HHS Public Access

Author manuscript

Nat Biotechnol. Author manuscript; available in PMC 2017 January 11.

Published in final edited form as:

Nat Biotechnol. 2016 August ; 34(8): 845–851. doi:10.1038/nbt.3586.

A recellularized human colon model identifies cancer driver genes

Huanhuan Joyce Chen[1,‡], **Zhubo Wei**[2,‡], **Jian Sun**[3], **Asmita Bhattacharya**[4], **David J Savage**[5], **Rita Serda**[5], **Yuri Mackeyev**[7], **Steven A. Curley**[6], **Pengcheng Bu**[1], **Lihua Wang**[8], **Shuibing Chen**[9], **Leona Cohen-Gould**[10], **Emina Huang**[11,12], **Xiling Shen**[1], **Steven M. Lipkin**[3], **Neal G. Copeland**[2], **Nancy A. Jenkins**[2,*], and **Michael L. Shuler**[1,*]

[1]Department of Biomedical Engineering, Cornell University, NY

[2]Cancer Research Program, Houston Methodist Research Institute, TX

[3]Departments of Medicine, Genetic Medicine, Weill Cornell Medical College, NY

[4]Genetics, Genomics and Development, Cornell University, NY

[5]University of Texas Medical School at Houston, TX

[6]Department of Surgery, Baylor College of Medicine, TX

[7]Department of Chemistry, Rice University, TX

[8]Department of Biological and Environmental Engineering, Cornell University, NY

[9]Chemical Biology in Surgery, Surgery, Weill Cornell Medical College, NY

[10]Department of Biochemistry, Weill Cornell Medical College, NY

[11]Department of Stem Cell Biology and Regenerative Medicine, Cleveland Clinic, OH

*To whom correspondence should be addressed: Michael L. Shuler PhD, Biomedical Engineering, 115 Weill Hall, Cornell University, Ithaca, NY. 14850, Phone: 607-255-7577, mls50@cornell.edu; Nancy A. Jenkins PhD, Cancer Biology Program, Houston Methodist Research Institute, 6670 Bertner Ave, Houston, TX 77030, Phone: 713-441-6623, njenkins2@houstonmethodist.org.

‡Equal contribution

Competing Financial Interests The authors declare no competing financial interests.

Author Contributions H.J.C, Z.W., N.G.C., N.A.J., and M.L.S. conceived the concept, designed the experiments, and co-wrote the manuscript; H.J.C, Z.W., A.B., D.J. S., P.B., L.W., and L.C. G., performed the experiments; J.S., Z.W., N. G. C., and N. A. J., contributed to bioinformatics analyses.

Accession ID: Illumina sequence data are deposited in NCBI SRA as SRR1634458 and SRR3342115.

Supplemental Information: Statistical analyses: Sample sizes for all figures and tables were estimated based on our previous studies [38, 49]. For mouse experiments, animals were randomly assigned in each experimental group and no animals were excluded from the analyses. For each set of experiments, samples were prepared for all experimental arms at the same time. All statistical tests are 2-sided. No adjustments were made for multiple comparisons. Both PIs (Shuler and Jenkins) and the Study Statistician (Sun) were blinded to experimental allocations among different experimental arms for all experiments. For all parametric statistical analyses, data were determined to be normally distributed by the D'Agostino-Pearson test. For all parametric and non-parametric tests, variances were similar between groups being compared. For comparison between experimental and control groups at a specific time point or tissue site in Figures 1, Figure 2, Figure 3, Figure 4, Supplementary Figures 4, 6, 8, 9, 10 and Supplementary Table 1, 2-sided Student t- tests, 2-sided Mann-Whitney (MW) tests and Fisher's exact tests were used. All cells were purchased from ATCC in the past 2 years or derived from patients within 10 passages and were negative for mycoplasma.

Study approval: Use of patient tissues was approved by the Institutional Review Board (IRB) of Weill Cornell Medical College and the IRB of Cleveland Clinic, and informed Consent obtained for each participant. All animal protocols in this study were approved by the IACUC committees of Weill Cornell Medical College, Cornell University.

[12]Department of Colorectal Surgery, Cleveland Clinic, OH

Abstract

Refined cancer models are needed to bridge the gap between cell-line, animal and clinical research. Here we describe the engineering of an organotypic colon cancer model by recellularization of a native human matrix that contains cell-populated mucosa and an intact muscularis mucosa layer. This *ex vivo* system recapitulates the pathophysiological progression from *APC*-mutant neoplasia to submucosal invasive tumor. We used it to perform a *Sleeping Beauty* transposon mutagenesis screen to identify genes that cooperate with mutant *APC* in driving invasive neoplasia. 38 candidate invasion driver genes were identified, 17 of which have been previously implicated in **colorectal cancer** progression, including *TCF7L2*, *TWIST2*, *MSH2*, *DCC* and *EPHB1/2*. Six invasion driver genes that to our knowledge have not been previously described were validated *in vitro* using cell proliferation, migration and invasion assays, and *ex vivo* using recellularized human colon. These results demonstrate the utility of our organoid model for studying cancer biology.

Clinical translation of basic cancer research is hindered by several obstacles[1]. One obstacle arises from the intrinsic limitations of experimental systems, which may not demonstrate concordance with human studies. For example, conventional two-dimensional (2D) cell culture models do not maintain the interactions of tumor cells with the extracellular matrix (ECM) and the tissue microenvironment, which are essential for tumor pathogenesis [2]. Similarly, animal models are expensive and time-consuming, and may not have the appropriate resolution and sensitivity required to track the dynamics of cancer progression. In addition, because of species variation, animal studies can differ from humans with regard to the requirements for oncogenic transformation. Another problem stems from the fact that tumors evolve heterogeneously, and the large number of passenger mutations they accumulate can confound the identification of driver genes [3, 4]. Although various high-throughput genetic approaches have facilitated the identification of genome-wide alterations in cancers[5, 6], testing the role of each gene in cancer pathogenesis in a native environment is a crucial, albeit difficult task, given the large number of low-frequency mutations found in most cancer genomes.

The technology of intestinal organoid has been largely advanced for modeling both normal and disease tissues [7-9]. However, these systems usually lack native tissue structures and ECM, thus are still far away from mimicking physiological conditions. Recent technical advances have made it possible to isolate natural cellular matrix with preserved ECM and normal three-dimensional (3D) tissue architecture, providing a potential new approach for producing more physiologically relevant models of cancer [10-12]. Ridky et al. have successfully established ex vivo systems recapitulating the normal-neoplasia-invasion sequence using modified human epidermis, oropharynx, esophagus and cervix epithelial cells in a human three-dimensional (3D) tissue environment [13]. Here we create an physiologically active ex vivo model of the human colon that mimics physiological conditions by first decellularizing normal human colon tissue under conditions that retain the tissue's complete geometry with a well-preserved ECM, a relatively integral vascular network and intact muscularis mucosa, followed by reseeding with primary colonic

epithelial cells, endothelial cells and myofibroblasts. We use the model to study colorectal cancer (CRC) progression by recellularizing the colon matrix with epithelial cells carrying mutations in genes known to be important in CRC progression, such as *APC* and *KRAS*. Similar to what occurs in CRC, this led to the development of neoplasia *in situ* with little submucosal invasion, whereas reseeding with cells carrying mutant *APC* and an activating mutation in *KRAS* (*KRAS*^{G12D})[14], in the context of reduced TGF-β signaling[15, 16], led to the development of invasive submucosal tumors. We used the *ex vivo* model to perform a *Sleeping Beauty* (*SB*) transposon mutagenesis screen designed to identify genes that cooperate with mutant *APC* in driving invasive neoplasia.

Results

Preparation of decellularized human colon

The cellular components from fresh normal human colon tissue (~5 cm^3) were removed by sodium dodecyl sulfate (SDS) treatment[10] followed by Triton-X100 washing, resulting in the complete decellularization of the colon tissue (Figure 1 A-D). The DNA content of these acellular colon scaffolds was decreased more than 95% compared to that found in normal colon (Supplementary Table 1), whereas no differences were observed in the quantities of the four main ECM proteins[17] – glycosaminoglycan (GAG), collagen type I, laminin and fibronectin (Supplementary Table 1, Figure 1H).

The removal of most cellular components was further confirmed by the observation that F-actin and cellular nuclei were undetectable in decellularized scaffolds by immunohistochemistry (Figure 1H). The decellularized scaffolds successfully preserved the tissue architecture, main vasculature and crypt niches of the original colon (Figure 1E, 1F, 1G).

Recellularization of the acellular human colon matrix

A key feature that makes the *ex vivo* colon model useful for identifying CRC driver genes is that it has a genetically defined epithelium, which is free of malignant origins and secondary genetic alterations. To recellularize the acellular colon matrix we first established primary cultures of the three main cellular components of the colon—epithelial cells, endothelial cells and myofibroblasts. Epithelial cells and myofibroblasts cells were obtained from routine colonoscopy patient samples, whereas endothelial cells were purchased from commercial sources. Epithelial cells were grown under conditions normally used for organoids[9] in order to help retrain their stem cell component and differentiation potential. These cells were then modified with the human telomerase reverse transcriptase (hTERT) to prevent pre-mature senescence and ensure their long-term growth potential (Supplementary Figure 1A, top panel). DNA sequencing of these primary colonic epithelial cells (hCECs) confirmed that they did not carry mutations in hotspot regions of *APC*, *KRAS* or *TP53* (data not shown). When grown in 3D culture, these hCECs formed organoids with microcrypt-like structures (Supplementary Figure 1A, bottom panel), and were capable of self-renewal and multilineage differentiation (Supplementary Figure 1B)[18].

Recellularization of the colon matrix was done in several discrete steps. First, the large bowel mucosa with intact crypt niches and muscularis mucosa was mechanically separated from the submucosa by forceps (Supplementary Figure 2A, 2B; Supplementary Video 1). This made it possible to reintroduce the hCECs, endothelial cells and myofibroblasts into their normal physiological locations within the acellular colon matrix (Figure 2B). For populating the mucosa, a microinjection pipette was filled with endothelial cells (Supplementary Figures 3B, 2C; Supplementary Video 2) and the cells were injected into the mucosa through the side of the tissue (0.1 million/cm^2) (Supplementary Figure 2D; Supplementary Video 3). After the endothelial cells grew out, the hCECs (0.2 million/cm^2) were placed on top of the crypt surface and seeded into the crypt niches through gravity precipitation. Finally, colonic myofibroblasts (Supplementary Figure 3A) (0.1 million/cm^2) were seeded and grown on the muscularis mucosa, which was exposed through separation from the submucosa. In addition to the muscularis mucosa, myofibroblasts can also be found in the stroma surrounding the colonic crypts in the normal colon. Due to technical limitations, however, we were unable to repopulate the mucosa pericryptal with myofibroblasts and focused instead on repopulating the muscularis mucosa layer with colonic myofibroblasts. The *ex vivo* colon was then generated by assembling together the mucosa and submucosa layers (Figure 2B; Supplementary Figure 2E, 2F). The process of decellularization and recellularization preserved native colon ECM and its mechanical and chemical properties, including intestinal fibers composed of laminin and collagen in the correct orientation, composition and microstructure (Supplementary Figure 4A). Notably, the recellularized colon matrix retained the ultrastructural elements of the muscularis mucosa, the native barrier for the invasion of the malignant CRC submucosa [19], which was not substantially different in thickness from that found in native tissues (Figure 1G; Supplementary Figure 4B, 4C). The recellularized *ex vivo* colon not only was viable but also contained Lgr5+ stem cells in addition to other differentiated cell types normally found in the colon (Supplementary Figure 5). Physiologically similar to native ones, our *ex vivo* colon crypts contained dynamic proliferation capacity (Supplementary Figure 5; Supplementary Figure 6B), actively reproduced WNT/β-catenin signaling that is essential for intestinal homeostasis (Supplementary Figure 6A), and expressed the major mucin protein MUC2 (Supplementary Figure 6C), which indicates the secretory activity of the recellularized crypts. We also detected the growth capacity of the 3 cell lines in recellularized mucosa by Ki67 staining and observed that endothelial cells were the first cell type undergoing growth dormancy/senescence after the reconstituted mucosa were cultured for 8-10 weeks (data not shown).

Establishment of a CRC progression model

CRC develops through the stepwise selection of genetic changes, with the initiating or gate-keeping mutation being bi-allelic loss of *APC*, which occurs in >80% of the cases[20]. This is followed by mutations in other genes, such as activating mutations in *KRAS* at the early adenoma to intermediate adenoma stage, and loss of function mutations in *SMAD4* and *TP53*, at the intermediate adenoma to adenocarcinoma stage[21, 22]. To determine whether it was possible to transform hCECs into tumor cells, we infected these cells with a lentivirus that expresses an *APC* small-hairpin RNA (shRNA) in addition to constitutively active *KRAS* (*$KRAS^{G12D}$*) (Supplementary Figure 7). These cells were then subcutaneously

injected into immunodeficient mice and two weeks later, TGF-β signaling was blocked through the intraperitoneal injection of a TGF-β receptor inhibitor (TGF-β RI Inhibitor III) [15]. This was done to mimic the inactivation of TGF-β signaling, which drives CRC progression, especially submucosal invasion that can be caused by a simple genetic alteration in this pathway along with mutated APC, as revealed by a recent study [23]. Within 6-8 weeks after injection, tumors were visible at 60-80% of the injection sites and presented with typical epithelial CRC-like features, including crypt-like lumen structures and microvasculature formation (Supplementary Figure 8). In control experiments, injection of parental hCECs did not induce any tumors (Supplementary Figure 8, Fisher's exact test, P=0.00025).

Based upon these findings, we asked whether we could model CRC progression using the *ex vivo* colon model (Figure 2A). Similar to normal human colon (Figure 2C), the acellular colon recellularized with unmodified hCECs formed a single-cellular layer in crypt niches, and the cells were tightly attached to the basal membrane and stromal ECM (Figure 2D). By contrast, when the acellular colon was recellularized with hCECs expressing *APC* shRNA, dysplasia was visible in the mucosa epithelium, a phenotype typically seen in early stage adenomas, with which cells underwent fast proliferation to form multi-cellular layers and distorted crypt structures with multi-luminal fusions (Figure 2E, 2F). These dysplastic lesions formed in only 3 weeks and were restricted to the mucosa layer (P<0.001, one-way ANOVA). The dysplasia grew larger without invasive events when the acellular colon was recellularized with *APC*-shRNA hCECs in addition to activated $KRAS^{G12D}$ (Figure 2G, 2H). However, when the acellular colon was recellularized with *APC*-shRNA hCECs combined with activated $KRAS^{G12D}$ and subsequent inhibition of TGF-β signaling pathway, large adenomas developed, broke through the muscularis mucosa, and invaded into the submucosa, a key feature of malignant CRC (Figure 2I-K) (P=0.0024, one-way ANOVA). This malignant transformation occurred within 4 weeks from the onset of the dysplastic lesions that formed with *APC*-shRNA hCECs.

Collectively, these results show that this type of organotypic colon cancer model, developed using human natural matrix and genetically defined primary colon cells, is capable of recapitulating the key features of CRC initiation, progression and malignant transition, from non-invasive neoplasia to invasive submucosal tumor. The different stages of oncogenic transformation can be generated relatively easily in this model, within weeks, correlating well with histological features and avoiding redundant passenger mutations. This *ex vivo* system can also provide single-cell resolution and time-lapse sensitivity for anatomically tracking and dissecting CRC progression. These qualities make the *ex vivo* CRC model potentially useful for high-throughput genetic or therapeutic studies.

Identification of genes driving invasive neoplasia

In previous studies, we used *SB* mutagenesis to model several types of cancer in mice, thereby identifying many novel cancer genes and cancer signaling pathways[24-27]. Here we extend the *SB* mutagenesis system to model cancer directly in human tissues. To create an active human *SB* transposition system covering a broad target range, we made improvements to the transposon (T2/Onc) we used in previous studies. We added GFP and

neomycin selection markers, which make it possible to directly select for cells carrying the transposon. In addition to these selection markers, the transposon carries a murine stem cell virus promoter and downstream splice acceptor site, which are used to activate expression of proto-oncogenes, as well as splice acceptor sites in both orientations and bi-directional polyA signals, which are used to inactivate tumor suppressor genes. This transposon is thus bi-functional and can both activate cellular proto-oncogenes and inactivate tumor suppressor genes. Finally, in these experiments we also used a more highly active *SB* transposase (*SB100X*) than used previously [28]. Using this modified transposition system, we found that 500 ng of transposon plasmid combined with 100 ng of transposase plasmid, produced the highest enzyme activity, while avoiding the effect of overproduction inhibition (OPI) in hCECs[28] (Supplementary Figure 9A). Using these conditions, each transfected hCEC contains on average 4 ± 3 transposon copies inserted in its genome after 4 weeks of antibiotic selection (Supplementary Figure 9B).

To determine whether we could use SB mutagenesis to identify genes driving invasive neoplasia in the *ex vivo* colon model, we used the modified *SB* transposition system to mutagenize *APC* shRNA-expressing hCECs, and then recellularized the acellular colon matrix with these mutagenized cells (Figure 3A). Within 6-7 weeks following recellularization, multiple tumors undergoing submucosal invasion could be seen in the recellularized colons (Figure 3D, 3E). This is in contrast to what occurs with unmutagenized *APC*-shRNA hCECs, where little submucosal invasion occurs (Figure 3B, 3C). The average number of tumors undergoing submucosal invasion formed in 10 cm^2 matrix was 5.5 ± 1.9 for mutagenized cells, which was significantly higher than the average number of 0.83 ± 0.75 for unmutagenized cells (Figure 3F) (P=0.002, Student t test). *SB*-based mutagenesis thus appears to have mutated genes that drive the invasion of *APC* mutant cells. Colon tumors in multiple disease stages were also observed. in this *ex vivo* system, and single-cell resolution could be obtained from time-lapse tracking (Figure 3G - I).

To identify genes mutated by *SB* that promote invasive neoplasia, we excised 21 invasive neoplasias from 15 recellularized colon matrixes by laser-capture microdissection, using a procedure that minimized the contamination from adjacent non-invasive tumor tissues. Transposon junction fragments were then amplified by ligation-mediated PCR and the amplification products sequenced by Illumina sequencing [29]. Based on the distribution of read depths at the insertion sites, it was clear that the majority of insertions represented subclonal insertions. We therefore considered only the top 10% of insertions with the highest read depth to be clonal insertions. We then mapped these clonal insertion sites to the human genome and picked the sites that were located within 1,000 bp of a known transcript in RefGene. We subsequently confirmed the sequences of these transposon insertion sites by Sanger DNA sequencing. Collectively, these studies identified 38 candidate invasion driver genes from 21 invasive neoplasias (Supplementary Table 2). Notably, 17 of the 38 genes have already been implicated as contributing to CRC progression, including *TCF7L2*, which functions in WNT1 signaling[30, 31]; *MSH2*, a DNA mismatch repair gene[5, 22]; *TWIST2*, a gene important for epithelial to mesenchymal transition[32]; *JAK1* and *STAT3*, which function in JAK-STAT signaling[33]; and *DCC*, a gene commonly deleted in CRC[34] (Table 1).

Most of these candidate invasion driver genes were mutated in a single invasive neoplasia (Supplementary Table 2). However, six genes (*CSTF3*, *GRM8*, *KDM2B*, *LATS2*, *PAX7*, and *TCF7L2*) were mutated in two invasive neoplasias, and one gene (*PTPRD*) was mutated in three invasive neoplasias. In the case of *CSTF3*, *GRM8*, *KDM2B*, *PAX7*, and *TCF7L2*, *SB* was inserted at the same nucleotide site in each invasive neoplasia, indicating that these invasive neoplasias were generated from the same clone of *SB*-mutagenized cells. For *LATS2* and *PTPRD*, two invasive neoplasias for each gene were identified in which SB was located at different nucleotide sites (Supplementary Table 2), suggesting that these neoplasias were derived from different clones of *SB*-mutagenized cells. One of these genes, *Drosophila* large tumor suppressor homolog 2 (*LATS2*), encodes a putative serine/threonine kinase that physically interacts with MDM2 to inhibit p53 ubiquitination and promote p53 activation[35, 36]. Expression of the other gene, protein-tyrosine phosphatase receptor delta (*PTPRD*), is down-regulated in highly invasive cancers, and has been shown to suppress colon cancer cell migration in coordination with CD44[37]. Moreover, 55% of the identified candidate invasion driver genes (21/38) are mutated in ≥5% human CRC patients and/or show a 2-fold mRNA levels expression changes in human CRC patients (Supplementary Table 2) (mutation frequency data: TCGA[5]; mRNA level data: TCGA RNA-seq, http://firebrowse.org). Collectively, these results suggest that insertional mutations in these genes were selected because of their influence on invasive cell growth.

Although the majority of genes identified by *SB* in solid tumors are tumor suppressor genes, genes such as *TWIST2, JAK1, ROCK1* and *STAT3* are more likely to be oncogenes based on published studies. However, due to the limited amounts of invasive tissue obtained from laser-capture microdissection and the potential RNA degradation the procedure may cause, the samples were only sufficient for making DNA in order to PCR amplify and sequence transposon insertion sites. It was therefore not possible to determine whether SB insertion in these genes resulted in activation or inactivation of their expression. Because these genes have already been implicated in colorectal cancer, we decided to focus our validation studies on the 21 genes we identified that have not previously been implicated in CRC.

In initial experiments, we used siRNA to knockdown the expression of these 21 genes in *APC*-shRNA hCECs or in SW480 colon cancer cells, and then measured the effect of the knockdown on cell proliferation, migration and invasion. We used siRNA knockdown because the majority of genes identified by *SB* in solid tumors are tumor suppressor genes, and we chose the SW480 cell line because it was derived from an early-stage adenocarcinoma that harbors a mutation in *APC*. As a positive control, we generated siRNA against *LATS2*, a known tumor suppressor identified in the screen, in addition to five randomly chosen negative control genes (shown in blue in Table 2). As might have been expected, *LATS2* scored positive *in vitro* cell proliferation and invasion assays, whereas none of the negative control genes scored positive in any assay (Table 2). Among the 21 candidate invasion driver genes, 9 scored positive in one or more assays in hCECs and/or SW480 cells (Table 2).

The 9 genes that scored positive were further validated using the *ex vivo* colon model, in addition to *LATS2,* which was again included as a positive control, and *CSTF3*, which was included as a negative control, because it did not score positive in any of the *in vitro* assays

(Table 2). shRNAs against these 11 genes were generated and used to stably downregulate their expression in *APC*-shRNA hCECs, and the hCECs were then used to recellularize acellular colon matrices. Among the 11 genes tested, 6 genes including *ASXL2*, *CAMTA1*, *DDX20*, *FXR1*, *MITF* and *PAX7*, along with the positive control gene *LATS2*, significantly promoted submucosal invasion in the recellularized colon model, whereas the negative control gene *CSTF3* scored negative (Supplementary Figure 10A, 10B). Notably, as shown in Supplementary Table 3, most invasive neoplasia samples (85%, 18/21) harbored at least one mutated candidate invasion driver gene that is previously implicated in CRC progression and/or functionally validated in this study. In addition, two or more driver genes that are previously implicated in CRC progression and/or functionally validated in this study were identified in 8 of the 21 samples, suggesting potential collaboration among these genes in the invasion events we observed. Collectively, these results suggest that *SB* mutagenesis in the *ex vivo* colon model can identify genes important for neoplastic cell invasion.

Discussion

We describe here a tissue-engineering method that makes it possible to perform unbiased forward genetics screens in human tissues under physiologically relevant conditions. The decellularized human native colon matrix provides major tissue-relevant elements, including complex tissue structure, cell-matrix interactions and physiological co-location of multiple types of differentiated cells. This makes this model more desirable than conventional assays of cell migration or invasion though synthetic matrigel or collagen layers, which does not mimic real tissues. Compared to the complex bioreactor and medium formula used to generate native functional organs for orthotopic transplantation, we used simpler systems that fulfill minimal requirements for developing tissue-level tumor models, enabling low-cost and large-scale cancer studies. Currently, only a few animal models have been developed to study malignant events in late-stage CRC [38]. By introducing pathologically paired genetic elements into this model, the *ex vivo* model was able to recapitulate a number of features associated with CRC progression. Malignant transformation also occurred rapidly, within 6-7 weeks post recellularization, further enhancing the value of this *ex vivo* model as a complement to current cancer models.

Transposon mutagenesis provides an unbiased high-throughput genetic tool for cancer gene discovery[24]. Although *SB* mutagenesis has been used previously to model different types of cancers in mice, to our knowledge no previous study has shown that *SB* is capable of inducing cancer under close to native human conditions. Identification of genes driving invasive neoplasia can identify new drug targets and diagnostic markers. However, identifying and studying genes driving malignancy still remains a challenge. The *ex vivo* colon model provides a tractable system that enables one to recapitulate the dynamics of malignant transformation.

In a forward genetic screen, we used this *ex vivo* model to identify 38 candidate genes driving CRC submucosal invasion through cooperation with mutant *APC*, demonstrating the effectiveness of the *ex vivo* model as a functional platform for identifying genes driving tumor cell invasion. One of the more notable genes we identified was additional sex combs-like 2 (*ASXL2*)[39]. ASXL family members are scaffolding proteins that function as

epigenetic regulators by recruiting polycomb-group repressor complexes (PRC) and trithorax-group (trxG) activator complexes to DNA. They also play roles in histone modification by assembling transcription factors to specific genetic domains. ASXL2, ASXL1 (another ASXL family member), BRCA1 and YY1 are the binding partners of BAP1, a nuclear deubiquitinating enzyme and tumor suppressor gene that mainly functions during metastasis. Truncation mutations of *ASXL2* have been correlated with poor prognosis in prostate cancer, pancreatic cancer and breast cancer. Additionally, *ASXL1* is involved in the malignant progression of CRC with microsatellite instability (MSI)[40].

Another gene we identified was *CAMTA1*. This is a putative tumor suppressor gene in neuronal cancers[41], whose expression decreases glioblastoma cell growth by stimulating expression of the anti-proliferative peptide NPPA. *CAMTA1* also regulates neuroblastoma cell mobility through increasing the expression of β3 tubulin (TUBB3), microtubule associated protein 2 (MAP2) and neurofilament light chain (NEFL). Notably, decreased expression of *CAMTA1* is observed frequently in CRC and reported to be an independent indicator for poor patient survival[42]. Consistent with previous findings, both our *in vitro* and *ex vivo* functional assays showed significantly increased cell proliferation and invasion in hCEC and SW480 lines with decreased *CAMTA1* expression, further indicating the potential tumor suppressor role of *CAMTA1* in CRC malignant transformation.

We observed that 8 of the 21 samples had mutations in two or more candidate invasion driver genes that are previously implicated in CRC progression and/or functionally validated in this study. These combinations of driver genes may explain how the invasion events occurred. In future studies, the potential cooperative relationships of the genes found in the same samples will be investigated for better understanding of CRC progression that might eventually help develop novel therapies.

Because of the heterogeneity generated during tumor evolution, and the genetic diversity that occurs between individuals, the 38 candidate genes identified here to some extent may be specific to the patients from whom the primary colonic epithelial cells were derived. Theoretically, this model system should enable the discovery of a broader range of novel cancer genes by deriving epithelial cells from multiple patient cell sources or by generating initial pools of mutagenized cells covering different mutation profiles through varying rounds of *SB* mutagenesis.

Although mutations in tumors are now routinely identified on a genome-wide scale, elucidating the roles of these genetic alterations in tumorigenesis is still challenging. The *ex vivo* colon cancer model complements existing cell line systems and animal models for multi-dimensionally testing the mechanistic roles of recurrent human cancer mutations. Our model and the recently developed tractable biomimetic systems such as organoid systems using human and mouse intestinal tissues[7, 8] will help us better understand the mechanisms of CRC progression at different stages and develop novel therapeutics. Our model also offers opportunities to create specialized physiological microenvironments for mimicking clinical diseases, for example, by incorporating engineering of vascular networks, the immune system and organ-specific microbes. In addition, our approach might be extended to the generation of other types of *ex vivo* cancer organs, such as lung, liver,

skin and kidney. It could also be applied in other areas of oncology, from developing biomarkers for diagnosis and prognosis to screening drugs, chemicals, pathogens and toxins for personalized medicine.

Methods

Decellularization of human colon tissues

All normal colon tissues in this study were taken from discarded surgical normal colon from CRC surgical resections. Pathological analysis was used to check the normal origin by tissue morphology. Briefly [10], fresh patient colon tissues were collected in Medium 199 supplemented with 200 U/ml penicillin and 200 mg/ml streptomycin, immediately after patient operative resection. Fat and blood clots were removed from tissues before being rinsed 10 times in sterile PBS. Samples were cut into 5 cm × 2 cm pieces and incubated in sterile 1% SDS (Fisher Scientifics) in deionized water for 4-6 hours at room temperature with gentle shaking. Sterile 1% Triton –X100 (Sigma) in deionized water was applied to rinse the tissues for 1 hour and the acellular matrix was then washed in sterile PBS containing penicillin/streptomycin/amphotericin at 37°C with changing the PBS once every 30 minutes for the first 5 hours and twice each day for the remaining 5 days. The decellularized matrix can be freshly used for the following recellularization or stored at -80 °C for up to 6 months. Molecular characterizations were performed using immunostaining in F-actin (Invitrogen, fluorescent phallotoxins) for cytoskeleton, DAPI (Invitrogen) for cell nuclei, collagen-I (Novus Biologicals, Cat# NB600-408), laminin (Sigma, Cat# HPA001895) and fibronectin (Sigma, Cat# F0916) for three main ECM proteins in native tissues and acellular matrices.

Culture of primary human colonic cells

Normal colon biopsies (0.5-1 cm^3) without visible adenomas by pathology were obtained from patients undergoing colonoscopy screening. The protocol of isolation and culture of human colonic epithelial cells (hCEC) were slightly modified from previous studies [9, 43]. Briefly, colonic specimens were immersed in cold X medium (HyClone) supplemented with 2% penicillin/streptomycin, immediately after the patient operative procedure and rinsed with sterile PBS with antibiotics/ antimycotic (Invitrogen) for 5 times. The tissues were minced into small pieces (~1 mm^3 in size) and crypts were gently extracted by digestion in X medium containing collagenase type XI (150 U/ml, Sigma, St. Louis, MO), dispase neutral protease (40 μg/ml, Roche Applied Science), with stirring at 37°C for 15-30 min. The crypt cells were then cultured in X medium with growth supplements of 5% FBS, EGF (25ng/ml R&D systems), insulin (5.0 μg/ml, Sigma), hydrocortisone (1.0 μg/ml Sigma), transferrin (2 μg/ml Sigma), BPE (50 μg/ml Sigma), B27 supplement (Invitrogen), R-spondin 1 (200 ng/ml, R&D systems) and Noggin (50ng/ml, Peprotech) in collagen-I coated flasks (BD scientific) and incubated at 33°C, with 5% CO2. After 48 hour culture, fibroblast inhibitory reagent (Human Colon FibrOut™ from CHI Scientific) was added to the culture medium for 2 days to reduce fibroblast growth. After cell colonization was observed, cells were transfected with retroviral hTERT (ATCC hTERT immortalization kit). The hCECs were then grown in 3D matrigel culture and checked for their ability to form organoids. The hCECs were also characterized for their expression of the stem cell marker Lgr5 (OriGene

Tech. Cat# TA503316) and differentiation markers (anti-Villin, Abcam, Cat# ab739; anti-MUC2, Sigma, Cat# HPA006179; anti-CHGA, Sigma, Cat# HPA017369; anti-A33, Sigma, Cat# HPA018858; anti- β-catenin, Santa Cruz, Cat# sc1496 (for immunocytochemistry); anti-pan-cytokeratin, Sigma, Cat# c2562; anti-Ki67, Cell Signaling, Cat# 9027; anit-β-Catenin, Abcam, Cat# ab16051 (for immunohistochemistry)). Co-culture with human colonic fibroblasts was required to develop tight-junction (anti-TJP1, Sigma, Cat# HPA001636) in hCECs. Using primers previously described [44-47], DNA sequencing of hCEC indicated no mutations in hotspot regions of *APC*, *KRAS* and *TP53* genes.

Human colonic myofibroblasts were cultured as below procedure. Briefly, biopsies were placed in ice-cold DMEM supplemented with 3% FBS, L-glutamine (2 mmol/l), sodium pyruvate (1 mmol/l), 2% penicillin and streptomycin, immediately after patient operative procedures and then rinsed with sterile PBS with antibiotics/ antimycotic (Invitrogen) for 5 times. Then biopsies were minced into small pieces of $1mm^3$ sizes and dispersed by mechanical testing with pipetting. The tissue mixture was washed 3 times in medium and followed with centrifugation at 2,000 rpm for 5 min. The tissue explants including the tissue fragments were then placed on culture dishes and covered with DMEM supplemented with 3% FBS, L-glutamine (2 mmol/l), sodium pyruvate (1 mmol/l), 1% penicillin and streptomycin, and maintained at 37°C in a 5% CO2 incubator. Once fibroblast-like cells appeared and began to colonize, the remaining tissue fragments were discarded. The fibroblast strains were characterized by immunostaining of human α-smooth muscle actin (Abcam, Cat# ab5694) and used for follow-up experiments within 5 passages.

Human colonic microvascular endothelial cells were purchased and maintained in culture according to the company instructions (Sciencell Research Laboratories, Cat# 2930). The primary cells were characterized for the endothelial marker CD31 (Sigma, Cat# WH0005175M1) and used within 10 passages.

Recellularization of acellular colon matrix

The mucosa layer of the decellularized colon matrix was physically separated from the submucosa using forceps, and the human endothelial cells were seeded by microinjection of 5,000 cells/side into four sides (the front, back, left and right sides) of a 0.25 cm^2 (0.5 cm × 0.5 cm) mucosa layer using Eppendorf TransMan NK micromanipulators under invert microscope (Nikon Diaphot). After 5 days culture with gentle shaking in endothelial cell medium, $5×10^4$ hCECs were seeded on the 0.25 cm^2 mucosa evenly with cells seated in crypt niches and maintained in culture for 10 days in 1:1 mixture of epithelial cell medium with endothelial cell basal medium. After the myofibroblasts were seeded in the opposite surface of mucosa (the muscularis mucosa layer), the tissues were turned over with muscularis mucosa layer up and continued to culture for another 15-20 days. On the other hand, this orientation with crypt niches facing bottom also facilitated crypt repopulation with epithelial cells by neutral gravity. Lastly, the mucosa layers were assembled with the submucosa part and the whole “sandwich tissues” being returned for continuous culture with mucosal crypts up in a 1:1 mixture of epithelial cell medium with endothelial cell basal medium until CRC in different stages developed.

LM-PCR amplification and preparation for Illumina sequencing

The transposon insertion sites were amplified by ligation-mediated PCR (LM-PCR) and prepared for Illumina sequencing as previously described [29]. Briefly, total genomic DNA was digested with *Nla*III (IRR) or AluI (IRL), and fragments were linked with double-stranded adaptor by ligation. Primers used to generate adaptors:

(IRDRR adaptor)

NlaIII linker+ 5′-GTAATACGACTCACTATAGGGCTCCGCTTAAGGGACCATG-3′

NlaIII linker- 5′-Phos-GTCCCTTAAGCGGAG-C3spacer-3′

(IRDRL adaptor)

AluI linker+ 5′-GTAATACGACTCACTATAGGGCTCCGCTTAAGGGAC-3′

AluI linker- 5′-Phos-GTCCCTTAAGCGGAG-C3spacer-3′

After digestion and ligation with *BamHI* to prevent the fragment from vector transposons, the IRR or IRL regions were amplified by PCR using the following primers:

For IRR: 5′GGATTAAATGTCAGGAATTGTGAAAA 3′ and linker primer

5′GTAATACGACTCACTATAGGGC 3′

For IRL: 5′AAATTTGTGGAGTAGTTGAAAAACGA 3′ and linker primer 5′GTAATACGACTCACTATAGGGC 3′

At last, secondary PCR was performed to amplified IRR and IRL by adding barcodes, which made it possible to sequence multiple samples together, and then the DNA samples are sent for Illumina sequencing (Illumina HiSeq 2500). Secondary PCR primers for IRR and IRL:

5′AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG CTCTTCCGATCT (barcode)TGTATGTAAACTTCCGACTTCAACTG;

5′CAAGCAGAAGACGGCATACGAGCTCTTCCGATCTAGGGCTCCGCTT AAGGGAC

Analysis of SB insertion sites in invasive neoplasia

For analysis of transposon insertion sites, ligation-mediated PCR (LM-PCR) was performed to specifically amply the integrated transposon-human genomic junction fragments for Illumina sequencing (HiSeq 2500, single-end reads). Raw read sequence data were first demultiplexed according to the 6 bp barcodes, then the LM-PCR primer was trimmed from the read sequences. The trimmed read sequences were then mapped to the reference human genome hg19 (UCSC Genome Browser website: http://hgdownload.cse.ucsc.edu/goldenPath/hg19/chromosomes/) using the BWA program (version 0.7.5a). Low mapping quality reads (MAPQ<30 MQ) were filtered out, and only those reads with their mate read with MAPQ>30 MQ were kept for subsequent analyses. Among the high mapping quality reads, about one quarter to one half of them were mapped at “TA” dinucleotide sites (varying between different samples). Further analysis was focused on these consensus SB insertion sites. From the distribution of read depth at the insertion sites, it was clear that the

majority of the insertion sites resulted from background insertion events or PCR artifacts. We considered the top 10% of the insertion sites with the deepest read counts as clonal insertion sites. We mapped these clonal insertion sites in each sample to the human genome annotation file, RefGene, and picked the sites that were located within 1000 bp of known transcripts in RefGene to cover both the core promoter and proximal promoter elements [48]. In addition to this, Sanger DNA sequencing of chain-terminating dideoxynucleotides (Genewiz, Inc.) was used to confirm the transposon insertion sites.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank other members of the Shuler laboratory and the Jenkins/Copeland Laboratory. We also thank Dr. Harold E. Varmus for providing laboratory resources and helpful discussions. This work was supported by NCI-PSOC Young Investigator trans-network grant (to H.J.C and Z.W.), NIH-UH2TR000516 (to M.L.S), NSF-1106153 (to M.L.S), NSF GRFP- 2011131053 (to H.J.C), NIH R01GM095990 (to X.S) Arnold O. Beckman Postdoctoral fellowship (to H.J.C) and the Cancer Prevention Research Institute of Texas (CPRIT) (N.G.C. and N.A.J.). N.G.C. and N.A.J. are also both CPRIT scholars in Cancer Research.

References

1. Arrowsmith J. Trial watch: Phase II failures: 2008-2010. Nat Rev Drug Discov. 2011; 10(5):328–9. [PubMed: 21532551]
2. Gout S, Huot J. Role of cancer microenvironment in metastasis: focus on colon cancer. Cancer Microenviron. 2008; 1(1):69–83. [PubMed: 19308686]
3. Bozic I, et al. Accumulation of driver and passenger mutations during tumor progression. Proc Natl Acad Sci U S A. 2010; 107(43):18545–50. [PubMed: 20876136]
4. Jones S, et al. Comparative lesion sequencing provides insights into tumor evolution. Proc Natl Acad Sci U S A. 2008; 105(11):4283–8. [PubMed: 18337506]
5. Cancer, Genome; Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. Nature. 2012; 487(7407):330–7. [PubMed: 22810696]
6. Copeland NG, Jenkins NA. Deciphering the genetic landscape of cancer--from genes to pathways. Trends Genet. 2009; 25(10):455–62. [PubMed: 19818523]
7. Li X, et al. Oncogenic transformation of diverse gastrointestinal tissues in primary organoid culture. Nat Med. 2014; 20(7):769–77. [PubMed: 24859528]
8. Matano M, et al. Modeling colorectal cancer using CRISPR-Cas9-mediated engineering of human intestinal organoids. Nat Med. 2015; 21(3):256–62. [PubMed: 25706875]
9. Jung P, et al. Isolation and in vitro expansion of human colonic stem cells. Nat Med. 2011; 17(10): 1225–7. [PubMed: 21892181]
10. Gilbert TW, Sellaro TL, Badylak SF. Decellularization of tissues and organs. Biomaterials. 2006; 27(19):3675–83. [PubMed: 16519932]
11. Ott HC, et al. Regeneration and orthotopic transplantation of a bioartificial lung. Nat Med. 2010; 16(8):927–33. [PubMed: 20628374]
12. Ott HC, et al. Perfusion-decellularized matrix: using nature's platform to engineer a bioartificial heart. Nat Med. 2008; 14(2):213–21. [PubMed: 18193059]
13. Ridky TW, et al. Invasive three-dimensional organotypic neoplasia from multiple normal human epithelia. Nat Med. 2010; 16(12):1450–5. [PubMed: 21102459]
14. Tuveson DA, et al. Endogenous oncogenic K-ras(G12D) stimulates proliferation and widespread neoplastic and developmental defects. Cancer Cell. 2004; 5(4):375–87. [PubMed: 15093544]

15. Biswas S, et al. Transforming growth factor beta receptor type II inactivation promotes the establishment and progression of colon cancer. Cancer Res. 2004; 64(14):4687–92. [PubMed: 15256431]
16. Trobridge P, et al. TGF-beta receptor inactivation and mutant Kras induce intestinal neoplasms in mice via a beta-catenin-independent pathway. Gastroenterology. 2009; 136(5):1680–8 e7. [PubMed: 19208363]
17. Lohi J, et al. Laminins, tenascin and type VII collagen in colorectal mucosa. Histochem J. 1996; 28(6):431–40. [PubMed: 8863048]
18. Vermeulen L, et al. Single-cell cloning of colon cancer stem cells reveals a multi-lineage differentiation capacity. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105(36):13427–32. [PubMed: 18765800]
19. Liotta LA. Tumor invasion and metastases: role of the basement membrane. Warner-Lambert Parke-Davis Award lecture. Am J Pathol. 1984; 117(3):339–48. [PubMed: 6095669]
20. de Lau W, Barker N, Clevers H. WNT signaling in the normal intestine and colorectal cancer. Frontiers in bioscience : a journal and virtual library. 2007; 12:471–91. [PubMed: 17127311]
21. Fearon ER, Vogelstein BA. genetic model for colorectal tumorigenesis. Cell. 1990; 61(5):759–67. [PubMed: 2188735]
22. Wood LD, et al. The genomic landscapes of human breast and colorectal cancers. Science. 2007; 318(5853):1108–13. [PubMed: 17932254]
23. Oshima H, et al. Suppressing TGFβ Signaling in Regenerating Epithelia in an Inflammatory Microenvironment Is Sufficient to Cause Invasive Intestinal Cancer. Cancer Research. 2015; 75(4): 766–776. [PubMed: 25687406]
24. Copeland NG, Jenkins NA. Harnessing transposons for cancer gene discovery. Nat Rev Cancer. 2010; 10(10):696–706. [PubMed: 20844553]
25. Dupuy AJ, et al. Mammalian mutagenesis using a highly mobile somatic Sleeping Beauty transposon system. Nature. 2005; 436(7048):221–6. [PubMed: 16015321]
26. March HN, et al. Insertional mutagenesis identifies multiple networks of cooperating genes driving intestinal tumorigenesis. Nat Genet. 2011; 43(12):1202–9. [PubMed: 22057237]
27. Takeda H, et al. Transposon mutagenesis identifies genes and evolutionary forces driving gastrointestinal tract tumor progression. Nat Genet. 2015; 47(2):142–50. [PubMed: 25559195]
28. Grabundzija I, et al. Comparative analysis of transposable element vector systems in human cells. Mol Ther. 2010; 18(6):1200–9. [PubMed: 20372108]
29. Brett BT, et al. Novel molecular and computational methods improve the accuracy of insertion site analysis in Sleeping Beauty-induced tumors. PLoS One. 2011; 6(9):e24668. [PubMed: 21931803]
30. Angus-Hill ML, et al. T-cell factor 4 functions as a tumor suppressor whose disruption modulates colon cell proliferation and tumorigenesis. Proc Natl Acad Sci U S A. 2011; 108(12):4914–9. [PubMed: 21383188]
31. Behrens J, et al. Functional interaction of an axin homolog, conductin, with beta-catenin, APC, and GSK3beta. Science. 1998; 280(5363):596–9. [PubMed: 9554852]
32. Hao, Yu; GZ, J.; Kai, Liu; Hui, Dong; Hua, Yu; Ji-Cheng, Duan; Zhe, Li; Wei, Dong; Wen-Ming, Cong; Jia-He, Yang. Twist2 is a valuable prognostic biomarker for colorectal cancer. World J Gastroenterol. 2013; 19(15)
33. Xiong H, et al. Inhibition of JAK1, 2/STAT3 signaling induces apoptosis, cell cycle arrest, and reduces tumor cell invasion in colorectal cancer cells. Neoplasia. 2008; 10(3):287–97. [PubMed: 18320073]
34. Forcet C, et al. The dependence receptor DCC (deleted in colorectal cancer) defines an alternative mechanism for caspase activation. Proc Natl Acad Sci U S A. 2001; 98(6):3416–21. [PubMed: 11248093]
35. Aylon Y, et al. A positive feedback loop between the p53 and Lats2 tumor suppressors prevents tetraploidization. Genes Dev. 2006; 20(19):2687–700. [PubMed: 17015431]
36. Yabuta N, et al. Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the Drosophila tumor suppressor gene lats/warts. Genomics. 2000; 63(2):263–70. [PubMed: 10673337]

37. Funato K, et al. Tyrosine phosphatase PTPRD suppresses colon cancer cell migration in coordination with CD44. Exp Ther Med. 2011; 2(3):457–463. [PubMed: 22977525]
38. Chen HJ, et al. Comprehensive models of human primary and metastatic colorectal tumors in immunodeficient and immunocompetent mice by chemokine targeting. Nat Biotechnol. 2015; 33(6):656–60. [PubMed: 26006007]
39. Katoh M. Functional and cancer genomics of ASXL family members. Br J Cancer. 2013; 109(2): 299–306. [PubMed: 23736028]
40. Williams DS, et al. Nonsense mediated decay resistant mutations are a source of expressed mutant proteins in colon cancer cell lines with microsatellite instability. PLoS One. 2010; 5(12):e16012. [PubMed: 21209843]
41. Henrich KO, et al. CAMTA1, a 1p36 tumor suppressor candidate, inhibits growth and activates differentiation programs in neuroblastoma cells. Cancer Res. 2011; 71(8):3142–51. [PubMed: 21385898]
42. Kim MY, et al. Recurrent genomic alterations with impact on survival in colorectal cancer identified by genome-wide array comparative genomic hybridization. Gastroenterology. 2006; 131(6):1913–24. [PubMed: 17087931]
43. Roig AI, et al. Immortalized epithelial cells derived from human colon biopsies express stem cell markers and differentiate in vitro. Gastroenterology. 2010; 138(3):1012–21 e1-5. [PubMed: 19962984]
44. Gomes CC, et al. Assessment of TP53 mutations in benign and malignant salivary gland neoplasms. PLoS One. 2012; 7(7):e41261. [PubMed: 22829934]
45. Groden J, et al. Identification and characterization of the familial adenomatous polyposis coli gene. Cell. 1991; 66(3):589–600. [PubMed: 1651174]
46. Lievre A, et al. KRAS mutation status is predictive of response to cetuximab therapy in colorectal cancer. Cancer Res. 2006; 66(8):3992–5. [PubMed: 16618717]
47. Ruiz-Ponte C, et al. Mutation analysis of the adenomatous polyposis coli (APC) gene in northwest Spanish patients with familial adenomatous polyposis (FAP) and sporadic colorectal cancer. Hum Mutat. 2001; 18(4):355. [PubMed: 11668620]
48. Maston GA, Evans SK, Green MR. Transcriptional regulatory elements in the human genome. Annu Rev Genomics Hum Genet. 2006; 7:29–59. [PubMed: 16719718]
49. Chen HJ, et al. Chemokine 25-induced signaling suppresses colon cancer invasion and metastasis. J Clin Invest. 2012; 122(9):3184–96. [PubMed: 22863617]

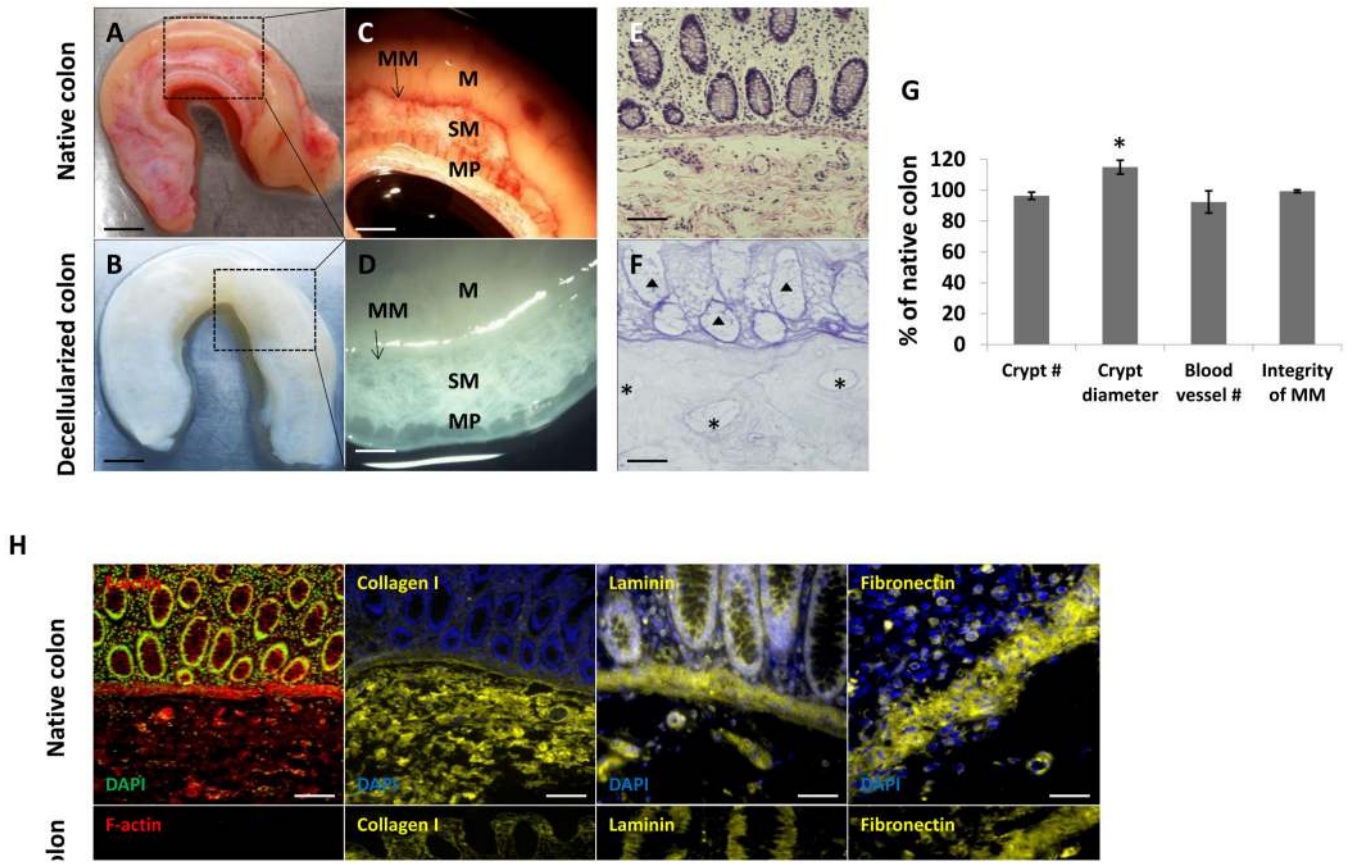

Figure 1. Preparation and characterization of acellular human colon matrix

(**A,C**) Representative images of normal human colon tissue and (**B,D**) decellularized human colon. **MM**, muscularis mucosa; **M**, mucosa area; **SM**, submucosal area; **MP**, muscularis propria. Scale bars, 0.5 cm **(A,B)** and 0.1 cm **(C,D). (E)** Hematoxylin and eosin (H+E) staining of native colon tissue and (**F**) decellularized colon matrix**.** The acellular colon matrix shows a well-preserved vasculature (asterisks) and crypt niches (triangles). Scale bars, 100 μ. **(G)** Quantification of crypt numbers and crypt diameters, blood vessel numbers and MM integrity in acellular matrix relative to native tissue. (n=8 of independent tissues; *P< 0.05 compared to native tissue by Mann-Whitney test). Error bars indicate S.E.M. (stand. error of the mean). **(H)** Representative immunostaining images including F-actin for cytoskeleton, DAPI for cell nuclei, collagen-I, laminin and fibronectin for ECM in native tissues (top panels) and acellular matrix (bottom panels). Note that DAPI and cytoskeleton protein was undetected in the acellular matrix, while the three main ECM proteins were well preserved. Scale bars, 100 μ.

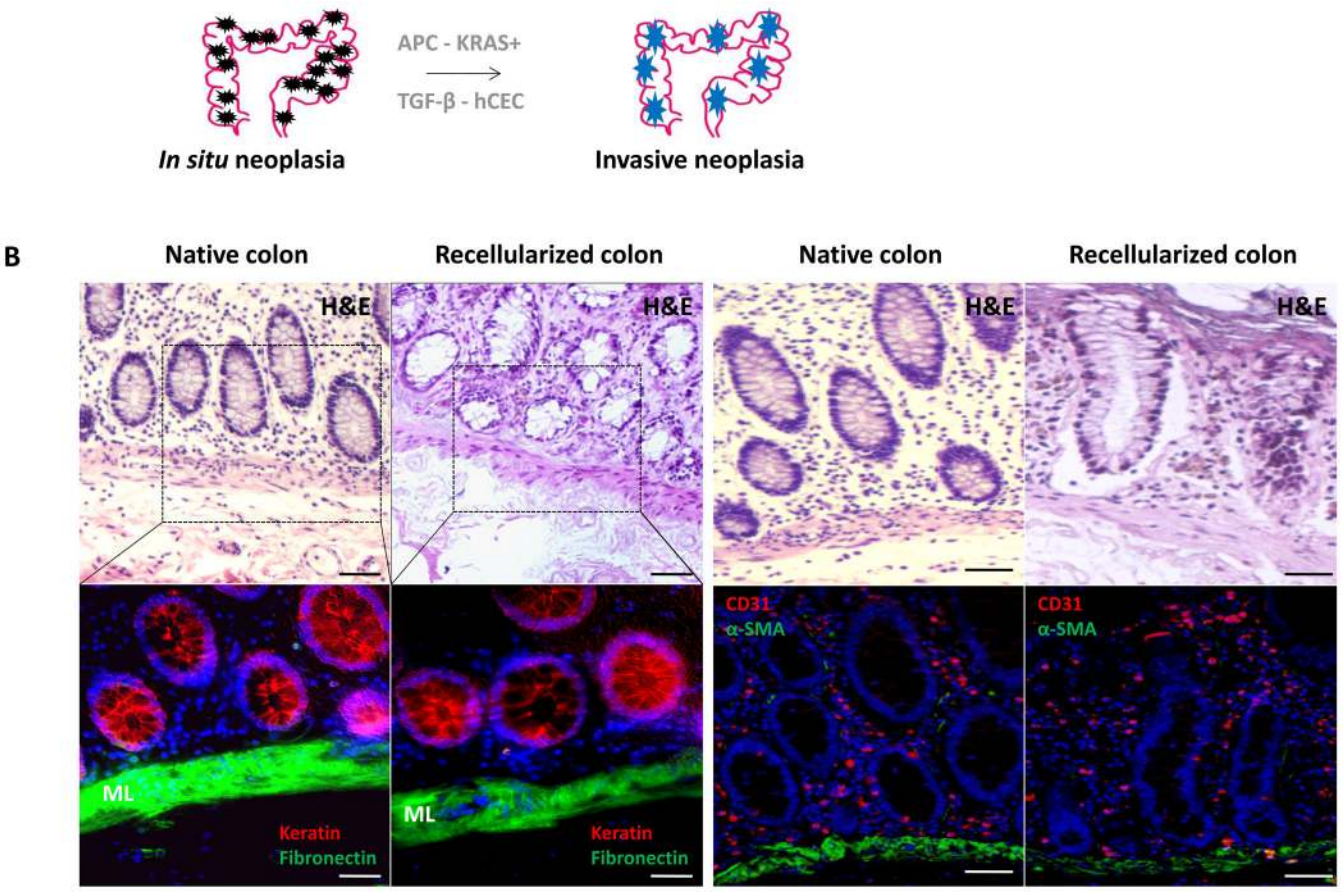

Figure 2. Creation of an *ex vivo* human colon cancer model. (A)
A schematic showing the different steps used in the creation of the *ex vivo* colon cancer model. **(B)** Representative H+E stained images (upper panel) and dual immunostained images of cytokeratin (epithelial cells) and fibronectin or CD31 (endothelial cells) and α-small muscle actin (α-SMA; myofibroblasts) in native or recellularized colon tissue (lower panel). **ML:** muscularis layer; Scale bars, 50μ. Representative H+E stained images of (**C**) native colon**, (D)** colon recellularized with parental hCECs**, (E,F)** colon recellularized with *APC* shRNA-expressing hCECs**,** (**G, H**) colon recellularized with hCECs expressing *APC* shRNA and activated $KRAS^{G12D}$, or **(I)** colon recellularized with hCECs expressing APC shRNA, activated $KRAS^{G12D}$ and incubated in medium containing 500 nmol/L TGF-β RI Inhibitor III and stained with H+E or **(J)** immunostained for cytokeratin and fibronectin. Dot lines designate edges of muscularis muscosa; black arrow indicated submucosal invasive neoplasia; **M**, mucosa area; **SM**, submucosal area; Blue nuclei, DAPI; Scale bars, 50μ. **(K)** Quantification of *in situ* neoplasia and invasive neoplasia formed with parental hCECs, *APC* shRNA-expressing hCECs (APC-), *APC* shRNA-expressing hCECs that also express $KRAS^{G12D}$ (APC-KRAS+), or *APC* shRNA-expressing hCECs that also express $KRAS^{G12D}$ and which subsequently were treated with a TGF-β RI inhibitor (APC-KRAS+TGF-β-). All of the recellularized tissues have been cultured for 7 weeks when they were terminated for quantification of tumorigenesis. *APC* knockdown combined with activated $KRAS^{G12D}$ and inactivated TGF-β signaling enhanced the formation of invasive neoplasia in this model

system. (n= 10 of independent matrix; difference analysis by one-way ANOVA). Error bars indicate S.E.M.

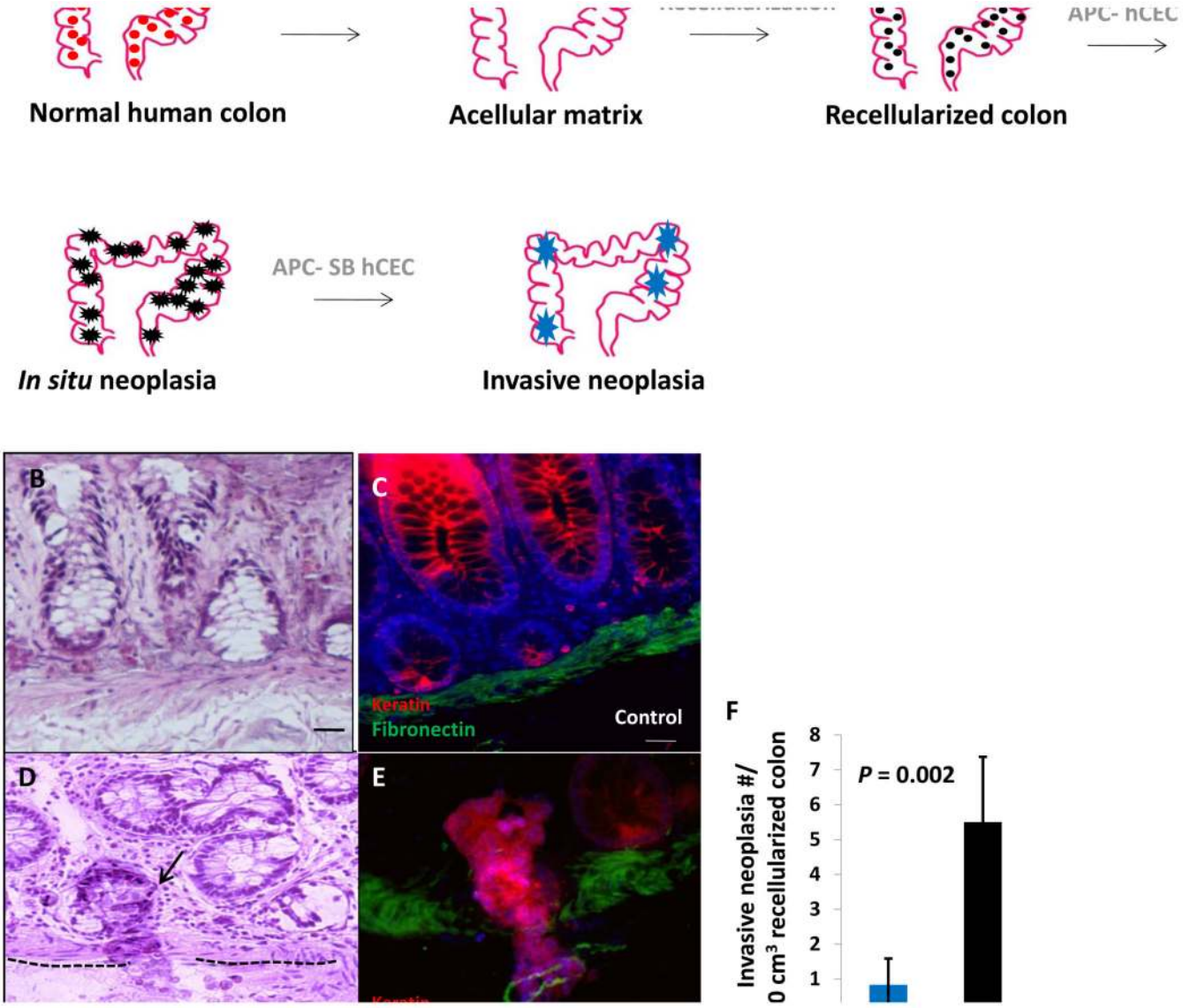

Figure 3. Invasive adenomas induced by *SB* mutagenesis. (A)

A schematic showing the different steps used in the creation of invasive adenomas by *SB*. (**B,D**) H+E stained images or (**C,E**) immunostained images of cytokeratin and fibronectin in CRC models recellularized with wild type hCECs mutagenized *SB* (upper panels) or *APC* shRNA-expressing hCECs mutagenized with *SB* (lower panels). Blue nuclei, DAPI; Scale bars, 50μ. **(F)** Quantification of invasive neoplasia in the two CRC models. All of the recellularized tissues had been cultured for 7 weeks when they were terminated for quantification of tumorigenesis. These data show that *SB*-mutagenesis promotes CRC progression in *APC* shRNA-expressing hCECs. (n=6 of independent matrix; difference analysis by 2-sided student *t* test). Error bars indicate S.E.M. **(G-I)** Dual immunostained images of cytokeratin and fibronectin showing progression from *in situ* neoplasia to submucosal invasion in *SB*-mutagenized hCECs. Blue nuclei, DAPI; Scale bars, 50μ.

Table 1

Seventeen candidate invasion driver genes previously implicated as contributing to CRC progression

Gene symbol	No. unique insertions*	No. invasive neoplasia with insertions**	Function
AKTIP	1	1	AKT signaling
DCC	1	1	Cell adhesion/apoptosis signaling
EPHB1	1	1	Cell migration and adhesion
EPHB2	1	1	Cell migration and adhesion
FSTL5	1	1	Calcium ion binding
JAK1	1	1	STAT signaling
LATS2	2	2	P53 signaling
MAML3	1	1	NOTCH signaling
MSH2	1	1	DNA mismatch repair
NRCAM	1	1	Cell adhesion
PTPRD	2	3	Cell growth and differentiation
ROCK1	1	1	Cell motility
STAT3	1	1	STAT signaling
TCF7L2	1	2	WNT signaling
TTN	1	1	Cell motility
TWIST2	1	1	Epithelial to mesenchymal transition
WNT9B	1	1	*WNT* signaling

*Number unique *SB* insertions in each gene;

**Number of independent invasive neoplasias with *SB* insertions in the same gene.

Nat Biotechnol. Author manuscript; available in PMC 2017 January 11.

Table 2

***In vitro* functional validation of candidate invasion driver genes**

Twenty-one candidate invasion driver genes, which were not previously implicated in CRC (black type), in addition to one positive control (red type) and 5 negative control (blue type) genes, were assayed for their ability to increase cell proliferation by WST-1 dye staining, or migration and invasion through matrigel in Boyden chambers, in hCEC or SW480 cells, following siRNA knockdown.

	hCEC			SW480		
Gene	**proliferation**	**migration**	**invasion**	**proliferation**	**migration**	**invasion**
LATS2			+	+		+
AFF3						
ANTXR1	--			--		
ASXL2	+			++		
CAMTA1	+		++	+		++
CSTF3						
DBF4						
DDX20		+		+	+	
DENND2C						
DMXL2						
FGF13						
FXR1	+		+		+	+
GRM8						
KDM2B			+			
LPP						
MITF	+	+				+
PAX7	+	+		++	+	
PDE4DIP		--			--	
PRKG1			+			
RPAP1			+			
SEMA6D						
SUPT3H	-- --	--		-- --	--	
CCDC113						

Gene	hCEC			SW480		
	proliferation	migration	invasion	proliferation	migration	invasion
KLHL10						
POMZP3						
RNPC3						
SMTNL2						

+two-fold increase

++three-fold increase

-two-fold decrease

- - -three-fold decrease, compared to *APC* shRNA-expressing hCECs or parental SW480 cells transduced with control oligos.