AWARD NUMBER
W81XWH-18-1-0112

TITLE: Preventing adverse patient responses to cancer chemotherapeutics.

PRINCIPAL INVESTIGATOR: Libusha Kelly

CONTRACTING ORGANIZATION: Albert Einstein College of Medicine

REPORT DATE: June 2020

TYPE OF REPORT: Annual Report

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT:

Approved for Public Release; Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
June 2020	Annual Report	06/01/2019 - 05/31/2020

4. TITLE AND SUBTITLE

Preventing adverse patient responses to cancer chemotherapeutics.

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-18-1-0112

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Libusha Kelly

E-Mail: libusha.kelly@einsteinmed.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Albert Einstein College of
Medicine
1300 Morris Park Ave,
Bronx, NY, 10461

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The purpose of this project is to prevent adverse patient responses to the cancer drug irinotecan by analyzing the gut microbiomes of patients. The scope of this project is to study irinotecan metabolism and the microbiome over time using fecal samples from healthy individuals and metastatic colorectal cancer patients. We have several major findings from the past year of work. We are happy to report a number of significant results for this year. We have successfully collected longitudinal samples for four colorectal cancer patients. We were invited to write a review for Annual Reviews in Pharmacology and Toxiciology about microbiome contributions to adverse events and we discuss the potential for microbiome interventions to improve drug and treatment safety and efficacy in colorectal cancer (Khan, Hauptman, and Kelly, Ann Rev Pharm Tox, in press). We wrote commentaries on harnessing the microbiome to improve drug therapy (Kelly, Clin Pharmacol Ther, 2019) and on microbial metabolism of L-dopa (Hitchings and Kelly, Cell Metab, 2019). We developed a novel computational approach to identify metastable states and state transitions in microbiome data that are linked to patient outcomes (Chang, VanInsberghe, and Kelly, npj Biofilms and Microbiomes, provisionally accepted). To better predict the likelihood of a patient suffering an adverse event based solely on his or her microbiome, we developed, tested, and validated three machine learning approaches to predict clinical outcomes based on microbiome data (Khan and Kelly, Pac Symp Biocomput, 2020). Each publication acknowledges the DoD.

15. SUBJECT TERMS

Colorectal cancer, drug metabolism, microbiome, carbohydrate active enzymes, phase II drug metabolism, metabolomics, metagenomics

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified		19b. TELEPHONE NUMBER *(include area code)*

TABLE OF CONTENTS

INTRODUCTION: *Narrative that briefly (one paragraph) describes the subject, purpose and scope of the research.*

The microbiome shapes the metabolic and immunological landscape of individuals in health and disease. Its plasticity can be leveraged for therapeutic interventions and to improve therapeutic outcomes. Recent studies have implicated gut microbiome metabolism at the gene and species level in driving the variability in patient drug response and toxicity. One of few therapeutic drugs for which we have a mechanistic understanding of how the gut microbiome influences drug metabolism is the colorectal cancer chemotherapeutic and prodrug irinotecan (CPT-11). CPT-11, in combination with fluorouracil and leucovorin, is one of three first- line treatments for metastatic colorectal cancer. Reactivation of the drug by beta-glucuronidases (BGs) in the gut can lead to severe diarrhea in patients. We hypothesize that individuals with high gut-driven turnover of SN-38G are at heightened risk for ADRs and can be identified via microbiome-based pretherapy analysis. Our overall objective is to identify patients at high risk for adverse events by non-invasive fecal sampling. The results will provide a clinical forecast for therapy in high-risk patients.

1. **KEYWORDS:** *Provide a brief list of keywords (limit to 20 words).*

Colorectal cancer, drug metabolism, microbiome, carbohydrate active enzymes, phase II drug metabolism, metabolomics, metagenomics

2. **ACCOMPLISHMENTS:** *The PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction.*

What were the major goals of the project?
List the major goals of the project as stated in the approved SOW. If the application listed milestones/target dates for important activities or phases of the project, identify these dates and show actual completion dates or the percentage of completion.

For all major tasks I have included only subtasks that were designed to be completed within the first 24 months of the project, I have not included tasks that were already completed in months 1-12 of the project.

Major Task 1: Quantify CPT-11 metabolites in healthy and metastatic colorectal cancer Patients

Subtask 4: Collect fecal samples from 20 healthy individuals and quantify metabolite production over time. We target 5 samples per individual. Samples will be used for both metabolite analysis and for metagenomic sequencing. For both Subtask 4 and Subtask 5, concentrations of SN-38G, SN-38 and the ISTD in the fecal extracts will be determined used, the Agilent G6490 Triple Quadrupole Mass Spectrometer. We will examine our mass spectrometry data for any additional, closely structurally related, metabolites of CPT-11 that have not been previously described (**~25% COMPLETE, we have collected 22 individual fecal samples from healthy individuals, we have not collected additional samples this year**)

Subtask 5: Collect fecal samples from 20 metastatic colorectal cancer patients, targeting 5 samples
per patient, and quantify metabolite production over time per Subtask 1 **(~27 % COMPLETE, we
have collected 27 samples total from four metastatic colorectal cancer patients)**

Subtask 6: Correlate adverse responses in metastatic colorectal cancer patients with metabolite production. Parametric and nonparametric tests will be used to identify significant differences between adverse events quantified as continuous outcomes (number of instances of diarrhea, severity of diarrhea), and chi-square tests will used to compare categorical outcomes (diarrhea/no diarrhea). All statistical tests will be two-tailed, and p values of less than 0.05 will be considered statistically significant. Analyses will be performed with the R statistical software package. (**~25 % COMPLETE, we have developed and tested R scripts for the statistical analysis)**

Milestone(s) Achieved: Characterization of variability in CPT-11 metabolite production in healthy individuals (~**10% COMPLETE**)

Major Task 2: Quantify beta-glucuronidase abundance and taxonomy in colorectal cancer patients over time

Subtask 1: Sequence fecal metagenomes of 20 metastatic colorectal cancer patients (from Major Task 1) using Illumina NextSeq sequencing, with a target of 3.5 M paired end reads and 1 Gb sequence per sample (**10% COMPLETE**)

Subtask 2: Correlate beta-glucuronidase abundance with adverse responses to CPT-11 by comparing all reads with our in-house database of beta-glucuronidases and comparing the relative abundance of specific beta-glucuronidases in patients who suffer adverse responses (diarrhea >= grade 3) with those patients who do not suffer adverse responses (**25% COMPLETE, we have collected adverse response data from patients for which we have 27 samples, we have extracted DNA from the 27 samples and prepared it for sequencing, and we have our computational analysis pipeline ready for analysis**)

Major Task 3: Activity-based protein profiling of functionally active human gut microbiome β-glucuronidases (**0% COMPLETE, the problems with this task are discussed below**)

Subtask 1: Optimize synthesis of a custom fluorescently labeled SN-38G probe in collaboration with the Einstein Chemical Synthesis core. We currently have low, impure yields that are not yet sufficient for our experiments

Subtask 2: Validate uptake of SN38-G labeled probe using positive and negative controls. Positive control: E. coli strain ATCC 25922 which can convert SN-38G to S38. Negative control: E. coli strain BW18812 (ΔuidA), which lacks the BG gene and thus should not convert SN-38G to S38G. Successfully sort labeled cells via flow cytometry

Subtask 3: Optimize flow sorting of fecal samples; specifically identify optimal sample

concentrations and buffer conditions to reliably sort these very heterogeneous samples. Optimize sorting of cells that uptake the labeled probe by defining appropriate parameterizations for sorting and by quantifying the populations of cells that have taken up the labeled probe

Major Task 4: Quantify microbiome gene expression during SN-38G exposure. (0% COMPLETE, the problems with this task are discussed below)

Subtask 1: Amend fresh fecal samples from the same 6 healthy volunteers referenced in Major Task 3, 3 high and 3 low metabolizers, with SN-38G, extract RNA at timepoints corresponding to known metabolism of SN- SN-38G. Sequence total RNA with a target of 12.5 M PE reads and 3.8 Gb of sequence per sample.

Subtask 2: Analyze RNASeq data to identify genes significantly associated with SN-38G -> SN38 conversion using the SAMSA pipeline, an open source tool which breaks down metatranscriptome data by organism and transcript function

What was accomplished under these goals?
For this reporting period describe: 1) major activities; 2) specific objectives; 3) significant results or key outcomes, including major findings, developments, or conclusions (both positive and negative); and/or 4) other achievements. Include a discussion of stated goals not met. Description shall include pertinent data and graphs in sufficient detail to explain any significant results achieved. A succinct description of the methodology used shall be provided. As the project progresses to completion, the emphasis in reporting in this section should shift from reporting activities to reporting accomplishments.

Major activities

- We have collected 27 total samples from colorectal cancer patients, and have extracted DNA and prepared the DNA for whole community metagenomic sequencing by Genewiz.

- We have collected and organized metatadata related to adverse events (diarrhea) for each of four colorectal patients for whom fecal samples were collected.

- During this funding cycle we have published one original research paper, one invited review, and one commentary and we have two papers provisionally accepted; one original research paper and one invited review. All papers acknowledge this funding from the Department of Defense. This brings the number of manuscripts that this grant has been acknowledged on since the commencement of funding to eight total publications.

Specific objectives

We conducted longitudinal sampling from colorectal cancer patients. Samples were prepared for sequencing and metabolomic analysis (**Major Task 1**, **Major Task 2**). DNA was extracted from patient samples and prepared for sequencing. We developed computational methods for analysis of temporal microbiome data and published this work (**Major Task 2**). We wrote an invited

review detailing the challenges and opportunities for engineering the microbiome to prevent adverse events. We continue to struggle with RNA extraction from incubated fecal samples but have reached out to experts at Rutgers and are testing a new protocol (**Major Task 4**). The COVID-19 pandemic has delayed our efforts to improve our approach to study microbial uptake of glucuronidated substrates (**Major Task 3**). Both of these issues are discussed in detail in "5 CHANGES/PROBLEMS".

Significant results

We are happy to report a number of significant results for this year. We have successfully collected longitudinal samples for four colorectal cancer patients. Two of four patients suffer from sporadic diarrhea during treatment. Two patients did not report diarrhea. Given these observations we are optimistic that we will be able to distinguish microbiome features associated with diarrheal events from those in patients who do not suffer diarrhea. We were invited to write a review for Annual Reviews in Pharmacology and Toxiciology about microbiome contributions to adverse events and we discuss the potential for microbiome interventions to improve drug and treatment safety and efficacy in colorectal cancer (Khan, Hauptman, and Kelly, *Ann Rev Pharm Tox*, in press). We were invited to write a commentary on harnessing the microbiome to improve drug therapy (Kelly, Clin *Pharmacol Ther*, 2019) and we were invited to write a commentary on a paper describing microbial metabolism of L-dopa for Cell Metabolism (Hitchings and Kelly, *Cell Metab*, 2019). In anticipation of our sequencing data for the temporal microbiome samples from colorectal cancer patients, we developed a novel computational approach borrowed from physics, topological data analysis, to identify metastable states and state transitions in microbiome data that are linked to patient outcomes (Chang, VanInsberghe, and Kelly, *npj Biofilms and Microbiomes*, provisionally accepted). To better predict the likelihood of a patient suffering an adverse event based solely on his or her microbiome, we developed, tested, and validated three machine learning approaches to predict clinical outcomes based on microbiome data (Khan and Kelly, *Pac Symp Biocomput*, 2020). Each of these five publications acknowledges funding from the DoD.

Collection and processing of samples and associated metadata from colorectal cancer patients.

We have now successfully collected longitudinal samples for four colorectal cancer patients. Two of four patients suffer from sporadic diarrhea during treatment. Importantly, the longitudinal sampling means that patients can act as their own controls, therefore even if there are not overlapping microbes between two patients that are associated with the propensity for diarrhea we will be able to ask in each patient, which microbes change when the patient suffers from diarrhea and which microbes are stable over time? Given these observations we are optimistic that we will be able to distinguish microbiome features associated with diarrheal events from those in patients who do not suffer diarrhea.

Development of novel computational approaches to link temporal variation in the microbiome to outcomes.

The dynamics of microbial ecosystems influence human and environmental health. Our publication "Topological analysis reveals state transitions in human gut and marine bacterial

communities", provisionally accepted in npj Biofilms and Microbiomes, identifies major compositional states in microbiomes and associates these states with clinical features and biological function using a novel method based on topological data analysis and network analysis.

Applying our method to environmental and human microbiome data we observed state transition dynamics corresponding to known clinical and environmental processes, such as recovery from cholera infection in the human gut and geochemical cycling in the oceans. More broadly, we propose a novel means of assessing microbial community stability and its relation to human health, and suggest topological analysis as a basis for powerful coarse-grained, quantitative modeling of microbial community dynamics that we will use for our analyses of colorectal cancer patient microbiomes.

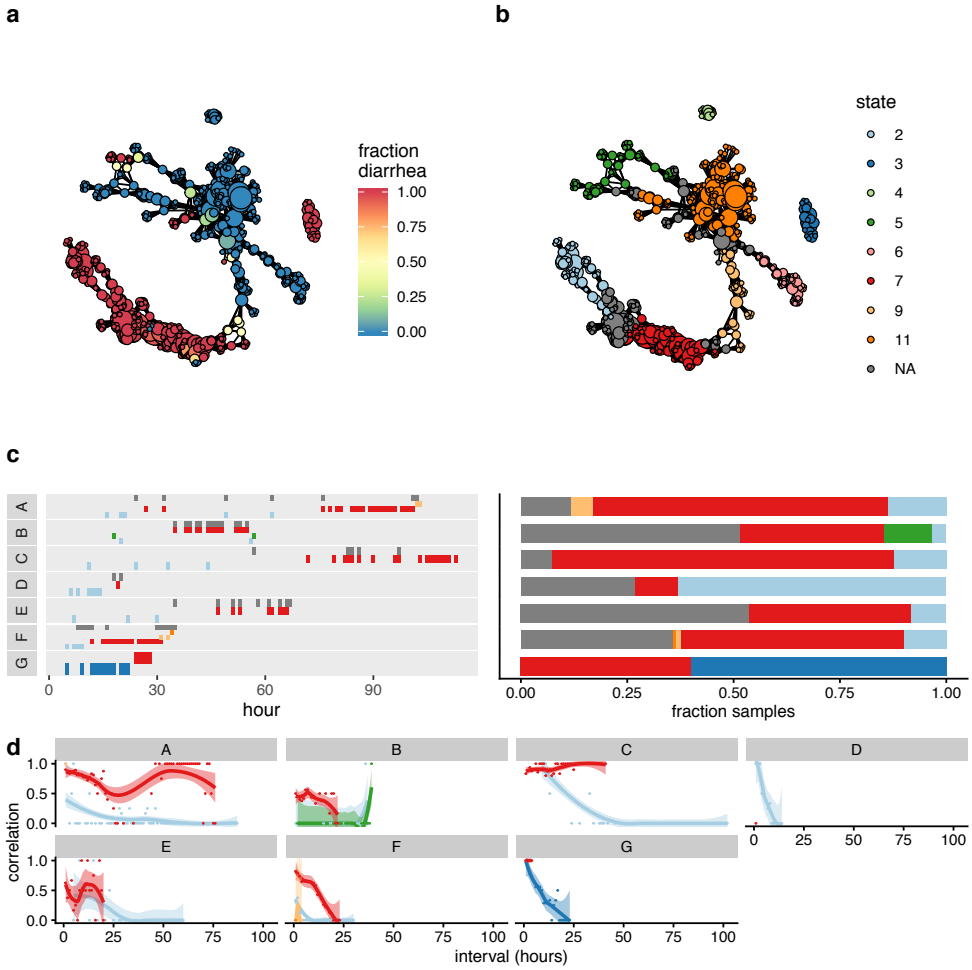

Figure 1: The phase space of the cholera gut microbiome. a) Mapper representation of the combined cholera data reveals disease- and healthy-associated neighborhoods of the phase space. Color: fraction of samples in each vertex associated with diarrhea. Connected components of the Mapper graph representing only one sample are not shown. Disjoint regions of phase space are represented as separate connected components. b) Partitioning of the phase space into metastable states. Vertices unassigned to any state are colored in grey. c) Left: progression of subject compositions during the diarrhea phase by state, showing persistence of states over time. Y axis and color indicate state index, with color indexing as in B. Where a sample was associated with multiple states, all were included. Right: frequency of samples associated with each states during the diarrhea phase for each subject with colors as in

B. d) Temporal correlation function for the diarrhea phase of each subject. Dots: raw values of f_X' for pairs of samples (see Methods). Lines: smoothed empirical mean of f_X'. Ribbons: standard error of the mean.

Our approach overcomes several challenges in microbial time series analysis: **1)** it makes minimal assumptions regarding biological mechanisms, and is applicable to systems as diverse as the human gut and the ocean; **2)** it uses all available information regarding the similarity between samples of microbial communities; and **3)** it inherently scales with the amount and dimensionality of data.

Applying our method to a published dataset monitoring the gut microbiomes of 7 cholera patients from disease (diarrhea) through recovery we identify previously unrecognized 'early' and 'late' disease-associated community states shared across multiple patients. We further show recurrent occupation of the 'late' state to be associated with longer time to recovery in 3 patients, and hypothesize that this reflects difficulty for these microbiomes in stably transitioning to a 'healthy' state (**Figure 1**). This analysis is directly relevant for our work in colorectal cancer patients; we will be able to map our patient metagenomes into the same space to ask if there are similar states and trajectories for patients when they suffer diarrhea.

For another published longitudinal dataset of the gut microbiomes of two mostly healthy adult human subjects, we found that the 'healthy' microbiome is dynamic, occupying a limited set of compositional states over time with fixed probability. We showed that one subject, after experiencing traveler's diarrhea, recovered the prior probability across states, showing full and robust recovery. By comparison, the other subject experienced *Salmonella* infection, after which their gut microbiome showed transient and non-recurrent occupation of different compositional states compared with pre-infection, showing instability even after clinical recovery.

Our method allows the identification of important compositional states and state transitions of microbial communities from time series data alone. Our method has the potential to **1)** facilitate analysis of the trajectories of complex diseases, such as bacterial infections, in the absence of detailed knowledge of the underlying biological mechanisms; and **2)** to detect impending changes in disease course, for example the diarrhea associated with irinotecan treatment, that could improve treatment plans.

Validation of three machine learning approaches to predict patient outcomes.

Genomic information encoded in the microbiome may predict phenotypes such as disease, and numerous case-control studies have been performed to quantify the differences between microbiomes in various disease states. We hypothesize that multi-disease comparisons may reveal unrecognized alterations in the microbiome that inform the underlying etiology of health and disease. Here, we utilize 5643 aggregated, annotated whole-community metagenomes from 19 different diseases to implement the first multiclass microbiome disease classifier at this scale. We compared three different machine learning models: random forests, deep neural nets, and a novel graph convolutional architecture which exploits the graph structure of phylogenetic trees as its input. We show that the graph convolutional model outperforms the standard neural net in terms of accuracy (achieving 75% average test-set accuracy), receiver-operator-characteristics (92.1% average AUC), and precision-recall (50% average AUPR). The graph convolutional

model performs on par with the random forest classifier in terms of accuracy but produces better receiver-operator-characteristics. Additionally, we are able to achieve 91.8% average top-3 accuracy (correct label is in top 3 predictions) and 96.5% top-5 accuracy with our model. Together, these results indicate that
there are predictive, disease specific signatures across microbiomes which could potentially be used for diagnostic purposes.

The key results of this work that are relevant to our colorectal cancer patient datasets include: **1)** the ability to reliably and robustly predict a patient outcome using machine learning approaches; **2)** the advantage of using hierarchical classifiers to represent the taxonomic structure of the microbiome, which supports our metagenomic sequencing efforts; **3)** the mapping and analysis of colorectal cancer patient microbiomes (**Figure 2**), from which we will have a foundational set within which we can map our own patient microbiomes to see **a)** if they are similar to or different from other colorectal cancer patient microbiomes; and **b)** to see how far our individual patient microbiomes move over time in colorectal cancer microbiome space.

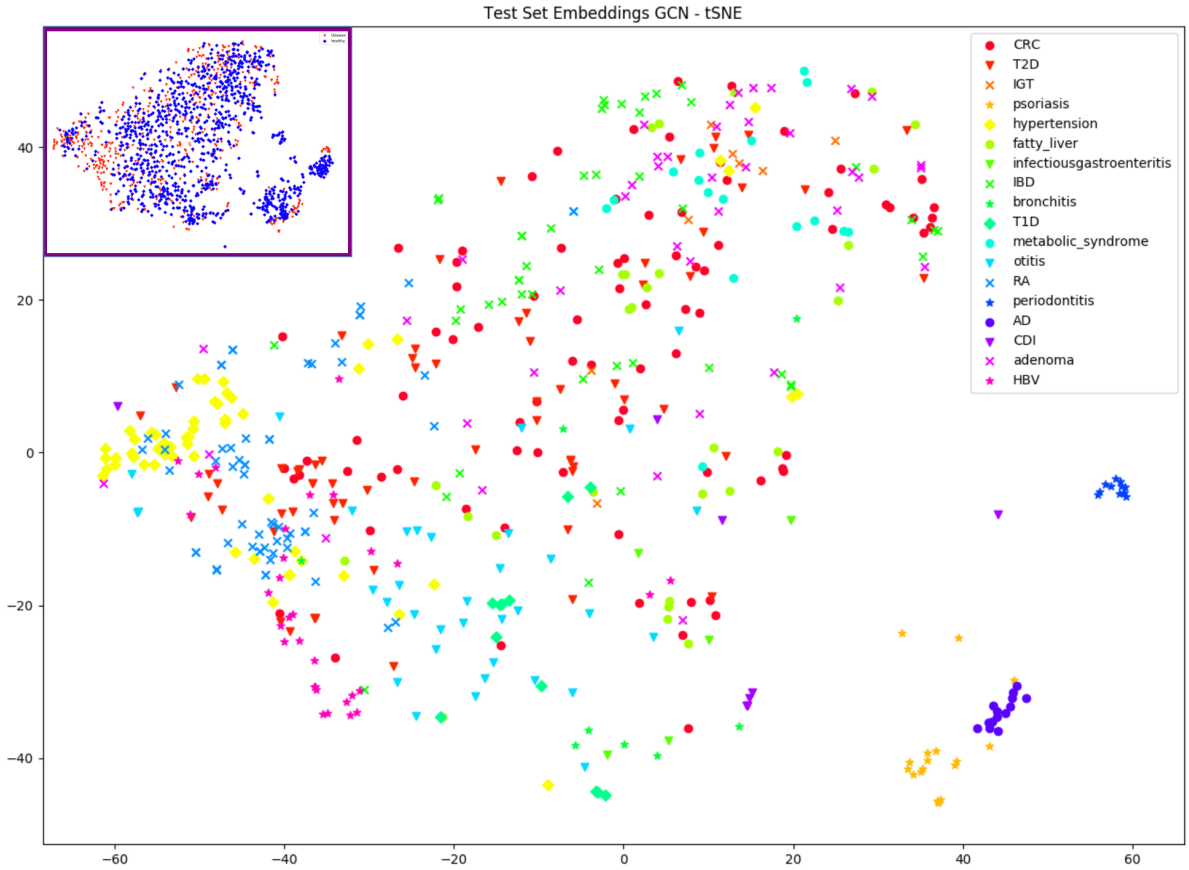

Figure 2. tSNE map of ~5000 microbiome samples. Disease states are represented in the large panel by color, the inset represents all disease (red) and healthy (blue) samples. Colorectal cancer samples are shown as red circles.

Proposing microbiome interventions to improve treatment outcomes for patients.
In the past decade of microbiome research, we have learned about numerous adverse interactions between the microbiome and medical interventions such as drugs, radiation, and surgery,

including the role of the microbiome in irinotecan-associated adverse events. What if we could alter our microbiomes to prevent these events? In an invited review for Annual Reviews in Pharmacology and Toxicology, we discuss potential routes to mitigate microbiome adverse events including applications from the emerging field of microbiome engineering. We highlight cases where the microbiome acts directly on a treatment, such as via differential drug metabolism, and cases where a treatment directly harms the microbiome, such as in radiation therapy. Understanding and preventing microbiome adverse events is a difficult challenge which will require a data driven approach involving causal statistics, multi-omics techniques, and a personalized approach to adverse event mitigation. Here, we propose research considerations to encourage productive work in preventing microbiome adverse events and we highlight the challenges and opportunities that await.

The main points we derive from the literature discussed and ideas proposed in this review are:

1. Microbiome adverse events (MAE) are a two-way street. Medical interventions can both perturb the microbiome and be adversely modified by it.

2. Microbiome engineering is at a very early stage. Many interventions, such as probiotics, can fail to actually achieve a significant perturbation to the native microbiome, and this should be kept in mind when evaluating MAE research.

3. Researchers should employ regular meta-omic sampling when conducting experiments and clinical trials. 'Omics provides a common basis for interpreting MAE research, and in particular is essential for understanding causal relationships between interventions and the microbiome.

4. Animal models should be carefully considered when studying MAE. Humanized mice may not always be the most appropriate model system.

5. MAE is an inherently personalized phenomenon and hence requires a personalized approach to diagnose and mitigate.

4) other achievements

- Dr. Kelly was promoted to associate professor at Albert Einstein College of Medicine.
- Dr. Kelly's MSTP student, Ruth Hauptman, who is supported as part of this proposal, successfully passed her qualifying exam. She will have her first thesis committee meeting in the fall.

What opportunities for training and professional development has the project provided?
If the project was not intended to provide training and professional development opportunities or there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe opportunities for training and professional development provided to anyone who worked on the project or anyone who was involved in the activities supported by the project. "Training" activities are those in which individuals with advanced professional skills and experience assist others in attaining greater proficiency. Training activities may include, for example, courses or one-on-one work with a mentor. "Professional development" activities

result in increased knowledge or skill in one's area of expertise and may include workshops, conferences, seminars, study groups, and individual study. Include participation in conferences, workshops, and seminars not listed under major activities.

Subtasks 1-4 were met in the first year of the grant.

Subtask 1: Attend a scientific research workshop. Dr. Kelly was one of 25 scientists selected to attend the National Cancer Institute's Innovation Lab: Systems Biology for the Cancer Microbiome from April 29-May 3, 2019 at the Beaver Hollow Conference Center in Java Center, NY. This intensive, weeklong workshop brought together experts from a wide variety of fields to form new collaborations, ideate and refine new projects, and identify opportunities to accelerate research on the influence of the microbiome in cancer using systems approaches. This workshop resulted in a group publication that Dr. Kelly is an author on this year in Trends in Cancer that is listed under "6 PRODUCTS".

Subtask 5: Attend a translational research workshop. Dr. Kelly attended the American Association for Cancer Research Translational Cancer Research for Basic Scientists Workshop in November 2019. This Workshop was an extraordinary opportunity to meet with clinical researchers, hear about cutting edge work in cancer research, and identify new potential collaborations. Dr. Kelly was very grateful for the opportunity to attend!

Subtask 6: Meet monthly with Dr. Mani. Dr. Kelly continues to meet monthly with Dr. Mani and is an author on one of Dr. Mani's papers assessing the role of bacterial swarming in inflammation. This paper is currently under review at Cell Host & Microbe.

How were the results disseminated to communities of interest?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the results were disseminated to communities of interest. Include any outreach activities that were undertaken to reach members of communities who are not usually aware of these project activities, for the purpose of enhancing public understanding and increasing interest in learning and careers in science, technology, and the humanities.

In addition to the publications and formal scientific talks described below, Dr. Kelly gave an invited seminar at the SLAC National Accelerator Laboratory at Stanford University as part of their Colloquium Series. There she was able to share my work with an audience of physicists and engineers at all levels who were not well versed in microbiome science. It

What do you plan to do during the next reporting period to accomplish the goals?
If this is the final report, state "Nothing to Report."

Describe briefly what you plan to do during the next reporting period to accomplish the goals and objectives.

- Recruit remaining colorectal cancer patients (**Major Task 1**). Current issues with the COVID-19 pandemic may slow recruitment but we are still aiming to meet our goal of 20 patients.
- Obtain metabolomic data for all samples. As described in the previous annual report, we have a protocol in place for determination of SN-38 and SN-38G concentrations in fecal samples (**Major Task 2).**
- Sequence and analyze colorectal cancer patient metagenomes (**Major Task 2**).
- Solve the problem of isolating sufficient RNA from fecal cultures (**Major Task 3**). (See "5 CHANGES/PROBLEMS")
- Determine the most appropriate probe to use to identify microbes that take up glucuronidated substrates into their cells (**Major Task 4**). (See "5 CHANGES/PROBLEMS")

4. **IMPACT:** D*escribe distinctive contributions, major accomplishments, innovations, successes, or any change in practice or behavior that has come about as a result of the project relative to:*

What was the impact on the development of the principal discipline(s) of the project?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how findings, results, techniques that were developed or extended, or other products from the project made an impact or are likely to make an impact on the base of knowledge, theory, and research in the principal disciplinary field(s) of the project. Summarize using language that an intelligent lay audience can understand (Scientific American style).

Our findings from this year will make an impact on the base of knowledge in the field of microbiome influences on patient outcomes as follows:

Our research demonstrating that machine learning/artificial intelligence approaches can reliably and robustly classify disease in a difficult, multi-class, classification setting lays the foundation for routine clinical use of the microbiome as a diagnostic. We demonstrate cases where the microbiome is a strong predictor of disease and contrast these with cases where current research does not support a strong influence of the microbiome on a given disease state. The tools used to develop the classifiers and the classifiers themselves from this work are publicly available and we hope the tools, approaches, and framework will be widely used by the community. We will use this classifier to predict adverse responses to irinotecan in our patient population using their microbiome data.

Our novel, topological data analysis-based approach to analyzing microbiome time-series data identifies 'states', that is, microbiome compositions, and state transitions in microbiomes that can be linked to clinical outcomes. As noted above, our approach overcomes several challenges in microbial time series analysis: **1)** it makes minimal assumptions regarding biological mechanisms, and is applicable to systems as diverse as the human gut and the ocean; **2)** it uses all available information regarding the similarity between samples of microbial communities; and **3)** it inherently scales with the amount and dimensionality of data. We therefore anticipate that many other microbiome studies will benefit from this approach. All code for this project is

publicly available. We will use this approach to understand whether there are particular microbiome compositions that are associated with adverse responses to irinotecan treatment.

Finally, our review on engineering the microbiome to improve treatment outcomes may serve as a starting point for clinicians who observe adverse outcomes in their patients and are interested in finding a basic research collaborator to explore the potential for microbiome modifications in specific patient settings.

What was the impact on other disciplines?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how the findings, results, or techniques that were developed or improved, or other products from the project made an impact or are likely to make an impact on other disciplines.

Both our machine learning and our topological data analysis approaches for classifying and visualizing microbiome data are relevant and applicable to any microbiome dataset. We note that in the topological data analysis paper (Chang, Van Insberghe, and Kelly, npj Biofilms and Microbiomes, provisionally accepted) we include analysis of two longitudinal marine datasets. We therefore think this approach will be useful in any discipline that considers microbial communities, from the human body, to the soil, to the oceans.

What was the impact on technology transfer?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe ways in which the project made an impact, or is likely to make an impact, on commercial technology or public use, including:

- *transfer of results to entities in government or industry;*
- *instances where the research has led to the initiation of a start-up company; or*
- *adoption of new practices.*

Nothing to report.

What was the impact on society beyond science and technology?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe how results from the project made an impact, or are likely to make an impact, beyond the bounds of science, engineering, and the academic world on areas such as:

- *improving public knowledge, attitudes, skills, and abilities;*
- *changing behavior, practices, decision making, policies (including regulatory policies), or social actions; or*
- *improving social, economic, civic, or environmental conditions.*

Nothing to report.

5. **CHANGES/PROBLEMS:** *The PD/PI is reminded that the recipient organization is required to obtain prior written approval from the awarding agency grants official whenever there are significant changes in the project or its direction. If not previously reported in writing, provide the following additional information or state, "Nothing to Report," if applicable:*

Changes in approach and reasons for change
Describe any changes in approach during the reporting period and reasons for these changes. Remember that significant changes in objectives and scope require prior approval of the agency.

Nothing to report.

Actual or anticipated problems or delays and actions or plans to resolve them
Describe problems or delays encountered during the reporting period and actions or plans to resolve them.

1. We have been unable to synthesize stable fluorescent probes for SN38-G (MAJOR TASK 3). As an alternative we intend to utilize the following probe, which can be purchased, to study microbial uptake of glucuronidated substrates with similarity to SN-38G.
   https://www.carbosynth.com/carbosynth/website.nsf/(w6productdisplay)/6322D24BA12D2AFB80256E19004DFE27

   The advent of the COVID-19 pandemic put this work on hold. Einstein is now reopening and we anticipate we will be able to re-start this work in the next month.

2. We have had difficulty extracting RNA from fecal samples (MAJOR TASK 4). We have attempted to work with Genewiz (http://www.genewiz/com) for mRNA extraction and purification from fecal samples; however while Genewiz has successfully extracted mRNA from our samples it was not of sufficient quality and quantity for mRNA sequencing. We next attempted a new protocol that has been used to extract mRNA from soil and fecal samples that was shared with us by colleagues at Rutgers. This approach has successfully extracted sufficient mRNA from test fecal samples, we are currently working on cleaning up the mRNA such that it will be appropriate for sequencing. We are confident that we can overcome this hurdle for fecal samples from this project.

3. Our colorectal cancer patient recruitment has taken longer than expected. One reason for this is that many of our patients at Einstein/Montefiore speak Spanish as their primary language. We have translated our informed consent documents and our description of the study into Spanish. We anticipate that this will expand our potential patient recruits.

4. Ruth Hauptmann, an MSTP student, joined the project in August, 2019. She successfully passed her qualifying exam this summer and has been collecting patient samples and getting up to speed on the experimental and computational protocols required for completion of the project. This transition period has led to some delays in getting the work done as anticipated.

Changes that had a significant impact on expenditures

Describe changes during the reporting period that may have had a significant impact on expenditures, for example, delays in hiring staff or favorable developments that enable meeting objectives at less cost than anticipated.

Nothing to report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Describe significant deviations, unexpected outcomes, or changes in approved protocols for the use or care of human subjects, vertebrate animals, biohazards, and/or select agents during the reporting period. If required, were these changes approved by the applicable institution committee (or equivalent) and reported to the agency? Also specify the applicable Institutional Review Board/Institutional Animal Care and Use Committee approval dates.

Significant changes in use or care of human subjects

Nothing to report.

Significant changes in use or care of vertebrate animals

Nothing to report.

Significant changes in use of biohazards and/or select agents

Nothing to report.

6. **PRODUCTS:** *List any products resulting from the project during the reporting period. If there is nothing to report under a particular item, state "Nothing to Report."*

- **Publications, conference papers, and presentations**
  Report only the major publication(s) resulting from the work under this award.

  Journal publications. *List peer-reviewed articles or papers appearing in scientific, technical, or professional journals. Identify for each publication: Author(s); title; journal; volume: year; page numbers; status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

  Authors who are members of the Kelly lab are **bolded**.

  a. Xavier JB, Young VB, Skufca J, Ginty F, Testerman T, Pearson AT, Macklin P, Mitchell A, Shmulevich I, Xie L, Caporaso JG, Crandall KA, Simone NL, Godoy-Vitorino F, Griffin TJ, Whiteson KL, Gustafson HH, Slade DJ, Schmidt TM, Walther-Antonio MRS, Korem T, Webb-Robertson BM, Styczynski MP, Johnson WE, Jobin C, Ridlon JM, Koh AY, Yu M, **Kelly L**, Wargo JA. The Cancer Microbiome: Distinguishing Direct and Indirect Effects Requires a Systemic View. Trends Cancer. 2020 Mar;6(3):192-204. doi: 10.1016/j.trecan.2020.01.004. Epub 2020 Feb 7. Review. PubMed PMID: 32101723; PubMed Central PMCID: PMC7098063.

  b. **Khan S, Kelly L**. Multiclass Disease Classification from Microbial Whole-Community Metagenomes. Pac Symp Biocomput. 2020;25:55-66. PubMed PMID: 31797586; PubMed Central PMCID: PMC7120658.

  c. **Hitchings R, Kelly L**. Drug Metabolism as a Community Effort. Cell Metab. 2019 Aug 6;30(2):235-237. doi: 10.1016/j.cmet.2019.07.005. PubMed PMID: 31390549.

  d. **Kelly L**. Harnessing the Microbiome to Improve Drug Therapy. Clin Pharmacol Ther. 2019 Aug;106(2):287-289. doi: 10.1002/cpt.1510. PubMed PMID: 31355459

e. **Khan S, Hauptman R, Kelly L.** Engineering the Microbiome to Prevent Adverse Events: Challenges and Opportunities. *Annual Reviews in Toxicology and Pharmacology*, in press.

f. **Chang W**, VanInsberghe D, **Kelly L.** Topological Analysis Reveals State Transitions in Human Gut and Marine Bacterial Communities. *npj Biofilms and Microbiomes*, provisionally accepted.

Books or other non-periodical, one-time publications. *Report any book, monograph, dissertation, abstract, or the like published as or in a separate publication, rather than a periodical or series. Include any significant publication in the proceedings of a one-time conference or in the report of a one-time study, commission, or the like. Identify for each one-time publication: author(s); title; editor; title of collection, if applicable; bibliographic information; year; type of publication (e.g., book, thesis or dissertation); status of publication (published; accepted, awaiting publication; submitted, under review; other); acknowledgement of federal support (yes/no).*

Nothing to report

Other publications, conference papers and presentations. *Identify any other publications, conference papers and/or presentations not reported above. Specify the status of the publication as noted above. List presentations made during the last year (international, national, local societies, military meetings, etc.). Use an asterisk (*) if presentation produced a manuscript.*

Conference presentations

1. Invited speaker, Eleventh International Workshop on Pharmacodynamics of Anticancer Agents (Sept 8-12, 2019, Chateau des Vigiers, Monestier, France).

External seminar and colloquia

1. Invited speaker, Pharmaceutical Sciences and Pharmacogenomics Seminar Series, University of California at San Francisco (Dec 4, 2019, San Francisco, California)

2. Invited speaker, SLAC National Accelerator Laboratory (Dec 2, 2019, Palo Alto, California)

- **Website(s) or other Internet site(s)**
List the URL for any Internet site(s) that disseminates the results of the research activities. A short description of each site should be provided. It is not necessary to include the publications already specified above in this section.

Kelly lab Github:
https://github.com/kellylab
This site is the repository for code and data used in all published analyses.

- **Technologies or techniques**
  Identify technologies or techniques that resulted from the research activities. Describe the technologies or techniques were shared.

  Nothing to report.

- **Inventions, patent applications, and/or licenses**
  Identify inventions, patent applications with date, and/or licenses that have resulted from the research. Submission of this information as part of an interim research performance progress report is not a substitute for any other invention reporting required under the terms and conditions of an award.

  Nothing to report.

- **Other Products**
  Identify any other reportable outcomes that were developed under this project. Reportable outcomes are defined as a research result that is or relates to a product, scientific advance, or research tool that makes a meaningful contribution toward the understanding, prevention, diagnosis, prognosis, treatment and /or rehabilitation of a disease, injury or condition, or to improve the quality of life. Examples include:
  - *data or databases;*
  - *physical collections;*
  - *audio or video products;*
  - *software;*
  - *models;*
  - *educational aids or curricula;*

- *instruments or equipment;*
- *research material (e.g., Germplasm; cell lines, DNA probes, animal models);*
- *clinical interventions;*
- *new business creation; and*
- *other.*

Nothing to report.

7 PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals have worked on the project?
Provide the following information for: (1) PDs/PIs; and (2) each person who has worked at least one person month per year on the project during the reporting period, regardless of the source of compensation (a person month equals approximately 160 hours of effort). If information is unchanged from a previous submission, provide the name only and indicate "no change".

Name: Libusha Kelly, no change.

Name: Ruth Hauptman

Project Role: Graduate Student

Researcher Identifier (e.g. ORCID ID):

Nearest person month worked: 8.75 (75%)

Contribution to Project: Ruth Hauptman is in charge of the sample collection, DNA extraction, metabolomics, and computational analyses.

Funding Support: No other support.

Has there been a change in the active other support of the PD/PI(s) or senior/key personnel since the last reporting period?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

If the active support has changed for the PD/PI(s) or senior/key personnel, then describe what the change has been. Changes may occur, for example, if a previously active grant has closed and/or if a previously pending grant is now active. Annotate this information so it is clear what has changed from the previous submission. Submission of other support information is not necessary for pending changes or for changes in the level of effort for active support reported

previously. The awarding agency may require prior written approval if a change in active other support significantly impacts the effort on the project that is the subject of the project report.

Nothing to report.

What other organizations were involved as partners?
If there is nothing significant to report during this reporting period, state "Nothing to Report."

Describe partner organizations – academic institutions, other nonprofits, industrial or commercial firms, state or local governments, schools or school systems, or other organizations (foreign or domestic) – that were involved with the project. Partner organizations may have provided financial or in-kind support, supplied facilities or equipment, collaborated in the research, exchanged personnel, or otherwise contributed.

Provide the following information for each partnership:
Organization Name:
Location of Organization: (if foreign location list country)
Partner's contribution to the project (identify one or more)

- *Financial support;*
- *In-kind support (e.g., partner makes software, computers, equipment, etc., available to project staff);*
- *Facilities (e.g., project staff use the partner's facilities for project activities);*
- *Collaboration (e.g., partner's staff work with project staff on the project);*
- *Personnel exchanges (e.g., project staff and/or partner's staff use each other's facilities, work at each other's site); and*
- *Other.*

Nothing to report.

8 SPECIAL REPORTING REQUIREMENTS

COLLABORATIVE AWARDS: *For collaborative awards, independent reports are required from BOTH the Initiating Principal Investigator (PI) and the Collaborating/Partnering PI. A duplicative report is acceptable; however, tasks shall be clearly marked with the responsible PI and research site. A report shall be submitted to* ***https://ers.amedd.army.mil*** *for each unique award.*

QUAD CHARTS: *If applicable, the Quad Chart (available on https://www.usamraa.army.mil) should be updated and submitted with attachments.*

9 **APPENDICES:** *Attach all appendices that contain information that supplements, clarifies or supports the text. Examples include original copies of journal articles, reprints of manuscripts and abstracts, a curriculum vitae, patent applications, study questionnaires, and surveys, etc.*