

RESEARCH ARTICLE

Open Access



# Epigenetic loss of heterozygosity of *Apc* and an inflammation-associated mutational signature detected in *Lrig1*<sup>+/-</sup>-driven murine colonic adenomas

Jessica L. Preston<sup>\*</sup>  and Nicholas Stiffler

## Abstract

**Background:** The loss of a single copy of adenomatous polyposis coli (*Apc*) in leucine-rich repeats and immunoglobulin-like domains 1 (*Lrig1*)-expressing colonic progenitor cells induces rapid growth of adenomas in mice with high penetrance and multiplicity. The tumors lack functional APC, and a genetic loss of heterozygosity of *Apc* was previously observed.

**Methods:** To identify genomic features of early tumorigenesis, and to profile intertumoral genetic heterogeneity, tumor exome DNA ( $n = 9$  tumors) and mRNA ( $n = 5$  tumors) sequences were compared with matched nontumoral colon tissue. Putative somatic mutations were called after stringent variant filtering. Somatic signatures of mutational processes were determined and splicing patterns were observed.

**Results:** The adenomas were found to be genetically heterogeneous and unexpectedly hypermutated, displaying a strong bias toward G:C > A:T mutations. A genetic loss of heterozygosity of *Apc* was not observed, however, an epigenetic loss of heterozygosity was apparent in the tumor transcriptomes. Complex splicing patterns characterized by a loss of intron retention were observed uniformly across tumors.

**Conclusion:** This study demonstrates that early tumors originating from intestinal stem cells with reduced *Lrig1* and *Apc* expression are highly mutated and genetically heterogeneous, with an inflammation-associated mutational signature and complex splicing patterns that are uniform across tumors.

**Keywords:** *Lrig1*, *Lgr5*, Colorectal cancer, Intestinal stem cells, Adenoma, Mutations

## Background

Human colorectal cancer (CRC) is the second leading cause of cancer death in the US, and ~25% of patients with CRC are incurable at the time of diagnosis [1, 2]. Genetic heterogeneity is inherent to this disease, providing tumor cells with the ability to rapidly adapt and resist treatments [3, 4]. Currently CRC is clinically segregated into four broad subcategories based on the expression of various biomarker molecules called consensus molecular subtypes (CMS1–4) [5]. Generally, CMS1 tumors display microsatellite instability and immune activation, CMS2 tumors are epithelial and display WNT pathway activation, CMS3 tumors are KRAS-

driven, and CMS4 tumors are mesenchymal with VEGF activation. However, this classification system oversimplifies the diversity and interrelatedness of cancer cell subtypes. The specific steps required to counteract the critical aspects of CRC progression remain poorly understood despite decades of research.

The cell-of-origin of CRC derives from a population of rapidly-dividing stem cells located at the base of the colonic epithelial crypts, which are identifiable based on the expression of Leucine-rich repeat containing G protein-coupled receptor 5 (*Lgr5*). *Lgr5* is the downstream target of *R-spondin* in the canonical Wnt/ $\beta$ -catenin pathway. Mutations in the tumor-suppressor gene adenomatous polyposis coli (*Apc*) and other members of the canonical Wnt pathway are the hallmark of CRC [6–

\* Correspondence: [jpreston@molbio.uoregon.edu](mailto:jpreston@molbio.uoregon.edu)  
Institute of Molecular Biology, University of Oregon, Eugene, OR 97403, USA



9]. Loss of heterozygosity (LOH) of *Apc* tends to occur during the early stages of human CRC tumorigenesis.

The standard mouse model used for CRC research, *Apc*<sup>Min/+</sup>, contains a truncating point mutation in one copy of *Apc*. It is believed that most *Apc*<sup>Min/+</sup> tumors have lost *Apc* function through a spontaneous genetic LOH [10]. *Apc*<sup>Min/+</sup> mice exhibit tumor formation predominantly in the small intestine rather than in the distal colon as observed in humans. Interestingly, Tanaka, T., et al. 2006 reported that dextran sodium sulfate (DSS)-induced inflammation increased the incidence of polyps in the distal colon of *Apc*<sup>Min/+</sup> mice [11]. Similarly, Yang, K., et al. 2008 found that reduced mucus production in *Apc*<sup>Min/+</sup> mice shifted tumor development toward the distal colon. In 2009 Ritchie, K., et al. crossed a glutathione S-transferase Pi (*Gstp*) null allele into the standard *Apc*<sup>Min/+</sup> mouse and reported a 6-fold increase in distal colorectal adenoma incidence and a 50-fold increase in adenoma multiplicity compared to *Apc*<sup>Min</sup> mice [12]. The authors also noted that the colons of the (*Gstp*) null *Apc*<sup>Min/+</sup> mice expressed higher levels of inflammatory molecules interleukin 4 (*IL4*), interleukin 6 (*IL6*), and nitric oxide synthase. Taken together, these results indicate an important role for mucus in reducing inflammation-associated tumors in the distal colon [13].

Leucine-rich repeats and immunoglobulin-like domains 1 (*Lrig1*) is a transmembrane feedback regulator of growth factor receptor tyrosine kinases that is expressed in the *Lgr5*<sup>+</sup> stem cell population present at the base of colonic crypts [14–17]. *Lrig1* acts as a tumor-suppressor gene in several contexts [18–23]. The *Lrig1-CreERT2/+;Apcfl/+* inducible mouse model of colonic adenoma is based on the conditional Cre-recombinase-driven loss of a single copy of *Apc* under the control of the *Lrig1* promoter [24, 25]. The colonic stem cells of these mice express one single wild type copy each of the *Lrig1* and *Apc* genes after the engineered recombination of *Apc* in *Lrig1*<sup>+/-</sup>-expressing stem cells. Within 100 days of the loss of one copy of *Apc*, rapidly growing adenomas appear in the distal colon with extremely high tumor penetrance and multiplicity. The *Lrig1-CreERT2/+;Apcfl/+* CRC model is very similar to the *Gstp*-null;*Apc*<sup>Min</sup> CRC model in terms of tumor onset, penetrance, multiplicity, anatomical location, and mortality. These findings imply that a common mechanism involving inflammation-induced tumor formation is responsible for tumorigenesis in the distal colon.

Human tumors often exhibit distinct patterns of mutations that can provide clues into the origin and mechanism of tumorigenesis. The ‘somatic signature of mutations’ of a tumor is based on the specific nucleotide alterations present and the background sequence context of the mutations [26–28]. Somatic signatures are influenced by specific carcinogenic agents and DNA repair

genes, and can therefore sometimes reveal sources of mutation and mechanisms of tumorigenesis. For example, C>A transversion point mutations typically occur in low frequencies compared to C>T transition point mutations. However, lung and esophageal tumors caused by tobacco tar often contain an abundance of C>A transversions due to the conjugation of nitrosamines to glutathione, which forms guanine adducts [29]. Stomach cancers caused by *H. pylori* infection also contain a high incidence of C>A transversions, presumably due to the inflammation-associated reactive oxygen and nitrogen species (ROS and RNS) [30].

This work sought to understand the genomic changes occurring during the early stages of tumorigenesis in rapidly-growing colonic adenomas in *Lrig1-CreERT2/+;Apcfl/+* mice. Exomic DNA and mRNA sequences from the adenomas were analyzed in order to detect the presence of transcriptomic and genomic alterations. Specifically, the genetic heterogeneity across tumors and somatic signatures of mouse colonic adenomas tumors were assessed using exome DNA profiling, and the prevalence of differential gene expression and splicing defects was assessed using mRNA-Seq.

## Materials and methods

### Generation of *Lrig1*-driven colonic adenomas

Generation of *Lrig1-CreERT2/+;Apcfl/+* tumors was previously described [25]. Briefly, *Lrig1-CreERT2/+* mice were crossed to *Apc580S/+* mice (Jackson Laboratory, Bar Harbor, ME, U.S.A) [31] to generate *Lrig1-CreERT2/+;Apcfl/+* mice [24]. Adult (6- to 8-week-old) *Lrig1-CreERT2/+;Apcfl/+* mice were intraperitoneally injected with 2 mg tamoxifen per mouse (Sigma-Aldrich, St. Louis, MO, U.S.A.) in corn oil for 3 consecutive days and multiple dysplastic colonic adenomas were extracted 100 days later. For the DNA studies, the control is nontumor tissue parts of the same *Lrig1-CreERT2/+;Apcfl/+* mouse, referred to as ‘nontumor’ in figures. For the RNA studies, the control is untreated wild type C57BL/6J mice, referred to as ‘wild type’ in figures.

### Exomic DNA sequencing

Exomic DNA sequencing ( $n = 9$  tumors) was performed by HudsonAlpha Labs (Huntsville, AL, U.S.A.) Exome capture using NimbleGen v3.0 captured 64 megabase-pair (Mbp) baits. A total of six gigabasepairs (Gbp) of exome data were sequenced per sample. Exomic tumor DNA of nine tumors from three mice was sequenced in parallel with adjacent normal colon ( $n = 3$ ) to 12x average read depth. Adjacent normal colon tissue was used for the ‘nontumor’ control.

### DNA read alignment and processing

Raw fastq reads were cleaned to remove low quality bases at the ends of the reads using Stacks (v 1.35) process\_shortreads [32]. Cleaned reads were aligned to the *Mus musculus* (house mouse) genome (2011 assembly, UCSC Genome Browser assembly ID mm10, Genome Reference Consortium Mouse Build 38, Accession GCA\_000001635.2) using Bowtie2 (v2.2.1), in default-sensitive mode [33]. Sam/bam files were sorted and indexed using SAMtools (v0.1.18) [34] and Picard Tools (v1.92) (Broad Institute, Cambridge, MA, U.S.A.). Base quality score recalibration and indel realignment were performed according to GATK best practices [35].

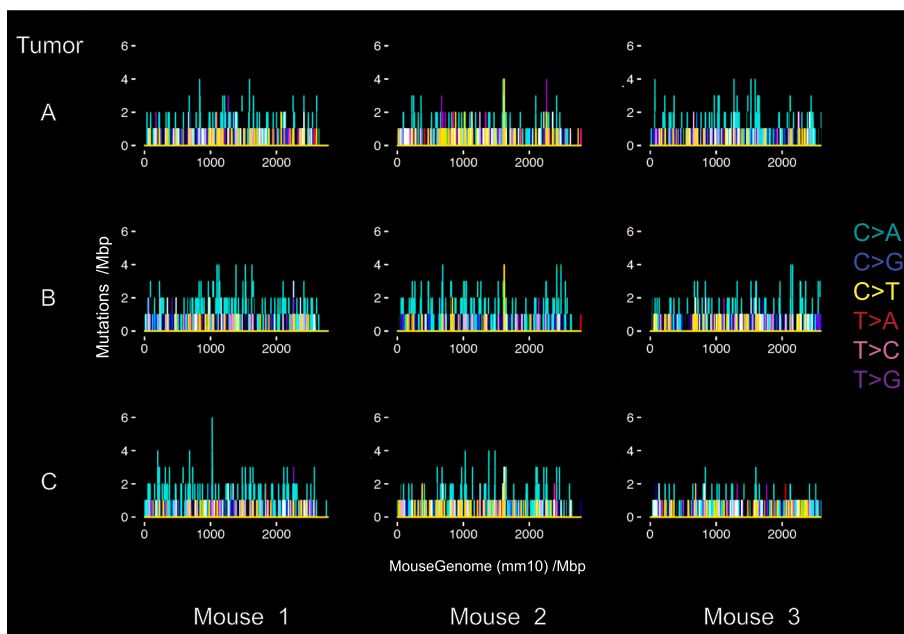
### Mutation calling, filtering, and visualization

Somatic mutations were called using SeuratSomatic (v2.5), a GATK module [36–38]. Somatic variants were called using a minimum variant coverage of 4 reads and minimum quality score of Q10. To ensure that the identified mutations were somatic in nature, matched normal tissue samples were used to filter germline variants. Residual germline polymorphisms were removed by filtering all germline variants reported on the National Center for Biotechnology Information (U.S.A.) dbSNP database [39]. To reduce the occurrence of sequencing errors in the data, variants were called only when present in both the forward and reverse DNA strand

[40]. To confirm that the discovered mutations were not due to alignment artifacts, the presence of representative variants was verified in the raw sequencing reads using grep and BLAST. The predicted effect of the mutations on protein function was determined with Ensembl's variant effect predictor (VEP) program [41] as well as the SNPeff variant annotation and effect prediction tool [42]. The somatic signature of the mutations was identified using R package SomaticSignatures [38]. Genomic data was visualized using R packages GenVisR and ggplot2 [43, 44] Negative control nontumor samples and mouse glioblastoma whole-genome DNA sequencing datasets ( $n = 3$ , provided by Hui Zong Lab, University of Virginia Medical School) were compared to the adenoma dataset; hypermutation or a clear somatic signature were not observed in the negative controls.

### RNA sequencing and analysis

Whole tumors ( $n = 5$ ) as well as normal wild type colonic tissue from untreated C57BL/6 J mice ( $n = 3$ ), were assayed. RNA was sequenced to 25x average read depth. Cleaned cDNA reads were aligned to the mouse mm10 genome (Accession GCA\_000001635.2) with STAR (v1.0) [45]. Changes in gene expression were identified using R package DESeq2 and visualized with R packages DESeq2, ggplot2, and Pathview. A false discovery rate (FDR) < 0.1 was used as the significance threshold.



**Fig. 1** Colonic adenomas resulting from inducible loss of adenomatous polyposis coli (*Apc*) in leucine-rich repeats and immunoglobulin-like domains 1 (*Lrig1*)<sup>+/-</sup> stem cells are hypermutated and genetically heterogeneous. Exome sequencing of adenomas from *Lrig1*-(*Apc*-Floxed) mice detected the presence of ~25 to 35 high-quality somatic mutations per megabase, with an abundance of C:G>A:T transversion point mutations. The number and type of somatic mutations from each adenoma exome is plotted. There were nine adenomas from three mice; each column plots three tumors from a single mouse. Similar mutational frequencies and motif distributions were observed across the nine tumor exomes sequenced

Custom bigwig tracks were created on the University of California at Santa Cruz Genome Browser in order to assess splicing patterns and copy number variations [46].

## Results

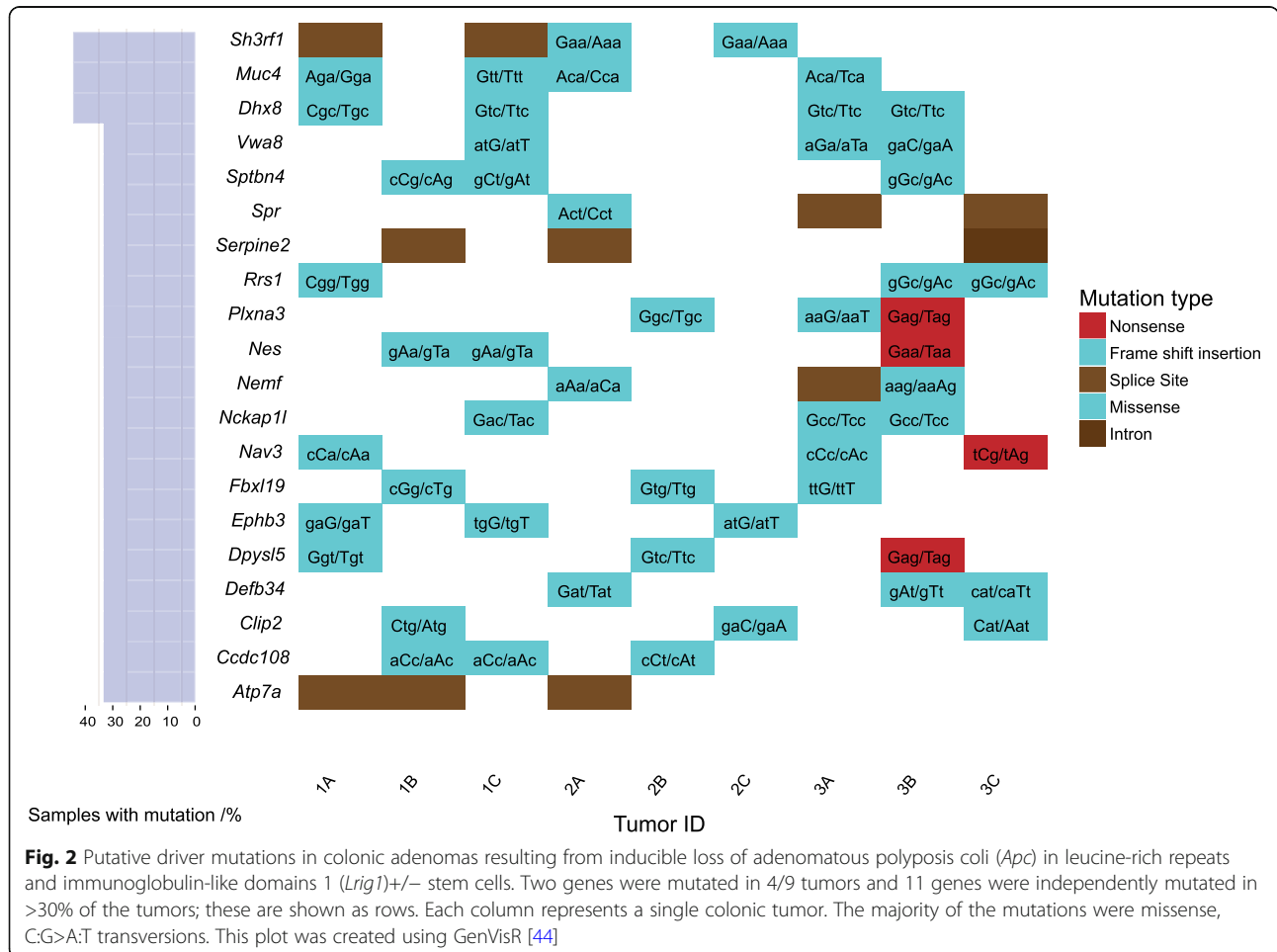
### Tumor exome mutations

Adenomas resulting from the inducible loss of one copy of *Apc* in *Lrig1<sup>+/-</sup>*-expressing colonic stem cells are hypermutated and genetically heterogeneous. Tumor exome DNA contained 930–1300 high-quality somatic mutations per tumor (~ 25–30 mutations per megabase) distributed uniformly throughout the genome. Variants were called only when present in both the forward and reverse DNA strand, with a minimum quality score of Q10 and a minimum variant coverage of 4 reads. The tumors are highly heterogeneous in terms of the specific genetic loci mutated, however nearly identical frequencies of mutated nucleotide motifs were observed across all nine tumors sequenced (Fig. 1). No large copy number variations or rearrangements were detected.

Each adenoma contained a unique profile of ~ 100 high-frequency and high-quality mutations predicted to

cause significant functional disruptions based on amino acid conservation (Table S1). Two genes were mutated in 4/9 tumors: mucin 4 (*Muc4*) and DEAH-Box Helicase 8 (*Dhx8*). *Dhx8* is involved in alternative splicing. Eleven genes were independently mutated in >30% of the tumors (Fig. 2, Table S2). The 40 most commonly mutated genes (>20%) are predicted to affect cell morphology and migration based on pathway analysis, and were found to be mutated in 10–20% of human CRC cases in The Cancer Genome Atlas (TCGA) database.

Mutated tumor-suppressor genes, such as phosphatidylinositol-4,5-bisphosphate 3-kinase catalytic subunit alpha (*Pi3kca*), Neurofibromin 1 (*Nf1*), and *Lrig1* were observed. The DNA repair genes *MutS* protein homolog 4 (*Msh4*), MutS protein homolog 6 (*Msh6*), and DNA polymerase delta 1 (*Pold1*) were also mutated. Unexpectedly, most tumors contained several additional mutations in genes of canonical and non-canonical *Wnt* pathways. The specific genes affected were different for each of the tumors.



### Tumor somatic signatures

The adenomas exhibited an overwhelming abundance of C to A transversion point mutations at remarkably similar frequencies. The somatic mutations displayed nearly identical patterns of background genomic sequence context across the nine tumors (Fig. 3). This type of mutational signature is often seen in smoking-induced lung cancer and stomach cancer induced by *Helicobacter pylori* infection, which are associated with an increase in the formation of guanine adducts [27, 28]. The high incidence of C to A transversions observed in the tumor exomes implies that the loss of one copy of *Apc* in *Lrig1*<sup>+/-</sup> colonic stem cells increased inflammation and led to the formation of DNA adducts due to the oxidation of guanine.

### Tumor transcriptomes

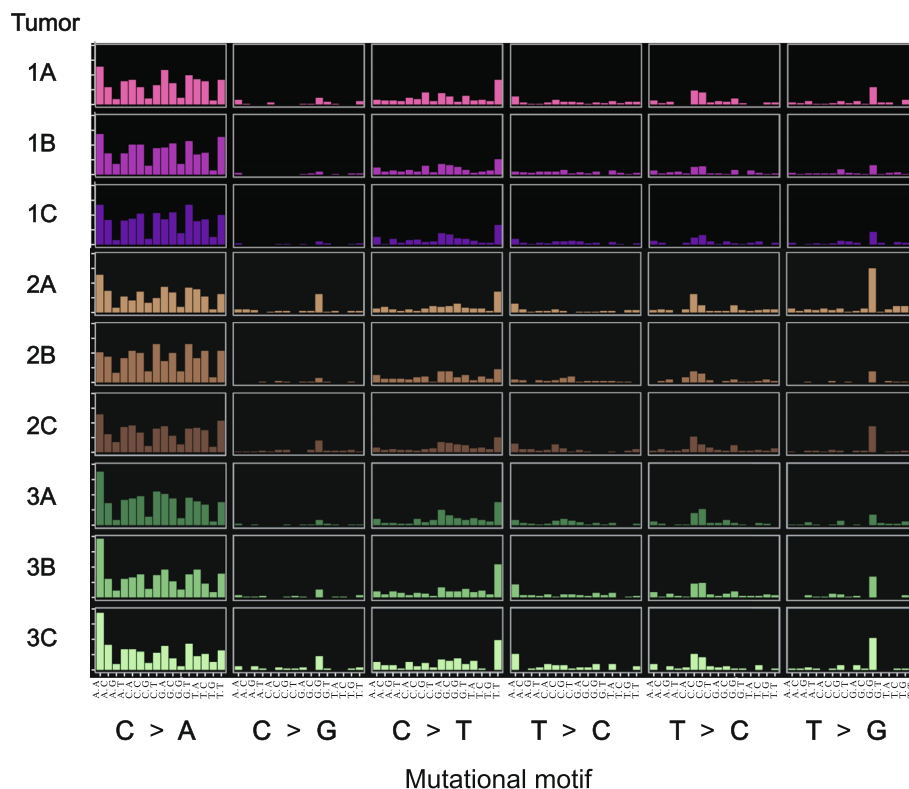
Despite their highly variable genomes, the tumors exhibited substantial transcriptome similarity to each other (Fig. 4). Down-regulation of the DNA repair genes *Msh3* (FDR = 0.57) and *Msh4* (FDR = 0.026) were observed (Table S3). In addition, eight UDP

glucuronosyltransferases were downregulated in the tumors (FDR < 0.0069).

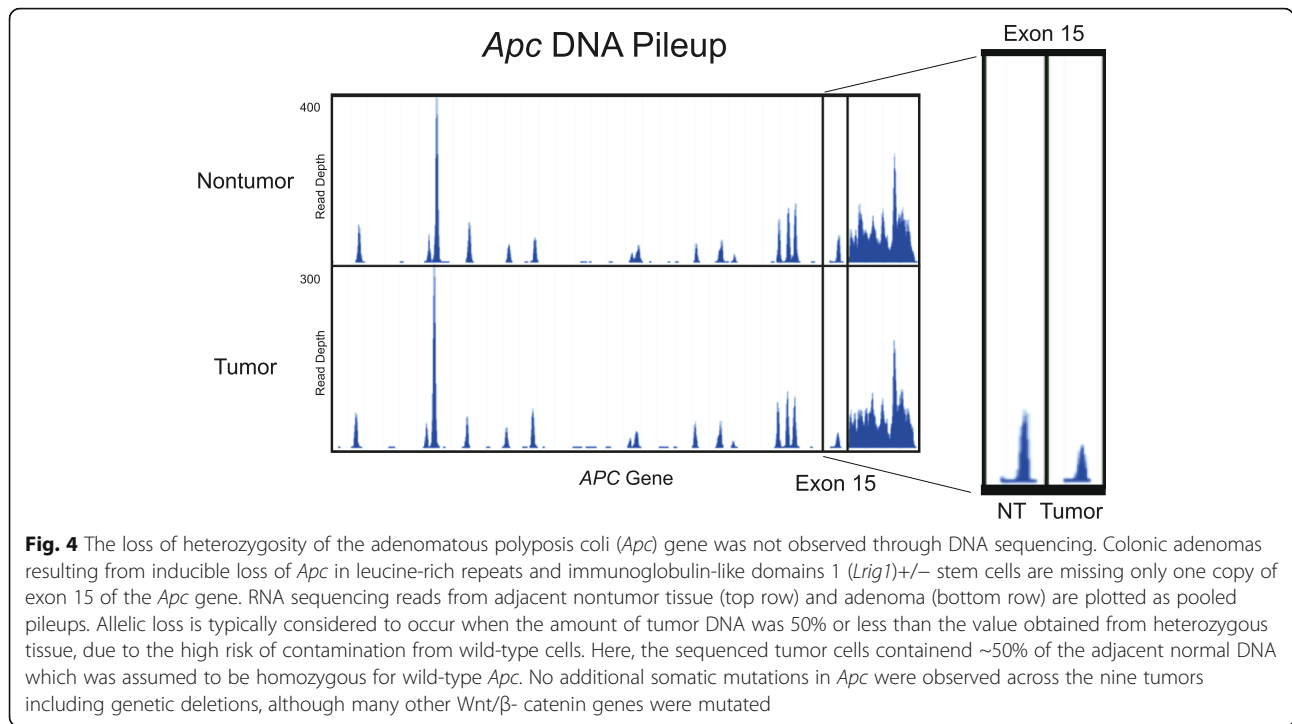
The tumors consistently displayed abnormal splicing patterns in several RNA-binding genes (Fig. 5). Compared to those in wild type mice, the tumor transcripts displayed a loss of intron retention in CRC-associated genes. Intron retention is a common feature of tumor genomes, but intron loss has been observed in breast cancer [47]. This observation suggests that modifications in post-transcriptional gene regulation are exploited by tumor cells in order to gain additional selective advantages against the host.

### Transcriptional silencing of wild type *Apc*

In contrast to previous reports, LOH of *Apc* was not observed in tumor exome DNA (Fig. 6). Allelic loss is typically considered to occur when the amount of tumor DNA is 50% or less of that obtained from heterozygous tissue due to the high risk of contamination from wild type cells. However, in this case, the tumor cells contained ~50% of the adjacent normal DNA, which was assumed to be homozygous for *Apc*

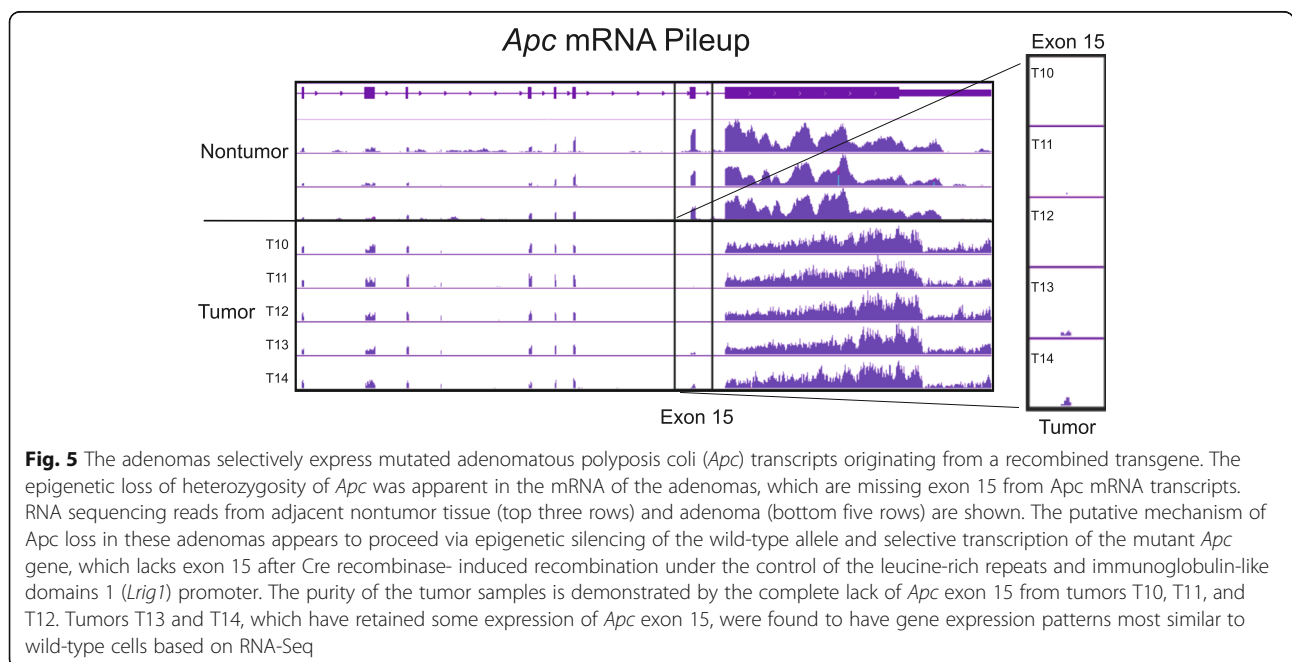


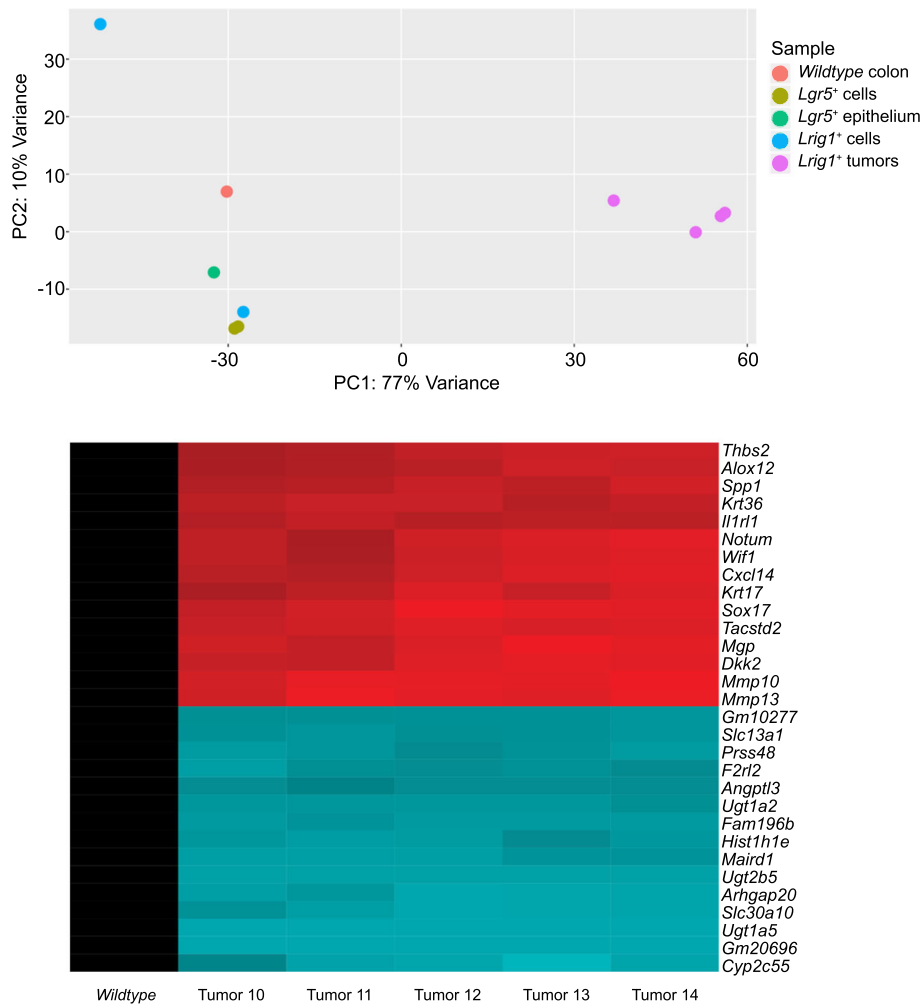
**Fig. 3** Colonic adenomas resulting from inducible loss of adenomatous polyposis coli (*Apc*) in leucine-rich repeats and immunoglobulin-like domains 1 (*Lrig1*)<sup>+/-</sup> stem cells have a distinct pattern of somatic mutations dominated by C:G>A:T transversion point mutations. The somatic signature of mutations of the adenomas are plotted as rows; there were nine adenomas from three mice. Each adenoma was found to contain a nearly identical pattern of somatic mutations based on motifs and background sequence context. The somatic signature is characterized by an abundance of C to A transversion point mutations, which are caused by defects in repairing DNA adducts of guanine and are frequently observed in cancer associated with tobacco tar and *Helicobacter pylori* infections



since tamoxifen-induced Cre recombination in the colonic epithelium of these mice is a very rare event. No spontaneous somatic mutations were identified in the *Apc* gene that were of meaningful quality or significance. No genetic copy number variations or insertions or deletions of *Apc*, or any other gene, were readily apparent in the data.

Unlike the exomic DNA, the tumor mRNA displayed a clear lack of functional *Apc* across the entire dataset ( $n = 5/5$  tumors). The adenomas expressed exclusively *Apc* transcripts lacking codon 580 in exon 15 (Fig. 7), which is the same sequence as the original *Apc* transgene after tamoxifen-induced recombination has occurred. These results imply that the tumors have





**Fig. 6** The tumors exhibited substantial transcriptome similarity to each other. Top: Principal components analysis of tumor transcriptome data compared to colons from untreated wild type mice. Tumor and wild type colon transcriptomes form distinct groups based on principal components analysis. Bottom: Heatmap visualization of highly significant differentially-expressed genes in the colonic adenomas. Rows are genes and columns are tumors. Genes colored red were up-regulated (top 15 rows) and genes colored blue were down-regulated (bottom 15 rows). Expression levels were normalized against that in the wild-type colon sample (black column on left). Many genes down-regulated are involved in detoxification of exogenous carcinogens

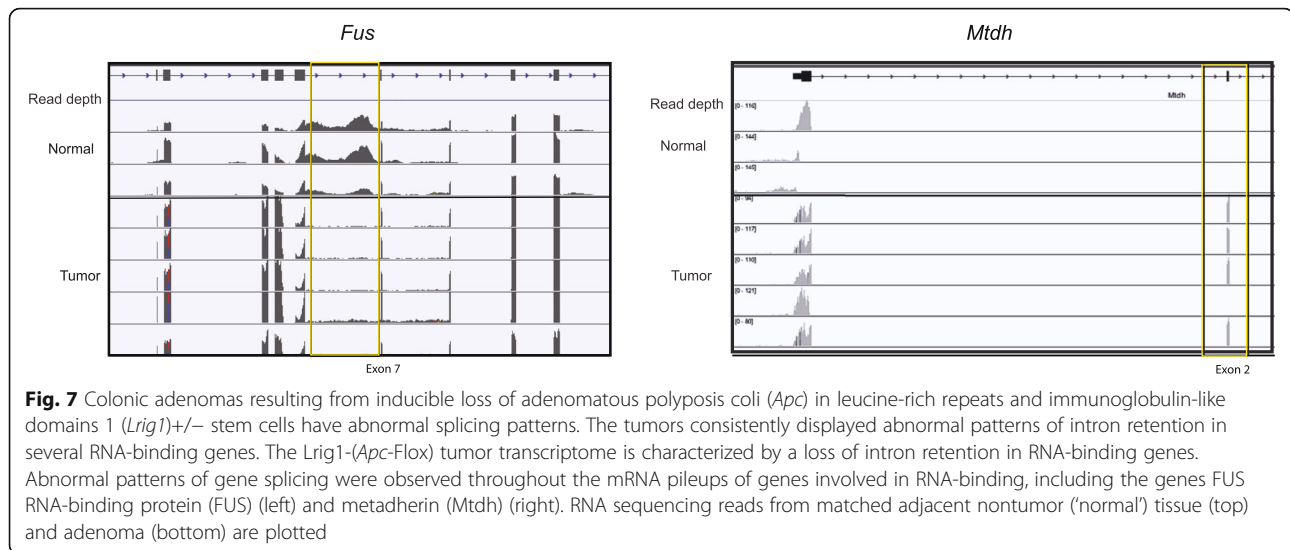
acquired the ability to selectively express nonfunctional *Apc* transcripts. This epigenetic phenomenon appears to have occurred across all tumors, implying that the LOH of *Apc* was an early event of adenoma formation.

**Discussion**

Evidence of early genomic changes were apparent throughout the exomes and transcriptomes of the tumors sequenced in this study. The *Lrig1-CreERT2/+;Apcfl/+* tumors displayed a hypermutated phenotype; each tumor exome contained a unique profile of ~100 mutated alleles predicted as ‘high impact’ based on amino acid sequences. The tumors contained a very high incidence of C:G > A:T transversion point mutations, which result from the oxidation of guanine to form

guanine adducts, and typically occur in smoking-induced lung cancer and stomach cancer caused by *H. pylori*.

The hypermutated phenotype of *Lrig1-CreERT2/+;Apcfl/+* tumors could have resulted from an increased rate of mutation or a reduced rate of DNA repair. The glucuronidation pathway was strongly downregulated in the tumor transcriptomes, and many inflammation markers were upregulated. It is unclear if these gene expression changes were the cause or result of the high incidence of mutations in the tumor. Microbial-produced reactive oxygen and nitrogen species (ROS and RNS) are a likely source of guanine adducts and C:G > A:T transversion point mutations in the colon. The high incidence of *Lrig1-CreERT2/+;Apcfl/+* adenomas localized to the distal colon might be explained by the high levels of ROS inherent to the distal colon and rectum.



Genetic heterogeneity was observed in tumors from the same mouse (intratumoral heterogeneity), as well as from sibling mice (intertumoral heterogeneity). Interestingly, in contrast to their variable genomes, the tumor transcriptomes and splicing patterns were strikingly uniform across all samples. The uniformity of the tumor transcriptomes suggests that the adaptation of an invasive phenotype resulted from convergent evolution. The lack of uniformity in the tumor genomes compared to the transcriptomes implies that expressed RNA and proteins may represent better drug targets than mutations, due to the fact that tumor transcriptomes are more uniform and predictable than tumor genomes.

The *Lrig1*-*CreERT2*<sup>+/+</sup>;*Apcfl*<sup>+/+</sup> mouse model of CRC is notable for its high penetrance and rapid onset after the loss of just a single copy of *Apc*. The extremely high tumor burden observed in *Lrig1*-*CreERT2*<sup>+/+</sup>;*Apcfl*<sup>+/+</sup> mice is especially striking considering that prior to the spontaneous loss of an *Apc* allele, tamoxifen-induced Cre-recombination of the other *Apc* allele, must first occur in the same cell under the control of the *Lrig1* promoter, which is also a rare event. Spontaneous LOH of *Apc* in humans and mice tends to occur much more slowly, generally in the order of several months to years. For example, human familial adenomatous polyposis intestinal tumors typically become malignant at around age 40 years.

In contrast to previous reports, a genetic loss of heterozygosity (LOH) of *Apc* was not observed in the *Lrig1*-*CreERT2*<sup>+/+</sup>;*Apcfl*<sup>+/+</sup> tumors. This fact is unsurprising when one considers the extremely low theoretical likelihood of multiple spontaneous somatic LOH events at a single genetic locus multiple times in one colon over the span of a few months. It is worth noting that in the original report on these tumors, a standard polymerase

chain reaction genotyping assay was used to detect the LOH of *Apc*, which is an archaic and much less sensitive technique for genotyping compared to DNA sequencing.

Despite the normal appearance of the remaining *Apc* alleles in the tumor exome DNA, there was clear evidence of an epigenetic ('above the genes') LOH of exon 15 of *Apc* observed in the tumor mRNA. Unlike the tumor DNA, the *Apc* mRNA transcripts expressed by the tumor cells clearly displayed a copy number loss of exon 15. The expressed *Apc* transcripts had a genetic sequence identical to that of the original *Apc* transgene following Cre-induced recombination, indicating that only the mutant transgene is being expressed by the tumor cells. The assumed mechanism responsible for the lack of wild type *Apc* expression is transcriptional silencing of the single wild type copy of *Apc* remaining after engineered recombination of the *Apc* transgene. It is possible that the double heterozygosity of *Lrig1* and *Apc* act synergistically with *Lrig1* haploinsufficiency to increase the possibility of transcriptional silencing of wild type *Apc*.

The proposed mechanism of colonic adenoma formation in the *Lrig1*-*CreERT2*<sup>+/+</sup>;*Apcfl*<sup>+/+</sup> model is the following: Transcriptional silencing of *Apc* leads to a decrease in goblet cells and a corresponding increase in inflammatory reactive oxygen species (ROS) in the distal colon. Somatic C > A mutations resulting from increased inflammation provide the raw genetic material required for adenoma formation and progression. Natural selection of growth-promoting mutations leads to convergent evolution towards a transcriptome marked by dysregulated splicing and altered gene expression resulting in an invasive phenotype.

The results of this study demonstrate that the genomic events leading to tumor growth and invasion can happen



early in tumorigenesis and that tumors can generate large pools of mutations early in the course of disease progression. The selection of advantageous mutations is a dynamic process, changing throughout the course of tumorigenesis. The ability of a tumor to create and store mutations in a reservoir for later use provides an advantage to the tumor by optimizing its ability to dynamically modulate the selection of oncogenic mutations in a context-dependent manner throughout disease progression. Mutational reservoir creation provides tumors with the ability to conditionally express somatic mutations which could ultimately lead to tumor metastasis and drug resistance.

In order to develop therapeutic treatments that are less susceptible to drug resistance, it is important to understand the mechanisms underlying tumor formation and evolution. Going forward, mutational processing mechanisms altered during early tumorigenesis make attractive therapeutic targets since they should theoretically be present in a large proportion of tumor cells and they are the drivers of tumor growth. Additional studies are needed in order to understand the mechanisms underlying the processes of mutation generation in tumors.

## Conclusions

Understanding the process of mutation generation during tumorigenesis can provide insights into the mechanisms underlying cancer cells' ability to evolve and grow. *Lrig1-CreERT2/+;Apcfl/+* tumors exhibited transcriptional silencing of *Apc* and contained a high incidence of C > A mutations caused by increased inflammation. In contrast to their variable genomes, the tumor transcripts and splicing patterns were strikingly uniform across all tumors, and likely represent more predictable cancer biomarkers.

## Supplementary information

**Supplementary information** accompanies this paper at <https://doi.org/10.1186/s12885-020-6616-y>.

**Additional file 1.** Somatic mutations identified in *Lrig1-Cre/+;Apcfl/+* colonic adenomas.

**Additional file 2.** Annotated RNA-Seq results for in *Lrig1-Cre/+;Apcfl/+* colonic adenomas vs. wildtype colon.

**Additional file 3.** Raw RNA-Seq results for in *Lrig1-Cre/+;Apcfl/+* colonic adenomas vs. wildtype colon.

## Abbreviations

CRC: Colorectal cancer; LOH: Loss of heterozygosity; ROS: Reactive oxygen species

## Acknowledgements

Bob Coffey provided the sequencing data used for this project. This work benefited from access to the University of Oregon's high performance computer, Talapas.

## Authors' contributions

JLP analyzed and interpreted the genomic datasets, generated the figures and wrote the paper. NS provided data analysis tools and bioinformatics support. All authors have read and approved this manuscript.

## Funding

Funding was provided by the University of Oregon for the study design, data collection, data analysis and interpretation, and in the writing of the manuscript.

## Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

## Ethics approval and consent to participate

All mouse experiments were approved by Institutional Animal Care and Use Committee (IACUC) at Vanderbilt University Medical Center. Animal studies were performed according to the recommendations in the Guide for the Care and Use of Laboratory Animals of the National Institutes of Health.

## Consent for publication

Not applicable.

## Competing interests

The authors declare that they have no competing interests.

Received: 23 September 2019 Accepted: 10 February 2020

Published online: 14 February 2020

## References

- Nitsche U, Stoss C, Stecher L, Wilhelm D, Friess H, Ceyhan GO. Meta-analysis of outcomes following resection of the primary tumour in patients presenting with metastatic colorectal cancer. *Br J Surg*. 2018;105(7):784–96.
- Advani S, Kopetz S. Ongoing and future directions in the management of metastatic colorectal cancer: update on clinical trials. *J Surg Oncol*. 2019; 119(5):642–52.
- Sasaki N, Clevers H. Studying cellular heterogeneity and drug sensitivity in colorectal cancer using organoid technology. *Curr Opin Genet Dev*. 2018;52: 117–22.
- Roerink SF, Sasaki N, Lee-Six H, Young MD, Alexandrov LB, Behjati S, Mitchell TJ, Grossmann S, Lightfoot H, Egan DA, et al. Intra-tumour diversification in colorectal cancer at the single-cell level. *Nature*. 2018;556(7702):457–62.
- Linnekamp JF, Hooff SRV, Prasetyanti PR, Kandimalla R, Buikhuisen JY, Fessler E, Ramesh P, Lee K, Bochove GGW, de Jong JH, et al. Consensus molecular subtypes of colorectal cancer are recapitulated in vitro and in vivo models. *Cell Death Differ*. 2018;25(3):616–33.
- Albuquerque C, Breukel C, van der Looij R, Fidalgo P, Lage P, Slors FJ, Leitao CN, Fodde R, Smits R. The 'just-right' signaling model: APC somatic mutations are selected based on a specific level of activation of the beta-catenin signaling cascade. *Hum Mol Genet*. 2002;11(13):1549–60.
- Smits R, Hofland N, Edelmann W, Geugien M, Jagmohan-Changur S, Albuquerque C, Breukel C, Kuchelapati R, Kielman MF, Fodde R. Somatic *Apc* mutations are selected upon their capacity to inactivate the beta-catenin downregulating activity. *Genes Chromosom Cancer*. 2000;29(3):229–39.
- Christie M, Jorissen RN, Mouradov D, Sakhianandeswari A, Li S, Day F, Tsui C, Lipton L, Desai J, Jones IT, et al. Different APC genotypes in proximal and distal sporadic colorectal cancers suggest distinct WNT/beta-catenin signaling thresholds for tumorigenesis. *Oncogene*. 2013;32(39):4675–82.
- Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell*. 1990;61(5):759–67.
- Luongo C, Moser AR, Gledhill S, Dove WF. Loss of *Apc+* in intestinal adenomas from min mice. *Cancer Res*. 1994;54(22):5947–52.
- Tanaka T, Kohno H, Suzuki R, Hata K, Sugie S, Niho N, Sakano K, Takahashi M, Wakabayashi K. Dextran sodium sulfate strongly promotes colorectal carcinogenesis in *Apc* (min/+) mice: inflammatory stimuli by dextran sodium sulfate results in development of multiple colonic neoplasms. *Int J Cancer*. 2006;118(1):25–34.
- Ritchie KJ, Walsh S, Sansom OJ, Henderson CJ, Wolf CR. Markedly enhanced colon tumorigenesis in *Apc* (min) mice lacking glutathione S-transferase pi. *Proc Natl Acad Sci U S A*. 2009;106(49):20859–64.

13. Yang K, Popova NV, Yang WC, Lozonschi I, Tadesse S, Kent S, Bancroft L, Matisse I, Cormier RT, Scherer SJ, et al. Interaction of Muc2 and Apc on Wnt signaling and in intestinal tumorigenesis: potential role of chronic inflammation. *Cancer Res.* 2008;68(18):7313–22.
14. Wong VW, Stange DE, Page ME, Buczacki S, Wabik A, Itami S, van de Wetering M, Poulsom R, Wright NA, Trotter MW, et al. Lrig1 controls intestinal stem-cell homeostasis by negative regulation of ErbB signalling. *Nat Cell Biol.* 2012;14(4):401–8.
15. Faraz M, Herdenberg C, Holmlund C, Henriksson R, Hedman H. A protein interaction network centered on leucine-rich repeats and immunoglobulin-like domains 1 (LRIG1) regulates growth factor receptors. *J Biol Chem.* 2018; 293(9):3421–35.
16. Neirinckx V, Hedman H, Nicloux SP. Harnessing LRIG1-mediated inhibition of receptor tyrosine kinases for cancer therapy. *Biochim Biophys Acta Rev Cancer.* 2017;1868(1):109–16.
17. Munoz J, Stange DE, Schepers AG, van de Wetering M, Koo BK, Itzkovitz S, Volckmann R, Kung KS, Koster J, Radulescu S, et al. The Lgr5 intestinal stem cell signature: robust expression of proposed quiescent ‘+4’ cell markers. *EMBO J.* 2012;31(14):3079–91.
18. Yu S, Yang M, Lim KM, Cho Y, Kim H, Lee K, Jeong SH, Coffey RJ, Goldenring JR, Nam KT. Expression of LRIG1, a negative regulator of EGFR, is dynamically altered during different stages of gastric carcinogenesis. *Am J Pathol.* 2018;188(12):2912–23.
19. Torigoe H, Yamamoto H, Sakaguchi M, Youyi C, Namba K, Sato H, Shien K, Soh J, Suzawa K, Tomida S, et al. Tumor-suppressive effect of LRIG1, a negative regulator of ErbB, in non-small cell lung cancer harboring mutant EGFR. *Carcinogenesis.* 2018;39(5):719–27.
20. Li W, Zhou Y. LRIG1 acts as a critical regulator of melanoma cell invasion, migration, and vasculogenic mimicry upon hypoxia by regulating EGFR/ERK-triggered epithelial-mesenchymal transition. *Biosci Rep.* 2019;39:1.
21. Fontao F, Barnes L, Kaya G, Saurat JH, Sorg O. From the cover: high susceptibility of Lrig1 sebaceous stem cells to TCDD in mice. *Toxicol Sci.* 2017;160(2):230–43.
22. Lindquist D, Alsina FC, Herdenberg C, Larsson C, Hoppener J, Wang N, Paratcha G, Tarjan M, Tot T, Henriksson R, et al. LRIG1 negatively regulates RET mutants and is downregulated in thyroid cancer. *Int J Oncol.* 2018; 52(4):1189–97.
23. Mao F, Holmlund C, Faraz M, Wang W, Bergenheim T, Kvarnbrink S, Johansson M, Henriksson R, Hedman H. Lrig1 is a haploinsufficient tumor suppressor gene in malignant glioma. *Oncogenesis.* 2018;7(2):13.
24. Powell AE, Wang Y, Li Y, Poulin EJ, Means AL, Washington MK, Higginbotham JN, Juchheim A, Prasad N, Levy SE, et al. The pan-ErbB negative regulator Lrig1 is an intestinal stem cell marker that functions as a tumor suppressor. *Cell.* 2012;149(1):146–58.
25. Powell AE, Vlachic G, Zhao ZY, McKinley ET, Washington MK, Manning HC, Coffey RJ. Inducible loss of one Apc allele in Lrig1-expressing progenitor cells results in multiple distal colonic tumors with features of familial adenomatous polyposis. *Am J Physiol Gastrointest Liver Physiol.* 2014;307(1):G16–23.
26. Alexandrov LB, Nik-Zainal S, Wedge DC, Campbell PJ, Stratton MR. Deciphering signatures of mutational processes operative in human cancer. *Cell Rep.* 2013;3(1):246–59.
27. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, Bignell GR, Bolli N, Borg A, Borresen-Dale AL, et al. Signatures of mutational processes in human cancer. *Nature.* 2013;500(7463):415–21.
28. Alexandrov LB, Stratton MR. Mutational signatures: the patterns of somatic mutations hidden in cancer genomes. *Curr Opin Genet Dev.* 2014;24:52–60.
29. Chatterjee A, Gupta S. The multifaceted role of glutathione S-transferases in cancer. *Cancer Lett.* 2018;433:33–42.
30. Nagarajan N, Bertrand D, Hillmer AM, Zang ZJ, Yao F, Jacques PE, Teo AS, Cutcutache I, Zhang Z, Lee WH, et al. Whole-genome reconstruction and mutational signatures in gastric cancer. *Genome Biol.* 2012;13(12):R115.
31. Haigis KM, Kendall KR, Wang Y, Cheung A, Haigis MC, Glickman JN, Niwa-Kawakita M, Sweet-Cordero A, Sebolt-Leopold J, Shannon KM, et al. Differential effects of oncogenic K-Ras and N-Ras on proliferation, differentiation and tumor progression in the colon. *Nat Genet.* 2008;40(5): 600–8.
32. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. *Mol Ecol.* 2013;22(11):3124–40.
33. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25.
34. 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(16):2078–9.
35. Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics.* 2013;43:11–33.
36. Christoforides A, Carpten JD, Weiss GJ, Demeure MJ, Von Hoff DD, Craig DW. Identification of somatic mutations in cancer through Bayesian-based analysis of sequenced genome pairs. *BMC Genomics.* 2013;14:302.
37. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernysky A, Garimella K, Altshuler D, Gabriel S, Daly M, et al. The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 2010;20(9):1297–303.
38. Gehring JS, Fischer B, Lawrence M, Huber W. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. *Bioinformatics.* 2015; 31(22):3673–5.
39. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, Sirotkin K. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* 2001; 29(1):308–11.
40. Preston JL, Royall AE, Randel MA, Sikkink KL, Phillips PC, Johnson EA. High-specificity detection of rare alleles with paired-end low error sequencing (PELE-Seq). *BMC Genomics.* 2016;17:464.
41. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GR, Thormann A, Flicek P, Cunningham F. The Ensembl variant effect predictor. *Genome Biol.* 2016; 17(1):122.
42. Cingolani P, Platts A, Wangle L, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin).* 2012;6(2):80–92.
43. Skidmore ZL, Wagner AH, Lesurf R, Campbell KM, Kunisaki J, Griffith OL, Griffith M. GenVisR: genomic visualizations in R. *Bioinformatics.* 2016;32(19): 3012–4.
44. Ginestet C. ggplot2: elegant graphics for data analysis. *J R Stat Soc A Stat.* 2011;174:245.
45. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15–21.
46. Raney BJ, Dreszer TR, Barber GP, Clawson H, Fujita PA, Wang T, Nguyen N, Paten B, Zweig AS, Karolchik D, et al. Track data hubs enable visualization of user-defined genome-wide annotations on the UCSC genome browser. *Bioinformatics.* 2014;30(7):1003–5.
47. Dvinge H, Bradley RK. Widespread intron retention diversifies most cancer transcriptomes. *Genome Med.* 2015;7(1):45.

## Publisher's Note

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

**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](https://www.biomedcentral.com/submissions)

