Gene Signature in Sessile Serrated Polyps Identifies Colon Cancer Subtype

Priyanka Kanth1, Mary P. Bronner2,3, Kenneth M. Boucher3,4, Randall W. Burt1,3, Deborah W. Neklason5, Curt H. Hagedorn1,6, and Don A. Delker1

Abstract

Sessile serrated colon adenoma/polyps (SSA/P) are found during routine screening colonoscopy and may account for 20% to 30% of colon cancers. However, differentiating SSA/Ps from hyperplastic polyps (HP) with little risk of cancer is challenging and complementary molecular markers are needed. In addition, the molecular mechanisms of colon cancer development from SSA/Ps are poorly understood. RNA sequencing (RNA-Seq) was performed on 21 SSA/Ps, 10 HPs, 10 adenomas, 21 uninvolved colon, and 20 control colon specimens. Differential expression and leave-one-out cross-validation methods were used to define a unique gene signature of SSA/Ps. Our SSA/P gene signature was evaluated in colon cancer RNA-Seq data from The Cancer Genome Atlas (TCGA) to identify a subtype of colon cancers that may develop from SSA/Ps. A total of 1,422 differentially expressed genes were found in SSA/Ps relative to controls. Serrated polyposis syndrome (n = 12) and sporadic SSA/Ps (n = 9) exhibited almost complete (96%) gene overlap. A 51-gene panel in SSA/P showed similar expression in a subset of TCGA colon cancers with high microsatellite instability. A smaller 7-gene panel showed high sensitivity and specificity in identifying BRAF-mutant, CpG island methylator phenotype high, and MLH1-silenced colon cancers. We describe a unique gene signature in SSA/Ps that identifies a subset of colon cancers likely to develop through the serrated pathway. These gene panels may be utilized for improved differentiation of SSA/Ps from HPs and provide insights into novel molecular pathways altered in colon cancer arising from the serrated pathway. Cancer Prev Res; 9(6); 456–65. ©2016 AACR.

Introduction

Colon cancer is the second leading cause of cancer-related deaths in United States and third most common cancer in men and women (1). Serrated colon polypos are found in 12% to 36% of patients undergoing routine screening colonoscopy (2–4). Serrated polyps are classified into three groups: hyperplastic polyps (HP), sessile serrated adenoma/polyps (SSA/P), and traditional serrated adenomas (TSA; ref. 5). Both SSA/Ps and relatively rare TSAs have malignant potential. Histologically, SSA/Ps often have basilar crypt dilation, which may present as an L-shaped or inverted T-shaped morphology. HPs lack these specific features (6). However, differentiating SSA/Ps from HPs by colonoscopy or histopathology remains difficult due to overlapping morphologic and pathologic features (7, 8).

The serrated polyposis syndrome (SPS) is an extreme phenotype, with patients presenting with multiple SSA/Ps, and has a high risk of colon cancer (9–11). So far, no inherited gene mutation has been found in SPS. The risk of SSA/Ps progressing to colon cancer is not unique to SPS patients and has also been described in patients with sporadic SSA/Ps (2, 12).

The "serrated polyph pathway" has been described as an underlying mechanism in the development of colon cancer from SSA/Ps and may account for 20% to 30% of sporadic colon cancers (6, 13–15). However, the molecular mechanisms or signaling pathways important in the progression of SSA/Ps to colon cancer are uncertain. DNA microsatellite instability (MSI), CpG island methylation, and BRAF mutations are possible underlying molecular mechanisms in the development of SSA/Ps (14–17). At least a subset of proximal colorectal cancers have the CpG island methylator phenotype (CIMP) and high MSI (MSI-H), suggesting similar molecular backgrounds in serrated polyps and proximal cancer (18).

There is limited information on gene expression profiles differentiating SSA/Ps from traditional HPs. Two prior studies have described gene expression in SSA/Ps using microarray technologies (19, 20). We recently identified >1,200 differentially expressed genes in SSA/Ps from patients with SPS using RNA sequencing (RNA-Seq) and developed several immunohistochemical markers specific for SSA/Ps (21). However, comprehensive RNA-Seq gene expression profiles have not been defined for sporadic SSA/Ps and HPs, and it is not known whether sporadic SSA/Ps differ from syndromic SSA/Ps that have a very high risk for progressing to colon cancer. The goals of our study were 2-fold:

1. Develop a unique gene signature in SSA/Ps from patients with serrated polyp syndrome (SPS) using RNA-Seq and leave-one-out cross-validation methods.
2. Evaluate the gene signature in colon cancer RNA-Seq data from The Cancer Genome Atlas (TCGA) to identify a subset of colon cancers likely to develop through the serrated pathway.
Materials and Methods

Patients

Samples were obtained from patients visiting University of Utah Health Care and George Wahlen Veterans Affairs Medical Center (Salt Lake City, UT) between ages 45 and 75 for routine screening, surveillance, or diagnostic colonoscopy. Patients with SPS were between 18 to 75 years of age. Subjects with family history of colon cancer, familial cancers, including familial adenomatous polyposis and Lynch syndrome, history of inflammatory bowel disease, and prior colonic resections were excluded. The samples were prospectively collected from 2008 to 2013 for RNA-Seq. All patients signed and agreed to informed consent as approved by the respective hospitals Institutional Review Boards. If polyps were found during colonoscopy, a biopsy of polyp tissue was collected in RNAlater for RNA-Seq. If a polyp was too small to obtain a biopsy for both histology and RNA-Seq, a tissue sample for RNA-Seq was not collected for the study.

Twelve sessile serrated polyps were obtained from 8 patients with serrated polyposis syndrome (ten right colon and two left colon; refs. 21). SSA/Ps from these patients were previously analyzed for specific mRNA changes by qPCR but not analyzed by RNA sequencing. Uninvolved mucosa from right and left colon was also collected. Right colon was defined as colonic region from splenic flexure to cecum.

Sporadic sessile serrated polyps (n = 9, six right colon, three left colon), HPs (n = 10, two right colon, eight left colon), and adenomatous polyps (n = 10, nine right colon, one left colon) were obtained along with uninvolved mucosa from patients undergoing routine colonoscopy. Normal colon tissue (n = 20, ten right colon, ten left colon) was obtained from patients undergoing screening colonoscopy, with no polyps found on exam. All samples were collected prospectively and placed in RNAAlater (InVitrogen) immediately after tissue removal, stored at −80°C prior to performing RNA isolation. The demographics of sporadic SSA/Ps and hyperplastic polyps are presented in Supplementary Table S1A and B, respectively. The demographics of patients with adenoma and control colon tissues (analyzed using qPCR) have been described previously (21, 22). HPs were not subdivided into microvesicular hyperplastic polyps (MVHP) and goblet cell hyperplastic polyps, as these classifications are not used clinically or discussed in the recent post-polypectomy colonoscopy surveillance guidelines (22). We decided to follow the classification that is most appropriate and practical in clinical practice with the aim to define clinically relevant and realistic gene signatures. Moreover, these two HP subtypes have not been shown to have different risks for the development of colon cancer.

Pathologic classification

All biopsy specimens were reviewed by an expert gastrointestinal pathologist. Serrated polyps were classified according to the recent recommendations of the Multi-Society Task Force on Colorectal Cancer for post-polypectomy surveillance and as described previously (21, 22). HPs were not subdivided into microvesicular hyperplastic polyps (MVHP) and goblet cell hyperplastic polyps, as these classifications are not used clinically or discussed in the recent post-polypectomy colonoscopy surveillance guidelines (22). We decided to follow the classification that is most appropriate and practical in clinical practice with the aim to define clinically relevant and realistic gene signatures. Moreover, these two HP subtypes have not been shown to have different risks for the development of colon cancer.

RNA isolation, RNA-Seq, and differential expression analysis

Total RNA was isolated using TRIzol (Invitrogen) and quality of RNA assessed by an Agilent 2000 Bioanalyzer as described previously (21, 23, 24). RNA-Seq was performed on 86 individual colon samples: 21 SSA/Ps (12 syndromic and 9 sporadic), 10 HPs, 10 adenomatous polyps, 21 uninvolved colon, 20 control colon, and 4 colon cancer samples. PCR-amplified cDNA sequencing libraries were prepared using oligo dT–selected RNA according to the Illumina TruSeq library protocol. Single-end 50 bp sequence reads were performed on an Illumina HiSeq 2000 instrument and aligned to the GRCh37/Hg19 human reference genome using the NovoAlign (Novocraft) application as described previously (21). Differentially expressed genes were determined using the USeq DefinedRegionsDifferentialSeq application and hierarchical clustering and principal component analysis of genes and samples performed using Cluster 3.0 as described previously (21). The RNA-Seq datasets described in this study have been deposited in the NCBI Gene Expression Omnibus (GEO) with accession number GSE76987.

Derivation of 51 SSA/P gene signature and 7-gene panel

A 27-gene signature was developed to include genes with high fold change expression in SSA/Ps compared with HPs (see Supplementary Table S2). A separate 28-gene signature was obtained using a leave-one-out cross-validation method (see selection and cross-validation of a 28-gene signature section below). Combining the two gene signatures (27 and 28) resulted in a 55-gene signature unique for SSA/Ps. Four of these 55 genes were not found in colon cancer RNA-Seq datasets from the TCGA database, resulting in a 51-gene signature to compare across all RNA-Seq datasets. We next looked at the correlation of increased expression of each of the 51 genes in BRAF-mutant, CIMP-H, and MLH1-silenced colon cancers (see Supplementary Tables S3 and S4). Seven of the 51 genes (ZIC5, SEMG1, TRNP1, MUC6, CRBY2A, FSCN1, and ZIC2) frequently overexpressed in MSI-H colon cancers were also frequently overexpressed in BRAF-mutant, CIMP-H, and/or MLH1-silenced colon cancers. We used this 7-gene panel for sensitivity and specificity calculations for identifying colon cancers that likely develop through the serrated pathway.

Selection and cross-validation of a 28-gene signature

Sequencing data from 10 HP and 21 SSA/P samples were used to construct and cross-validate a gene signature. Prior to analysis, genes differentially expressed between left and right colon (≥2-fold change, FDR < 0.01) were removed. An
"unpaired" analysis was then performed on all 31 serrated polypl samples using DESeq2 negative binomial statistics with histology as the only predictor. The FDR threshold for the signature genes was set at 0.01. Twenty-eight genes met these criteria and were used for cross-validation. The average of log (count + 0.5) for the selected genes was used to form separate signatures for HP and serrated colon polypl samples. A normalized Euclidean distance measure was constructed from the selected genes. SDs of the Euclidean distance measure, and classification was cross-validated. A principal component analysis was performed using Cluster 3.0 and a 3D plot constructed using the "rgl" package in R.

Analysis of signature genes in published microarray data of serrated polypl

No previously published RNA-Seq data of serrated polypl are available for comparison with our datasets. We evaluated the expression of each of our 51 signature genes in a previously published microarray dataset (GEO number GSE43841; ref. 19). See Supplementary Methods.

Comparison with TCGA colon cancer RNA-Seq datasets

Fifty-one SSA/P signature genes were used to interrogate 68 colon cancer RNA-Seq datasets from TCGA [36 specimens from Christiana Healthcare and 32 from Memorial Sloan Kettering (New York, NY)] and four from the University of Utah (26). Raw sequencing data for each colon cancer dataset was downloaded from the TCGA database (27) and normalized by number of reads per kilobase of gene length per million of total reads (RPKM). There was expression data for 18,130 unique RefSeq genes in both the TCGA and University of Utah RNA-Seq datasets. A total of 195 TCGA colon cancer datasets were also evaluated for mRNA expression in the 51 signature genes using the cbioPortal for Cancer Genomics (28, 29).

Mutual exclusivity and cooccurrence analysis

Mutual exclusivity and cooccurrence of genomic alterations in each of our 51 signature genes and incidence of BRAF mutations was evaluated using the cbioPortal for Cancer Genomics. This analysis uses a previously published statistical method, mutual exclusivity modules, to identify genes that may be involved in the same cancer pathway (30).

Sensitivity and specificity of a 7-gene panel

The sensitivity and specificity of a 7-gene panel was evaluated in 182 TCGA colon cancer samples with gene expression, methylation, and BRAF mutation data available. There were 31 MLH1-silenced, CIMP-H, and/or BRAF-mutant samples out of 182 regarded as positive and the rest as negative. Cut-off values for each gene were set at twice the average expression of all samples. K-fold cross-validation was used to get an estimate of sensitivity and specificity. In addition to individual expression, we also investigated panels of genes. For the panels, we considered the count of the number of genes above the 2-fold threshold as a predictor (see details in Supplementary Methods). PCR validation was performed on 4 of these genes FSCN1, ZIC3, SEMG1, and MUC6 (see Supplementary Methods).

Results

Differential gene expression analysis

RNA-Seq was performed on 86 colon specimens with a mean sequence depth of 14.7 million mapped reads per sample. Comparing syndromic (n = 12) and sporadic (n = 9) SSA/P RNA-Seq datasets to control right colon (n = 10), we identified 1,422 differentially expressed annotated genes (≥2-fold change, FDR < 0.05) by negative binomial statistical analysis (Fig. 1A; Supplementary Table S2). Comparing HPs (n = 10) to control left colon (n = 10), we identified 711 differentially expressed genes using the same fold change and FDR cut-off value. A total of 475 genes were differentially expressed in both SSA/Ps and HPs. In the RNAs that were differentially expressed in SSA/Ps, 1,095 (77%) were protein coding and 327 (23%) were non-coding (Fig. 1B). A similar percentage of protein coding (80%) and non-coding (20%) RNAs was also significantly differentially expressed in HPs relative to control colon.

To determine if sporadic SSA/Ps had a gene expression profile similar to syndromic SSA/Ps, we compared differentially
preserved genes with a >2- and 4-fold change in each group (Supplementary Fig. S1A and Fig. 2A, respectively). Greater than 89% (>2-fold) and 96% (>4-fold) of the differentially expressed genes observed in sporadic SSA/Ps were also differentially expressed in syndromic SSA/Ps. We are not aware of another gene expression comparison of sporadic and syndromic SSA/Ps, and these results describe major molecular similarities in SSA/Ps from these two very different patient cohorts. A total of 215 genes (77%) were uniquely differentially expressed >2-fold in SSA/Ps as compared with HPs (Fig. 2A), whereas nearly 86% of the differentially expressed genes in HPs overlapped with SSA/Ps and only 10 genes (14%) were uniquely differentially expressed >4-fold in HPs. This suggests that the molecular phenotype in HPs (considered at little or no risk for progression to colon cancer) is surprisingly similar to that of SSA/Ps (considered high risk). One notable difference between SSA/Ps and HPs was the magnitude of fold change in many differentially expressed genes. Hierarchical clustering of 27 protein-coding genes with average increased expression >13 fold in SSA/Ps illustrates what was shared in gene expression changes among all but two of the SSA/Ps (Fig. 2B; Supplementary Table S2). It should be noted that 2 of 10 (20%) HPs and 5 of 21 (24%) SSA/Ps were from right and left colon, respectively. Although our numbers of HPs from right colon and SSA/Ps from left colon are small, we did not see appreciable differences in gene expression between left and right HPs or SSA/Ps. Increased expression of these 27 genes was not observed in adenomatous polyposis RNA-Seq datasets (Fig. 2C).

We also compared gene expression in the uninvolved colon (n = 10) of SPs patients and patients with sporadic SSA/Ps, with the control right colon (n = 10) of patients undergoing screening colonoscopy with no polyps (Supplementary Fig. S2). Surprisingly, 1,922 genes were differentially expressed between the uninvolved colon of patients with SSA/Ps and control colon (≥2-fold change, FDR < 0.01). A significant overlap in the gene expression profile of uninvolved colon from patients with SPS and sporadic SSA/Ps was observed. However, the magnitude of fold change was small for most genes (<3 fold), and the genes differentially expressed were not common to genes differentially expressed in SSA/Ps.

Selection and cross-validation of a gene signature that differentiates SSA/Ps from HPs

Count data from 31 serrated polyps (21 SSA/Ps and 10 HPs) were used in a leave-one-out cross-validation analysis. Twenty-eight genes with an FDR < 0.01 and ≥2-fold change (SSA/Ps vs. HPs) defined the signature (Supplementary Table S2). Twenty-eight of 31 serrated polyps were classified correctly for a nominal error rate of 10%. After cross validating four times, the cross-validated error rate was 18%. Principal component analysis of the gene expression of each of the 28 genes in all 31 serrated polyps is shown in Fig. 3A, which demonstrates the misclassification of two SSA/Ps and one HP. The relative expression of each of the 28 genes in SSA/Ps and HPs is shown in Fig. 3B. Six genes were overexpressed and 22 underexpressed in SSA/Ps relative to HPs.

Evaluation of gene signature in published microarray data of serrated polyps

We compared the relative expression of each of our 51-gene signature in SSA/Ps, MVHPs and normal colon (left and right) from a previously published microarray study (19). Clear separation of SSA/Ps from MVHPs and control colon was observed by hierarchical clustering (Supplementary Fig. S3). In fact, 5 of 6 MVHPs, showed gene expression patterns more closely resembling control colon than SSA/Ps.

Identifying colon cancers with the SSA/P gene signature in TCGA

We compared our 51-gene SSA/P signature with 68 colon cancer RNA-Seq datasets available in TCGA and four colon cancers obtained from the University of Utah (Fig. 4A; Supplementary Table S2). RNA-Seq data from 4 of the 55 genes were not available in the TCGA datasets. We performed RNA sequencing on four colon cancers from the University of Utah to identify potential laboratory/batch effect differences in gene expression between our RNA-Seq datasets and the TCGA datasets. The 51-gene SSA/P signature showed similar expression patterns between syndromic and sporadic SSA/Ps and the MSI-H subset of colon cancers. No batch effects were observed between our colon cancer datasets and the TCGA datasets. Sixty-three of 72 cancers had data on their MSI status, with 11 cancers being MSI-H (MSI status unknown for 9 colon cancers). Eighteen colon cancers clustered with SSA/Ps and 8 of the 18 colon cancers (44%) were MSI-H. This is a significant finding, as of the remaining 54 colon cancers that did not cluster with SSA/Ps, only 3 were MSI-H (6%). This suggests that our SSA/P signature identifies MSI-H cancers.

We also evaluated mRNA expression of each of our 51 SSA/P signature genes in 195 TCGA colon cancers using the cBioPortal for Cancer Genomics. Thirteen of the 51 signature genes had frequent increased mRNA expression in ≥10% of hypermutated colon cancers but not in nonhypermutated cancers (Table 1). Seven of these genes (FSCN1, ZIC2, ZIC5, CRYBA2, MUC6, TRNP1, and SEMG1) had increased mRNA expression in 13% to 30% of hypermutated and only 0% to 3% of nonhypermutated colon cancers with Fisher exact P < 0.01 (Table 1). Twenty-two of the 30 (73%) hypermutated colon cancers showed increased expression of at least one of the 7-gene panel. Seventeen of the 22 (77%) hypermutated colon cancers showed increased expression of at least one of the 7-gene panel also showed MLH1 silencing. (Fig. 4B). Eleven of 51 genes showed frequent overexpression in CIMP-H and/or MLH1-silenced colon cancers, including all 7 that showed frequent increased expression in hypermutated cancers (Supplementary Table S3). We did not observe frequent increased expression of previous SSA/P markers (Annexin A10, ANXA10 and claudin 1, CLDN1) in hypermutated, CIMP-H and/or MLH1-silenced colon cancers (Table 1 and Supplementary Table S3; refs. 19, 31).

Mutual exclusivity and cooccurrence analysis

Using the cBioPortal, we evaluated concurrent genomic alterations (RNA expression and somatic mutation) in each of our 51-gene panel and two genes from previous microarray studies (ANXA10 and CLDN1) with alterations in BRAF (19, 31). Thirteen of 51 genes showed statistically significant associations with BRAF mutation both by Fisher exact test and log OR (Supplementary Table S4). Six of these genes (FSCN1, ZIC5, CRYBA2, MUC6, TRNP1, and SEMG1) were common to genes frequently overexpressed in hypermutated, CIMP-H and MLH1-silenced colon cancers. ZIC2 and CLDN1 did not show significant associations with BRAF mutation, and ANXA10 showed a positive association by log OR but not the Fisher exact test.
Sensitivity and specificity of a 7-gene panel

Using a 7-gene panel (FSCN1, ZIC2, ZIC5, CRYBA2, MUC6, TRNP1, and SEMG1), we determined the sensitivity and specificity of each gene in identifying 31 BRAF-mutant, CIMP-H and/or MLH1-silenced colon cancers out of 182 total colon cancers from the TCGA database (Table 2A). The specificity of each gene in identifying this subset of cancers was very high, between 85% and 99%. SSA/P RNA markers ANXA10 and CLDN1 showed similar
compared with HPs (range 2.2 to 6.7).)

specificity to our 7-gene panel. In contrast, the sensitivity of each gene in identifying BRAF-mutant, CIMP-H and/or MLH1-silenced colon cancers was more variable between genes (26%–68%), with ZIC5 showing the highest sensitivity at 68%. The two previously identified RNA markers for SSA/Ps were lower, with 19% and 6% sensitivity for ANXA10 and CLDN1, respectively. Using a 7-gene panel, our sensitivity increased to 94% if at least one of the seven genes showed a 2-fold increase in expression (Table 2B). Using ANXA10 or CLDN1 with our 7-gene panel, the sensitivity was 97% and 94%, and the specificity was 72% and 63%, respectively (Supplementary Table S5). qPCR validation was performed on 4 genes (FSCN1, ZIC5, SEMG1, and MUC6) and showed high expression in SSA/Ps compared with HPs, uninvolved or control colon consistent with our RNA-Seq data (Supplementary Fig. S4).

**Discussion**

SSA/Ps are now recognized as polyps with malignant potential, with SSA/Ps originating in the serrated polyposis syndrome having the highest risk for progression to colon cancer. Recent cancer surveillance guidelines recommend earlier follow-up for patients with sporadic SSA/Ps almost at par with individuals with adenomatous polyps (22). Nevertheless, differentiating SSA/Ps from HPs by histopathology and identifying patients with SSA/Ps have some challenges in clinical practice. The RNA-Seq datasets we describe identify 51 differentially expressed genes in SSA/Ps that molecularly distinguish them from HPs. These genes are also differentially expressed in sporadic microsatellite unstable (MSI-H) colon cancers. We further refined our panel to seven genes that also show frequent overexpression in BRAF-mutant, CIMP-H, and MLH1-silenced colon cancers. Our data provide clear evidence that RNA expression changes in BRAF-mutant, CIMP-H, and MLH1-silenced colon cancers are observed in early SSA/Ps and that these new gene expression markers may lead to improved diagnostics for SSA/Ps. Moreover, our data demonstrate similar gene expression profiles of SSA/Ps in the SPS and sporadic SSA/Ps, indicating that common mechanisms of progression to cancer are operating in both.

Comparing the transcriptome of SSA/Ps and HPs produced findings that raise some critical questions about these two subtypes of serrated polyps with very different potentials for progression to colon cancer. It is unclear whether serrated adenocarcinoma originates directly through SSA/Ps or whether genetic alterations in certain hyperplastic polyps found in right colon lead to the development of SSA/Ps and eventually to colon cancer. SSA/Ps, especially in the SPS, have a significant risk for progression to cancer (9–11), whereas HPs have a negligible risk (32, 33). The finding that most of the genes found differentially expressed in HPs were also found in SSA/Ps at least partly explains why both types of polyps have a similar morphologic appearance. On the other hand, there were many uniquely and highly differentially expressed genes in SSA/Ps compared with HPs. The unique SSA/Ps gene signature established in this study provides an opportunity to identify critical pathways that may explain these differences in cancer risk.

Our 7-gene panel (FSCN1, ZIC2, ZIC5, CRYBA2, MUC6, TRNP1, and SEMG1) identified BRAF-mutant, CIMP-H, and MLH1-silenced colon cancers with high sensitivity and specificity. In comparison with other gene markers described for SSA/Ps (ANXA10 and CLDN1), our 7-gene panel showed increased sensitivity and similar specificity. This increase in sensitivity might be related to the use of RNA-Seq versus microarray technology. RNA-Seq provides a more quantitative analysis of transcript abundance and is not dependent on previously defined gene annotation. Also, the analysis of SSA/Ps from SPS patients, known to have high colon cancer risk, may have further increased our
ability to identify a gene signature more closely associated with sporadic colon cancer developing from the serrated pathway.

Three genes (FSCN1, TRNP1, and ZIC2) of our 7-gene panel were previously identified to be overexpressed in BRAF-positive colon cancers in a European patient cohort (34). These genes were part of a 64-gene expression classifier for BRAF-positive colon cancers with poor prognosis. Another study classifying colon cancers into four consensus molecular subtypes with subtype 1 (CMS1) consisting of microsatellite unstable, CIMP-H, and BRAF-positive tumors identified one of our 7-gene panel (ZIC2) as a marker of serrated cancers (35, 36). ZIC proteins play a role in regulating the sonic hedgehog and Wnt/B-catenin signaling pathways (37, 38). ZIC2 expression has been associated with multiple cancers, including brain, ovarian, and cervical cancer (39, 40). FSCN1 is an actin-binding protein frequently overexpressed in a variety of cancers, including colon cancer, and predicts poor prognosis (41). FSCN1 is also highly expressed in serrated colon cancers (42). TMF-regulated nuclear protein (TRNP1) is a nuclear protein that plays a role in mammalian brain cortex development (43). The significance of TRNP1 overexpression in colon cancer remains unknown. Our study reinforces the importance of these genes in serrated colon cancers, providing the first evidence that these mRNA changes occur early in the cancer process in pre-neoplastic serrated lesions (SSA/Ps).

Other genes described in our 7-gene panel may also participate in colon cancer progression. MUC6 is a gastric mucin protein, shown to have increased expression in SSA/Ps compared with HPs (44). Increased expression of MUC6 has been documented in hypermethylated colon cancers, suggesting its possible role in serrated pathway (45). Data lack about the role of SEMG1 and CRYBA2 in colon cancer. SEMG1 is a seminal vesicle protein that has been studied as a biomarker for the detection of prostate
to understand the functions of these key genes in the serrated pathway.

A significant number of the genes that were differently expressed in the uninvolved colonic mucosa of patients with syndromic (SPS) and sporadic SSA/Ps, relative to normal colon (patients with no polyps), overlapped and suggest a field effect may be present in the colonic mucosa of patients with SSA/Ps. These genes were different from those found common to syndromic and sporadic SSA/Ps and had smaller fold changes relative to controls. A "field cancerization" effect has been reported in studies of sporadic colon cancer (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.

MSI, CIMP, and the inactivation of MLH1 and BRAF mutations have all been implicated as underlying events in the serrated pathway to colon cancer (14–18, 51). A recent study showed MLH1 silencing in a subgroup of hypermutated colon cancers that had increased BRAF and decreased APC and KRAS mutations. The authors concluded that MLH1 silencing occurred through a different pathway, suggestive of the serrated pathway (52). However, not all SSA/Ps have these changes, and it remains uncertain whether they are absolute requirements for progression to cancer. A recent large serrated polyp study only identified MLH1 methylation in 11% of SSA/Ps (53). We report a new set of 51 genes that are differently expressed in most SSA/Ps and sporadic MSI-high cancers (48, 49). There are also limited studies investigating possible field effects in patients with colon polyps, particularly SSA/Ps (50). Our data raise important questions regarding the origin of such changes. The question of predictive value of field effect will require studies with larger numbers of patients, which are underway at this time.
sporadic SSA/Ps and the serrated polyposis syndrome in the general population. Even with this limitation, this is the largest RNA-Seq study performed to characterize the transcriptome of SSA/Ps. Finally, our gene panel was not validated in a separate RNA-Seq study of serrated polyps because these datasets are not publicly available. However, our gene signature did accurately classify SSA/Ps from MVHPs using expression data from a previous microarray study. Future validation studies are currently being designed and are beyond the scope of this study.

In summary, this report provides a comprehensive gene expression comparison of SSA/Ps with HPs, which share many histopathologic similarities but differ markedly in the risk of progression to colon cancer. Despite many similarities in gene expression in SSA/Ps and HPs, both sporadic and syndromic SSA/Ps have a unique gene signature with a number of highly differentially expressed genes of interest relative to oncogenesis. The identification of a set of novel genes uniquely differentially expressed in SSA/Ps and BRAF-mutant, CIMP-H, and MLH1-silenced colon cancers provides additional leads for further understanding the molecular pathways leading to cancer progression via the serrated pathway. This may lead to the development of a gene panel that can be used in clinical practice to stratify patients with increased colon cancer risk from serrated polyps. This could be especially helpful in identifying patients with serrated polyposis syndrome in whom no currently recognized genetic mutation has been identified.

Disclosure of Potential Conflicts of Interest
R.W. Burt has received speakers bureau honoraria from Myriad Genetics and Thriss Pharma. D.A. Delker and P. Kanth reports having ownership interest (including patents). No potential conflicts of interest were disclosed by the other authors.

References

Authors’ Contributions
Development of methodology: P. Kanth, M. P. Bronner, D.A. Delker
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): P. Kanth, M. P. Bronner, R.W. Burt, D.W. Neklason, C.H. Hagedorn, D.A. Delker
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P. Kanth, M. P. Bronner, K.M. Boucher, C.H. Hagedorn, D.A. Delker
Writing, review, and/or revision of the manuscript: P. Kanth, M. P. Bronner, K.M. Boucher, R.W. Burt, D.W. Neklason, C.H. Hagedorn, D.A. Delker
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): P. Kanth, M. P. Bronner, R.W. Burt, D.A. Delker
Study supervision: M. P. Bronner, D.A. Delker

Acknowledgments
The authors thank Kathleen Boynton and Michelle Done for their assistance in sample collection and Mark Hazel for qPCR validation experiments.

Grant Support
This study was supported by a pilot clinical research award from American College of Gastroenterology (to P. Kanth), University of Utah Personalized Medicine Program Seed Grant (to C.H. Hagedorn), NIH grants CA176130 and D430666 (to C.H. Hagedorn) and CA073992 and CA146329 (to R.W. Burt). This study was also supported by Cancer Center Support Grant P30-CA42014 and National Center for Advancing Translational Sciences of the NIH Award UL1TR001067 and Huntsman Cancer Foundation.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received October 15, 2015; revised March 3, 2016; accepted March 22, 2016; published OnlineFirst March 29, 2016.
Serrated Pathway Gene Panel


Gene Signature in Sessile Serrated Polyps Identifies Colon Cancer Subtype

Priyanka Kanth, Mary P. Bronner, Kenneth M. Boucher, et al.


Updated version  Access the most recent version of this article at: doi:10.1158/1940-6207.CAPR-15-0363

Supplementary Material  Access the most recent supplemental material at: http://cancerpreventionresearch.aacrjournals.org/content/suppl/2016/03/29/1940-6207.CAPR-15-0363.DC1

Cited articles  This article cites 50 articles, 11 of which you can access for free at: http://cancerpreventionresearch.aacrjournals.org/content/9/6/456.full#ref-list-1

Citing articles  This article has been cited by 4 HighWire-hosted articles. Access the articles at: http://cancerpreventionresearch.aacrjournals.org/content/9/6/456.full#related-urls

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, use this link http://cancerpreventionresearch.aacrjournals.org/content/9/6/456. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.