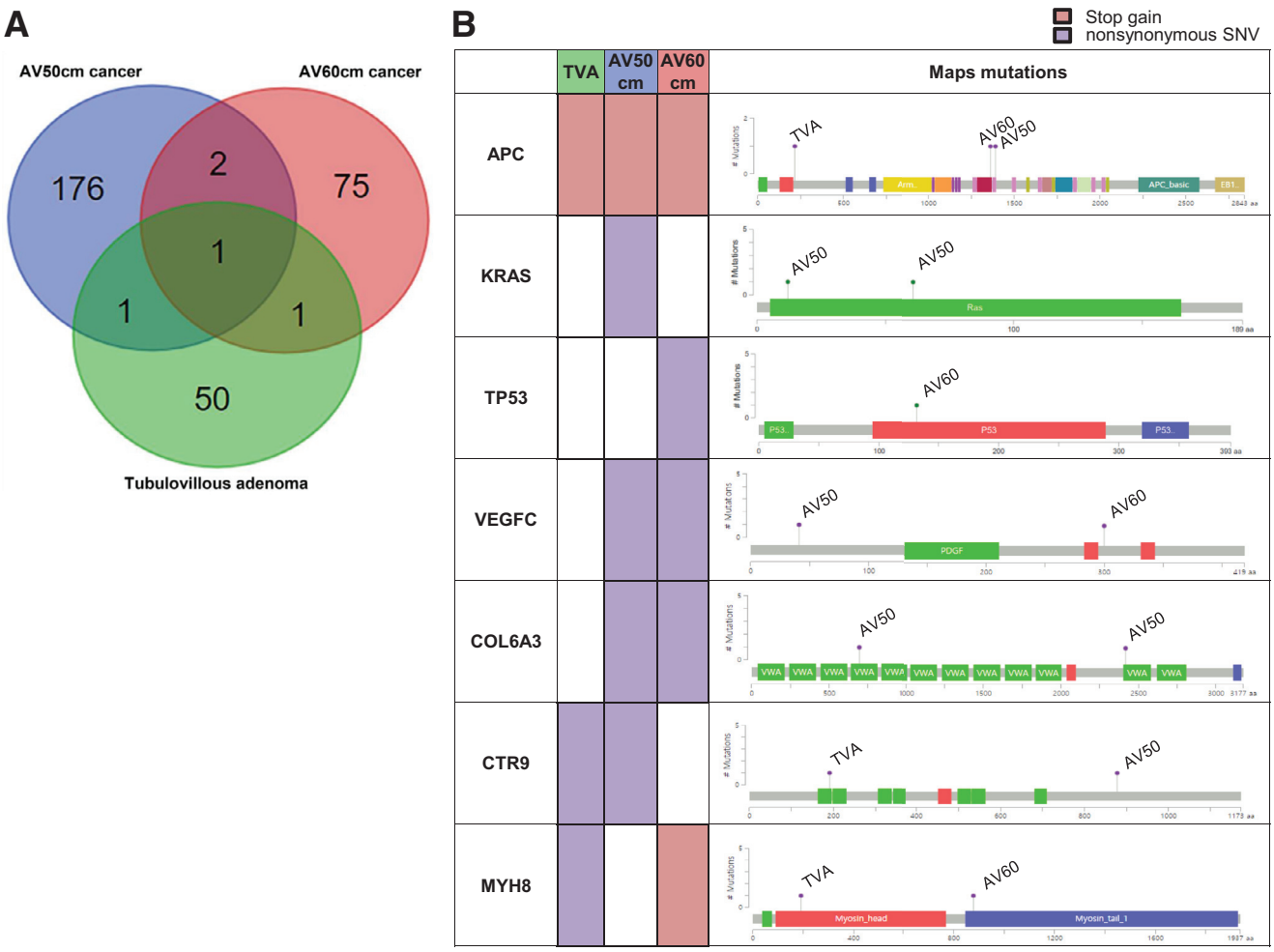

Fig. 3 (A) Comparison of gene alterations in the adenoma and 2 colon cancers. (B) Common genes acquired non-silent somatic mutations in the adenoma and 2 colon cancers. Whole exome sequencing in the adenoma and 2 colon cancers. Mutation landscape with the mutation map of the 5 common mutated gene and 2 mutated oncogene. Each row represents a gene and each column represents sample.

Previously, such research was conducted only in mainly, familial CRCs study. Unlike the other research studied CRCs, we sequenced and founded these mutations from three lesions from a non-familial CRC patient, interestingly.

Multiple cancers may arise simultaneously from regions of normal tissue containing certain genomic alterations. This phenomenon is exemplified by the 'field cancerization' caused by exposure to carcinogens [26,27]. In contrast with genetic alterations found in isolated cases and in only minor fractions of normal tissue cells [26,28], somatic epigenetic alterations are common in normal tissues adjacent to various cancers [29–31]. Although considerable genomic data has been produced for colon cancers, whole exome sequencing has rarely been applied to synchronous lesions. The aims of our study were: first, we attempted to identify somatic mutations and genome-wide MAFs for synchronous colon cancers. Secondly, we attempted to detect genomic differences between TVA and sporadic colon cancers that might drive TVA to cancer progression.

We found that genomic alterations for TVA were comparable to those for synchronous cancers in quantity (i.e., total mutations), but that driver alterations for TVA were less common than those for synchronous colon cancer (i.e., number of driver mutations and co-occurrence of mutations). Our data indicated that TVA may have qualitatively less aggressive genomes that may need further driver hits to develop into colon cancer genomes. To find critical determinants for TVA progression to colon cancer, the analysis for driver genes identified that synchronous TVA-cancer harbored many more drivers than TVA (Fig. 2). However, we could not pinpoint recurrent determinants for the progression. These data indicate that there may be neither a single driver nor a recurrent group of drivers for the progression. *APC* (gain) somatic mutation was the only one identified in synchronous TVA-AV 50 cm-AV 60 cm in this study [3].

Our study demonstrated that *APC* is an early genomic event. Despite the lower prevalence of driver mutations in TVA than synchronous cancers, even TVA harbored at least one driver such as *APC, CTR9*, and *MYH8* mutations, suggesting that these drivers may be essential for the early phase of cancer development, and that gradual accumulation of driver mutations might be required for its progression.

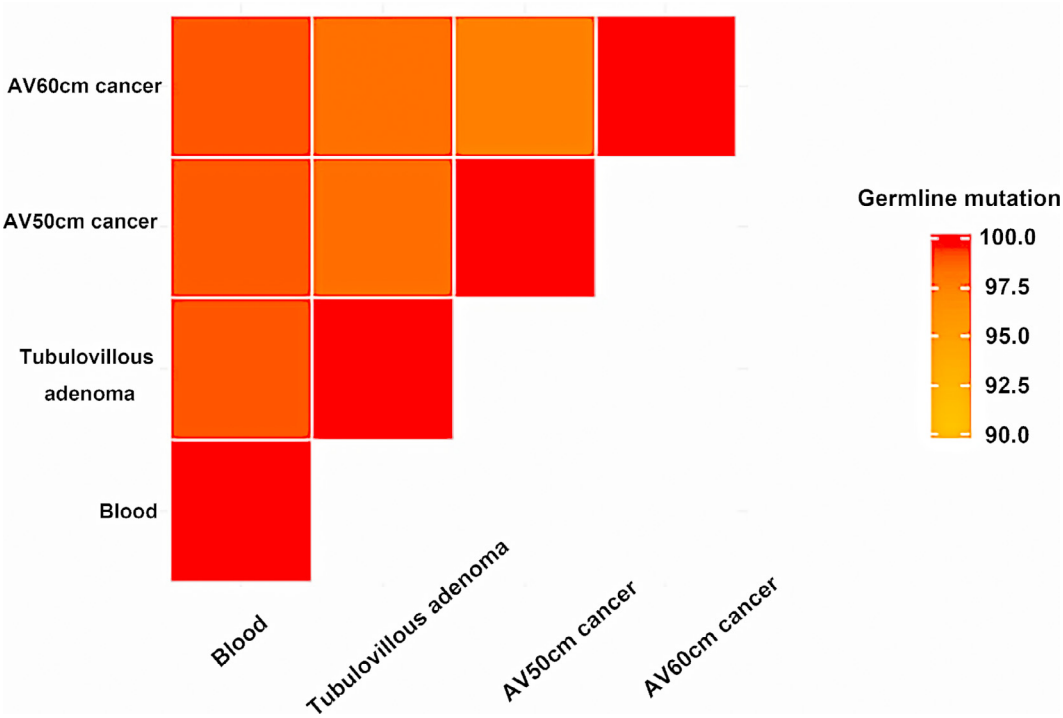

Fig. 4 Heatmap indicated the degree to which germline mutations are overlapped between two samples. Based on the samples on the *X*-axis, it was shown how many of the overlapped germline mutations between *X*-axis sample and *Y*-axis sample.

CTR9 is located at 11p15.3 and encodes 1173 amino acid proteins. It is widely expressed, including in fetal and adult kidney, and shows evolutionary conservation throughout eukaryotes [32]. *CTR9* contains multiple tetratricopeptide repeats (TPRs), a versatile protein–protein interaction domain that can act as an interaction scaffold in multi-protein complexes involved in diverse cellular processes, and is a core component of polymerase-associated factor complex (PAF1c), which has multiple roles in RNA Polymerase II regulation [33]. Mutations in cancer predisposition genes markedly contribute to both familial and non-familial cases, whereas for others the contribution to non-familial cases is small. To evaluate the contribution of *CTR9* to non-familial TVA to colon cancer progression is warranted. *CTR9* mutations seem to be a very rare cause of colon cancer and need to be analyzed via further investigations.

Some genes displayed alterations in both TVA and AV 50 cm and AV 60 cm, indicating their roles in both the initiation and progression/maintenance of colon cancers.

For example, the two mutations shared by AV 50 cm and AV 60 cm in our study were VEGFC and COL6A3. Collagen 6A3 (*COL6A3*), a component of the extracellular matrix, is often up-regulated in tumors and is believed to play a pro-oncogenic role. It could promote tumor growth by modulating Hippo and Wnt signaling [34].

The study involved a typical genetic change from TVA to colon cancer. Also, it confirmed tumor heterogeneity by genetic difference between colon cancers. Furthermore, in addition to the typical colon cancer mutation, characteristics for random mutations can be seen.

Role of the funder/sponsor

The funders had no role in the design and conduct of the study.

Author contributions

Study concept and design: Won-Suk Lee and Jeeyun Lee.

Acquisition analysis or interpretation of data: Kyung Kim, Su-Hye Choi and Jeeyun Lee.

Draft of the manuscript: Won-Suk Lee and Kyung Kim.

Won-Suk Lee and Kyung Kim had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Conflicts of interest

The authors declare that there is no conflict of interest.

Acknowledgment and funding

This work was supported by funding from the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1A6A3A11932444) to K. Kim and NRF-2017R1D1A1B03035453 to W.-S. Lee. This study was supported by a grant from the Korean Health Technology R&D Project, Ministry of Health & Welfare, Republic of Korea (HI14C3418)

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cancergen.2019.01.004.

References

[1] Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. Cell 1990;61:759–67.

[2] Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Nakamura Y, White R, Smits AM, Bos JL. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525–32.

[3] Powell SM, Zilz N, Beazer-Barclay Y, Bryan TM, Hamilton SR, Thibodeau SN, Vogelstein B, Kinzler KW. APC mutations occur early during colorectal tumorigenesis. Nature 1992;359:235–7.

[4] Smith G, Carey FA, Beattie J, Wilkie MJV, Lightfoot TJ, Coxhead J, Garner RC, Steele RJC, Wolf CR. Mutations in APC, Kirsten-ras, and p53—alternative genetic pathways to colorectal cancer. Proc Natl Acad Sci 2002;99:9433–8.

[5] Poehlmann A, Kuester D, Meyer F, Lippert H, Roessner A, Schneider-Stock R. K-ras mutation detection in colorectal cancer using the Pyrosequencing technique. Pathol Res Pract 2007;203:489–97.

[6] Raponi M, Winkler H, Dracopoli NC. KRAS mutations predict response to EGFR inhibitors. Curr Opin Pharmacol 2008;8:413–418.

[7] De Roock W, Jonker DJ, Di Nicolantonio F, Sartore-Bianchi A, Tu D, Siena S, Lamba S, Arena S, Frattini M, Piessevaux H, Van Cutsem E, O'Callaghan CJ, Khambata-Ford S, Zalcberg JR, Simes J, Karapetis CS, Bardelli A, Tejpar S. Association of KRAS p.G13D mutation with outcome in patients with chemotherapy-refractory metastatic colorectal cancer treated with cetuximab. JAMA 2010;304:1812–20.

[8] Karapetis CS, Khambata-Ford S, Jonker DJ, O'Callaghan CJ, Tu D, Tebbutt NC, Simes RJ, Chalchal H, Shapiro JD, Robitaille S, Price TJ, Shepherd L, Au HJ, Langer C, Moore MJ, Zalcberg JR. K-ras mutations and benefit from cetuximab in advanced colorectal cancer. N Engl J Med 2008;359:1757–65.

[9] Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, Fisher GA, Crocenzi TS, Lee JJ, Duffy SM, Goldberg RM, de la Chapelle A, Koshiji M, Bhaijee F, Huebner T, Hruban RH, Wood LD, Cuka N, Pardoll DM, Papadopoulos N, Kinzler KW, Zhou S, Cornish TC, Taube JM, Anders RA, Eshleman JR, Vogelstein B, Diaz LA. PD-1 blockade in tumors with mismatch-repair deficiency. N Engl J Med 2015;372:2509–20.

[10] Pai SG, Carneiro BA, Chae YK, Costa RL, Kalyan A, Shah HA, Helenowski I, Rademaker AW, Mahalingam D, Giles FJ. Correlation of tumor mutational burden and treatment outcomes in patients with colorectal cancer. J Gastrointest Oncol 2017;8:858–866.

[11] Campbell CD, Eichler EE. Properties and rates of germline mutations in humans. Trends Genet 2013;29:575–84.

[12] Li H, Durbin R. Fast and accurate short read alignment with Burrows–Wheeler transform. Bioinformatics 2009;25:1754–60.

[13] McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 2010;20:1297–303.

[14] Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, Cole CG, Ward S, Dawson E, Ponting L, Stefancsik R, Harsha B, Kok CY, Jia M, Jubb H, Sondka Z, Thompson S, De T, Campbell PJ. COSMIC: somatic cancer genetics at high-resolution. Nucleic Acids Res 2017;45:D777–DD83.

[15] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. The sequence alignment/map format and SAMtools. Bioinformatics 2009;25:2078–9.

[16] Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol 2013;31:213–19.

[17] DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, Philippakis AA, del Angel G, Rivas MA, Hanna M, McKenna A, Fennell TJ, Kernytsky AM, Sivachenko AY, Cibulskis K, Gabriel SB, Altshuler D, Daly MJ. A framework for variation discovery and genotyping using next-generation DNA sequencing data. Nat Genet 2011;43:491–8.

[18] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res 2010;38:e164.

[19] Tyner C, Barber GP, Casper J, Clawson H, Diekhans M, Eisenhart C, Fischer CM, Gibson D, Gonzalez JN, Guruvadoo L, Haeussler M, Heitner S, Hinrichs AS, Karolchik D, Lee BT, Lee CM, Nejad P, Raney BJ, Rosenbloom KR, Speir ML, Villarreal C, Vivian J, Zweig AS, Haussler D, Kuhn RM, Kent WJ. The UCSC genome browser database: 2017 update. Nucleic Acids Res 2017;45:D626–DD34.

[20] Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 2009;4:1073–81.

[21] Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR. A method and server for predicting damaging missense mutations. Nat Methods 2010;7:248–9.

[22] Sherry ST, Ward M, Sirotkin K. dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome Res 1999;9:677–9.

[23] Genomes Project C, Abecasis GR, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. Nature 2010;467:1061–73.

[24] Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, Tukiainen T, Birnbaum DP, Kosmicki JA, Duncan LE, Estrada K, Zhao F, Zou J, Pierce-Hoffman E, Berghout J, Cooper DN, Deflaux N, DePristo M, Do R, Flannick J, Fromer M, Gauthier L, Goldstein J, Gupta N, Howrigan D, Kiezun A, Kurki MI, Moonshine AL, Natarajan P, Orozco L, Peloso GM, Poplin R, Rivas MA, Ruano-Rubio V, Rose SA, Ruderfer DM, Shakir K, Stenson PD, Stevens C, Thomas BP, Tiao G, Tusie-Luna MT, Weisburd B, Won HH, Yu D, Altshuler DM, Ardissino D, Boehnke M, Danesh J, Donnelly S, Elosua R, Florez JC, Gabriel SB, Getz G, Glatt SJ, Hultman CM, Kathiresan S, Laakso M, McCarroll S, McCarthy MI, McGovern D, McPherson R, Neale BM, Palotie A, Purcell SM, Saleheen D, Scharf JM, Sklar P, Sullivan PF, Tuomilehto J, Tsuang MT, Watkins HC, Wilson JG, Daly MJ, MacArthur DG. Exome aggregation C. Analysis of protein-coding genetic variation in 60,706 humans. Nature 2016;536:285–91.

[25] Vogelstein B, Fearon ER, Hamilton SR, Kern SE, Preisinger AC, Leppert M, Smits AMM, Bos JL. Genetic alterations during colorectal-tumor development. N Engl J Med 1988;319:525–32.

[26] Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter's concept of field cancerization: evidence and clinical implications. Cancer Res 2003;63:1727–30.

[27] Slaughter DP, Southwick HW, Smejkal W. Field cancerization in oral stratified squamous epithelium; clinical implications of multicentric origin. Cancer 1953;6:963–8.

[28] Deng G, Lu Y, Zlotnikov G, Thor AD, Smith HS. Loss of heterozygosity in normal tissue adjacent to breast carcinomas. Science (New York, NY) 1996;274:2057–9.

[29] Ahlquist T, Lind GE, Costa VL, Meling GI, Vatn M, Hoff GS, Rognum TO, Skotheim RI, Thiis-Evensen E, Lothe RA. Gene methylation profiles of normal mucosa, and benign and malignant colorectal tumors identify early onset markers. Mol Cancer 2008;7:94.

[30] Belshaw NJ, Pal N, Tapp HS, Dainty JR, Lewis MP, Williams MR, Lund EK, Johnson IT. Patterns of DNA methylation in individual colonic crypts reveal aging and cancer-related field defects in the morphologically normal mucosa. Carcinogenesis 2010;31:1158–63.

[31] Konishi K, Shen L, Jelinek J, Watanabe Y, Ahmed S, Kaneko K, Kogo M, Takano T, Imawari M, Hamilton SR, Issa JP. Concordant DNA methylation in synchronous colorectal carcinomas. Cancer Prev Res 2009;2:814–22.

[32] Tomson BN, Arndt KM. The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states. Biochim Biophys Acta 2013;1829:116–26.

[33] Hanks S, Perdeaux ER, Seal S, Ruark E, Mahamdallie SS, Murray A, Ramsay E, Del Vecchio Duarte S, Zachariou A, de Souza B, Warren-Perry M, Elliott A, Davidson A, Price H, Stiller C, Pritchard-Jones K, Rahman N. Germline mutations in the PAF1 complex gene CTR9 predispose to Wilms tumour. Nat Commun 2014;5:4398.

[34] Martianov I, Cler E, Duluc I, Vicaire S, Philipps M, Freund JN, Davidson I. TAF4 inactivation reveals the 3 dimensional growth promoting activities of collagen 6A3. PLoS One 2014;9:e87365.