

LETTER TO THE EDITOR

Open Access



Comparative proteogenomic analysis of right-sided colon cancer, left-sided colon cancer and rectal cancer reveals distinct mutational profiles

Robin Imperial¹, Zaheer Ahmed¹, Omer M. Toor^{1,2}, Cihat Erdoğan³, Ateeq Khaliq², Paul Case², James Case⁴, Kevin Kennedy⁵, Lee S. Cummings⁶, Niklas Melton⁷, Shahzad Raza^{1,2}, Banu Diri⁸, Ramzi Mohammad⁹, Bassel El-Rayes¹⁰, Timothy Pluard^{1,2}, Arif Hussain^{11,12}, Janakiraman Subramanian^{1,2*} and Ashiq Masood^{13*}

Abstract

Right-sided colon cancer (RCC) has worse prognosis compared to left-sided colon cancer (LCC) and rectal cancer. The reason for this difference in outcomes is not well understood. We performed comparative somatic and proteomic analyses of RCC, LCC and rectal cancers to understand the unique molecular features of each tumor sub-types. Utilizing a novel in silico clonal evolution algorithm, we identified common tumor-initiating events involving APC, KRAS and TP53 genes in RCC, LCC and rectal cancers. However, the individual role-played by each event, their order in tumor development and selection of downstream somatic alterations were distinct in all three anatomical locations. Some similarities were noted between LCC and rectal cancer. Hotspot mutation analysis identified a nonsense mutation, APC R1450* specific to RCC. In addition, we discovered new significantly mutated genes at each tumor location, Further in silico proteomic analysis, developed by our group, found distinct central or hub proteins with unique interactomes among each location. Our study revealed significant differences between RCC, LCC and rectal cancers not only at somatic but also at proteomic level that may have therapeutic relevance in these highly complex and heterogeneous tumors.

Keywords: Right-sided colon cancer, Left-sided colon cancer, Rectal cancers, Clonal evolution, Proteomics, Hotspot mutations

Often grouped as one disease, right-sided colon cancer (RCC, originating from cecum, ascending colon, hepatic flexure) and left-sided colon cancer (LCC, originating from splenic flexure, descending colon, sigmoid colon) represent clinically distinct entities with significant differences in their prognosis and treatment outcomes [1, 2]. Therefore, given their anatomic continuity, the reason for these clinical differences presumably lie at the molecular level delineated by embryological origin. Previous studies have sought to identify these differences by

analyzing significantly mutated genes and RNA expression [3, 4]. However, molecular differences including significant specific amino acid alterations (hot spots), proteomic differences and order of mutations in clonal evolution of these tumors have not been studied. We used somatic and proteomic data of colorectal cancers from The Cancer Genome Atlas (TCGA) [4, 5], Memorial Sloan Kettering Cancer Center (MSKCC) [6] and The Cancer Proteome Atlas (TCPA) [7] to study proteogenomic differences in these tumors (See Additional files 1 and 2).

Results and discussion

Clonal evolution trajectories

Understanding the mutational timing and evolutionary trajectory of tumors is key to investigate the molecular

* Correspondence: jsubramanian@saint-lukes.org; ashiq_masood@rush.edu

¹Department of Medicine, University of Missouri-Kansas City School of Medicine, Kansas City, MO 64108, USA

¹³Division of Hematology/Oncology and Cell Therapy, Rush University Medical Center, Chicago, IL 60612, USA

Full list of author information is available at the end of the article



(See figure on previous page.)

Fig. 1 shows ensemble-level clonal evolution trajectories in colorectal cancer using CAPRI algorithm. The events of the model are connected by dashed lines where red dotted lines denote hard and orange denotes soft exclusivity. Algorithm uses both Bayesian information criterion 'BIC' and Akaike information criterion 'AIC' as a regularization. Non-parametric bootstrap scores (NPB) are shown in the figure with hypergeometric test p -value cutoff of < 0.05 . Other relations including temporal priority, probability raising are shown in Fig. 1a, b, and c and reported data in Additional file 4. 1a) clonal evolution in RCC, 1b) clonal evolution in LCC and 1c) clonal evolution in rectal cancers

underpinnings of cancer development and progression. Thus, we applied the PiCnIc (Pipeline for Cancer Inference) algorithm to our data to study ensemble-level cancer progression models and predict the evolutionary mutational trajectories between RCC, LCC and rectal cancers in the TCGA cohort (see Additional file 3). All three cancer locations had mutations in APC, TP53 and KRAS, possibly reflecting common initiating somatic events (Fig. 1). However, there were differences in the hierarchical groupings of mutations that surrounded these events.

In RCC ($n = 135$; Fig. 1a; Additional file 4: Table S1), APC somatic mutations and TP53 somatic mutations were independent events. APC somatic mutations 'selected' for KRAS mutations or amplifications. APC somatic mutations also 'selected' for SMAD4 somatic mutations or deletions, BRAF mutations and amplification. KRAS and BRAF showed mutual exclusivity. Interestingly, alterations in FBWX7, TCF7L2, and SMAD2 clustered in RCC tumors harboring APC and PIK3CA mutations. With respect to TP53, alterations in this location were associated with CTNNB1, MYC or/and BRCA2 mutations.

In LCC ($n = 143$; Fig. 1b; Additional file 4: Table S2), KRAS somatic mutations 'selected' for BRCA2 amplification, PTEN deletions or somatic mutations, PIK3CA somatic mutations, IGF2 amplification or somatic mutations and ERBB2 amplification or somatic mutations. Unlike RCC, alterations in PIK3CA were a late event in LCC and IGF2 amplification via CTNNB1. APC seemed to 'select' for TP53, but this did not reach statistical significance ($p = 0.06$). Similarly, APC somatic mutations 'selected' for BRCA2 mutations and TCF7L2 somatic mutations or deletions, but this association also did not reach statistical significance ($p = 0.3$ and $p = 0.2$, respectively).

In rectal cancers ($n = 76$; Fig. 1c; Additional file 4: Table S3), key initial mutations are split between TP53 and KRAS. TP53 'selects' for MYC amplification, SMAD4 deletion and BRCA2 somatic mutation or amplification. KRAS 'selects' for PTEN deletion or somatic mutations, PIK3CA somatic mutations, IGF2 amplification and ERBB2 amplification or somatic mutations. Among rectal cancer patients with AURKA mutations there is clustering of NRAS amplifications.

Our model shows significant differences in the mutational profiles of genes between RCC and LCC; the early

common somatic gene mutations are associated with the 'selection' of different subsequent genomic events in RCC compared to LCC. Our results suggest that although LCC and rectal cancers have some similarities in the tumor progression model wherein KRAS 'selected' for several genes in common (such as PIK3CA, IGF2, and ERBB2 alterations), significant differences were also noted between these two sites. Taken altogether, our results show non-adherence to the established Vogelstein linear progression model of colorectal cancer progression from normal mucosa to adenoma to carcinoma [8]. Further, our data suggest that RCC, LCC and rectal cancers have distinct mutational behavior in the context of their evolutionary trajectories, mutational timing during cancer development and progression. However, initial events such as mutation in the gatekeeper gene, APC, appear to be similar in colorectal cancers irrespective of location.

Mutation hotspot analysis

We studied somatic mutations at the residue sites that can disrupt functional protein domains leading to tumorigenesis and clonal evolution via selective pressure (see Methods in Additional file 3). We found APC R1450* to be a significant mutation specifically enriched in RCC (12–15%) compared to LCC (1%) and rectal tumors (1%) in both the TCGA and MSKCC datasets (all $p < 0.001$, Fig. 2a). To our knowledge, this is the first report to describe the APC R1450* mutation as being predominantly located in RCC. This particular hotspot in APC is exclusively a truncation mutation and lies within the MCR domain (residues 1282–1581; [9]) of the protein, which is a highly mutated area. The resulting truncated mutant conserves beta-catenin binding sites (15 AA repeats) but loses all three axin-binding sites (SAMP repeats) and microtubule interaction via EB1 and PdZ domains. Unlike APC R1450*, the frequency of other mutations within this region is relatively similar among the TCGA and MSKCC data sets. The relative frequencies of non-R1450* mutations within the MCR domain of APC for RCC were 63 and 64% in the TCGA and MSKCC data sets, respectively, for LCC 52 and 51%, respectively, and for rectal cancers 64% vs 58% (which did not meet statistical significance, $p = 0.35$). APC R1450* mutations are mutually exclusive from β -catenin destruction complex genes suggesting that they may be

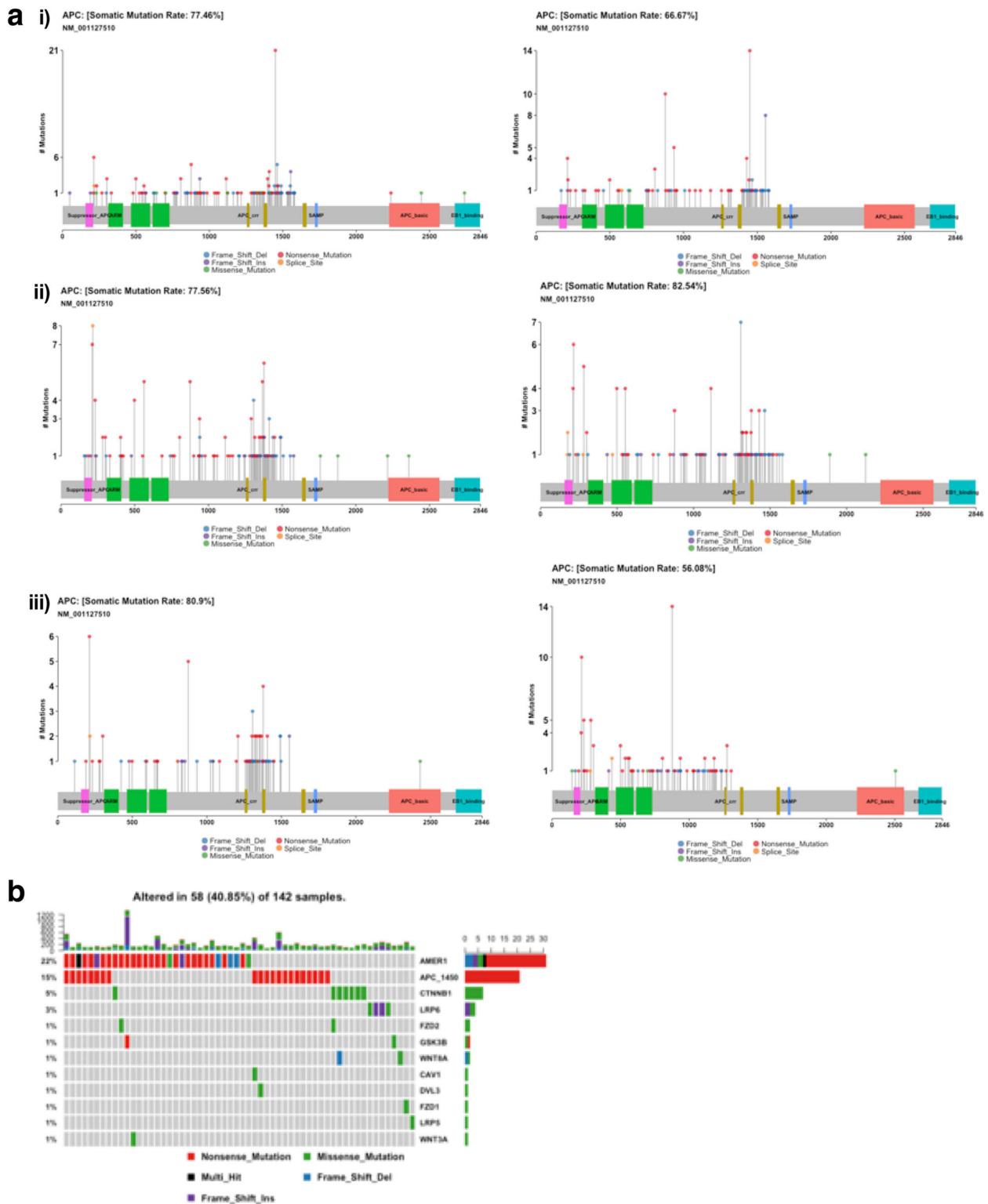


Fig. 2 **a** shows the frequency of APC hotspot the R1450 residue in (i) right-sided colon cancers, (ii) left-sided colon cancers and (iii) rectal cancers in TCGA (left) and MSKCC (right) datasets. Y-axis represent total number of mutations at each residue. **b** shows the mutual exclusivity of APC R1450* (APC_1450) compared to other genes of β -Catenin destruction complex in RCC. "APC_MCR" represents other APC mutations within the MCR region that are not at the 1450 residue. The bar plot above the oncoplot represents total mutations in each sample

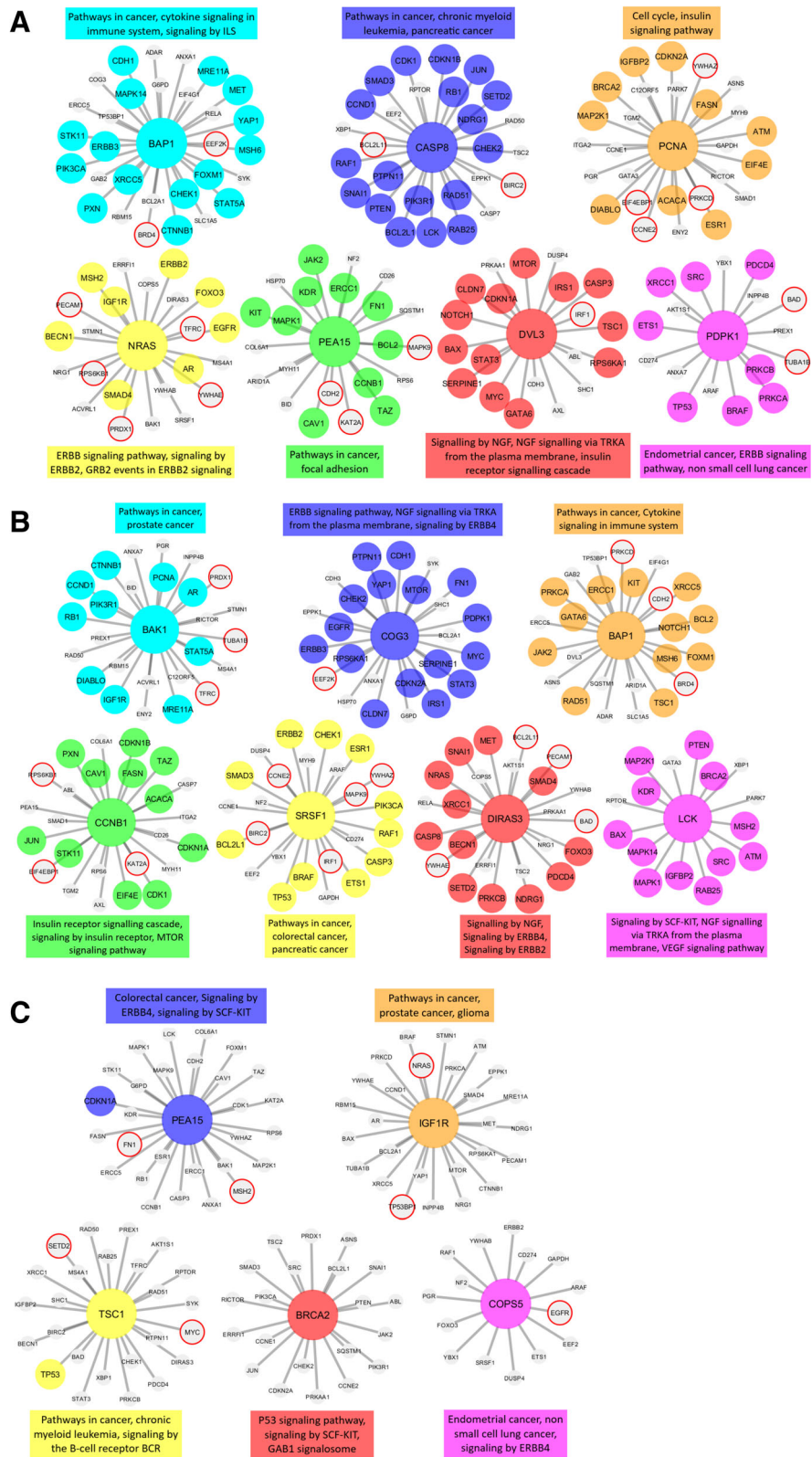


Fig. 3 (See legend on next page.)

(See figure on previous page.)

Fig. 3 shows the hub genes and neighbors in the disease-related sub-networks obtained by the most successful KDE method (in terms of precision score) in **a** RCC, **b** LCC and **c** rectal cancers. The genes registered in DisGeNET and experimentally confirmed for the diseases are shown with colored and larger nodes. Among these, genes that are not colored but have a red frame have a PMID value of one (e.g. have one supporting publication). There is no entry in DisGeNET for the grey colored nodes. Also, the most associated top three biological pathways, to which each module is related, are given above or below the relevant module to annotate each module

early events in right-sided colon cancer tumorigenesis (Fig. 2b). Given the recent findings by Zhang et al. demonstrating the efficacy of TASIN-1 (small molecule inhibitor) in a murine xenograft model of human colorectal cancer harboring a truncation mutation (A1309*) similar to APC R1450* suggests that this mutation may be viable therapeutic target [10]. In addition, we performed significantly mutated gene analysis and discovered newer driver genes at each location (see Additional file 5: Figure S1).

Proteogenomic analysis

Using The Cancer Proteome Atlas (TCPA) data, we examined RCC, LCC and rectal cancers by proteomic cancer co-expression subnetworks using association estimators methodology previously described by our group (see Methods in Additional file 3). Interestingly, no common protein emerged as having a centralized role (hub protein) across all 3 cancer locations. Within protein-protein interaction networks, several hub proteins, and their respective interactomes, were found to be unique to each of the locations (Fig. 3, see Additional file 6: Table S4).

Several hub proteins that might have a major role in RCC were identified: BAP1 (tumor suppressor gene) CASP8 (apoptosis) PCNA (DNA repair) NRAS (RTK-RAS pathway) PEA15 (apoptosis and RET signaling) DVL3 (cell proliferation and ATM-dependent DNA damage response) and PDPK1 (growth regulation) (Fig. 3a). The potentially significant hub proteins in LCC were: BAP1, BAK1 (apoptosis and prognostic in breast cancer) COG3 (protein glycosylation/golgi function) CCNB1 (mitosis and prognosis in breast cancer) SRSF1 (RNA splicing and prognosis in small cell lung cancer) DIRAS3 (tumor suppressor gene) and LCK (resistance to apoptosis) (Fig. 3b). Hub proteins unique to rectal cancers were: IGF1R (proliferation, invasion, migration), TSC1 (cell growth) BRCA2 (DNA repair) and COPS5 (multiple pathways) (Fig. 3c).

BAP1 was found to have a prominent role in both RCC and LCC. Although there are several conserved interactions, the BAP1 interactome of LCC diverges from that of RCC. Among the conserved interacting proteins are: BRD4, ADAR, GAB2, SLC1A5, EIF4G1, ERCC5 and TP53BP1, BRD4, ADAR, MSH6, FOXM1 and XRCC5. Specific to LCC, BAP1 showed interactions with ERCC1,

PRKCA, GATA6, JAK2, RAD51, TSC1, RSC1, NOTCH1, BCL2, KIT, PRKCD, CDH2, ARID1A, ASNS, SQSTM1 and DVL3. Specific to RCC, BAP1 was noted to interact with CDH1, MAPK14, MRE11A, MET, YAP1, STK11, ERBB3, PIK3CA, PXN, CHEK1, CTNBNB1, STAT5A, EEF2K, G6PD, COG3, RBM15, BCL2A1, SYK, RELA and ANXA1.

Our results suggest BAP1 may have an essential role in carcinogenesis of colon cancer with conserved as well as divergent evolutionary interactions with other proteins in RCC and LCC that are largely absent in rectal cancers.

A somewhat surprising observation from this analysis is that the protein hubs and their interactomes are distinct for each of the anatomically defined tumor sites examined. Further, these protein signatures are not necessarily concordant with the somatic tumor profiles. Identifying alterations in tumor DNA and RNA have been of paramount importance. Clarifying post-transcriptional events and protein-protein interactions will also be highly relevant to understanding the variations in tumor biology and clinical behavior of these tumors. Prospective studies are needed to validate our findings and their implications in the clinical outcomes.

Additional files

Additional file 1: Patients demographics from TCGA. (DOCX 19 kb)

Additional file 2: Inclusion and exclusion criteria for somatic mutation analysis. MSI-H, POLE mutation samples, rectosigmoid and transverse colon cancers were excluded for analysis (highlighted green). (TIF 666 kb)

Additional file 3: Methods section. (DOCX 52.9 kb)

Additional file 4: Tables S1-S3. PicNlc statistics (bic and aic) for RCC, LCC and rectal cancers. (XLSX 43 kb)

Additional file 5: Somatic mutation analysis for RCC, LCC and rectal cancers. (DOCX 433 kb)

Additional file 6: Table S4. Proteomics pathway and gene level analysis results. (XLSX 47 kb)

Abbreviations

AIC: Akaike information criterion; BIC: Bayesian information criterion; CAPRI: Cancer progression inference algorithm; KDE: Kernel density estimator; LCC: Left colon cancer; MSKCC: Memorial Sloan Kettering cancer center; PicNlc: Pipeline for cancer inference; PMID: Pubmed identifier; RCC: Right colon cancer; TCGA: The cancer genome atlas; TCPA: The cancer proteome atlas

Acknowledgements

Part of A.H.'s time was supported by a Merit Review Award (I01 BX000545), Medical Research Service, Department of Veterans Affairs.

Funding

Not applicable.

Availability of data and materials

TCGA somatic mutation data for colorectal cancers can be obtained from the GDC legacy archive and Broad GDAC Firehose. Proteomic data can be downloaded from The Cancer Proteome Atlas (TCPA). The datasets used and analyzed in the current study are also available from the corresponding author in response to reasonable requests.

Authors' contributions

RI and AM contributed to the design of study, involved in acquisition, analysis and interpretation of data; involved in drafting the manuscript and revising it critically for important intellectual content. ZA, OMT, CE, AK, PC, JC, KK, NM, BD involved in data acquisition, analysis and interpretation of data and involved in drafting the manuscript. LSC, SR, RM, BE, TP, AH, JS were involved in revising the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

Ashiq Masood Advisory board and speaker Bureau Bristol-Myers Squibb and Boehringer Ingelheim, Honorarium Biocept. Janakiraman Subramanian Advisory board - Astra Zeneca, Pfizer, Boehringer Ingelheim, Alexion, Paradigm, Bristol-Myers Squibb Speakers Bureau - Astra Zeneca, Boehringer Ingelheim, Lilly Research Support - Biocept and Paradigm. Arif Hussain Advisory board - Novartis, Bayer, Astra Zeneca, Consultant - Bristol-Myers-Squibb. All other authors declare that they have no competing interest.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Department of Medicine, University of Missouri-Kansas City School of Medicine, Kansas City, MO 64108, USA. ²Division of Oncology, Saint Luke's Cancer Institute, University of Missouri School of Medicine, 4321 Washington St, Kansas City, MO 64111, USA. ³Department of Computer Engineering, Namik Kemal University, Tekirdag, Turkey. ⁴ASPIRE Foundation, Saint Luke's Health System of Kansas City, Kansas City, MO 64111, USA. ⁵Division of Cardiovascular Research, Saint Luke's Hospital, Kansas City, MO 64111, USA. ⁶Department of Surgery, University of Missouri-Kansas City, Kansas City, MO 64108, USA. ⁷Department of Computer Sciences, Missouri University of Science and Technology, Rolla, MO 65409, USA. ⁸Department of Computer Engineering, Yildiz Technical University, Istanbul, Turkey. ⁹Wayne State University, Karmanos Cancer Institute, Detroit, MI 48201, USA. ¹⁰Department of Hematology and Medical Oncology, Winship Cancer Institute, Emory University, Atlanta, GA 30322, USA. ¹¹Division of Oncology, University of Maryland Greenebaum Comprehensive Cancer Center, Baltimore, MD 20201, USA. ¹²Baltimore Veterans Affairs Medical Center, Baltimore, MD 21201, USA. ¹³Division of Hematology/Oncology and Cell Therapy, Rush University Medical Center, Chicago, IL 60612, USA.

Received: 28 August 2018 Accepted: 30 November 2018

Published online: 21 December 2018

References

- Venook AP, Niedzwiecki D, Lenz H-J, et al. CALGB/SWOG 80405: phase III trial of irinotecan/5-FU/leucovorin (FOLFIRI) or oxaliplatin/5-FU/leucovorin (mFOLFOX6) with bevacizumab (BV) or cetuximab (CET) for patients (pts) with KRAS wild-type (wt) untreated metastatic adenocarcinoma of the colon or rectum (MCR). *J Clin Oncol*. 2014;32:LBA3. https://doi.org/10.1200/jco.2014.32.18_suppl.lba3.
- Hong TS, Clark JW, Haigis KM. Cancers of the Colon and Rectum: identical or fraternal twins? *Cancer Discov*. 2012;2:117–21. <https://doi.org/10.1158/2159-8290.CD-11-0315>.
- Hu W, Yang Y, Li X, et al. Multi-omics approach reveals distinct differences in left- and right-sided Colon Cancer. *Mol Cancer Res*. 2018;16:476–85. <https://doi.org/10.1158/1541-7786.MCR-17-0483>.
- Muzny DM, Bainbridge MN, Chang K, et al. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487:330–7. <https://doi.org/10.1038/nature11252>.
- Grossman RL, Heath AP, Ferretti V, et al. Toward a shared vision for Cancer genomic data. *N Engl J Med*. 2016;375:1109–12. <https://doi.org/10.1056/NEJMp1607591>.
- Yaeger R, Chatila WK, Lipsyc MD, et al. Clinical sequencing defines the genomic landscape of metastatic colorectal cancer. *Cancer Cell*. 2018;33:125–136.e3. <https://doi.org/10.1016/j.ccell.2017.12.004>.
- Li J, Lu Y, Akbani R, et al. TCPA: a resource for cancer functional proteomics data. *Nat Methods*. 2013;10:1046–7. <https://doi.org/10.1038/nmeth.2650>.
- Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. *Science* (80-). 2013;339:1546–58. <https://doi.org/10.1126/science.1235122>.
- Miyoshi Y, Nagase H, Ando H, et al. Somatic mutations of the APC gene in colorectal tumors: mutation cluster region in the APC gene. *Hum Mol Genet*. 1992;1:229–33.
- Zhang L, Theodoropoulos PC, Eskiocak U, et al. Selective targeting of mutant adenomatous polyposis coli (APC) in colorectal cancer. *Sci Transl Med*. 2016;8:361ra140. <https://doi.org/10.1126/scitranslmed.aaf8127>.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

