Differences in the Stool Microbiome Before and After Colorectal Cancer Treatment

\begin{center}

Running Title: Human Microbiome and Colorectal Cancer

Marc A Sze[1], Nielson T Baxter[2], Mack T Ruffin IV[3], Mary AM Rogers[4], and Patrick D Schloss[1†]

† To whom correspondence should be addressed: pschloss@umich.edu

[1]Department of Microbiology and Immunology, University of Michigan, Ann Arbor, MI

[2]Department of Internal Medicine, Division of Infectious Disease, University of Michigan, Ann Arbor, MI

[3]Department of Family Medicine and Community Medicine, Penn State Hershey Medical Center, Hershey, PA

[4]Department of Internal Medicine, University of Michigan, Ann Arbor, MI \end{center}

Abstract

Background: Colorectal cancer (CRC) continues to be a worldwide health problem with early detection being used as a key component in mitigating deaths due to the disease. Previous research suggests a link between stool bacterial microbiome and CRC. The overall objective was to investigate the changes in the bacterial microbiome after surgery in patients with lesion (i.e. adenoma or carcinoma). Specifically, we wanted to identify what within the community was different within those undergoing surgical removal of lesion. We also investigated the use of the bacterial microbiome and Fecal Immunoglobulin Test (FIT) to build models to either classify individuals as having a lesion or whether, based on the bacterial microbiome, the sample could be classified correctly as before or after surgery.

Results: Adenoma individual's bacterial microbiome were more similar to their pre-surgery sample then those with carcinoma (P-value = 0.00198) and this was also reflected in FIT as well (P-value = 2.15e-05). There was no significant difference in any indivdiual OTU between samples before and after surgery (P-value > 0.125). A model with a total of 37 variables was able to classify lesion with an AUC range of 0.847 to 0.791 while the model to classify samples as before and after had 33 with an AUC range of 0.79 to 0.651 for 100 20 repeated 10-fold cross-validated runs. Both models had a significant decrease in the positive probability of a lesion between individual's before versus after surgery samples (P-value = 1.91e-11 and 6.72e-12). In total there were 14 OTUs that were common to both models and were mostly commensals with largest representation from OTUs belonging to Bacteroides, Blautia, Streptococcus, and Clostridiales.

Conclusions: Our data suggests that treatment not only significantly reduces the probability of having a colonic lesion but also causes detectable changes in the bacterial microbiome. Further surveillance of these individuals will enable us to determine whether models such as the one we present here can also be used to predict recurrence of colorectal cancer.

Keywords

bacterial microbiome; colorectal cancer; polyps; FIT; detection; risk factors

Background

Colorectal cancer (CRC) continues to be a leading cause of cancer related deaths and is the second most common cancer death among men aged 40-79 years of age [1,2]. Over the last few years death due to the disease has seen a significant decrease, thanks mainly to improvements in screening [1]. However, despite this improvement there are still approximately 50,000 deaths from the disease a year [2]. It is estimated that around 5-10% of all CRCs can be explained by autosomal dominant inheritance [3]. The vast majority of CRCs are not inherited and the exact etiology of the disease has not been well worked out [2]. Although many risk factors have been identified [2] and non-invasive screening techniques have started to be put into consistent use [4,5] there has been an consistent additional increase in the incidence of CRC in the younger population.

This increased incidence of CRC in the younger population is concerning since having either an adenoma or carcinoma increases ones risk for future adenomas or carcinomas [6–8]. This increased risk can also carry with it an increased chance of mortality due to this recurrence [9,10]. Therefore there has been a great amount of interest in early risk stratification tools [11,12] that can help identify those they may be at most susceptibility to reccurence. Concurrently, there has also been a lot of interest in new areas that could have a role in disease pathogenesis, such as the gut bacterial microbiome.

There has been promising work on the bacterial microbiome and it's ability to be able to complement existing screening methods such as Fecal Immunoglobulin Test (FIT) or act alone as a screening tool [13,14]. There has also been research into how this microbiome could be altered directly on tumor tissue itself [15]. A few studies have now even shown how this microbiome [16] or specific members within it [17] could be directly involved with the pathogenesis of CRC. These studies have helped to provide a tantilzing link between the bacterial microbiome and CRC. However, at this present time there remains limited

information on the bacterial microbiome before and after successful surgery for removal of the adenoma or carcinoma and whether it changes at all.

In this study we investigated what happened to the bacterial microbiome before and after surgery for both adenoma and carcinoma individuals. Our anlaysis includes both alpha and beta diversity analysis along with investigation of individual operational taxonomic units (OTUs). We also utilized Random Forest models and observed how these models well as specific OTUs within this model performed on before (initial) and after (follow up) surgery samples. We used these models to look for similar important OTUs to identify the most important ones for not only classifying initial and follow up samples but also lesion or normal.

Results

Bacterial Community and Fit Changes before and after Treatment Based on thetayc distance metrics, comparing the initial to the follow up samples, there was a significant difference between the adenoma and carcinoma groups (P-value = 0.00198) [Figure 1a]. There was also a significant difference in FIT between initial and follow up samples with the carcinoma group having a significant decrease in FIT versus the adenoma group (P-value = 2.15e-05) [Figure 1b]. The whole community structure before and after surgery are visualized on NMDS graphs for both adenoma [Figure1c] (PERMANOVA = 0.002) and carcinoma [Figure 1d] (PERMANOVA = 0.997). When all initial and follow up samples were compared to each other there was no significant overall difference between them (PERMANOVA = 0.085). There was also no significant difference between initial and follow up samples for observed OTUs, Shannon diversity, or evenness after correction for multiple comparisons [Table S1]. There was no significant difference between initial and follow up samples for any single OTU [Figure S1].

Differences in Adenoma and Carcinoma of Previously Associated Cancer Bacteria We next examined whether there were differences in previously well described carcinoma associated OTUs. These included the OTUs that aligned with Porphyromonas asaccharolytica (Otu000153), Fusobacterium nucleatum (Otu000226), Parvimonas micra (Otu000460), and Peptostreptococcus stomatis (Otu000653). First, the carcinoma samples showed a significant difference between initial and follow up samples for Peptostreptococcus stomatis (P-value = 0.0183) and Porphyromonas asaccharolytica (P-value = 0.0154) whereas there were no signficant differences in any of these OTUs in the adenoma samples [Table S4]. Second, when these OTUs were present, there was a clear magnitude difference in these specific OTUs based on whether they were from adenoma or carcinoma individuals [Figure 2]. However, only a small percentage of those with adenoma or carcinoma were positive for any of these OTUs.

Full and Reduced Model Results Since differences were observed between initial and follow up samples and only a small number of individuals were positive for previously associated CRC bacteria we next investigated if we could create a model that could adequately classify and adjust lesion probability based on the bacterial community and FIT. The lesion model had an AUC range of 0.723 to 0.795 versus the initial follow up model which had and AUC range of 0.451 to 0.67 after 100 20 repeated 10-fold cross validations. Interestingly, identification of the most important variables and reducing the variables considered to only these increased the AUC in the lesion model (0.791 - 0.847) and initial and follow up model (0.651 - 0.79).

The test set AUC range for the full and reduced lesion model were similar to that reported for the training set AUC ranges and the ROC curve ranges overlap each other [Figure 3a]. The ROC curve for the final lesion model used falls within the range of both the full and reduced lesion model [Figure 3a]. Interestingly, the test set AUC range for the initial and follow up performed much better then the training set AUCs. Both the full and reduced initial and follow up models overlapped with each other [Figure 3b] there was a marked decrease in the ROC curve for the final model used.

Most Important Variables to the Models The reduced models were built based on the most important variables to the respective full model. For the lesion model there were a total of 37 variables [Figure S2] whereas for the initial and follow up model there were a total of 33 variables [Figure S3]. For both models FIT resulted in the largest decrease in MDA [Figure S2a & S3a].

Positive Probability Prediction after Surgical Removal of Adenoma or Carcinoma Regardless of model used there was a significant decrease in the positive probability of either the sample being lesion or an initial sample on follow up [Figure 4 & S4] (full lesion P-value = 1.11e-11, reduced lesion P-value = 1.91e-11, initial and follow up P-value = 6.71e-12, and reduced initial and follow up P-value = 6.72e-12).

For the full and reduced lesion model there was a significant difference in the classification for the lesion model between predicted and actual (P-value = 4.19e-10 and 6.98e-10, respectively) but not for the initial follow up model (P-value = 1.00 and 1.00). However, the lesion model correctly kept the one individual who still had a carcinoma on follow up above the cut off point [Figure 4a & S4a] for a positive call while the initial and follow up models did not [Figure 4b & S4b].

Common OTUs to both Lesion and Initial and Follow Up Models There were a total of 14 OTUs that were common to both models. Of these OTUs the most common taxonomic identifications were to Blautia, Bacteroides, Streptococcus, and Clostridiales. The majority of these OTUs had taxonomic identification to bacteria typically thought of as commensal [Table S2].

Treatment and Time Differences

There was no difference in the amount of change in positive probability for either the full or reduced lesion model for either chemotherapy (P-value = 0.821 and 0.821) or radiation therapy (P-value = 0.69 and 0.981). Although the initial follow up model was similar there was a significant decrease in positive probability for those treated with chemotherapy (P-value = 7.04e-04 and 5.07e-03Time of follow up sample from initial sampling, did not have a significant difference between adenoma and carcinoma (uncorrected P-value = 0.784).

Discussion

In our training set we show that the overall community structure as measured by different alpha diversity metrics, shows very little change between controls and those with either adenoma or carcinoma [Table S1]. With respect to our test set there was very little difference in magnitude of change in the thetayc distance metric between those with adenoma or carcinoma [Figure 1a]. In contrast, FIT had a large change in the initial and follow up samples in the carcinoma group versus the adenoma [Figure 1b]. An NMDS showed that there was very little observable change between initial and follow up for the adenoma group but there was one for the carcinoma group [Figure 1c & 1d]. This cursory information is suggestive that treatment of carcinoma, had the largest response.

We next created a model that incorporated both patient metadata, FIT, and the bacterial microbiome to be able to predict lesions (adenoma or carcinoma). Our middle training model, based on AUC, from 100 80/20 (train/test) splits was similar to the full training data model. It's 10-fold cross validated AUC was similar to it's test set AUC which was not the case for both the best and worse training model [Figure 3]. Using the full training data model we predicted the probability of a lesion in the inital and follow up samples [Figure 4]. There was a signficant decrease in positive probability regardless of whether the sample was a carcinoma or adenoma. The overall sensitivity for lesion detection in the intial samples was and for follow ups was . Although there was a decrease in overall probability of an adenoma or carcinoma only were below the 0.5 threshold out of the total individuals who were diagnosed as not having a carcinoma on follow up.

We then investigated which OTUs could potentially be more important in our model [Figure 5 & Table S3]. Many of the OTUs identified classified to normal flora bacterium [Table S3]. Only a single OTU though was significant after multiple comparison correction and the lowest taxonomic identification of was to . Although there was a difference in the relative

abundance at initial and follow up these values were not drastically different from the relative abundance values observed in the control individuals of the training set [Figure 5]. Although we were interested in what we could use to classify those with either adenoma or carcinoma versus normal. We found that the traditional bacteria associated with CRC were higher in magnitude in the carcinoma group and there were significant differences in some of these OTUs between the initial and follow up samples [Figure 5 and Tabl S5]. This research provides evidence that it is possible to use bacterial microbiome data to create a highily sensitive model, that is reactive to therapy, for detection of adenoma or carcinoma. It accomplishes this by using a unique sample set in which before and after surgery stool samples are available for assessment. By using these types of samples we are not only able to show sensitivity of lesion prediction but also able to show that this model is reactive. That is to say that after surgery for removal of the adenoma or carcinoma it decreases the positive probability to reflect a lower likelihood of the individual having an adenoma or carcinoma.

This study builds upon previous work from numerous labs that have looked into the bacterial microbiome as a potential screening tool (**insert citation**). Based on previous work by Jobin, et al. (**insert citation**) it may not be surprising to see E.coli in the top 5 OTUs for this model. Similarily, Porphyromonas has also been implicated in colorectal cancer (**insert citation**). Interestingly, many of the other OTUs had taxonomic identification for resident gut microbes. This could suggest that changes to the resident microbiome are important to the initiation of adenoma or carcinoma formation (**insert citation**) and provide support for the hypothesis that an initial change in the bacterial microbiome could pave the way for more inflammatory species: whether by creation of a new niche for oral microbes (**insert citation**) or allowing for a bloom of existing pro-inflammatory residents (**insert citation**).

Naturally, it is curious that normal staples of many screening studies such as Fusobacterium, Parvimonas, and Peptostreptococcus were not present in the majority

of the training models. One potential explanation for this is that FIT provides the same information to the model as these three organisms and so the model uses FIT preferentially over them. This has been suggested to be the case in a previous study (**insert Baxter Study**). It is also possible that these specific bacteria play a major role in the progression to carcinoma but may not be as important in the initiation of an adenoma, which would be supported by our data [Figure 5]. Regardless, our study does not argue against the importance of these bacterium in CRC initiation or pathogenesis but rather that the model does not utilize these specific bacteria for prediction purposes. Another potential reason why we did not identify the “usual suspects”" is that these bacteria may not change much between initial and follow up samples in those with an identified lesion. That is to say that the bacteria are consistently present even after removal of the lesion by surgery. Finally, it is likely that within our test set there was not enough indviduals in which detection was made or relative abundance high enough for these bacteria to be significant using a paired wilcoxson test.

One limitation in this study is that we do not know whether individuals in our test set eventually had a subsequent CRC diagnosis. This information would help to strengthen the case for our Random Forest based model keeping a number of individuals above the cutoff threshold even though at follow up they were diagnosed as no longer having a lesion. Another limitation is that we do not know if adding modern tests such as the stool DNA test (**insert citation**) could help improve our overall AUC. Another limitation is that this study drew heavily from those with caucasian ancestry. The results may not be immediately representative of those with either Asian or African ancestry. Finally, although our training and test set are relatively large we still run the risk of overfitting or having a model that may not be immediately extrapolateable to other populations. We’ve done our best to safeguard against this by not only running 10-fold cross validation but also having over 100 different 80/20 splits to try and mimic the type of variation that might be expected to occur.

By adding patient data such as age, BMI, etc. to the model and showing that it can successfully help to predict both carcinoma and adenoma our study provides further data that these patient factors in conjunction with the bacterial microbiome could potentially influence CRC and perhaps have a role in formation of adenomas. Further studies need to be carried out to verify our findings since not only are we dealing with stool, which could be very different than the communities present on the actual tissue, but also are dealing with correlations that may not be representative of the true pathogensis of disease.

Despite these limitations we think that these findings significantly add to the existing scientific knowledge on CRC and the bacterial microbiome. The ability for machine learning algorithms to take bacterial microbiome data and successfully lower positive probability after either adenoma or carcinoma removal provides evidence that there are specific signatures associated with these lesions. It also shows that these algorithms can not only successfully react to successful treatment regimens but also may be able to one day diagnose CRC with a high level of accuracy.

Methods

Study Design and Patient Sampling The sampling and design of the study was similar to that reported in Baxter, et al [13]. In brief, study exclusion involved those who had already undergone surgery, radiation, or chemotherapy, had colorectal cancer before a baseline stool sample could be obtained, had IBD, a known hereditary non-polyposis colorectal cancer, or Familial adenomatous polyposis. Samples used to build the model used for prediction were collected either prior to a colonoscopy or between 1 - 2 weeks after. The bacterial microbiome has been shown to nomralize within this time period [18]. Kept apart from this training set were a total of 67 individuals that not only had a sample as described previoulsy but also a follow up sample between 188 - 546 days after surgery and treatment had been completed. This study was approved by the University of Michigan Institutional Review Board. All study participants provided informed consent and the study itself conformed to the guidelines set out by the Helsinki Declaration.

Fecal Immunochemical Test and 16S rRNA Gene Sequencing FIT was analyzed as previously published using both OC FIT-CHEK and OC-Auto Micro 80 automated system (Polymedco Inc.) [19]. 16S rRNA gene sequencing was completed as previously described by Kozich, et al. [20]. In brief, DNA extraction used the 96 well Soil DNA isolation kit (MO BIO Laboratories) and an epMotion 5075 automated pipetting system (Eppendorf). The V4 variable region was amplified and the resulting product was split between three sequencing runs with control, adenoma, and carcinoma evenly represented on each run. Each group was randomly assigned to avoid biases based on sample collection location.

Sequence Processing The mothur software package (v1.37.5) was used to process the 16S rRNA gene sequences. This process has been previously described [20]. The general processing workflow using mothur is as follows: Paired-end reads were first merged into contigs, quality filtered, aligned to the SILVA database, screening for chimeras,

classified with a naive Bayesian classifier using the Ribosomal Database Project (RDP), and clustered into Operational Taxonomic Units (OTUs) using a 97% similarity cutoff with an average neighbor clustering algorithm. The numer of sequences for each sample was rarified to 10521 in an attempt to minimize uneven sampling.

Lesion Model Creation The Random Forest [21] algorithm was used to create the model used for prediction of lesion (adenoma or carcinoma) for the 67 individuals with follow up samples. The model included data on FIT and the bacterial microbiome. Non-binary data was checked for near zero variance and auto correlation. Data columns that had near zero variance were removed. Columns that were correlated with each other over a Spearman correlation coefficient of 0.75 had one of the two columns removed. This pre-processing was performed with the R package caret (v6.0.73). Optimization of the mtry hyperparameter involved taking the samples and making 100 80/20 (train/test) splits in the data where control and lesion were equally represented in the 80 and 20 split, respectively. This 80% portion was then split again into an 80/20 split, and run through 20 repeated 10-fold cross validations to optimzie the model's AUC (Area Under the Curve of the Receiver Operator Characteristic). This resulting model was then tested on the 20% of the data that was originally held out from this overall process. Once the ideal mtry was found the entire 490 sample set was used to create the final Random Forest model on which testing on the 67-person cohort was completed. The default cutoff of 0.5 was used as the threshold to classify individuals as positive or negative for lesion. The hyperparameter, mtry, defines the number of variables to investigate at each split before a new division of the data is created.

Initial Follow Up Model Creation We also investigated whether a model could be created that could identify before and after surgery samples. The training set utilized the 67-person cohort that was previously used for testing of the lesion modeld. The creation of this model and optimization of the mtry hyperparameter was completed using the same procedure

that was used to create the lesion model.

Selection of Important OTUs In order to assess which variables were most central to all the models we counted the number of times a variable was present in the top 10% of mean decrease in accuracy (MDA) for each different 80/20 split model and then filtered this list to variables that were only present more than 50% of the time. This final collated list of variables was what was considered the most important for the lesion or initial follow up models.

Statistical Analysis The R software package (v3.3.2) was used for all statisitical analysis. Comparisons between bacterial community structure utilized PERMANOVA [22] in the vegan package (v2.4.1) while comparisons between ROC curves utilized the method by DeLong et al. [23] executed by the pROC (v1.8) package. Comparisons between probabilities as well as overall amount of OTU between initial and follow up samples utilized a paired wilcoxson ranked sum test. Where multiple comparison testing was needed a Benjamini-Hochberg (BH) correction was applied [24] and a corrected P-value of less than 0.05 was considered significant. Unless otherwise stated the P-values reported are those of the BH corrected ones.

Analysis Overview Differences in FIT between initial and follow ups for either adenoma or carcinoma were investigated. Next, initial and follow up samples were analyzed for differences in alpha and beta diversity. All OTUs used in the lesion model were also analyzed using a paired wilcoxson test. The lesion model was then tested for accuracy in prediction and whether it reduced the positive probability of lesion after surgery. The most important OTUs for this were used to build an updated model and this reduced feature model was assessed for it's similarity to the original model. We then used the initial follow up model to assess whether this model could classify samples better then the lesion model. The most important OTUs were then identified from this model and used to create a reduced feature initial follow up model. This reduced feature model, as was

done with the lesion model, was compared to the full model for loss of accuracy. Finally, in order to investigate the relative abundance of specific bacteria, that have been previously associated with CRC, we selected OTUs that taxonomically classified to Fusobacterium Nucleatum, Parvimonas Micra, Peptostreptococcus Assacharolytica, and Porphyromonas Stomatis. Specifically, we wanted to test if there were any differences based on whether the individual had an adenoma or carcinoma.

Reproducible methods. A detailed and reproducible description of how the data were processed and analyzed can be found at https://github.com/SchlossLab/Sze_followUps_2017.

Figure 1: Changes and differences between the adenoma or carcinoma group. A) No significant difference was found between the adenoma and carcinoma group for thetayc (P-value = 0.00198). B) A significant difference was found between the adenoma and carcinoma group for FIT (P-value = 2.15e-05). C) NMDS of the intial and follow up samples for the Adenoma group. D) NMDS of the initial and follow up samples for the Carcinoma group. For C) and D) the teal represents initial samples and the pink represents follow up samples.

Figure 2: Previously Associated CRC Bacteria in Initial and Follow up Samples. A) Carcinoma initial and follow up samples. There was a significant difference in initial and follow up sample for the OTUs classfied as Peptostreptococcus stomatis (P-value = 0.0183) and Porphyromonas asaccharolytica (P-value = 0.0154). B) Adenoma initial and follow up samples. There were no significant differences between initial and follow up.

Figure 3: Graph of the Receiver Operating Characteristic Curve for lesion and Initial/Follow up models. The shaded areas represents the range of values of a 100 different 80/20 splits of the test set data using either all variables (grey) or reduced variable (red) models. The blue line represents the reduced variable model using 100% of the data set. A) Lesion model. B) Initial/Follow up model

Figure 4: Breakdown by Carcinoma and Adenoma of Prediction Results for Lesion and Initial and Follow Up for Reduced Variable Models A) Positive probability adjustment of those with carcinoma from intial to follow up sample B) Positive probability adjustment of those with adenoma as well as those with SRN and the probability adjustment from initial to follow up sample. The dotted line represents the threshold used to make the decision of whether a sample was lesion positive or not.

Figure S1: Distribution of P-values from Paired Wilcoxson Analysis of OTUs in Initial versus Follow Up

Figure S2: Summary of Important Variables in the Lesion Model A) MDA of the most important variables in the lesion model. The black point represents the median and the different colors are the different runs up to 100. B) The total number of appearances of each variable in the 100 different lesion models. The cutoff of 50% was used to assess importance.

Figure S3: Summary of Important Variables in Initial Follow Up Model A) MDA of the most important variables in the lesion model. The black point represents the median and the different colors are the different runs up to 100. B) The total number of appearances of each variable in the 100 different lesion models. The cutoff of 50% was used to assess importance.

Figure S4: Breakdown by Carcinoma and Adenoma of Prediction Results for Lesion and Initial and Follow Up for Full Variable Model A) Positive probability adjustment of those with carcinoma from intial to follow up sample B) Positive probability adjustment of those with adenoma as well as those with SRN and the probability adjustment from initial to follow up sample. The dotted line represents the threshold used to make the decision of whether a sample was lesion positive or not.

Figure S5: Thetayc Graphed Against Time of Follow up Sample from Initial

Declarations

Ethics approval and consent to participate

Consent for publication

Availability of data and material

Competing Interests

All authors declare that they do not have any relevent competing interests to report.

Funding

This study was supported by funding from the National Institutes of Health to P. Schloss (R01GM099514, P30DK034933) and to the Early Detection Research Network (U01CA86400).

Authors' contributions

All authors were involved in the conception and design of the study. MAS analyzed the data. NTB processed samples and analyzed the data. All authors interpreted the data. MAS and PDS wrote the manuscript. All authors reviewed and revised the manuscript. All authors read and approved the final manuscript.

Acknowledgements

The authors thank the Great Lakes-New England Early Detection Research Network for providing the fecal samples that were used in this study.

References

1. Jemal A, Siegel R, Xu J, Ward E. Cancer statistics, 2010. CA: a cancer journal for clinicians. 2010;60:277–300.

2. Haggar FA, Boushey RP. Colorectal cancer epidemiology: Incidence, mortality, survival, and risk factors. Clinics in Colon and Rectal Surgery. 2009;22:191–7.

3. Green RC, Green JS, Buehler SK, Robb JD, Daftary D, Gallinger S, et al. Very high incidence of familial colorectal cancer in Newfoundland: A comparison with Ontario and 13 other population-based studies. Familial Cancer. 2007;6:53–62.

4. Liao C-S, Lin Y-M, Chang H-C, Chen Y-H, Chong L-W, Chen C-H, et al. Application of quantitative estimates of fecal hemoglobin concentration for risk prediction of colorectal neoplasia. World Journal of Gastroenterology. 2013;19:8366–72.

5. Johnson DH, Kisiel JB, Burger KN, Mahoney DW, Devens ME, Ahlquist DA, et al. Multi-target stool DNA test: Clinical performance and impact on yield and quality of colonoscopy for colorectal cancer screening. Gastrointestinal Endoscopy. 2016;

6. Laiyemo AO, Doubeni C, Brim H, Ashktorab H, Schoen RE, Gupta S, et al. Short- and long-term risk of colorectal adenoma recurrence among whites and blacks. Gastrointestinal Endoscopy. 2013;77:447–54.

7. Matsuda T, Fujii T, Sano Y, Kudo S-e, Oda Y, Igarashi M, et al. Five-year incidence of advanced neoplasia after initial colonoscopy in Japan: A multicenter retrospective cohort study. Japanese Journal of Clinical Oncology. 2009;39:435–42.

8. Ren J, Kirkness CS, Kim M, Asche CV, Puli S. Long-term risk of colorectal cancer by gender after positive colonoscopy: Population-based cohort study. Current Medical

394 Research and Opinion. 2016;32:1367–74.

395 9. Løberg M, Kalager M, Holme Ø, Hoff G, Adami H-O, Bretthauer M. Long-term
396 colorectal-cancer mortality after adenoma removal. The New England Journal of Medicine.
397 2014;371:799–807.

398 10. Freeman HJ. Natural history and long-term outcomes of patients treated for early
399 stage colorectal cancer. Canadian Journal of Gastroenterology = Journal Canadien De
400 Gastroenterologie. 2013;27:409–13.

401 11. Lee JH, Lee JL, Park IJ, Lim S-B, Yu CS, Kim JC. Identification of Recurrence-Predictive
402 Indicators in Stage I Colorectal Cancer. World Journal of Surgery. 2016;

403 12. Richards CH, Ventham NT, Mansouri D, Wilson M, Ramsay G, Mackay CD, et al. An
404 evidence-based treatment algorithm for colorectal polyp cancers: Results from the Scottish
405 Screen-detected Polyp Cancer Study (SSPoCS). Gut. 2016;

406 13. Baxter NT, Ruffin MT, Rogers MAM, Schloss PD. Microbiota-based model improves the
407 sensitivity of fecal immunochemical test for detecting colonic lesions. Genome Medicine.
408 2016;8:37.

409 14. Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, et al. Potential of
410 fecal microbiota for early-stage detection of colorectal cancer. Molecular Systems Biology.
411 2014;10:766.

412 15. Dejea CM, Wick EC, Hechenbleikner EM, White JR, Mark Welch JL, Rossetti BJ, et al.
413 Microbiota organization is a distinct feature of proximal colorectal cancers. Proceedings of
414 the National Academy of Sciences of the United States of America. 2014;111:18321–6.

415 16. Zackular JP, Baxter NT, Chen GY, Schloss PD. Manipulation of the Gut Microbiota

Reveals Role in Colon Tumorigenesis. mSphere. 2016;1.

17. Arthur JC, Gharaibeh RZ, Mühlbauer M, Perez-Chanona E, Uronis JM, McCafferty J, et al. Microbial genomic analysis reveals the essential role of inflammation in bacteria-induced colorectal cancer. Nature Communications. 2014;5:4724.

18. O'Brien CL, Allison GE, Grimpen F, Pavli P. Impact of colonoscopy bowel preparation on intestinal microbiota. PloS One. 2013;8:e62815.

19. Daly JM, Bay CP, Levy BT. Evaluation of fecal immunochemical tests for colorectal cancer screening. Journal of Primary Care & Community Health. 2013;4:245–50.

20. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Applied and Environmental Microbiology. 2013;79:5112–20.

21. Breiman L. Random Forests. Machine Learning [Internet]. 2001 [cited 2013 Feb 7];45:5–32. Available from: http://link.springer.com/article/10.1023/A%3A1010933404324 http://link.springer.com/article/10.1023%2FA%3A1010933404324?LI=true

22. Anderson MJ, Walsh DCI. PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? Ecological Monographs [Internet]. 2013 [cited 2017 Jan 5];83:557–74. Available from: http://doi.wiley.com/10.1890/12-2010.1

23. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. Biometrics. 1988;44:837–45.

24. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and

powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological). 1995;57:289–300.