Author’s response to reviews

Title: E. coli diversity: low in colorectal cancer

Authors:
Le Tang (le.tang@ucalgary.ca)
Yu-Jie Zhou (yjzhou@ccmu.edu.cn)
Songling Zhu (zhu.maxim@gmail.com)
Gong-Da Liang (809759150@qq.com)
He Zhuang (369144445@qq.com)
Man-Fei Zhao (zhaomanfei001@163.com)
Xiao-Yun Chang (1304198038@qq.com)
Hai-Ning Li (874819195@qq.com)
Zheng Liu (liuzheng_915@163.com)
Zhi-Rong Guo (guozhirong0604@yahoo.com.cn)
Wei-Qiao Liu (weiqiao.liu@gmail.com)
Xiaoyan He (hexiaoyan1970@163.com)
Chun-Xiao Wang (chunxiao_wang@foxmail.com)
Dan-Dan Zhao (1131066269@qq.com)
Jia-Jing Li (jiajing1218@163.com)
Xiao-Qin Mu (muxiaoqin126@126.com)
Bing-Qing Yao (rao.bingqing@163.com)
Xia Li (363549688@qq.com)
Yong-Guo Li (liyongguodoctor@163.com)
Li-Bo Duo (duolibo@163.com)
Author’s response to reviews:

Dear Editor:

Many thanks to you for your work on our manuscript. We have made the corrections accordingly as detailed below.

Best regards,

Shu-Lin Liu

Harbin Medical University
China
MGNM-D-19-00140R1
E. coli diversity: low in colorectal cancer
Le Tang; Yu-Jie Zhou; Songling Zhu; Gong-Da Liang; He Zhuang; Man-Fei Zhao; Xiao-Yun Chang; Hai-Ning Li; Zheng Liu; Zhi-Rong Guo; Wei-Qiao Liu; Xiaoyan He; Chun-Xiao Wang; Dan-Dan Zhao; Jia-Jing Li; Xiao-Qin Mu; Bing-Qing Yao; Xia Li; Yong-Guo Li; Li-Bo Duo; Li Wang; Randal N. Johnston; Jin Zhou; Jing-Bo Zhao; Gui-Rong Liu; Shu-Lin Liu

BMC Medical Genomics

FORMATTING CHANGES:

1) Please could you provide abstract sections.
Response: We have divided the abstract into four sections (Lines 40-58).

2) Kindly provide List of Abbreviations after conclusions.
Response: We have provided a List of Abbreviations after Conclusions (Lines 418-422).

3) Please provide Declarations heading.
Response: We have provided the Declarations heading (Line 423), with sub-headings and contents (Lines 424-464).

4) Please provide Availability of Data and Materials on declarations section.
Response: We have provided Availability of Data and Materials (Lines 439-441).

5) We noticed that in your Figure 4 provided, there is a multi-panel “A” and “B”. Kindly indicate multi-panel “A” and “B” in your figure legends.
Response: We have modified the legend and indicated A and B.

Dear Editor:

Thank you again for your time and work handling our manuscript. We have made the revision accordingly as detailed below. We highly appreciate the comments/suggestions from you and the reviewers on our manuscript, which have greatly helped us in the revision for much better readability of the article.

Best regards,

Shu-Lin Liu

Harbin Medical University
China

MGNM-D-19-00140

E. coli diversity associated with age and health status of the host: low in colorectal cancer
Reviewer reports:

Thomas H. Hampton, Ph.D (Reviewer 1): Tang et al have assessed strain diversity in over 2,000 E coli strains sampled from over 100 fecal donors classified as children, students, young adults or colon cancer patients. They provide evidenced that children and cancer patients have similar numbers of genetically distinct E coli strains, and that young adults and seniors are colonized by more E coli strains than either children or cancer patients. In addition, they performed competition assays which they interpret as evidence that strains from colon cancer patients can outcompete strains from healthy donors. Putting the reduced diversity of E coli from cancer patients together with their supposed competitive advantage over E coli from healthy donors, the authors hypothesize that reduced E coli strain diversity is caused by super-competitive strains that "facilitate the pathogenesis of diseases such as CRC".

I find their research interesting, but their conclusion difficult to follow based on the data and accompanying text as follows.

1) The claim that there were group differences in the number of unique strains needs to be attached to a fully described statistic and supported by a figure in the main text. The figure should show a point for each of the 100+ observations, so that the reader can understand the distribution of the data and the appropriateness of the statistic used given the data. A table should be provided describing the ages and sex of the fecal donors in each group, as well as any other available information such as antibiotic use.

Response: This is a very good point, since detailed information on the 100+ participants is very important for the readers to fully comprehend this piece of work. Therefore, we have refined the information table (originally Supplementary Table 1) and moved Supplementary Table 1 to the main text as Table 1 for the revised manuscript. As suggested by the reviewer, we added the key statistic in the main text for a more convenient reading and comprehension of the research results, including a figure to present the statistics (Fig 2b). We also indicted the fact that no antibiotic was used by the participants during the period of six months prior to the specimen collection time point.

2) The resolution of Figures 2, 3 and 5 was too low for me to read them.

Response: Figures 2, 3 and 5 have been refined or replaced by new versions (Original Fig 2 now is Fig 2a) for higher resolution.
3) Figure 5 does not provide visual or statistical evidence that E coli from cancer patients outcompetes E coli from healthy donors in general. A figure and statistics are needed.

Response: We revised the figure legend to make the information in Figure 5 clearer and easier to comprehend as follows “…2, ccpm6195; 3, ccpm5172; 4, ccpm5062; 5, a mixture of ccpm6195, 5172 and 5062 before the competition assay; 6, a mixture of ccpm6195, 5172 and 5062 cultured in LB broth at the end of the competition assay; 7, a mixture of ccpm6195, 5172 and 5062 cultured in M9 medium at the end of the competition assay. Remarkably, after culture for 10 days in M9 medium, only ccpm6195 (lane 7, the E. coli strain from a CRC patient, in which the genomic cleavage pattern is indistinguishable with that in lane 2) survived. These growth competition assays demonstrate that when nutrient was ample (as in LB broth), the three E. coli strains did not interfere with one another for growth; but when the nutrient was limited, the E. coli strains from a CRC patient had greater capabilities to compete for nutrient to grow (Lines 709-715)”.

4) The assertion that low diversity of E coli strains in and of itself is a risk factor for colon cancer is difficult to believe on two fronts.

First, assuming that certain strains of E coli are protective against cancer and others predispose an individual to cancer (as the authors suggest) it follows that being colonized by one strain would be protective as long as it were the right strain. All things being equal, one would predict that some protective E coli would be capable of dominating other E coli. Perhaps what is going on is that the conditions that lead to low E coli diversity are conditions that lead to colorectal cancer. I find that much more plausible.

Response: Although the observed association between low E. coli diversity and CRC does not exclude the possibility that the conditions that lead to low E. coli diversity are conditions that lead to colorectal cancer, we are more inclined to believe low E. coli diversity, caused by the purging capabilities of certain non-benign E. coli lineages, to be a novel risk factor for CRC based on the results we obtained in this study, especially the phylogenetically distinct and biologically aggressive E. coli lineages. We have added such discussions in the revised manuscript (Lines 297-303).

Second, a great deal of research has gone into studying CRC, as the authors must be aware, including specific E coli strains: https://doi.org/10.1080/1040841X.2018.1481013. I think factors such as host genetics, and the gut microbiome in general are well known to be associated with risk: https://doi.org/10.1371/journal.pone.0020447, DOI: 10.1126/science.aaw2367. E coli does not seem to play a pivotal role in general, let alone strain diversity of E coli. If the authors would liker to put forth their view that E coli strain diversity plays any role in CRC risk whatsoever, they have introduce this in the context of what we already know, rather than pointing out that different strains of E coli differ in overall pathogenicity, which is all that they cover in their current introduction.
Response: Our emphasis on the findings about the association between low E. coli diversity and CRC is that dysbiosis, reflected in this study by low E. coli diversity, represents a separate kind of risk factors for CRC in addition to genetic and environmental CRC risk factors. Probing dysbiosis in the intestinal environment is a complicated endeavor, but the determination of E. coli diversity can be a convenient, low cost and reliable method to reflect the structural changes of the intestinal microbiota or microbiome. Specific strains of E. coli or other bacteria have been reported in the literature as well as in our study as associated with CRC, but the insight of our findings in this study into CRC pathogenesis is the low E. coli diversity in CRC patients compared to healthy controls rather than actual causation of CRC by the specific bacteria themselves; we believe that under the theme of low diversity of E. coli or other commensal bacteria in association with the occurrence and development of CRC, a great deal of concrete discoveries will be made in the near future. We appreciate this comment by this reviewer and have revised the manuscript to make it more concise and clearer.

Ana Lucia Fachin (Reviewer 2):

Title: E. coli diversity associated with age and health status of the host: low in colorectal cancer

The title should be revised to clarify the major idea

My suggestion: E. coli diversity is low in colorectal compared healthy host

Response: We like this title and have adopted it with a minor alteration in the revised manuscript (E. coli diversity: low in colorectal cancer).

Abstract:

Background: Escherichia coli are mostly commensals but also contain pathogenic lineages. It is largely unclear whether the commensal E. coli as the potential origins of pathogenic lineages may consist of monophyletic or polyphyletic populations.

In the background: Please insert the aims of manuscript. The background is confused and not showed the relation among, diversity and pathogenicity or colorectal cancer.

Response: We have revised the Abstract as follows “Escherichia coli are mostly commensals but also contain pathogenic lineages. It is largely unclear whether the commensal E. coli as the potential origins of pathogenic lineages may consist of monophyletic or polyphyletic populations, elucidation of which is expected to lead to novel insights into the associations of microbial diversity with human diseases…(Lines 40-43)”. We appreciate the reviewer for this important point.

gel electrophoresis (PFGE) techniques was not cited in the abstract.
Response: We added this “…to probe the commensal E. coli populations for their diversity by genomic sequencing and pulsed field gel electrophoresis techniques (Lines 46-47).”

Results:

Why did you use the pulsed field gel electrophoresis (PFGE) techniques to observed genomic differences among the E. coli strains? Because you have a DNAseq data is more reliable.

Response: We analyzed over 2000 E. coli strains, delineation of which by PFGE is more cost- and time-effective than by DNAseq and is equally reliable, although for gene level analysis, we did use the DNAseq data to compare selected representatives.

Please give more details about:

Probing the pan-genomes

140 The pan-genomes and core-genomes analysis was done by using home-made Perl scripts.

Response: We added the following “…We determined the genes common to all compared strains and used the genes as the “core-genome” for the strains compared; we added all non-redundant genes of the bacterial strains in comparison to the core-genome to obtain the “pan-genome”. The analysis of pan-genomes and core-genomes…(Lines 131-133).”

In MM there isn’t the descrition about DNA seq bioinformatic analyse.

Response: We provided the relevant citations and related information in the revised manuscript (Lines 121-122).

About the diversity and DNA sequencing: The authors need clarify how they proved the relation between bacterial diversity using the DNA seq and (PFGE)

Response: We added the information “…to the protocols published previously [25-27]. The endonuclease I-CeuI recognizes phylogenetic diversity of the bacteria from the genus and up levels [27-29] and cleavage data from the CTAG-recognizing endonucleases reflect bacterial diversity at the species level, which are consistent with genomic sequence data [30-33]…(lines 116-119)”.

The quality of figure 5 should be improved.

Response: The quality of figure 5 was improved in the revised manuscript.