Genomic Landscape of Colorectal Mucosa and Adenomas

Ester Borras1, F. Anthony San Lucas2,3,4, Kyle Chang1,4, Ruoji Zhou1,4, Gita Masand1, Jerry Fowler2, Maureen E. Mork2, Y. Nancy You5,6, Melissa W. Taggart7, Florencia McAllister1,4,8, David A. Jones9, Gareth E. Davies10, Winfried Edelmann11, Erik A. Ehli10, Patrick M. Lynch5,12, Ernest T. Hawk1, Gabriel Capella13, Paul Scheet2,4, and Eduardo Vilar1,4,5,8

Abstract

The molecular basis of the adenoma-to-carcinoma transition has been deduced using comparative analysis of genetic alterations observed through the sequential steps of intestinal carcinogenesis. However, comprehensive genomic analyses of adenomas and at-risk mucosa are still lacking. Therefore, our aim was to characterize the genomic landscape of colonic at-risk mucosa and adenomas. We analyzed the mutation profile and copy number changes of 25 adenomas and adjacent mucosa from 12 familial adenomatous polyposis patients using whole-exome sequencing and validated allelic imbalances (AI) in 37 adenomas using SNP arrays. We assessed for evidence of clonality and performed estimations on the proportions of driver and passenger mutations using a systems biology approach. Adenomas had lower mutational rates than did colorectal cancers and showed recurrent alterations in known cancer driver genes (APC, KRAS, FBXW7, TCF7L2) and AIs in chromosomes 5, 7, and 13. Moreover, 80% of adenomas had somatic alterations in WNT pathway genes. Adenomas displayed evidence of multiclonality similar to stage I carcinomas. Strong correlations between mutational rate and patient age were observed in at-risk mucosa and adenomas. Our data indicate that at least 23% of somatic mutations are present in at-risk mucosa prior to adenoma initiation. The genomic profiles of at-risk mucosa and adenomas illustrate the evolution from normal tissue to carcinoma via greater resolution of molecular changes at the inflection point of premalignant lesions. Furthermore, substantial genomic variation exists in at-risk mucosa before adenoma formation, and deregulation of the WNT pathway is required to foster carcinogenesis.


Introduction

Colorectal cancer is the third most common cancer diagnosed in both men and women in the United States and the second leading cause of cancer-related deaths when both sexes are combined (1). Colorectal cancer arises from adenomatous polyps, also known as adenomas, benign tumors with altered differentiation features found in 25% of men by age 50 years in the United States (2, 3). The critical transition from adenoma to carcinoma is fostered by the sequential acquisition of molecular events in driver genes and is known as the “adenoma-to-carcinoma” progression model. This model was deduced from comparison of genetic alterations observed among normal mucosa, adenomas of progressively larger size, and carcinomas. The key somatically acquired genomic events include chromosomal gains (13 and 20) and losses (18q and 17p), as well as point mutations and small insertions/deletions in driver genes, such as APC, KRAS, and TP53 (2). In fact, somatic mutations in APC have been described in 80% of colorectal adenomas and carcinomas (4). Moreover, germline mutations in APC result in a genetic condition known as familial adenomatous polyposis (FAP), in which patients have a 100% lifetime risk of developing colorectal cancer, with the vast majority developing carcinomas by age 35 years secondary to the development of hundreds of adenomas (2, 5). Although FAP is responsible for approximately only 1% of colorectal cancer diagnoses, this disease represents an attractive model for characterizing the genomic profile of adenomas and investigating the initial events that cooperate with APC loss in intestinal carcinogenesis. In fact,
FAP is the molecular model for 70% to 85% of colorectal cancers that progress through alterations in the WNT pathway and which are known as chromosomal unstable tumors (2, 6). Whole-exome sequencing (WES) analyses of colorectal cancers have provided a comprehensive view of the genomic make-up of this disease and have demonstrated the utility of this approach in identifying novel driver mutations (7, 8). Despite this wealth of information, however, there has been limited success in obtaining a comprehensive genomic analysis of colorectal adenomas and APC-mediated carcinogenesis in at-risk mucosa.

Furthermore, large-scale genomic studies based on high-throughput technologies have systematically ignored the contribution of the adjacent surrounding mucosa to intestinal carcinogenesis by using it as the normal reference to establish the somatic status of alterations. This approach could potentially bias results, due to the fact that preexisting mutations in the mucosa are carried through the development of adenomas due to the fact that they are secondary to exposure to carcinogens, aging, and other factors. Therefore, these mutations have been regarded as germline events and ignored as important molecular changes. Interestingly, recent systems biology analyses have determined that half of the somatic mutations identified in carcinomas originate prior to tumor initiation (9), thus making the assessment of genetic variation in at-risk mucosa critical for a global understanding of intestinal carcinogenesis.

In this study, we sought to characterize the mutational landscape of colorectal adenomas to identify novel candidate genes that cooperate with APC in intestinal carcinogenesis, to determine the presence of multiclonality in premalignant intestinal lesions, and to study the somatic variation accumulated in at-risk colonic mucosa using samples from FAP patients.

Materials and Methods

Subjects and samples

WES was performed in a total of 25 colorectal adenomas, 10 adjacent mucosa, and 12 blood samples from 12 patients diagnosed with FAP at the Catalan Institute of Oncology (Barcelona, Spain) and The University of Texas MD Anderson Cancer Center (Houston, TX). Furthermore, SNP arrays were performed on an additional 37 colorectal and matched blood samples from 14 FAP patients collected at MD Anderson Cancer Center (Supplementary Table S1). Informed consent was obtained from all individuals, and the Institutional Review Boards of both institutions approved this study (MD Anderson protocol ID numbers: PA12-0327 and PA13-0178; Catalan Institute of Oncology ID number: 30/06). Details on tissue preservation, nucleic extraction, and pathologic analyses are provided in the Supplementary Methods (Supplementary Table S2).

Illumina sequencing

Samples were sequenced on an Illumina HiSeq 2000 sequencer with 76-base paired end reads at the MD Anderson Sequencing Core Facility. Germline DNA extracted from blood samples was used as the genomic reference for the analysis of both adenomas and at-risk mucosa. A strategy for categorizing mutations into 5 functional tiers was applied to somatic events. Validation of exonic somatic mutations and indels detected in adenomas by using Illumina sequencing was performed using Sanger sequencing (see details in Supplementary Methods).

TCGA colorectal sample mutation calling

To minimize batch effects in our mutation comparisons between FAP adenomas and The Cancer Genome Atlas (TCGA; ref. 7) colorectal cancer samples, we downloaded available WES data for 107 carcinoma normal pairs from the TCGA project (stages I–V) where samples were sequenced on the Illumina HiSeq 2000 sequencer. We then analyzed these colorectal cancer exome samples using the same bioinformatics pipelines that we used in the analysis of our FAP samples to identify somatic point mutations and generate mutation reports that were applied to the colorectal cancer exomes. Samples with >10 mutations per megabase (Mb) were classified as hypermutated and those with <10 mutations/Mb were considered nonhypermutated.

AI analysis

HapLOH (10) was applied to our WES data for detection of chromosomal allelic imbalances (AI) in adenomas (implemented for sequencing data as hapLOHseq; scheet.org/software), where events identified in adenomas with no overlapping event in the paired blood sample were characterized as somatic AIs. Subsequently, these events were classified as amplifications, deletions, or copy neutral LOH (cn-LOH) by comparing WES read depth between adenomas and their paired blood samples within the event regions using a permutation-based algorithm. To globally validate the frequency of AI, purified DNA samples from 37 adenomas were genotyped using the HumanOmnimExpressExome beadChip array following the manufacturer’s instructions (Illumina). Details are provided in Supplementary Methods.

Comparison of FAP AIs with TCGA colorectal cancer copy number variation events

To compare our calls with the TCGA, copy number variation data were downloaded from the TCGA data portal, where copy number variation calls were made off Affymetrix SNP 6 micro-arrays by the TCGA consortium. We then downloaded the RefSeq transcript definitions from the UCSC Genome Browser and identified a set of candidate genes that have genomic coordinates that recurrently intersect the FAP adenoma AIs. We reported all recurrent AI events for colorectal cancer genes that have been reported to be significantly mutated (7, 8, 11) or significantly amplified or deleted in previous genomic annotations of the landscape of colorectal cancers (Supplementary Tables S3 and S4; refs. 7, 12) and in the genes classified in tiers 1 and 2 from our mutation analyses (Supplementary Table S5).

Pathway and clonality analysis

In regards to pathway analysis, genes harboring mutations classified as damaging using our tier classification system and those located in AI regions that have been reported before as significantly mutated, amplified, or deleted in colorectal cancer (Supplementary Table S6) were uploaded to DAVID Bioinformatics Resources 6.7 (13) and analyzed with the Functional Annotation Tool. Annotation was then performed using the KEGG pathways, and the resulting pathways with their corresponding nominal P values and FDRs from the functional annotation chart were recorded.

The ABSOLUTE computational method (14) was used to quantify the levels of tumor purity and ploidy and to estimate the cancer cell fraction that harbors each mutation. ABSOLUTE inputs are allele frequencies from somatic point mutations with
Results

Assessment of the mutation landscape of adenomas by WES

We performed WES of 25 colorectal adenomas and matched germline DNA samples that served as the genomic reference to establish the somatic mutational status. Samples were collected from 12 patients diagnosed with FAP with clinically confirmed germline APC alterations (Supplementary Table S1 and S2). Exomes were sequenced to a mean depth of 65 reads, with 86% of the targeted regions covered by at least 50 high-quality reads (Supplementary Tables S7–11). In adenomas, we detected a total of 237 indels (20 exonic), 10 splicing variants, and 2,067 somatic mutations (1,067 exonic), with the somatic mutations consisting of 750 nonsilent, 317 silent, and 1,000 noncoding mutations (Fig. 1A; Supplementary Tables S8–11). The most common base substitutions detected were C>T transitions (51%), similar to nonhypermutator colorectal cancers from TCGA (Fig. 1B and Supplementary Fig. S1; Supplementary Table S12; refs. 7, 15). In addition, we analyzed mutational signatures using 6 substitution subtypes (i.e., C>A, C>G, C>T, T>A, T>C, and T>G) and their 5’ and 3’ bases adjacent to the mutation site, generating 96 combinations of mutations (Supplementary Fig. S2; Supplementary Table S13). Previous studies have identified a total of 27 distinct mutation signatures across various types of human cancers (16). Our analysis showed that adenomas were closely related to signature 1, which is the most frequently identified among colorectal cancers as being associated with age of diagnosis (16) and which mirrors the pattern displayed by nonhypermutator colorectal cancers (Supplementary Fig. S2).

The average median allelic frequency of the mutations detected per adenoma was 15% (range 8%–25%), indicative of low aberrant DNA proportions (Fig. 1C; Supplementary Table S9–11). On average, the mean somatic mutation frequency was 83 mutations per adenoma (range 9–186), resulting in a mean mutation rate of 1.75 mutations/Mb (range 0.2–4.1; Fig. 1C and D; Supplementary Table S14). The overall adenoma mutation rate was compared with 107 colorectal cancers from the TCGA. Adenomas had lower mutation rates than carcinomas but did exhibit some degree of overlap with nonhypermutator colorectal cancers (1.75 mutations/Mb in adenomas compared with 4.26 in nonhypermutator carcinomas and 50.88 for hypermutators, both p < 0.0001; Fig. 1D; Supplementary Tables S14 and S15).

Profiling of AIs in colorectal adenomas and stage I carcinomas

We used WES data to characterize chromosomal alterations that result in somatic AIs. We detected AI in 14 adenomas (36%) of our samples (Fig. 2 and Supplementary Fig. S3; Supplementary Table S2). We then compared the AI profiles in adenomas with those for stage I nonhypermutated colorectal cancers from TCGA (7). Regions found to be aberrant in both adenomas and stage I carcinomas are suggestive of events occurring early in colonic carcinogenesis, such as loss of 5q, which we observed in 5 adenomas (20%), and amplifications of chromosomes 7 (20%), 13 (16%), 19 (12%), and 20 (16%; Fig. 2; Supplementary Table S16). On the other hand, recurrent events found only in carcinomas, such as deletions of 17p and 18q (observed in 60% and 70% of stage I colorectal cancers, respectively), are candidate events for association with progression to carcinomas. We validated the frequency of AIs using SNP array technology in an additional cohort of 37 adenomas from 14 FAP patients. AIs in chromosomes 5 (11%), 7 (14%), and 13 (16%) were highly frequent, thus confirming the importance of these events in early carcinogenesis (Supplementary Table S17). Interestingly, we noted that adenomas with AI in 20q had a mutational rate higher than the mean mutation rate in all adenomas (3.23 compared with 1.75, P = 0.01) and similar to the rate observed in nonhypermutated colorectal cancers, thus supporting the idea that these adenomas have progressed further along the adenoma-to-carcinoma transition and also implicating genes in 20q for a role in progression.

Detection of candidate genes related to intestinal carcinogenesis

We developed a system that categorized somatic mutations based on their predicted deleterious effects to discover candidate genes related to APC-driven intestinal carcinogenesis (see Materials and Methods for details). A total of 100 exonic mutations (9%) were classified as damaging events, 35 (33%) as recurrent potentially damaging, 359 (33%) as potentially damaging, 129 (12%) as variants of unknown significance, and 474 (43%) as neutral events (Supplementary Tables S11, S18 and S19; Supplementary Fig. S3). Damaging events resided in well-known cancer genes, such as APC, KRAS, FBXW7, TCF7L2, and BRCA2 (Fig. 3). Among the recurrent potentially damaging mutations, several novel genes emerged as candidates for intestinal carcinogenesis, such as ALK and CNOT3 (Fig. 3; Supplementary Tables S11, S18, and S19). Moreover, the following 11 genes harboring mutations classified as either damaging or recurrent potentially damaging were found altered in at least 2 adenomas: APC, ARID1A, CDC27, CNOT3, EWSRI, FBXW7, GNAQ, KRAS, PCMTD1, TCF7L2, and TTN (Supplementary Table S20; Supplementary Fig. S3). In addition, we annotated AIs to known copy number aberrant colorectal cancer genes and novel candidate genes from our somatic mutation analysis and found recurrent AIs in well-known genes such as APC, EPHB6, MAP1B, MAP2K7, and MET and novel candidates such as CNOT3 (Fig. 3; Supplementary Table S21; refs. 7, 8).

Pathway analysis and detection of somatic second hit in APC in colonic adenomas

Integrated analysis of damaging mutations and AIs in adenomas revealed enrichment in the WNT, MAPK, VEGF, and ErbB pathways, thus suggesting them as future avenues for chemoprevention (Supplementary Table S22). Among these pathways, WNT stood out as deregulated in 80% of all adenomas, with 20 of 25 harboring somatic mutations in at least 1 gene (Supplementary Fig. S4). In fact, cooccurrence of mutations in more than 1 gene of the pathway was observed in 7 adenomas (28%; Supplementary Table S23). The most frequently mutated gene in the WNT pathway was APC (18 adenomas, 72%), harboring a somatic second hit either secondary to LOH of chromosome 5q (5 adenomas, 28%) or point mutations (13 adenomas, 44%). The majority of adenomas with somatic loss of 5q were detected in a patient with a germline mutation identified within the mutation cluster region (MCR) of APC (amino acids 1286–1514). Among the point mutations detected, 10 were nonsense and 3 were small insertions or deletions (Supplementary Fig. S5; Supplementary Table S24). Four of the samples without somatic APC mutations.
lacked any alteration in other WNT signaling pathway genes (Supplementary Table S23) and were characterized by lower mutational rates compared with the 84% of adenomas with identified somatic WNT pathway alterations (0.511 vs. 1.99 mutations/Mb). Interestingly, in those samples harboring somatic APC mutations, the allelic frequency of the APC events was higher compared with the median allele frequency of the sample (Fig. 1C; Supplementary Tables S11, S16, S17, and S24).

Evidence of multiclonality in adenomas

We searched for evidence of clonality in our APC-driven adenomas by visually inspecting the distribution of mutations ranked by allelic frequency. We observed 2 clusters of mutation allelic frequencies in 5 samples, suggesting the presence of 1 major clone and other minor subclones (Fig. 4A and Supplementary Fig. S6). However, this type of analysis could be confounded by tumor purity and copy number variation. To mitigate the effect of variations in these 2 factors, we applied computational tools to transform the frequency distribution of mutations in each sample and to infer the number of clones. We compared in silico estimations of purity, ploidy, and number of clones between adenomas and stage I colorectal cancer. The purity was significantly lower in adenomas compared with stage I colorectal cancers (0.35 vs. 0.58, \( P < 0.0001 \); Fig. 4B; Supplementary Table S25), as was the ploidy (1.93 vs. 2.61, \( P < 0.0001 \); Fig. 4C; Supplementary Table S25). Then, our analysis revealed the presence of multiclonality in 18 (72%) of 25 adenomas, with at least 1 major and 1 minor subclone per lesion. Moreover, more than 50% of the polyps from the same individual is estimated to have multiclonality. Interestingly, the number of clones estimated in adenomas was not significantly different from that for stage I CRC (1.72 vs. 2.06, \( P = 0.165 \); Fig. 4D; Supplementary Table S25).

Assessment of the mutation landscape of at-risk colorectal mucosa by WES

To explore the contribution of the at-risk mucosa to intestinal APC-driven carcinogenesis, we performed WES of matched mucosa samples in 10 of the 12 patients previously analyzed using the
germline DNA calls as genomic reference and compared the mutational profiles with their corresponding adenomas (Supplementary Tables S1 and S2). Colorectal mucosa exomes were sequenced to a mean depth of 62 reads, with 82% of the targeted regions covered by at least 50 high-quality reads (Supplementary Table S7). We identified a total of 11 indels (4 exonic) and 268 somatic mutations (115 exonic), with the somatic mutations consisting of 36 nonsilent, 78 silent, and 154 noncoding mutations (Fig. 5A; Supplementary Tables S26–28). The vast majority (86%) of exonic mutations was classified as neutral or variants of unknown significance, and only a few (total of 6%) were identified as damaging (Supplementary Tables S29 and S30). The mutations classified as damaging were detected in genes such as HDAC10, PLEKHN1, PHOX2B, CDC27, and PPP1CC, which have not been previously described as related to intestinal carcinogenesis, thus suggesting they are not leading to a known route of initiation for carcinogenesis.

The most common base substitutions detected were C>T transitions (37%), similar to adenomas and nonhypermutator colorectal cancers (Fig. 5B; Supplementary Table S31; ref. 15). Analysis of mutation signatures showed that colorectal mucosas were closely related to signature 5, which has been observed in 14.4% of colorectal cancers but has unknown etiology (Supplementary Fig. S2; Supplementary Table S32; ref. 16). In addition, the average median allelic frequency of the mutations detected was 10% (range 5%–17%), which is lower than was observed for adenomas (Fig. 5C; Supplementary Tables S27 and S28). The mean number of somatic mutations observed was 27 (range 8–90), resulting in a mean mutation rate of 0.49 (range 0.1–1.5), significantly lower than what was observed for adenomas (0.49 vs. 1.75 mutations/Mb, P < 0.01; Fig. 5C and D; Supplementary Tables S14 and S33).

Interestingly, 3 adenomas that had mutational rates similar to their matched mucosa did not harbor somatic events in APC or in other genes involved in the WNT pathway (Supplementary Fig. S7A; Supplementary Tables S14, S23, and S33).

Proportion of driver and passenger mutations and the role of aging in intestinal carcinogenesis

We analyzed different factors that could influence the mutational rate observed in adenomas, such as age (>40 vs. ≤40 years),
sex (male vs. female), and the type (tubular vs. tubulovillous), location (colon vs. rectum), and size (>10 vs. ≤10 mm) of the adenoma lesions and the type of somatic second hit in \(\text{APC}\) (LOH vs. point mutation). Age seemed to be the most significant factor \(\left( P < 0.0015 \right)\), showing a positive correlation with the mutational rate \(\left( R = 0.63, P = .005; \text{Fig. 6A} \right)\). Likewise, age was correlated with the mutational rate in at-risk mucosa \(\left( R = 0.74, P = .01; \text{Fig. 6B} \right)\), consistent with observations of mutations accumulating as individuals age. Then, to study the relation between the occurrence of driver and passenger mutations, we recategorized the events observed in both adenomas and at-risk mucosa into 2 categories: drivers (those events classified as damaging and recurrent potentially damaging) and passengers (the rest). We observed a positive correlation between the number of passengers and age in the at-risk mucosa samples \(\left( R = 0.74, P = .01; \text{Supplementary Fig. S7B} \right)\); however, the lack of correlation between the number of passenger and driver events in the at-risk mucosa \(\left( R = 0.23, P = .5; \text{Supplementary Fig. S7C} \right)\) suggests that mutations classified as drivers by bioinformatics tools in the mucosa were not functional and did not lead to an effective increase of passenger aberrations (Supplementary Tables S1 and S34). In those adenomas that harbored second hits in \(\text{APC}\), we observed a correlation between the number of driver and passenger mutations \(\left( R = 0.56, P = .01; \text{Fig. 6C} \right)\) and between the number of mutations and age \(\left( R = 0.53, P = .02; \text{Fig. 6D}; \text{Supplementary Tables S1 and S34} \right)\). Moreover, comparison of the slope of the linear regression between passenger mutations and age between at-risk mucosa and adenomas \(\left( 0.94 \text{ vs. } 1.36 \right)\) demonstrated that adenomas harboring second hits in \(\text{APC}\) accumulated more passenger mutations (Fig. 6E), thus indicating that haploinsufficiency in \(\text{APC}\) (i.e.,
presence of one hit in APC in at-risk FAP normal mucosa) is not sufficient to foster carcinogenesis.

Finally, the average numbers of exonic mutations per sample were 12 (range 1-56) in at-risk mucosa and 51 (range 2-101) in adenomas with a somatic hit in APC (Supplementary Table S30). On the basis of this data, we extrapolated that 23% of the mutations existed before carcinogenesis started in a context of germline APC alterations. To further explore the relationship of drivers and passengers, we compared our results with estimates proposed by Bozic and colleagues (17) using mathematical models describing a logarithmic relation between drivers and accumulation of passengers. We observed a significant correlation between estimated and observed numbers of passengers in adenomas (7.6 vs. 6.9 passenger mutations per driver, respectively, \( R = 0.59, P = 0.01 \); Fig. 6F; Supplementary Table S34), indicating that this model provides accurate estimates based on our data (17).

**Discussion**

We have here provided the first comprehensive annotation of the genomic landscape of APC-driven adenomas from FAP patients using next-generation sequencing technologies. Our classification system of mutation by tiers based on bioinformatic prediction tools helps to characterize a set of recurrent genes involved in early tumor progression. Somatic hits in APC, KRAS, or FBXW7 were recurrently detected in our samples and classified into tiers 1 and 2, which is consistent with the assignment of a driver status, thus validating our tier classification. In addition, we found novel candidate genes that merit further investigation as potential cooperators of APC deficiency in early stages of carcinogenesis, including EWSR1, CNOT3, and PCMTD1, among others. Some of these genes have been shown previously to harbor somatic mutations in other cancer types. For example, CNOT3 has been reported in T-cell acute lymphoblastic leukemia, pancreatic cancer, and bladder cancer (18, 19) and ARID1A in ovarian and endometrial cancers (15). In addition, we observed recurrent mutations in TTN and CDC27 that have been previously reported as spurious with respect to driver status, due to their genomic size or accumulation of systematic technical artifacts (20).

Our comparative analysis of genomic imbalances in adenomas and stage I carcinomas provides an opportunity to detect aberrant regions in both types of samples, suggestive of their importance in occurring early in the development of colorectal cancer. We observed losses in 5q and events in chromosomes 7, 13, and 20 in both adenomas and carcinomas. Loss of 5q is a common
The gene most frequently mutated in our series was APC (chr18: 48556582–48611411). This observation is related to a recent observation that loss of APC is associated with poor survival in colorectal cancer patients (23).

Finally, we identified AI in 56% of adenomas, a proportion lower than previously reported by others (24) and which could be explained by the use of a different technical approach, such as comparative genomic hybridization array or by sample characteristics. The gene most frequently mutated in our series was APC, for which 72% of the samples presented a somatic second hit, which is in agreement with the frequency previously observed by other groups (25). We were not able to detect a second hit in APC or in a gene involved in the WNT signaling pathway in 4 adenomas. These adenomas presented with lower mutational rates compared with the rest that harbored somatic WNT pathway alterations and were also similar to the mutational rates detected in normal mucosa samples, thus suggesting that these lesions were at a very early stage of carcinogenesis. In line with this hypothesis, these 4 lesions would have a higher than the average percentage of normal mucosa, thus masking the presence of a somatic APC alteration. Higher depth in the sequencing would have helped to increase the resolution and uncover an APC mutation in a sparse but upcoming clones.

The majority of somatic APC alterations observed in our cohort were truncating mutations localized in the MCR coupled with germline events spread through other regions of the gene, thus highlighting the dependency between germline and somatic APC events (25–27). Moreover, results from the pathway enrichment analysis indicated that deregulation of the WNT pathway was the most prominent derangement, which is in line with the results of the TCGA analysis of colorectal cancer (7). Indeed, APC is an essential component of the WNT pathway, regulating β-catenin, migration, and polarity and having a central role in maintaining the stem cell fate of the intestinal crypt (28, 29). In addition, we detected cooccurrent mutations in more than 1 gene in other well-known WNT pathway genes, such as TCF7L2, FBXW7, ARID1A, CTNNB1, AXIN2, and SOX9 in 28% of adenomas. Interestingly, we did...
not observe other major oncogenic pathways deregulated in significant numbers of our adenomas, thus highlighting the almost exclusive dependency on WNT pathway events at this stage of carcinogenesis. Therefore, these findings further support the development of therapeutic strategies targeting WNT inhibitors for colorectal cancer prevention through adenoma interception.

To further explore the dynamics of tumor growth, we used computational tools to determine and characterize the presence of multiclonality in colorectal adenomas. Our findings indicated the presence of multiple clones in 72% of adenomas. The fact that we did not detect a significant difference between adenomas and stage I colorectal cancer in the number of clones suggests that clonality originates very early in carcinogenesis and not from late clonal expansions. These observations could suggest that the expansion of an APC-driven clone constitutes the founder progenitor of the tumor cell population early in carcinogenesis. The mutational profile of this founder clone contains the catalog of “public” mutations that are thus subsequently present in all tumor cells and include those cooperating with APC. The progeny of this major clone will then acquire additional mutations that are the source of private low-frequency events that remain less abundant, as these minor subclones are marginally distributed within the geography of the tumor mass. Our analysis is not adequate to observe geographic variations of genomic alterations in adenomas. However, it provides indirect support to this model of colorectal carcinogenesis (30).

To further explore the dynamics of tumor growth, we used WES of at-risk mucosa samples. It has been proposed by others that the number of mutations in tumors of self-renewing tissues, such as the large intestine, should be positively correlated with the age of the patient and that a large fraction (at least 50%) of the somatic mutations arises before tumor initiation (9). Our data confirmed this correlation between the number of mutations in at-risk mucosa (and adenomas) and age, even in the context of accelerated carcinogenesis due to the presence of germline APC mutation. Moreover, our sequencing results for at-risk mucosa have allowed us to calculate that 23% of mutations in adenomas were present prior to acquisition of the first initiating somatic driver event, which is the loss of APC (second hit). A possible explanation for the discrepancy between our calculations and the fraction reported previously by others (9) is that our samples already harbor a germline hit in APC, thus requiring only the acquisition of 1 somatic event to fully initiate tumorigenesis. Therefore, the timeline for adenoma formation is much shorter in FAP adenomas, and a much lower number of somatic alterations will be present in the at-risk mucosa prior to the acquisition of the tumor initiation event. This fact also suggests that only one hit in APC is not sufficient to accelerate the accumulation of mutations in the normal mucosa. Additional analysis comparing normal mucosa samples from sporadic individuals (general population) and FAP patients will be necessary to investigate the timeline of acquisition, role, and function of these somatic alterations accumulating in a normal tissue. Furthermore, we observed the presence of a relatively lower proportion of driver events in at-risk mucosa compared with adenomas, and these events were identified in genes previously reported to be spurious (20) or not related to colorectal carcinogenesis (such as CDC27, PLEKHN1, and PHOX2B), thus suggesting that these drivers detected in at-risk mucosa are not biologically effective or may be “false positives” (as drivers). Only 2 genes found mutated in the at-risk mucosa samples, WNT11 (31) and ARID1B (32), have been previously reported to play a role in colorectal carcinogenesis, thus indicating that these mucosa samples have already acquired a growth advantage.

In summary, we performed a comprehensive analysis of the genomic landscape of colorectal adenomas and the first such examination of the surrounding mucosa as an independent at-risk tissue to assess the contribution of the accumulated somatic variation to carcinogenesis. Our data show that the majority of recurrently mutated genes identified in adenomas are involved in the WNT pathway, confirming that APC loss is sufficient for the early growth of adenomas (33) and suggesting that future chemopreventive strategies should focus on developing drugs that can effectively target this pathway. Finally, using in silico analyses, we identified the presence of clonality in adenomas and estimated that 23% of somatic mutations in APC-driven adenomas were present before tumor initiation, thus highlighting the fact that the occurrence of at least 1 driver event, in this case the acquisition of a somatic APC aberration (second hit in a tumor suppressor), is necessary to trigger the accumulation of additional driver and passenger mutations that will ultimately propel carcinogenesis.

Disclosure of Potential Conflicts of Interest
No potential conflict of interest were disclosed.

Authors’ Contributions
Conception and design: E. Borras, D.A. Jones, E.T. Hawk, E. Vilar
Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): E. Borras, F.A. San Lucas, R. Zhou, G. Masand, M.E. Mork, Y.N. You, M.W. Taggart, D.A. Jones, G.E. Davies, E.A. Ehli, P.M. Lynch, G. Capella, E. Vilar
Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): E. Borras, F.A. San Lucas, K. Chang, J. Fowler, G. Capella, P. Scheet, E. Vilar
Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): F.A. San Lucas, K. Chang, M.E. Mork, Y.N. You, E.A. Ehli, M.W. Taggart
Study supervision: P. Scheet, E. Vilar
Other (proofreading): J. Fowler

Acknowledgments
The authors thank the patients and their families for their participation and the staff of the Sequencing and Microarray Facility at The University of Texas MD Anderson Cancer Center for the generation of the exome sequencing data. The authors also thank the Clinical Cancer Prevention Research Core for their assistance in obtaining the samples for this study.

Grant Support
This work was supported by grants R03CA176788 (NIH/NCI), the MD Anderson Cancer Center Institutional Research Grant (IRG) Program, and a gift from the Feinberg Family (to E. Vilar); ROI1GM92666 and R01HG005859 (NIH, to P. Scheet); the Janice Davis Gordon Memorial Postdoctoral Fellowship in Colorectal Cancer Prevention (Division of Cancer Prevention/MD Anderson Cancer Center, to E. Borras); the Schissler Foundation Fellowship (The University of Texas Graduate School of Biomedical Sciences) and Translational
Molecular Pathology Fellowship (M.D. Anderson Cancer Center; to F.A. San Lucas); Cancer Prevention Educational Award (R25T CA057730, NIH/NCI; to K. Chang); SAI2012-33636 (Spanish Ministry of Economy and Competitiveness) and the Scientific Foundation Asociacion Española Contra el Cancer (to G. Capella); R01CA76329 (NIH/NCI; to W. Edelmann); and P3DC016672 (NIH/NCI) to the M.D. Anderson Cancer Center Core Support Grant.

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 March 21, 2016; accepted April 7, 2016; published OnlineFirst May 24, 2016.

References

Genomic Landscape of Colorectal Mucosa and Adenomas

Ester Borras, F. Anthony San Lucas, Kyle Chang, et al.


Access the most recent version of this article at:
doi:10.1158/1940-6207.CAPR-16-0081

Access the most recent supplemental material at:
http://cancerpreventionresearch.aacrjournals.org/content/suppl/2021/03/13/1940-6207.CAPR-16-0081.DC1

This article cites 32 articles, 9 of which you can access for free at:
http://cancerpreventionresearch.aacrjournals.org/content/9/6/417.full#ref-list-1

This article has been cited by 16 HighWire-hosted articles. Access the articles at:
http://cancerpreventionresearch.aacrjournals.org/content/9/6/417.full#related-urls

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

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

To request permission to re-use all or part of this article, use this link
http://cancerpreventionresearch.aacrjournals.org/content/9/6/417.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.