---------- REVIEW 1 ----------
Authors cast recovery of microbial dynamics under perturbation as a Markov Decision Process learning. Microbial community states are clustered into a small set of representative clusters and finite set of perturbations/actions move the system between these states. The learned conditional probabilities are investigated on several longitudinal microbiome datasets. The key drawback is that actions are treated as independent, and hence there is no way to disentangle co-occurring causes.

It is an interesting idea although some of the datasets seem to be ill-suited due to lack of meta-data. There is no quantitative analysis of model fits. This sets a bad precedent for future users of this technique.

Authors should check robustness of their multi-step clustering approach. Even more importantly, they should run cross-validation on the models.

Qualitative analysis is interesting. Some results, analysis of David et al 2014, are left without support of existing literature or any evidence.

The manuscript needs more work. Language should be cleaned up, methodology clarified, additional computational experiments run.

Higher level questions and suggestions:

When is it useful to aggregate taxa? Did you aggregate taxa?

Define what well-populated cluster means. The clustering approaches' interplay is not clear. Is second approach used when any cluster has more than 5 samples? Or are you resorting to approach two when you get a single cluster?

Merging procedure seems unnecessarily complicated. Is it simply merging small clusters to nearest cluster with size > threshold?

Streamline clustering section by providing a high-level algorithm. Do not want clusters < 5. Do not want single cluster. 3 step procedure: run approach 1, if single cluster, run approach 2, merge small clusters (<5) to closest large cluster (>5).
Need robustness analysis for this algorithm. Rerun on different subsets of data and see how often are two samples co-clustered.

Mention of primary contribution should be moved out of section 2.3.3 to introduction ...

unclear: "that one within more samples after dysbiosis than before,"

Relabel q1 and q2 as mature and immature. In fact, label all clusters with something other than numbers.

The part of discussion of simulating from MDP for power analysis is fine, but overly long exposition for a power analysis. Wet-lab validation of the trained MDP models' prediction is a must before power analysis is meaningful.

Large number of minor typos:

an common -> a common

what does it mean to "remove OTUs whose sum is 0"

include pseudo-count -> add pseudo-count

for quotes use `` and '' instead of "

missing citation for Arumugam tutorial ?

"Conversely" might not have been used correctly, since choices outlined in that sentence are not quite in contrast to choices listed in the previous sentence. Use "On the other hand"

Repeated statement of interest in more than two microbiome states:
"identify seveal (>2) well-populated microbiome states …"
and
"we are interested in more than two relatively well populated clusters …"

Merge sequence of citations (Chadès et al., 2014) (Puterman, 1994) - > (Chadès et al., 2014; Puterman, 1994),

"This dataset came from" -> "We analyzed a dataset from" ..

Choose terminology and stick with it “Meta data of their dietary intake” vs “nutritional meta-data”
Both mentions of the nutrition data talk about the same thing that it satisfies the data requirement.
Simply say that the dataset also contains the nutritional meta-data which is used to define actions.

d0 sampling -> day 0 sampling

Salmonella vaccine was given at the beginning,
before the d0 sampling, only to the groups of chicks with vaccine
->
Salmonella vaccine was given at the beginning,
before the day 0 sampling, to chicks in the group ‘sc’.
Probiotic
was mixed with the food, on every day of the experiment, only to the groups ‘cp’ and ‘sp’

metric,because -> metric, because

defining actions use cartesian product. for example
{low,medium,high} x {calcium,calorie, carb … }

Sentence fragment: and to label our clusters/MDP states, S ->

“apparently useful patterns of
knowledge.” -> “biologically informative patterns”
Patterns of knowledge talks about patterns in knowledge, not knowledge reflected in patterns of data.

inmature -> immature

Use \citet for textual and \citep for parenthetical citations. When using as a subject or object use
\citet when simply providing a reference parenthetically use \citep. So, for example:
Jim et al (1990) show that
rather than
(Jim et al 1999) show that

Same goes for:
 from Jim et al (1990)

What does Probability(Pr)=0 mean?

in-isolation, -> in isolation

keeps you in the same state -> keeps the system in the same state
you will persist in the same state -> the system will persist in the same state
…

Unclear: “possible prediction of perturbation
inducing patterns of interaction between specific microbes present in two
states...” Example that follows makes no reference to microbes.

what does width of red ribbon signify? Translate into biological language rather than using reference to a graphical object whose interpretation is not given but requires reading the caption.

---------- REVIEW 2 ----------
The authors propose to model time series microbiome data using a Markov Decision Process framework. In particular, they consider the scenario whereby time series data is available under known perturbations (which could vary between consecutive time points). In this scenario, the goal is to identify perturbations that lead to a particular “state” of microbiome, or alternatively to predict the “state” of the microbiome after a particular perturbation.

I find the framework interesting, however I think there are major biological and statistical questions that are unanswered, deterring from the applicability or interpretability of this approach.

First, microbiom data is clustered to define plausible “states”. There is no science of finding the optional number of clusters and so one must resort to standard model selection metrics as done here. However, I find the use of clustering, and estimating the number of clusters as done, in this settting highly dissatisfactory, because the rest of the analysis is dependent on this initial cluster number. Related to this, do you think that any intervention changes the entire landscape of microbiome composition? I would think that some interventions are very “local”, only small number of microbes will vary, and therefore clustering the data with all quantifiable microbes doesn’t seem biologically optimal. Also, it would be desirable to have some bootstrap estimate of the overall sensitivity of the results to the various modeling choices.

---------- REVIEW 3 ----------
This paper address a very timely and important problem: analyzing time-series microbiome samples to predict the effect of external perturbations and to infer optimal sequences of perturbations to achieve “desirable” microbiome states. The proposed method models time-series microbiome metagenomics data as Markov Decision Process (MDP), where the states correspond to the microbiome samples and the probabilities of transitions between states are computed based on the external perturbations responsible for the transition between successive time points in the input time-series data. Previous metagenomics studies have made use of state transition diagrams, but did not integrate perturbations into these diagrams and did not correlate perturbations with state transitions.

The paper is well organized and easy to read, though there are several grammatical and other errors. The choice of MDPs to represent the changes in microbiome states appears reasonable but the proposed approach has several limitations. As the authors themselves acknowledge, identifying a reasonable set of states/clusters to be used by the MDP can be nontrivial, and the current handling of transitions in which multiple perturbations have been concurrently applied can be problematic. Still, the proposed method may prove useful in analyzing time-series microbiome data.

The authors have applied their method to three time-series datasets and the results seem biologically meaningful, but it is not clear how sensitive the results are to various factors. For example, how much do the results vary as you vary the number of states/clusters, and what is the impact of using different time-series samples from the same microbiome? Basic sensitivity analysis should not be too difficult to perform and will strengthen the paper substantially.