

RESEARCH

Open Access



Virulence genes are a signature of the microbiome in the colorectal tumor microenvironment

Michael B. Burns^{1,2*}, Joshua Lynch^{1,2}, Timothy K. Starr^{1,3,4}, Dan Knights^{5,6} and Ran Blekhman^{1,2*}

Abstract

Background: The human gut microbiome is associated with the development of colon cancer, and recent studies have found changes in the microbiome in cancer patients compared to healthy controls. Studying the microbial communities in the tumor microenvironment may shed light on the role of host–bacteria interactions in colorectal cancer. Here, we highlight the major shifts in the colorectal tumor microbiome relative to that of matched normal colon tissue from the same individual, allowing us to survey the microbial communities in the tumor microenvironment and providing intrinsic control for environmental and host genetic effects on the microbiome.

Methods: We sequenced the microbiome in 44 primary tumor and 44 patient-matched normal colon tissue samples to determine differentially abundant microbial taxa. These data were also used to functionally characterize the microbiome of the cancer and normal sample pairs and identify functional pathways enriched in the tumor-associated microbiota.

Results: We find that tumors harbor distinct microbial communities compared to nearby healthy tissue. Our results show increased microbial diversity in the tumor microenvironment, with changes in the abundances of commensal and pathogenic bacterial taxa, including *Fusobacterium* and *Providencia*. While *Fusobacterium* has previously been implicated in colorectal cancer, *Providencia* is a novel tumor-associated agent which has not been identified in previous studies. Additionally, we identified a clear, significant enrichment of predicted virulence-associated genes in the colorectal cancer microenvironment, likely dependent upon the genomes of *Fusobacterium* and *Providencia*.

Conclusions: This work identifies bacterial taxa significantly correlated with colorectal cancer, including a novel finding of an elevated abundance of *Providencia* in the tumor microenvironment. We also describe the predicted metabolic pathways and enzymes differentially present in the tumor-associated microbiome, and show an enrichment of virulence-associated bacterial genes in the tumor microenvironment. This predicted virulence enrichment supports the hypothesis that the microbiome plays an active role in colorectal cancer development and/or progression. Our results provide a starting point for future prognostic and therapeutic research with the potential to improve patient outcomes.

Background

Colorectal cancer (CRC) is the second most commonly diagnosed cancer in females and the third in males worldwide [1]. The microbial communities present in the intestinal tract have known associations with colon health, though until recently researchers were limited to the study of microbes that were amenable to in vitro culturing. As a result of recent advances in culture-independent measurements of microbial communities, we know that the

human gut is host to roughly a thousand different bacterial species [2]. Alterations of this bacterial community are correlated with host health, including diseases ranging from diabetes and obesity to Crohn's disease and arteriosclerosis [3]. The composition of the gut microbiome also has a known association with CRC, although the direction of causality remains unclear [4–10]. A recent report demonstrated that analysis of the microbiome can be used as a pre-screening test for CRC that dramatically outperforms the current standard methods [11]. These analyses have identified significant shifts in the relative abundances of specific bacterial taxa in CRC cancer patients' colon mucosa and stool microbiomes. For instance, bacteria in the

* Correspondence: michaelburns@umn.edu; blekhman@umn.edu

¹Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN, USA

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

genus *Fusobacterium* are enriched in some CRC patients' microbiomes [7, 8, 10, 12]. *Fusobacterium* are thought to elicit a pro-inflammatory microenvironment around the tumor, driving tumor formation and/or progression [7]. More specifically, a recent study has demonstrated that the FadA protein, a virulence factor expressed by *Fusobacterium nucleatum*, can signal epithelial cells via E-cadherin, a cell-surface molecule important for CRC metastasis as well as a component of the WNT/ β -catenin signaling pathway, the most commonly mutated pathway in CRC [13, 14]. Other cancer-associated bacterial taxa have been identified, including *Escherichia coli* strain NC101 and *Bacteroides fragilis*, each with a proposed mechanism of interaction with colon cancer [15]. The species *Akkermansia muciniphila*, a bacterium that has known associations with obesity, has also been implicated as a cancer-associated agent, with its mucin-degrading activity as a proposed mechanism to drive inflammation contributing to cancer genesis and/or progression [15].

In addition to defining the set of bacteria significantly associated with CRC, several groups have used measurements of microbiome diversity to compare cancer patients with normal subjects. There are distinct differences in these results that depend on the sources of the samples used to assess the microbiome (e.g., stool samples versus mucosal or tissue samples, longitudinal versus cross-sectional sampling). For instance, in a study that used stool samples to compare CRC patients with normal controls, the researchers showed decreased alpha-diversity among the microbial communities found in the CRC patients' stools compared with the control [16]. Another, recent study, also focusing on fecal samples, was unable to detect differences in microbial community diversity or richness between normal and cancer-associated microbiomes [17]. However, in a study that used tissue samples from patients with colon adenomas and compared them with patient-matched normal tissues, the alpha-diversity present at the site of the lesion was actually increased [18]. This finding was repeated by Mira-Pascual et al. [19], who performed side-by-side analyses of tissue and stool samples. They found that stool samples in general had roughly twice the microbial diversity when compared with tissue-associated microbiomes, with stools from cancer patients showing lower diversity than those from normal controls. When comparing the microbial diversity between tissue samples, however, the tumor microbiome was more diverse than the normal microbiome. It is likely that this is a function of the stool samples harboring microbes from the entire colonic environment, including species that are not directly related to the tumor microenvironment, adding noise to the taxonomic results acquired from assessment of stool samples relative to direct measurements of tissues. These findings suggest that in order to detect differences specific to the cancer-

associated microbiome, samples taken directly from the tumor microenvironment are preferable, at least at the initial characterization phase, to bulk stool samples, which are not likely to have the discriminatory power required to measure small, yet significant, effects [19].

The use of traditional case-control studies of the colon cancer microbiome makes it difficult to control for all of the external effects on the microbiome. For example, the composition of gut microbial communities is strongly affected by diet [20]. Host genetic variation is also expected to control variation in the gut microbiome [21], through differences in host immune response among other genetic mechanisms [22]. The large effects these factors have on microbiome composition are likely to confound traditional case-control studies. By using tumor and normal tissue samples taken from the same individual, our study controls for these variables internally, providing a more accurate view of the tumor-associated shifts in the microbiome. Several previous studies have utilized this strategy to assess changes in the cancer-associated microbiome in a variety of populations [7, 9, 10, 12, 19, 23, 24]. However, most used far fewer samples than would be necessitated by using unmatched tissue or stool samples. Another caveat with these analyses, even with the use of patient-matched samples, is that the normal tissue itself may possibly be affected by the changes elicited in the organ by the tumor. There has been a report indicating that, in some cases, biofilms may form in otherwise normal areas of the colon of patients who also have tumors [24]. In the event that this is the case, the overall differences seen between tumor and normal matched tissue microbiomes might be diminished as the tumor is producing a "field effect" that influences surrounding tissue.

In addition to measuring bacterial taxa levels in colon cancer, it is also important to take into account associated factors, such as host genetics and gene expression, as well as the microenvironmental metabolome. Independent research groups have attempted to uncover pertinent alterations in these factors and how they correlate with cancer state [25–27]. Of note, analysis of the CRC-associated metabolome highlighted differences in the biochemical composition of cancer patients' stools. CRC patients were found to have higher levels of some amino acids and alterations in the levels of some short chain fatty acids (SCFAs) in their stools when compared with controls [25]. Butyrate, a SCFA with known anti-cancer properties, was depleted in CRC patient stool samples, as were several genera of butyrate-producing bacteria [28]. Our work continues this effort by expanding the analysis of the CRC-associated microbiome to include virtual metagenomic profiling of the enzymes and pathways present in the CRC microbiome, with specific attention paid to assessing the presence of known

virulence-associated genes [29]. Previous work has been performed attempting to identify transcripts from 22 different, potentially cancer-related bacterial toxins in metatranscriptomic data, though with limited success [30]. As this study focuses on a comparison of fully developed tumors and compares them with patient-matched normal tissue, it is important to highlight that this work is not an assessment of the initiation of cancer, but rather a characterization of cancer after it has transformed into a malignancy.

Methods

Tissue samples and DNA isolation

We used 88 tissue samples from 44 individuals, with one tumor and one normal sample from each individual. These de-identified samples were obtained from the University of Minnesota Biological Materials Procurement Network (Bionet), a facility that archives research samples from patients who have provided written, informed consent. All research conformed to the Helsinki Declaration and was approved by the University of Minnesota Institutional Review Board, protocol 1310E44403; see Additional file 1 for detailed information on the samples used in this study. Tissue pairs were resected concurrently, rinsed with sterile water, flash frozen in liquid nitrogen, and characterized by staff pathologists. The criteria for selection were limited to the availability of patient-matched normal and tumor tissue specimens. The specific site of the tumor within the intestinal tract was recorded and can be found in Additional file 1. Total DNA was isolated from the flash-frozen tissue samples and their associated microbiomes by adapting an established nucleic acid extraction protocol [31]. Briefly, approximately 100 mg of flash-frozen tissue were physically disrupted by placing the tissue in 1 mL of Qiazol lysis solution and sonicating in a heated (65 °C) ultrasonic water bath for 1–2 h. The efficiency of this approach was verified by observing high abundances of Gram-negative bacteria across all samples, including those from the phylum Firmicutes. Additionally, sequences from the notoriously difficult to lyse bacterial genera *Mycobacterium* and/or *Bacillus* [32] were detected in the majority of samples, also indicating a rigorous and efficient lysis. DNA was purified from the lysate using the Qiagen All-prep kit (Qiagen Inc., Valencia, CA, USA).

16S rRNA sequencing

Briefly, DNA isolated from colon samples was quantified by quantitative PCR (qPCR), and the V5-V6 regions of the 16S rRNA gene were PCR amplified with the addition of barcodes for multiplexing. The forward and reverse primers were the V5F and V6R sets from Cai et al. [33]. The PCR conditions were as follows. Amplification was carried out in a 25 µL PCR reaction with 5 µL of template DNA with an initial denaturation step at

95 °C for 5 min followed by 30 cycles of denaturation (50 s at 94 °C), annealing (30 s at 40 °C), and elongation (30 s at 72 °C). Amplified samples were then diluted 1:100 in water for input into library tailing PCR. This PCR reaction was similar to initial amplification except the PCR conditions consisted of an initial denaturation at 95 °C for 5 min followed by 15 cycles of denaturation (50 s at 94 °C), annealing (30 s at 40 °C), and elongation (1 min at 72 °C). Quantification of PCR products was performed using the Quant-iT PicoGreen dsDNA Assay Kit (Life Technologies, Grand Island, NY, USA). A subset of the amplified libraries was spot-checked on a Bioanalyzer High-Sensitivity DNA Chip (Agilent Technologies, Santa Clara, CA, USA) to ascertain if the amplicons were the predicted size. These samples were each normalized to 2 nM and pooled. The total volume of the libraries was reduced using a SpeedVac and amplicons were size-selected at 420 bp $\pm$ 20 % using the Caliper XT (Perkin Elmer, Waltham, MA, USA). The pooled libraries were cleaned with 1.8× AMPureXP beads (Beckman Coulter, Brea, CA, USA) and eluted with water. DNA concentration in the final pool was assayed with PicoGreen and normalized to 2 nM for input into Illumina MiSeq (v3 Kit) to produce 2 × 250 bp sequencing products. Clustering was performed at 10 pM with a 5 % spike of PhiX. A single lane on an Illumina MiSeq instrument was used to generate the 16S rRNA gene sequences. Raw sequencing data have been submitted to the NCBI Sequence Read Archive under project accession PRJNA284355.

PCR and qPCR

Quantitative real-time PCR was performed to assess the abundance of the *FadA* gene present in a subset of normal and tumor tissue pairs. DNA from the ATCC control strain of *F. nucleatum* 25586 was used as a positive control. *FadA* abundances were normalized relative to pan-eubacteria abundance per sample. Primers *FadA*-F (5'-GAAGAAAGAGCACAAAGCTGA-3') and *FadA*-R (5'-GCTTGAAGTCTTTGAGCTCT-3') were used to measure *FadA* [14], and primers for universal eubacteria 16S (5'-GGTGAATACGTTCCCGG-3') and (5'-TACGGCTA CCTTGTTACGACTT-3') were used to determine the total eubacterial abundance per sample [34]. The analysis was performed using 10 ng of DNA in a 20 µL reaction containing FastStart Universal SYBR Green Master Mix (Rox; Roche Diagnostics, Indianapolis, IN, USA) on an Applied Biosystems 7300 Real Time PCR system. Reactions were performed in triplicate. *FadA* relative abundances were calculated as per the Δ CT method [35]. Relative fold differences were calculated by dividing the *FadA* abundance from the normal samples by that of the tumor sample.

Fusobacterium genus-specific PCR was performed on a subset of samples using previously characterized primers: forward (5'-GGATTTATTTGGGCGTAAAGC-

3') and reverse (5'-GGCATTCTACAAATATCTACGAA-3') [34, 36]. The PCR was carried out using Accusart Taq polymerase (Quanta Biosciences, Gaithersburg, MD, USA) following the manufacturer's protocol for 30 cycles with an annealing temperature of 55 °C. DNA from the ATCC control strain *F. nucleatum* 25586 was used as a positive control.

Providencia genus-specific PCR was performed using a previously published protocol and primer set: sp16s-F1 (5'-ACCGCATAATCTCTTAGG-3') and Psp16s-R2 (5'-CTACACATGGAATTCTAC-3'), with the following modifications [37]. The PCR was carried out using Accusart Taq polymerase (Quanta Biosciences, Gaithersburg, MD, USA) following the manufacturer's protocol for 30 cycles with an annealing temperature of 50 °C. The ATCC control strain *Providencia alcalifaciens* 9886 was used as a positive control. Amplicons were resolved in 2 % agarose TAE gel.

Sequence analysis

The sequence data contained approximately 21.4 million reads passing quality filtering in total, inclusive of forward and reverse reads, with a mean value of 242,940 quality reads per sample. The forward and reverse read pairs were merged using the USEARCH v7 program 'fastq_mergepairs', allowing stagger, with no mismatches allowed [38]. Merged reads were quality trimmed, again using USEARCH, to truncate reads at any quality scores of 20 or less. Following merging and trimming, there were an average of 62,100 high quality reads per sample (median 48,817; range 6559–173,471). The fasta sequence headers were renamed using a custom script to conform to QIIME standards.

The merged and filtered reads were used to pick operational taxonomic units (OTUs) with QIIME v.1.7.0 using 'pick_otus.py', with the closed-reference usearch_ref OTU picking protocol against the Greengenes database (August 2013 release) at 97 % similarity [39–41]. Reverse read matching was enabled, while reference-based chimera calling was disabled. Rarefaction was performed on the OTU table at 5000 reads prior to subsequent analyses.

The final OTU table was used to generate a phylogenetic tree by including only taxa with at least 0.1 % relative abundance in at least half of all samples. Starting with the full reference tree provided by the Greengenes database (August 2013 release, file 97_otus_unannotated.tree), a smaller tree file that contained only this limited set of taxa was generated using a custom pipeline (Sycamore from the Alm laboratory at MIT). The output of this pipeline was visualized with the Interactive Tree of Life [41, 42]. See Additional file 2 for the OTU table used in this study.

We used a linear model to correct for several patient and tumor covariates, individually as well as in combination,

including patient age, sex, tumor stage, and tumor site. None of these factors, alone or in combination, were found to have a significant impact in this sample set. We also performed principle coordinate analysis using the difference between the tumor and normal abundances for each taxon. Using this unsupervised approach, there was no clear segregation of the patients by age, sex, tumor stage, site, or microsatellite instable (MSI) status. Additionally, we focused specifically on *Providencia* and *Fusobacterium*, and while there was a slight trend toward higher tumor stage with increases in these two genera at the tumor site, it was not statistically significant. We note that microsatellite instable/microsatellite stable (MSI/MSS) statuses were only available for 13 of the 44 patients.

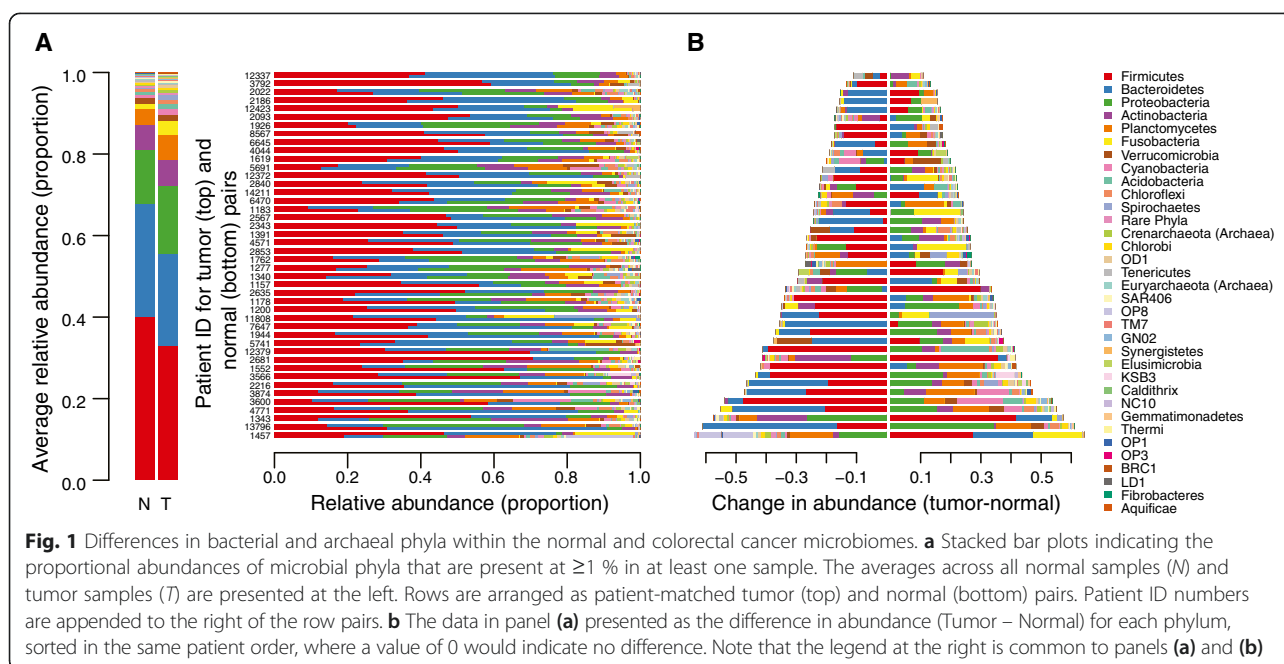
Correlation analysis was performed using SparCC, available at [43] from Jonathan Friedman at MIT, on the complete OTU table collapsed to the genus level [44]. Pseudo *p* values were inferred using 100 randomized sets. Correlations with pseudo *p* values ≤ 0.05 that were within two degrees of separation from *Providencia* or *Fusobacterium* with absolute correlations of 0.05 or more were visualized using Cytoscape v.3.1.0 [45].

The PICRUSt v.1.0.0 pipeline was used to generate a virtual metagenome using the OTU table containing raw counts generated in the previous analyses by QIIME [29, 39, 40]. Pathways and enzymes were assigned using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database options built into the pipeline. Virulence genes were identified by mapping the data in the PICRUSt enzyme abundance table to MViDB using the UniProt database file, idmapping.dat, available from [46], as a key. See Additional files 3 and 4 for the metabolic enzyme and pathway abundance tables, respectively.

Results

Tumor microenvironments harbor microbiomes distinct from those of normal tissue microenvironments

We obtained patient-matched normal and tumor colon tissue samples from the University of Minnesota Biological Materials Procurement Network (BioNet) from 44 patients (see Additional file 1 for sample information). We assessed the microbiome associated with each sample by Illumina sequencing across the V5-V6 hypervariable regions of the 16S rRNA gene (see "Methods" for details). This analysis showed variation in the bacterial phyla abundance when comparing the matched normal and tumor tissues (Fig. 1a). This variability is consistent with previous reports and demonstrates that there is indeed a cancer-associated signature in the tumor microbiome [6, 10, 16, 18, 34, 47, 48]. At the level of the phyla, each sample was dominated by Firmicutes, Bacteroidetes, and Proteobacteria. There were clear and significant changes in these phyla between the normal and cancer



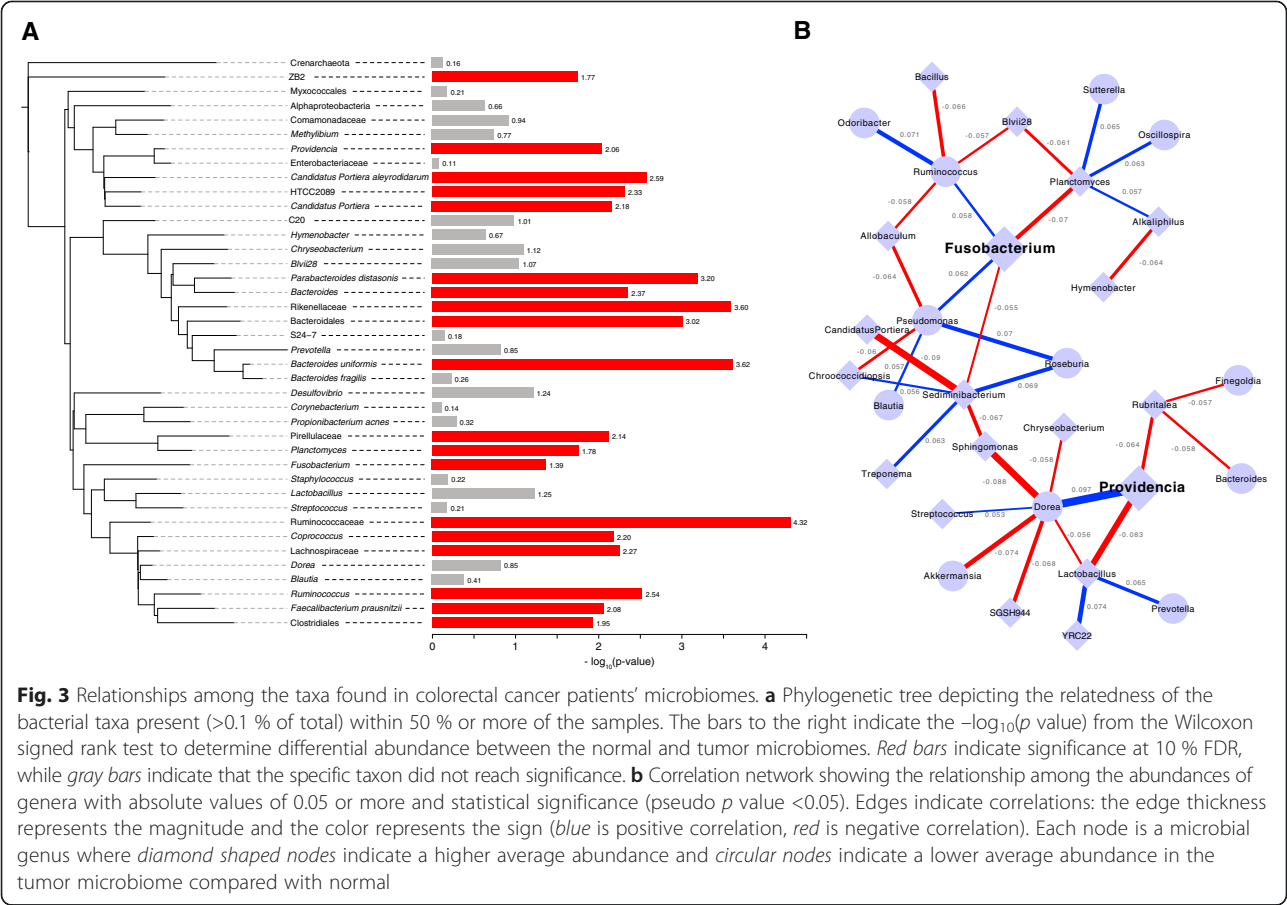
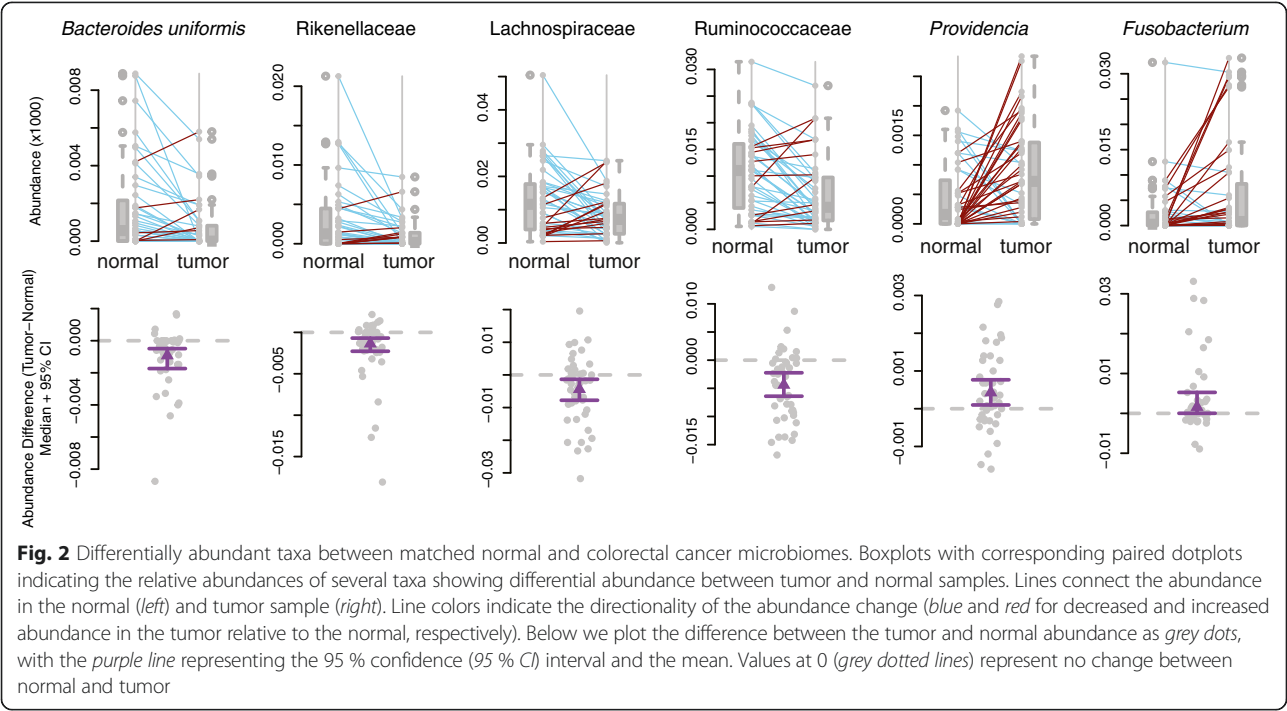
states, with the tumors showing an enrichment of Proteobacteria and a depletion of Firmicutes and Bacteroidetes (Fig. 1b). Also consistent with previous reports, we saw an increase in the phylum Fusobacteria in the tumor-associated microbiome (two-sided Wilcoxon signed rank test $q \leq 0.1$ after false discovery rate (FDR) correction for multiple tests) [7, 10, 12, 34].

When we assessed the differences at the level of OTUs we discovered numerous changes between the normal and tumor microbiomes with significant differences in the abundances of 19 different taxa (Wilcoxon signed rank test $q \leq 0.1$ after FDR correction; Additional file 5). Of note, the tumors showed decreases in the abundances of several taxa within the order Clostridiales, namely, Lachnospiraceae, Ruminococcaceae, and *Faecalibacterium prausnitzii*, as well as several members within the order Bacteroidales, including *Bacteroides*, Rikenellaceae, and *Bacteroides uniformis* (Figs. 2 and 3a). Taxa that were enriched in the tumor microbiomes included *Fusobacterium* and several Proteobacteria genera, including *Candidatus*, *Portiera* and *Providencia* (Figs. 2 and 3a; Additional file 6). Both *Fusobacterium* and *Providencia* are known pathogens, and when a correlation network is generated, it is clear that there are correlated abundance changes in the microbiome as a function of their presence (Fig. 3b). For instance, *Fusobacterium* species have been shown to have a mutualistic relationship with some *Pseudomonas* species at abscesses [49]. This co-occurrence is seen in our data as a positive correlation between the abundances of the two genera (Fig. 3b). Other specific interactions between different bacterial taxa remain speculative. In the case of *Lactobacillus* in

the human microbiome, it has been demonstrated that there can be reciprocal interference between species in this genus and other bacterial species in the form of competition for epithelial cell adhesion. As both *Lactobacillus* and *Providencia* utilize cell adhesion in their colonization of the human body, this may explain the negative correlation between the two genera in our dataset (Fig. 3b). While there was not a significant correlation between the relative abundances of *Fusobacterium* and *Providencia* in this analysis, we assessed the overlap among patients who showed increased levels of these genera at the tumor sites. Taken individually, *Fusobacterium* and *Providencia* were more abundant in the tumor microenvironment of 23 out of 44 and 28 out of 44 patients, respectively. Nineteen out of 44 patients showed increases in both of the genera in their tumor microenvironments with respect to their normal matched tissue microbiomes.

CRC-associated microbiome diversity

We calculated alpha-diversity using a variety of metrics within each of the samples using QIIME [39]. Alpha diversity metrics that account for phylogenetic relationships between the OTUs show that the tumor microbiomes exhibited higher alpha diversity than those of the normal, patient-matched microbiomes ($p = 0.029$ by two-sided Wilcoxon signed rank test). This is also true when using alternative measures of diversity such as the Shannon's index or the Inverse Simpson's ($p = 0.020$ and 0.024 , respectively, by Wilcoxon signed rank test; Additional file 7).



Variation in the functional pathways and enzymes in the tumor microbiome

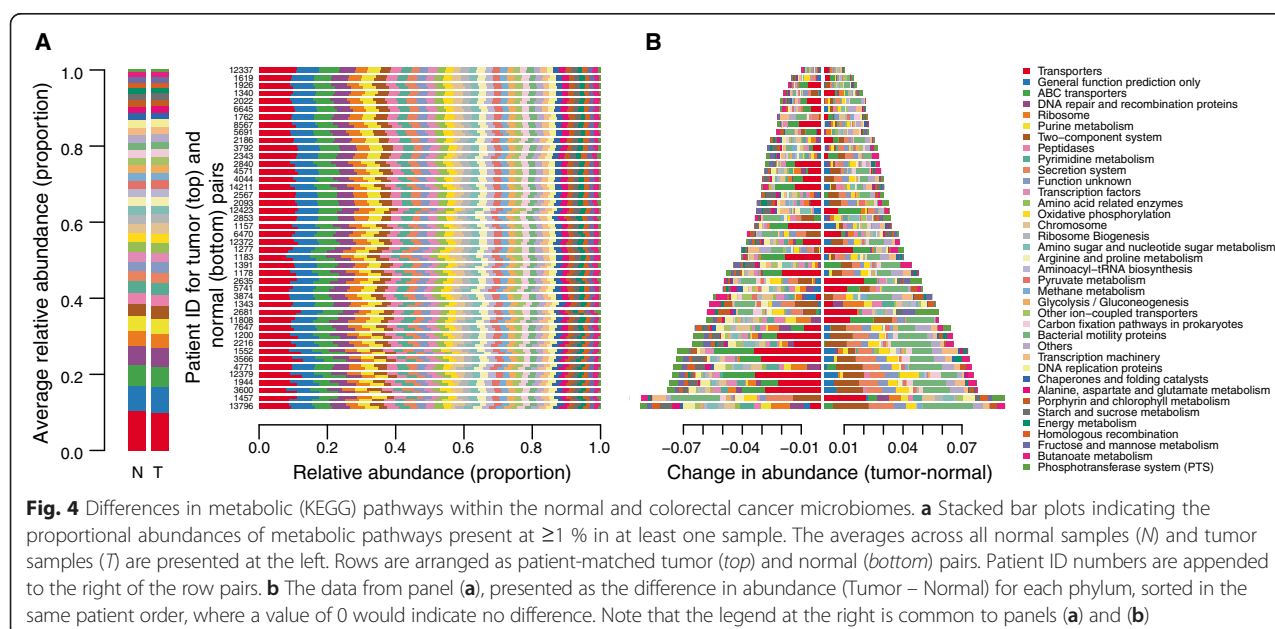
A recent report presented the validation of a pipeline that leverages knowledge of the human gut reference genomes to predict general microbiome function and enzyme composition from 16S rRNA gene sequencing data [29]. While this approach is not suitable for making conclusive statements regarding single, specific enzymes, it is appropriate for general functional comparisons between groups of samples, as is the case in this report. Using this validated pipeline — Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt) — we constructed a virtual metagenome for each of the samples' microbiomes [29]. The KEGG database was used as a reference to determine the abundances of metabolic pathways and enzymes within the virtual metagenomes [50, 51]. As with the bacterial phyla, we saw significant variation in the predicted functional pathways represented within each of the sampled microbiomes (Fig. 4a), though, as expected from previous studies, we find that the variability in phylum abundances is far greater than the variability in the functional pathways (Fig. 4b) [2, 52]. It is important to note that the results of this analysis are predictions only and not direct measurements of sequences that correspond to pathway member or enzyme genes. Despite the validation of this prediction approach, it is possible that this method biases the predictions toward microbial genomes that are well documented to the exclusion of other, unknown or poorly documented taxa.

These observations highlight the substantial functional redundancy across the phyla. In other words, diversity among the taxa within a given microbiome can mask the

functional similarities — taxonomically distinct microbes in the gut can operate identically, or nearly so, at the functional level. The patient-by-patient variability in phyla does not perfectly correspond to that seen at the functional pathway level, though, as expected, analysis at the level of enzymes and pathways provides insights that analyses of the taxa alone may not be able to identify, due to the high level of between-sample variation. In general, the differences seen at the pathway level are roughly an order of magnitude less than the differences seen at the phylum level (compare Figs. 1b and 4b). Although the pathway differences are smaller than those at the level of the phylum, there remain physiologically relevant, statistically significant changes between the normal and tumor metagenomes.

Twenty pathways (as defined by KEGG, level 3) were found to be differentially abundant between the tumor and normal tissue. Alanine, aspartate, and glutamate metabolism, DNA replication proteins, and starch and sucrose metabolism were significantly depleted in the tumor microbiome ($q \leq 0.01$ for each pathway by two-sided Wilcoxon signed rank test after FDR correction; Fig. 5a). Conversely, secretion system, two-component system, and bacterial motility protein pathways were significantly enriched in the tumor microbiome ($q \leq 0.04$ for each pathway by two-sided Wilcoxon signed rank test after FDR correction; Fig. 5a).

To more closely examine the variation in the microbiome virulence potential, we used information regarding virulence association from MVirDB to annotate the predicted enzymes [53]. We found that the tumor-associated microbiome is significantly enriched with enzymes related to microbial virulence. We found this enrichment to be



significant when including all possible virulence categories ($p = 0.0046$ by Fisher's exact test; Fig. 5b, c). Additionally, when assessing enrichment for virulence-related genes by known functional categories in MVirDB, the tumors were significantly enriched for genes encoding general virulence proteins ($p = 5.8 \times 10^{-5}$, by Fisher's exact test). Genes encoding bacterial toxins were also found at higher abundance in the tumor, but the enrichment was not

statistically significant ($p = 0.17$ by Fisher's exact test), likely due to low total gene counts for some categories (e.g., protein toxins).

We predicted that this virtual enrichment was driven by known pathogenic bacteria within the microbiome, i.e., *Providencia* and *Fusobacterium*. In fact, when a comparison is made among the virulence-associated genes significantly differentially found in the tumor microenvironment

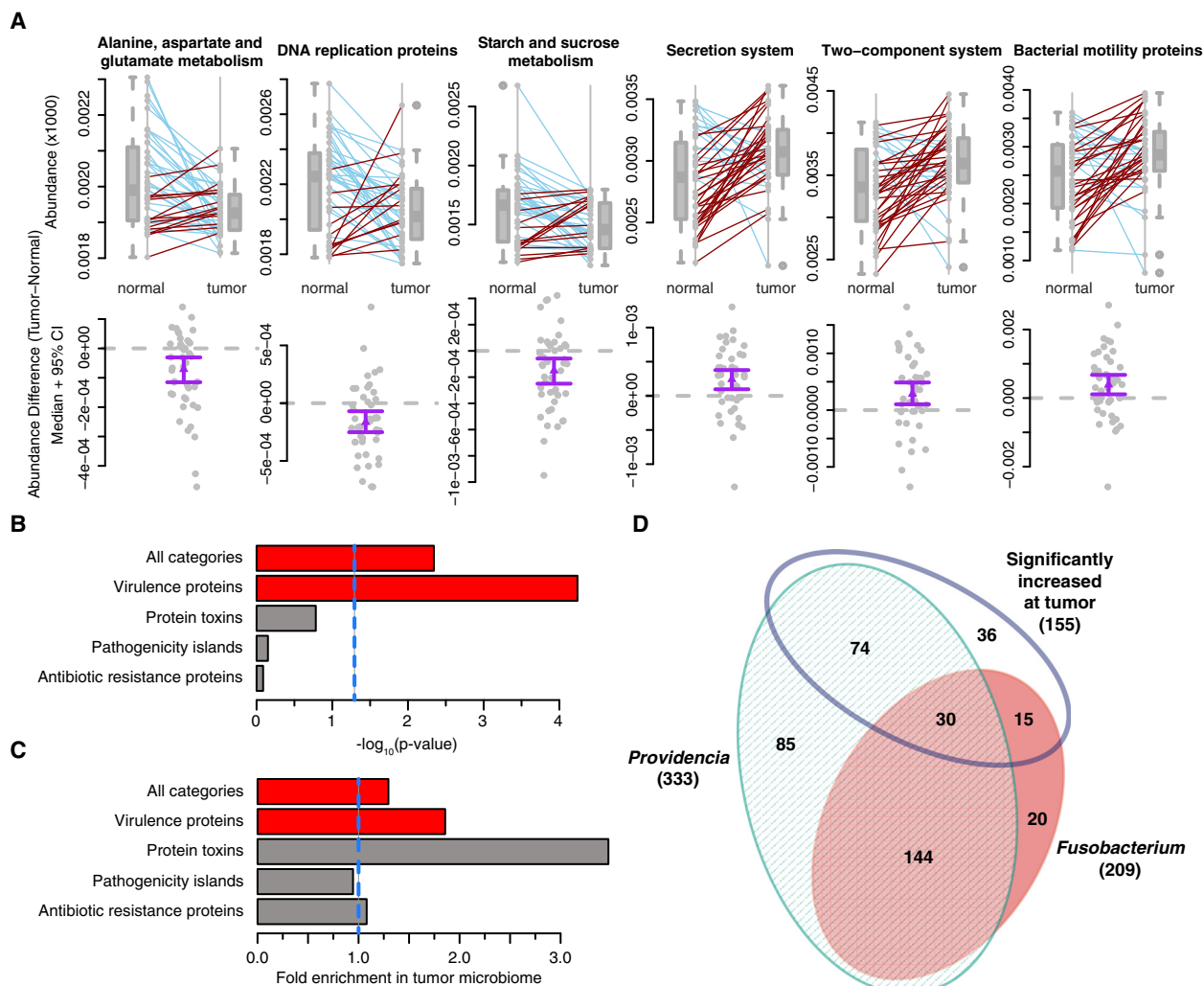


Fig. 5 Differentially abundant pathways and enzyme classes between matched normal and colorectal tissue microbiomes. **a** Boxplots with corresponding paired dotplots indicating the relative abundances of several pathways showing differential abundance between tumor and normal samples. Lines connect the abundance in the normal (left) and tumor sample (right). Line colors indicate the directionality of the abundance change (blue and red for decreased and increased abundance in the tumor relative to the normal, respectively). Below we plot the difference between the tumor and normal abundance as grey dots, with the purple line representing the 95 % confidence interval (95 % CI) and the mean. Values at 0 (grey dotted lines) represent no change between normal and tumor. **b** Bar chart showing the p values ($-\log_{10}$ transformed) obtained from Fisher's exact test used to determine virulence category enrichment in the tumor-associated microbiome on the x-axis with the gene categories labeled on the y-axis. Red bars indicate significance by Fisher's exact test ($p < 0.005$) and gray bars indicate no statistical significance. The blue dashed line indicates the standard significance cutoff of $p = 0.05$. **c** Bar chart from the analysis in panel (b), demonstrating the fold-enrichment of virulence protein-encoding genes in the tumor-associated microbiome. The x-axis is the fold enrichment of the different virulence enzyme classes within the tumor microbiome relative to the normal microbiome. The vertical blue dotted line placed at 1 indicates the point where there is no difference between the normal and tumor microbiomes. **d** Venn diagram indicating the numbers of shared virulence-associated genes among *Providencia*, *Fusobacterium*, and the set of statistically significantly increased abundance genes at the tumor

(155), the virulence genes associated with *Providencia* (333), and the virulence genes associated with *Fusobacterium* (209), there is substantial overlap among and between the three different groups (Fig. 5d). In order to exclude the possibility that the virulence enrichment and corresponding overlap between the virulence genes found at the tumor site were the result of non-specific effects rather than due to the potential contributions of *Providencia* and *Fusobacterium*, the entire PICRUSt pipeline and subsequent enrichment analyses were repeated using an OTU table from QIIME with both *Fusobacterium* and *Providencia* explicitly excluded. In this case, while there were still 123 virulence genes from other taxa associated with the tumor microbiome, the enrichment is not significant compared with the background of normal tissue from the same cancer patients (Fisher's exact test one-sided p value >0.9). This demonstrates that the virulence signature in the microenvironment is dependent on *Fusobacterium* and *Providencia*.

To highlight the differences seen when assessing patient-matched tissue samples compared with assessing case and control stool samples, we performed a comparison of our results with those of Zackular et al. [11]. Zackular et al. performed an assessment of the microbiomes of CRC patients' stool samples relative to those of normal patients or patients with colorectal adenomas [11]. In comparison with the data presented here, a prediction would be that both *Fusobacterium* and *Providencia* would be found at increased abundances in CRC patients relative to normal controls. While the data from Zackular et al. show a statistically significant increase in *Fusobacterium* in CRC patients' stools, *Providencia* at the genus level was not detected in the stool samples. However, there was a clear trend showing a doubling of the abundance of Enterobacteriaceae (the family to which *Providencia* belongs) when looking at stools from normal patients in comparison with stools from patients with adenomas and yet again when comparing adenoma patients and CRC patients. A likely explanation for the difference between our study and Zackular et al. is that data in the latter study were collected from stools and were not patient-matched. Thus, inter-individual variability likely decreased the ability to identify some tumor-associated taxa, such as *Providencia*.

Discussion

At the phylum level, the differences seen between the normal and tumor tissue-associated microbiomes are consistent with many previous reports [6, 10, 16, 18, 34, 47, 48]. When assessing the data using information that accounts for more fine-grained detail with respect to taxonomy, we have made several important findings. Two of the genera we found to be enriched in the tumor microbiome, *Providencia* and *Fusobacteria*, are known to be pathogenic;

Fusobacteria has been implicated previously in CRC [7, 10, 12, 34].

Species belonging to the genus *Providencia* have been implicated as infectious agents causing urinary tract infections, ocular infections, and gastroenteritis [37, 54–60]. In addition, it is a genus in which several sub-strains have acquired resistance to commonly used antibiotics [56, 61, 62]. *Fusobacterium* is a genus that encompasses several species known to be pathogenic in humans; they are obligate anaerobes, with known sites of infection in the oral cavity as well as in the gastrointestinal tract [63, 64]. The finding that these particular genera are prevalent in the tumor microenvironment suggests several alternative, though not mutually exclusive, hypotheses. One possibility is that these bacteria are causative in oncogenesis or tumor progression; another possibility is that these species are being enriched as the tumor has formed a niche that favors these bacteria. In the case of *Fusobacteria*, the results from several different studies, both correlative and mechanistic, indicate that it is likely a cancer driver [7, 14, 34]. In the case of *Providencia*, there are as yet no definitive studies that implicate this genus as a contributor to CRC. The discovery of *Providencia* in the tumor microbiome is interesting as, similar to *Fusobacteria*, it encodes a potent, immunogenic lipopolysaccharide [58, 65]. In fact, several virulence genes responsible for lipopolysaccharide biosynthesis are shared by both genera and are also significantly increased in the tumor microenvironment (Fig. 5d). A recent study, using *Drosophila* as a model system, performed a genomic comparison of four different species of *Providencia* isolated from the human gut [66]. These researchers demonstrated that these four species share common sets of virulence-related genes, including a type 3 secretion system and genes for cell adhesion. Additionally, *Providencia* has been shown to disrupt the epithelial membrane in the intestines, though the mechanism by which this is accomplished is still unclear [55, 58, 67, 68]. These factors manifest phenotypically as gastroenteritis, though with our discovery of its association with the cancer microenvironment, it is a promising candidate cancer-promoting pathogen [60].

From a diagnostic and therapeutic perspective, assessing the CRC-associated microbiome by testing for differentially abundant taxa is an eminently worthwhile endeavor as it is the logical location to look for specific taxa that could be biomarkers and/or targets for intervention in CRC. However, it is possible that the search for specific taxa might miss the larger perspective. For instance, as described above, *Fusobacteria* and *Providencia* share many important phenotypic characteristics — potent, immunogenic lipopolysaccharide and the ability to damage colorectal tissue. These similarities might be better assessed using metagenomic or metatranscriptomic approaches, virtual or otherwise, as these key features are undoubtedly

reflected in the genes that these particular bacteria encode, many of which are shared virulence factors with known detrimental properties.

Defining a clear set of virulence factors a priori as targets of interest and assessing their relative expression is a promising approach to CRC therapy. This report proposes such an approach by showing the striking predicted enrichment of virulence genes in the tumor-associated microbiome, potentially driven by *Fusobacteria* and *Providencia*. The fact that virulence proteins are predicted to be enriched in the tumor-associated microbiome lends support to the hypothesis that the microbiome is an active contributor to CRC and not just a passive byproduct of the changes the tumor makes in the organ. In the case of *F. nucleatum*, it is clear that there is a direct functional link between the bacteria and cancer development, though more work using cell culture and model organisms will be needed in the future to empirically assess the mechanistic interplay between colorectal tissues and specific components of the microbiome [14]. It is important to note that this clear enrichment is likely underestimated because MViDB, while expansive, does not currently encompass all known virulence genes in the microbiome, and, as the field of medical microbial genomics advances, new virulence genes will undoubtedly be discovered. For instance, the FadA protein from *F. nucleatum* has been reported as a critical virulence factor, yet as it is a recent report, this finding has not yet made its way into MViDB as of this submission [14, 53].

It is important to note that this research uses 16S rRNA gene sequences as the starting point. Although this approach has obvious benefits in terms of resource expenditures and computational processing, there are several potential disadvantages. First, the microbiome functional assessment presented here uses a prediction method that, while validated and robust when applied to human gut samples, remains a prediction and may not necessarily perfectly reflect the biological reality [29]. Another concern, as with all DNA-based approaches, is that even when a gene is predicted in a sample, it may still not be expressed or active. Additional metatranscriptomic research will undoubtedly shed light on this situation in the future. We also note that although our results show microbiome patterns with potential roles in cancer, we cannot determine causality as part of this study. We expect follow-up studies to focus on assessing the causal role of the microbiome in colon cancer using animal models and cell culture systems.

Conclusions

It is clear that there are numerous taxa in the CRC microbiome that are correlated with the disease. Here, in addition to the previously reported genus *Fusobacterium*, we report the discovery of another genus with similar

pathogenic features, *Providencia*. This manuscript also presents an analysis that incorporates predicted information at the functional (e.g., virulence potential) level to assess differences between the normal and cancer-associated microbiomes. It is important to note that these two approaches (taxonomy-based and function-based) provide different, yet interdependent information about the microbes in the tissue microenvironment. Our work demonstrates that utilizing this combined approach can provide researchers with specific taxa as biomarkers and/or therapeutic targets while also looking globally at the predicted pathogenic potential of the microbiome and showing a clear predicted enrichment of virulence-associated microbial genes present in the CRC microbiome. As with the bacterial genera associated with the disease, these virulence genes may provide researchers and clinicians with targets for therapeutic intervention to improve patient outcomes.

Additional files

Additional file 1: Sample and patient clinical information and metadata.

A tab-delimited file containing a sample map containing patient and tumor clinical information as well as the specific sample numbers that will allow proper identification of tumor and normal matched pairs.

Additional file 2: Unfiltered OTU table. A tab-delimited file containing the unfiltered OTU table with proportional abundances of the taxa output from the QIIME pipeline.

Additional file 3: Enzyme abundances. A tab-delimited file containing the unfiltered proportional abundances of enzymes detected in the virtual metagenomes of each sample using PICRUSt. The file also contains an additional column with an annotation from MViDB, when available, as well as a column that indicates if the predicted enzyme was significantly more abundant at the site of the tumor.

Additional file 4: Pathway abundances. A tab-delimited file containing the unfiltered level 3 KEGG pathway proportional abundances for each sample generated using PICRUSt.

Additional file 5: Taxon significance testing. A tab-delimited file containing *p* values, *q* values, and the site at which the taxa are more abundant for the taxa tested for significance after correcting for multiple tests. Those included in this analysis were those taxa that were present in at least 50 % of all samples at an abundance of >0.1 %.

Additional file 6: Genus-specific PCR and FadA qPCR. **a** Genus-specific PCR for *Fusobacterium*. Genus-specific PCR was carried out for a subset of samples for normal and tumor-matched DNA samples. A PCR reaction containing only water rather than DNA was used as a negative control, while DNA from the ATCC control strain *Fusobacterium nucleatum* 25586 was used as a positive control. Bands on the 2 % agarose TAE gel are visible for both the positive control as well as for the tumor samples at the expected amplicon size of 162 base pairs. **b** Genus-specific PCR for *Providencia*. Genus-specific PCR was carried out for a subset of samples for normal and tumor matched DNA samples. A PCR reaction containing only water rather than DNA was used as a negative control, while DNA from the ATCC control strain *Providencia alcalifaciens* 9886 was used as a positive control. Bands on the 2 % agarose TAE gel are present for both the positive control as well as for the tumor samples at the expected amplicon size of 515 base pairs. **c** FadA qPCR. A subset of samples was used to determine the relative abundance of the FadA gene in tumor samples relative to normal samples. Both the normal and tumor samples were normalized internally to the total abundance of eubacteria. The value for the normalized tumor FadA abundance was divided by the normalized value for normalized normal FadA abundance to arrive at the fold differences indicated.

Additional file 7: Microbial diversity within normal and tumor-associated microbiomes. Paired line plots show the phylogenetic diversity, Shannon's Index, and Inverse Simpson's Index (alpha diversity metrics) for the microbiomes associated with normal and patient-matched tumor samples. The colors of the lines represent the direction of the change for each matched pair (*blue lines* indicate a decrease in diversity from normal to tumor, while *red lines* indicate an increase in diversity from the normal to tumor). *P* values were calculated using a two-sided Wilcoxon signed rank test.

Abbreviations

CRC: colorectal cancer; FDR: false discovery rate; KEGG: Kyoto Encyclopedia of Genes and Genomes; OTU: operational taxonomic unit; PCR: polymerase chain reaction; PICRUST: Phylogenetic Investigation of Communities by Reconstruction of Unobserved States; qPCR: quantitative PCR; SCFA: short chain fatty acid.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

MBB and RB conceived of the study and study design. MBB carried out DNA extraction, sequence analyses, figure generation, and manuscript preparation. RB provided guidance with statistical analyses and assisted in figure generation. JL assisted in statistical analyses. TKS provided support with the experimental design related to patient sample inclusion and data interpretation. DK provided guidance with respect to the PICRUST analysis and experimental design. All authors contributed to manuscript revisions and read and approved the final version.

Acknowledgements

The authors thank the members of the Blekman lab for useful discussions, the Huttenhower lab at Harvard University for providing a publicly available Galaxy server for use with PICRUST software, and the Alm lab at MIT for making their Sycamore pipeline available. This work was carried out, in part, using computing resources at the Minnesota Supercomputing Institute.

Author details

¹Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, MN, USA. ²Department of Ecology, Evolution, and Behavior, University of Minnesota, Minneapolis, MN, USA. ³Masonic Cancer Center, University of Minnesota, Minneapolis, MN, USA. ⁴Department of Obstetrics, Gynecology and Women's Health, University of Minnesota, Minneapolis, MN, USA. ⁵Department of Computer Science and Engineering, University of Minnesota, Minneapolis, MN, USA. ⁶Biotechnology Institute, University of Minnesota, Minneapolis, MN, USA.

Published online: 24 June 2015

References

- Jemal A, Bray F, Center MM, Ferlay J, Ward E, Forman D. Global cancer statistics. *CA Cancer J Clin*. 2011;61:69–90.
- The Human Microbiome Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*. 2012;486:207–14.
- Jones ML, Ganopolsky JG, Martoni CJ, Labbé A, Prakash S. Emerging science of the human microbiome. *Gut Microbes*. 2014;5:446–57.
- Konstantinov SR, Kuipers EJ, Peppelenbosch MP. Functional genomic analyses of the gut microbiota for CRC screening. *Nat Rev Gastroenterol Hepatol*. 2013;10:741–5.
- Sobhani I, Tap J, Roudot-Thoraval F, Roperch JP, Letulle S, Langella P, et al. Microbial dysbiosis in colorectal cancer (CRC) patients. *PLoS One*. 2011;6:e16393.
- Bonnet M, Buc E, Sauvanet P, Darcha C, Dubois D, Pereira B, et al. Colonization of the human gut by *E. coli* and colorectal cancer risk. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2014;20:859–67.
- Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, Strauss J, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*. 2012;22:299–306.
- Tahara T, Yamamoto E, Suzuki H, Maruyama R, Chung W, Garriga J, et al. *Fusobacterium* in colonic flora and molecular features of colorectal carcinoma. *Cancer Res*. 2014;74:1311–8.
- Chen W, Liu F, Ling Z, Tong X, Xiang C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS One*. 2012;7:e39743.
- Kostic AD, Gevers D, Pedamallu CS, Michaud M, Duke F, Earl AM, et al. Genomic analysis identifies association of *Fusobacterium* with colorectal carcinoma. *Genome Res*. 2012;22:292–8.
- Zackular JP, Rogers MAM, Ruffin MT, Schloss PD. The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev Res (Phila)*. 2014;7:1112–21.
- Marchesi JR, Dutilh BE, Hall N, Peters WHM, Roelofs R, Boleij A, et al. Towards the human colorectal cancer microbiome. *PLoS One*. 2011;6:e20447.
- Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. *Nature*. 2012;487:330–7.
- Rubinstein MR, Wang X, Liu W, Hao Y, Cai G, Han YW. *Fusobacterium nucleatum* promotes colorectal carcinogenesis by modulating E-cadherin/β-catenin signaling via its FadA adhesin. *Cell Host Microbe*. 2013;14:195–206.
- Irrazabal T, Belcheva A, Girardin SE, Martin A, Philpott DJ. The multifaceted role of the intestinal microbiota in colon cancer. *Mol Cell*. 2014;54:309–20.
- Ahn J, Sinha R, Pei Z, Dominianni C, Wu J, Shi J, et al. Human gut microbiome and risk for colorectal cancer. *J Natl Cancer Inst*. 2013;105:1907–11.
- Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol*. 2014;10:766.
- Shen XJ, Rawls JF, Randall T, Burcal L, Mpande CN, Jenkins N, et al. Molecular characterization of mucosal adherent bacteria and associations with colorectal adenomas. *Gut Microbes*. 2010;1:138–47.
- Mira-Pascual L, Cabrera-Rubio R, Ocon S, Costales P, Parra A, Suarez A, et al. Microbial mucosal colonic shifts associated with the development of colorectal cancer reveal the presence of different bacterial and archaeal biomarkers. *J Gastroenterol*. 2014;50:167–79.
- David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505:559–63.
- Cullender TC, Chassaing B, Janzon A, Kumar K, Muller CE, Werner JJ, et al. Innate and adaptive immunity interact to quench microbiome flagellar motility in the gut. *Cell Host Microbe*. 2013;14:571–81.
- Knights D, Lassen KG, Xavier RJ. Advances in inflammatory bowel disease pathogenesis: linking host genetics and the microbiome. *Gut*. 2013;62:1505–10.
- Geng J, Fan H, Tang X, Zhai H, Zhang Z. Diversified pattern of the human colorectal cancer microbiome. *Gut Pathog*. 2013;5:2.
- Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ, et al. Microbiota organization is a distinct feature of proximal colorectal cancers. *Proc Natl Acad Sci U S A*. 2014;111:18321–6.
- Weir TL, Manter DK, Sheflin AM, Barnett BA, Heuberger AL, Ryan EP. Stool microbiome and metabolome differences between colorectal cancer patients and healthy adults. *PLoS One*. 2013;8:e70803.
- Warren RL, Freeman DJ, Pleasance S, Watson P, Moore RA, Cochrane K, et al. Co-occurrence of anaerobic bacteria in colorectal carcinomas. *Microbiome*. 2013;1:16.
- Ohgashi S, Sudo K, Kobayashi D, Takahashi O, Takahashi T, Asahara T, et al. Changes of the intestinal microbiota, short chain fatty acids, and fecal pH in patients with colorectal cancer. *Dig Dis Sci*. 2013;58:1717–26.
- Weaver GA, Krause JA, Miller TL, Wolin MJ. Short chain fatty acid distributions of enema samples from a sigmoidoscopy population: an association of high acetate and low butyrate ratios with adenomatous polyps and colon cancer. *Gut*. 1988;29:1539–43.
- Langille MGJ, Zaneveld J, Caporaso JG, McDonald D, Knights D, Reyes JA, et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol*. 2013;31:814–21.
- Dutilh BE, Backus L, van Hijing SAFT, Tjalsma H. Screening metatranscriptomes for toxin genes as functional drivers of human colorectal cancer. *Best Pract Res Clin Gastroenterol*. 2013;27:85–99.
- Burns MB, Lackey L, Carpenter MA, Rathore A, Land AM, Leonard B, et al. APOBEC3B is an enzymatic source of mutation in breast cancer. *Nature*. 2013;494:366–70.
- Vandevanter PE, Weigel KM, Salazar J, Erwin B, Irvine B, Doeblner R, et al. Mechanical disruption of lysis-resistant bacterial cells by use of a miniature, low-power, disposable device. *J Clin Microbiol*. 2011;49:2533–9.

33. Cai L, Ye L, Tong AHY, Lok S, Zhang T. Biased diversity metrics revealed by bacterial 16S pyrotags derived from different primer sets. *PLoS One*. 2013;8, e53649.
34. Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. 2013;14:207–15.
35. Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res*. 2001;29:e45–5.
36. Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CMJE, Savelkoul PHM. Periodontal pathogens: a quantitative comparison of anaerobic culture and real-time PCR. *FEMS Immunol Med Microbiol*. 2005;45:191–9.
37. Shima A, Hinenoya A, Asakura M, Nagita A, Yamasaki S. Prevalence of *Providencia* strains among children with diarrhea in Japan. *Jpn J Infect Dis*. 2012;65:545–7.
38. Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*. 2010;26:2460–1.
39. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Methods*. 2010;7:335–6.
40. Navas-Molina JA, Peralta-Sánchez JM, González A, McMurdie PJ, Vázquez-Baeza Y, Xu Z, et al. Advancing our understanding of the human microbiome using QIIME. *Methods Enzymol*. 2013;531:371–444.
41. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006;72:5069–72.
42. Letunic I, Bork P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res*. 2011;39:W475–8.
43. SparCC. <https://bitbucket.org/yonatanf/sparcc>.
44. Friedman J, Alm EJ. Inferring correlation networks from genomic survey data. *PLoS Comput Biol*. 2012;8, e1002687.
45. Lopes CT, Franz M, Kazi F, Donaldson SL, Morris Q, Bader GD. Cytoscape Web: an interactive web-based network browser. *Bioinforma Oxf Engl*. 2010;26:2347–8.
46. Uniprot. <http://www.uniprot.org>.
47. Arthur JC, Perez-Chanona E, Mühlbauer M, Tomkovich S, Uronis JM, Fan T-J, et al. Intestinal inflammation targets cancer-inducing activity of the microbiota. *Science*. 2012;338:120–3.
48. Wu N, Yang X, Zhang R, Li J, Xiao X, Hu Y, et al. Dysbiosis signature of fecal microbiota in colorectal cancer patients. *Microb Ecol*. 2013;66:462–70.
49. Brook I, Walker RI. The relationship between *Fusobacterium* species and other flora in mixed infection. *J Med Microbiol*. 1986;21:93–100.
50. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res*. 2014;42:D199–205.
51. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28:27–30.
52. Carroll IM, Chang Y-H, Park J, Sartor RB, Ringel Y. Luminal and mucosal-associated intestinal microbiota in patients with diarrhea-predominant irritable bowel syndrome. *Gut Pathog*. 2010;2:19.
53. Zhou CE, Smith J, Lam M, Zemla A, Dyer MD, Slezak T. MvirDB—a microbial database of protein toxins, virulence factors and antibiotic resistance genes for bio-defence applications. *Nucleic Acids Res*. 2007;35:D391–4.
54. Koreishi AF, Schechter BA, Karp CL. Ocular infections caused by *Providencia rettgeri*. *Ophthalmology*. 2006;113:1463–6.
55. Murata T, Iida T, Shiomi Y, Tagomori K, Akeda Y, Yanagihara I, et al. A large outbreak of foodborne infection attributed to *Providencia alcalifaciens*. *J Infect Dis*. 2001;184:1050–5.
56. Lau S-M, Peng M-Y, Chang F-Y. Resistance rates to commonly used antimicrobials among pathogens of both bacteremic and non-bacteremic community-acquired urinary tract infection. *J Microbiol Immunol Infect Wei Mian Yu Gan Ran Za Zhi*. 2004;37:185–91.
57. Kholodkova EV, Kriukov IM, Baturo AP, Lifshits MB, Glebovskaia MA. Etiologic role of bacteria of the genus *Providencia* in acute intestinal diseases. *Zh Mikrobiol Epidemiol Immunobiol*. 1977;1:20–3.
58. Asakura H, Momose Y, Ryu C-H, Kasuga F, Yamamoto S, Kumagai S, et al. *Providencia alcalifaciens* causes barrier dysfunction and apoptosis in tissue cell culture: potent role of lipopolysaccharides on diarrheagenicity. *Food Addit Amp Contam Part A Chem Anal Control Expo Risk Assess*. 2013;30:1459–66.
59. Shima A, Hinenoya A, Asakura M, Sugimoto N, Tsukamoto T, Ito H, et al. Molecular characterizations of cytolethal distending toxin produced by *Providencia alcalifaciens* strains isolated from patients with diarrhea. *Infect Immun*. 2012;80:1323–32.
60. Yoh M, Matsuyama J, Ohnishi M, Takagi K, Miyagi H, Mori K, et al. Importance of *Providencia* species as a major cause of travellers' diarrhoea. *J Med Microbiol*. 2005;54:1077–82.
61. Lee H-W, Kang H-Y, Shin K-S, Kim J. Multidrug-resistant *Providencia* isolates carrying blaPER-1, blaVIM-2, and armA. *J Microbiol Seoul Korea*. 2007;45:272–4.
62. Luzzaro F, Mezzatesta M, Mugnaioli C, Perilli M, Stefani S, Amicosante G, et al. Trends in production of extended-spectrum beta-lactamases among enterobacteria of medical interest: report of the second Italian nationwide survey. *J Clin Microbiol*. 2006;44:1659–64.
63. Kolenbrander PE. Oral microbial communities: biofilms, interactions, and genetic systems. *Annu Rev Microbiol*. 2000;54:413–37.
64. Allen-Vercoe E, Strauss J, Chadee K. *Fusobacterium nucleatum*: an emerging gut pathogen? *Gut Microbes*. 2011;2:294–8.
65. Onoue S, Niwa M, Isshiki Y, Kawahara K. Extraction and characterization of the smooth-type lipopolysaccharide from *Fusobacterium nucleatum* JCM 8532 and its biological activities. *Microbiol Immunol*. 1996;40:323–31.
66. Galac MR, Lazzaro BP. Comparative genomics of bacteria in the genus *Providencia* isolated from wild *Drosophila melanogaster*. *BMC Genomics*. 2012;13:612.
67. Albert MJ, Alam K, Ansaruzzaman M, Islam MM, Rahman AS, Haider K, et al. Pathogenesis of *Providencia alcalifaciens*-induced diarrhea. *Infect Immun*. 1992;60:5017–24.
68. Guth BE, Perrella E. Prevalence of invasive ability and other virulence-associated characteristics in *Providencia alcalifaciens* strains isolated in São Paulo, Brazil. *J Med Microbiol*. 1996;45:459–62.