GigaScience

Female reproductive tract microbiota influence egg production in layer chickens

--Manuscript Draft--

Manuscript Number:	GIGA-D-21-00132R1
Full Title:	Female reproductive tract microbiota influence egg production in layer chickens
Article Type:	Research
Funding Information:	Department of Science and Technology of Sichuan Province (2019JDTD0009) — Pro. Diyan Li Fok Ying Tong Education Foundation (161026) — Pro. Diyan Li Department of Science and Technology of Sichuan Province (2020YFH0138) — Dr. Zhongxian Xu Department of Science and Technology of Sichuan Province (2021YFYZ0009) — Pro. Mingzhou Li
Abstract:	Background: The microbiota of the female reproductive tract is increasingly recognized as having fundamental roles in animal reproduction. To explore the relative contribution of reproductive tract microbiomes on egg production in chickens, we investigated the microbiota in multiple reproductive and digestive tract sites from 128 female layer (egg-producing) chickens in comparable environments. Results: We identified substantial differences between the diversity, composition, and predicted function of site-associated microbiota. Reproductive tract microbiota were more profoundly affected egg production than that in the digestive tract. We identified four reproductive tract microbial species, Bacteroides fragilis , B. salanitronis , B. barnesiae, and Clostridium leptum , which were related to immune function and potentially contribute to enhanced egg production. Conclusions: These findings provide insights into the diverse microbiota characteristics of reproductive and digestive tracts, and may help design strategies for controlling and manipulating chicken reproductive tract microbiota to improve egg production.
Corresponding Author:	Diyan Li CHINA
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	
Corresponding Author's Secondary Institution:	
First Author:	Yuan Su
First Author Secondary Information:	
Order of Authors:	Yuan Su
	Shilin Tian
	Diyan Li
	Wei Zhu
	Tao Wang
	Shailendra Kumar Mishra
	Ranlei Wei
	Zhongxian Xu

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

	Mengnan He
	Xiaoling Zhao
	Huadong Yin
	Xiaolan Fan
	Bo Zeng
	Mingyao Yang
	Deying Yang
	Qingyong Ni
	Yan Li
	Mingwang Zhang
	Qing Zhu
	Mingzhou Li
Order of Authors Secondary Information:	
Response to Reviewers:	Dear editor and reviewers, Thank you very much for reviewing our manuscript "Female reproductive tract microbiota influence egg production in layer chickens" (ID: GIGA-D-21-00132) for possible publication in GigaScience. We sincerely thank the editor and two reviewers for their valuable feedback that we have used to improve the quality of our manuscript. According to the reviewers' comments, we have made the relevant modifications to our manuscript. All modifications are marked in red in the revised manuscript and a point-by-point response to the reviewers' comments follows. We hope these revisions meet your satisfaction and make our manuscript acceptable for publication in your journal. We look forward to hearing a positive response from you. Best regards, Diyan Li, Professor, Ph.D. Address: Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu 611130, China. E-mail: diyanli@sicau.edu.cn Detailed responses to reviewers: All comments provided by reviewers are in black, and our responses are in red. This following text is also included in the uploaded file "Response letter": Reviewer #1: Dear authors, the work described in the manuscript is very extensive! I have not yet seen an analysis of host genetics, microbiota composition and host transcriptomics coupled with egg production data. Also while not perfect the manuscript is written well. Response: Thank you for your positive comments. Main problems: 1.It is clear you found an association between specific bacteria and egg production. You have also shown that some form of inflammation is associated with the changes in microbiota and egg production. However, you assume that the bacteria affect inflammation which affect egg production. But other options also exist. For example, it is possible that waning egg production, from other causes, changes the conditions in the oviduct so that the microbiota would change. Or inflammation, for example because of a pathogen, might modulate both egg production as well as microbiota composition. Please change the text so that the readers understand you are speculating and briefly mention the other options. Response: As suggested, we have added discussion of the other options affecting egg production in the "Results and Discussion" sections (Lines 355-363). "Multiple factors, especially host species, potential pathogens, and immune status of the host, all play a major role in the female reproductive organs adversely interfering

with the egg industry in laying flocks [53]. Additionally, the digestive tract environment of low-egg producing hens is fragile and susceptible to the influence of exogenous microorganisms [41]. Pathogenic infection, room temperature fluctuations, management systems, and other sudden changes to various factors can alter the composition of microbiota [54, 55]. These alterations may cause a significant degradation in production performance. Here, our results indicate that the reproductive tract microbiota play an important role in egg production."

2.While it is too late to change, it is a pity that you did not characterize the cecum community. Of all of the intestinal communities the cecum is by far the biggest and the most likely to affect the nutrition of the hen, thereby possibly affecting egg production.
Response: Thank you for your comments and understanding. We have included some information regarding this in the Discussion (Lines 288-291).
"The cecum has distinct microbial community profiles [44-46] that were not explored in this study. Microbial community analysis of the cecum microbiota in chickens exhibiting different egg production performances requires further investigation."

3.Please make sure to reference and discuss relevant literature. I quickly identified Elokil A. A. Animal 2020 which performed a limited but similar analysis. Please perform a literature search and make sure to reference and discuss relevant work.
Response: Thank you for your helpful suggestions. As suggested, we have added and discussed additionally relevant literature references (Lines 280-288).
"In accordance with previous findings, there is clear evidence of the role of fecal microbiomes in low and high egg-laying performance in hens; Elokil et al [16] demonstrated a significantly positive association between the microbial genus Lactobacillus and egg-laying performance ($P < 0.05$). Likewise, Wang et al [41] reported that Lactobacillus was also abundant in the feces of high-yield hens. The genus Lactobacillus produces growth promoters and exhibits antimicrobial activity against pathogenic microbes [42, 43] which may explain why the increasing abundance of Lactobacillus in the high-yield group is beneficial to egg-laying performance."

Minor problems:
Lines 48-49 - something is wrong with this sentence.
Response: Thank you for the pointing this out. We have revised the sentence in the manuscript (Lines 46-48).
"An abnormal vaginal microbiota may predispose individuals to increased microbial invasion of the amniotic cavity and fetal damage [11, 12]."

Lines 67-69 - it is not clear what you are trying to say here.
Response: As commented, the sentence was not clear and had limited relevance to the manuscript so we deleted the sentence.

Line 71 - 16S rDNA sequencing and not whole-genome sequencing.
Response: In this study, we conducted whole-genome sequencing on 128 laying hens and 16S rDNA sequencing on 768 samples from six sites. We have revised the sentence in the manuscript (Lines 68-71).
"Here, we performed 16S rDNA sequencing on 768 samples from three reproductive (vagina, uterus, and isthmus) and three digestive (crop, gizzard, and small intestine) tract sites and whole-genome sequencing of 128 laying hens."

Line 96 - please be more specific regarding which part of the small intestine was sampled.
Response: The jejunum of the small intestine was sampled. As suggested, we provide a more specific explanation of which part of the small intestine was sampled in the revised manuscript (Lines 420-421).
"A 12-cm-long fixed mid-region of the small intestine (jejunum) was collected from each bird."

Lines 100-108 - description of host DNA extraction is not appropriate another the heading "microbial genomic DNA extraction".
Response: "Microbial genomic DNA extraction" has been changed to "Host and microbial genomic DNA extraction" in the revised manuscript (Line 426).

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Line 109 and elsewhere - 16S rDNA and not 16S rRNA.
Response: "16S rRNA" was changed to "16S rDNA" throughout the whole manuscript.

Line 154 - why are you using OTUs with 97% identity and not 100% identical amplicon sequence variants (ASVs)?
Response: As an alternative to OTUs, ASVs have been proposed as a way to adapt the thresholds suggested by genome sequencing to microbial community analysis using 16S rDNA sequences. Meanwhile, the OTU approach is still one of the primarily used methods for analyzing 16S rDNA-seq data (Dvergedal et al., 2020; Pinna et al., 2021; Wen et al., 2021). Schloss PD recently evaluated the clustering risk among ASV and OTU methods, and reached a conclusion that ASVs and the use of overly narrow thresholds to identify OTUs increase the risk of splitting a single genome into separate clusters (Schloss, 2021).
Since there is no consensus for a biological definition of a bacterial species (Sanford et al., 2021), microbiologists should accept that how bacterial species are named is biased and that taxonomic rules are not applied in a consistent manner. This makes it impossible to fit a distance threshold that matches a set of species names (Konstantinidis and Tiedje, 2005). Furthermore, the 16S rDNA sequence does not evolve at the same rate across all bacterial lineages (Schloss and Westcott, 2011), which limits the biological interpretation of a common OTU definition. A distance-based definition of a taxonomic unit based on the 16S rDNA or full-genome sequences is operational and not necessarily grounded in biological theory (Yarza et al., 2014; Barco et al., 2020). One benefit of a distance-based OTU definition is the ability to mask residual sequencing errors. The sequences generated in microbiome studies can harbor PCR and sequencing errors. These errors would only exacerbate the inflated number of ASVs.
Although there are multiple reasons why proponents favor ASVs, we feel the significant risk of artificially splitting genomes into separate clusters is too high to warrant their use.

References
Barco, R., G. Garrity, J. Scott, J. Amend, K. Nealson, and D. Emerson. 2020. A genus definition for bacteria and archaea based on a standard genome relatedness index. mBio 11:e02475-02419.
Dvergedal, H., S. R. Sandve, I. Angell, G. Klemetsdal, and K. Rudi. 2020. Association of gut microbiota with metabolism in juvenile Atlantic salmon. Microbiome 8:160.
Konstantinidis, K. T., and J. M. Tiedje. 2005. Towards a genome-based taxonomy for prokaryotes. Journal of Bacteriology 187:6258-6264.
Pinna, N. K., R. M. Anjana, S. Saxena, A. Dutta, V. Gnanaprakash, G. Rameshkumar, S. Aswath, S. Raghavan, C. S. S. Rani, and V. Radha. 2021. Trans-ethnic gut microbial signatures of prediabetic subjects from India and Denmark. Genome Medicine 13:36.
Sanford, R. A., K. G. Lloyd, K. T. Konstantinidis, and F. E. Löffler. 2021. Microbial taxonomy run amok. Trends in Microbiology 29:394-404.
Schloss, P. D. 2021. Amplicon sequence variants artificially split bacterial genomes into separate clusters. mSphere Jul 21:e0019121. doi: 10.1128/mSphere.00191-21.
Schloss, P. D., and S. L. Westcott. 2011. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Applied and Environmental Microbiology 77:3219-3226.
Wen, C., W. Yan, C. Mai, Z. Duan, J. Zheng, C. Sun, and N. Yang. 2021. Joint contributions of the gut microbiota and host genetics to feed efficiency in chickens. Microbiome 9:126.
Yarza, P., P. Yilmaz, E. Pruesse, F. O. Glöckner, W. Ludwig, K.-H. Schleifer, W. B. Whitman, J. Euzéby, R. Amann, and R. Rosselló-Móra. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nature Reviews Microbiology 12:635-645.

Line 157 - what are "singleton OTUs"? OTUs with only one read? OTUs found in only one sample?
Response: "singleton OTUs" represent OTUs found in only one sample (Lines 483-484).
"Singleton OTUs (OTUs found in only one sample) that did not match the reference database were removed."

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Lines 159-161 - why were low abundance OTUs discarded? Does it affect your analysis? In general, it is better not to modify the data base extensively. Specifically, some phylogenetic groups are represented by multiple low abundant strains with a high total abundance whereas other phylogenetic groups are represented by a single highly abundant strain. By getting rid of such lower abundance OTUs you might be creating an artifact.
Response: Thank you for the helpful comments. Errors introduced by next-generation amplicon sequencing tend to induce spurious OTUs and spurious counts in OTU tables, both of which are especially prevalent at low abundances. Despite the power of NGS and the progress achieved, generated data is imperfect, being subject to different types of errors, including those inherent to PCR amplification (substitutions and chimeric sequence formation) and sequencing-specific biases which are characteristic of each sequencing technology. Sequencing errors are predominantly caused by base substitutions, although base deletions, low-quality reads, variable read lengths and non-target amplification are also known error sources which may result in false species affiliation. Undetected chimeric sequences, caused by the hybridization of DNA fragments from different species also reduce the reliability of the 16S rDNA sequence-based phylogenetic composition of microbial communities. Together, these different errors generate a high number of lower-abundance sequences, which lead to overestimations of actual community diversity and the creation of many false taxa. Those spurious sequences are hard to filter out using current mainstream pipelines implementing error correction, denoising, and stringent filtration of chimeric sequences, contaminants and non-bacterial contents. Previous study have reported that although the overall abundance of these pseudo sequences was low, introducing them into analysis increased the total number of features to 10 times higher than expected and enlarged the divergence of the alpha and beta diversity analyses among the different methods (Wang et al., 2018). Lower-abundance and lower-quality sequences were observed to surround higher-abundance, biologically real sequences, forming error clouds (Bokulich et al., 2013; Edgar, 2013). Various researchers have developed different approaches to remove these pseudo sequences.

In our work, we described the numerical distribution of OTUs with different existing ratios in samples, after which the slope fluctuation is estimated. We considered there are a lot of false positives for these microbiotas when fluctuations are great. Therefore, we selected OTUs that existed in more than 20% of samples according to the slope distribution curve (Supplementary Fig. S1a, b). This filtering rule will improve the stability and accuracy of further analyses. Similar filtering rules have also been reported in previous analyses (Zierer et al., 2018). In addition, we have cited Supplementary Fig. S1a, b in the revised manuscript (Lines 488-489).

References
Bokulich, N. A., S. Subramanian, J. J. Faith, D. Gevers, J. I. Gordon, R. Knight, D. A. Mills, and J. G. Caporaso. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nature Methods 10:57-59.
Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10:996-998.
Wang, J., Q. Zhang, G. Wu, C. Zhang, M. Zhang, and L. Zhao. 2018. Minimizing spurious features in 16S rRNA gene amplicon sequencing. PeerJ Preprints 6:e26872v1 https://doi.org/10.7287/peerj.preprints.26872v1
Zierer, J., M. A. Jackson, G. Kastenmüller, M. Mangino, T. Long, A. Telenti, R. P. Mohney, K. S. Small, J. T. Bell, and C. J. Steves. 2018. The fecal metabolome as a functional readout of the gut microbiome. Nature Genetics 50:790-795.

Lines 378-382 - this is an interesting idea that you did not establish well. Could you determine unifrac (or any other metric) distances of samples from different organs. i.e. compare all vagina samples to all small intestine samples and determine average Unifrac distances and stDev, and do so for each pair of organ sites. Perhaps you will gain a new figure showing how similar or different are the communities in each site.
Response: We are grateful for the suggestion. As suggested by the reviewer, we added a new figure (Supplementary Fig. S2j) showing the weighted UniFrac distances among the six sites.

Line 387 - cyanobacteria are not likely gut or reproductive tract inhabitants. This is more likely an artifact of identifying feed derived chloroplast DNA. You might want to mention this.

Response: Thanks for your helpful suggestions. The chickens used in our study are fed using corn-soybean-based diets, so the Cyanobacteria (likely an artifact of feed-derived chloroplast DNA) were misidentified as microbes. Additionally, we have revised the sentence in the revised manuscript (Lines 118-120).
"Firmicutes, Proteobacteria, and Cyanobacteria (likely an artifact of feed-derived chloroplast DNA) accounted for 71.45% - 97.86% of all OTUs."

Lines 398-402 - what is the connection between the "well-documented protective role of lactobacillus…" and differences in the levels of lactobacillus between gut and reproductive tract? Does the reproductive tract not deserve protection? Please stick to the data. Do not attach unsupported conclusions. Instead connect the levels of Lactobacillus and pH.
Response: As suggested, we have connected the levels of Lactobacillus and pH in the revised manuscript (Lines 134-139).
"Lactobacillus is thought to inhibit pathogenic bacteria by lowering the environmental pH through lactic acid and hydrogen peroxide production [24]. This genus was highly abundant in the digestive tracts which were characterized by low pH values which strongly limits the growth of most pathogens [25, 26]. In contrast, Lactobacillus was less abundant in the reproductive tract where an alkaline pH is needed to maintain sperm motility [27, 28]."

Lines 407-411 - if I understand correctly, you are implying a connection possibly caused because of the flow of material from different sites. You may want to note that.
Response: Thank you for the helpful comments. Yes, we want to note that, and we have added a sentence in our revised manuscript (Lines 146-147).
"These results imply there is a connection of microbiome communities possibly caused because by the flow of material from different sites."

Line 419-421 - you have reduced a whole field of study to just one "previous study". And in any case I am not sure what you wanted to write here.
Response: We have now included additional references in the revised manuscript (Line 158).

Line 442 - not chloroplast and mitochondria-like microbes but rather true chloroplasts and mitochondria which you misidentified as microbes.
Response: Thanks for your helpful suggestion. We have revised the sentence in the manuscript (Lines 179-180).
"Unidentified Chloroplast (2.94%) and mitochondria-like (2.37%) materials from plant consumption."

Line 447 - does Helicobactor maintain near natural pH and a microaerophilic environment? Is that its role in the gut? Please do not bring up bits and pieces of information if this is not really relevant to your results. What is important is that a possible pathogen was found in your chickens.
Response: Thanks for the comments. In order to infect the gastric mucosa, Helicobacter pylori has to survive in the gastric acidic pH, and Helicobacter pylori has well developed mechanisms to neutralize the effects of acidic pH (Ansari, S. and Y. Yamaoka. 2017. Survival of Helicobacter pylori in gastric acidic territory, helicobacter 22:e12386), that is its role in the gastric tract. In addition, we have reorganized the sentence in the revised manuscript (Lines 183-185).
"As a possible pathogen, Helicobacter specifically inhabits the small intestine in chickens, and may be involved in inflammation, metabolism, and neutralization of gastric acid [35-37]."

Lines 450-463 - this whole paragraph contains a lot of disjoined bits and pieces of information. Does it really matter that some bacteria were changed in immune suppressed honey bees? What are you trying to say? Consider taking this whole paragraph off.
Response: Thanks for the comment. This whole paragraph mainly displays the results of the site-associated bacterial taxa in the three reproductive tract sites identified by LEfSe and discussed association with the previous literature. Additionally, we have deleted the disjoined bits and pieces of information and reorganized the sentence in the paragraph in our revised manuscript (Lines 186-194).
"Unidentified Erysipelotrichaceae showed higher abundance in the three reproductive

tract sites (1.83% - 2.40%). Bacteria associated with the isthmus and uterus both showed higher abundances than in the other sites. Several genera (typically, Romboutsia, Fusobacterium, and Clostridium sensu stricto 1) were dominant in the vagina (> 25% of the microbiota) but had lower abundances in the other sites (Fig. 2b). Among these, vaginal Romboutsia could be employed as a predictor for egg number in laying hens [8]. Six Bacteroidetes bacterial taxa were isthmus-associated; Bacteroides species live on host mucus-secreted polysaccharides and this flexible foraging behavior contributes to diversity and stability [38]."

Lines 464-466 - do your results support this speculation? If yes please expend, if not please remove.
Response: Thank you for the suggestion. Lines 464-466 have been removed from the revised manuscript.

Line 492 - 0.21% of total abundance? Of total species?
Response: Yes, it is "of total species". We have reorganized the sentence and clarified in the manuscript (Line 223).
"accounting for 0.21% of the microbiota species tested in the small intestine".

Lines 560-568 - this is a good example of discussion! Here you brought relevant information and created a data based speculation.
Response: Thank you for your positive remarks about this discussion.

Line 646 - what do you mean in "microbiome succession"?
Response: "microbiome succession" has been changed to "microbial community" in the revised manuscript (Line 390).

Supplementary figure 2 contains two separate topics - functional sequencing data and bacterial community analysis. Please divide into separate figures.
Response: As suggested, we have divided Supplementary figure 2 into two separate figures (Supplementary Fig. S2 and S11).

Reviewer #2: This is a very good study to which I have only a few comments, mostly to microbiota part in which I am stronger than in chicken genetics. The only weaker part, but not weak, is the fact that the study was performed with only a single flock and was not repeated with different hens at all. I understand that it is impossible to repeat the study in another flock but for example you could have set up Cl. leptum PCR and check for its presence in reproductive tract in completely different birds.
Response: We appreciate the positive feedback and sincerely thank the reviewer for the thoughtful and supportive recommendations, which are of great help in improving the quality of this manuscript.

Please consider the following points.
line 94, indeed laparotomy?
Response: "After laparotomy" has been changed to "After the abdomen was opened" in the revised manuscript (Line 417).

l.111 and 114, I do not understand wha PCR kit you used. Please, reword.
Response: The PCR amplification was performed using Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, Ma, USA). We have revised the sentence in the revised manuscript (Lines 436-439).
"The V4 hypervariable region of the bacterial 16S rDNA was amplified using Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, Ma, USA) and the universal primers 515 F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806 R (5′-GGACTACHVGGGTWTCTAAT -3′) [56]."

l.253, 254, this sentence is somewhat compromised. Please check and reword.
Response: Thank you for the comments, and we have revised the sentence in the revised manuscript (Lines 578-579).
"In addition, we further removed the SNPs with adjacent distances ≤ 5."

l.342, rather verified, or confirmed, then detected. By the way, I am not sure whether you used data from the qPCR in the rest of the manuscript
Response: We have changed "detected" to "verified" in the revised manuscript (Line

	664). We only used qPCR data in the "RNA sequencing (RNA-seq) analysis" section. l.388, these are not Cyanobacteria, this is chloroplast DNA from plants in the feed, you may check very recent paper Volf et al. Eggshell and Feed Microbiota Do Not Represent Major Sources of Gut Anaerobes for Chickens in Commercial Production. Microorganisms 2021, 9, 1480. Response: Thanks for the valuable suggestions, and we have read the above reference carefully. The chickens used in our study are fed using corn-soybean-based diets, so the Cyanobacteria (likely an artifact of feed derived-chloroplast DNA) was misidentified as microbes. Additionally, we have revised the sentence in the manuscript (Lines 121-122). "Cyanobacteria (likely an artifact of feed-derived chloroplast DNA) was the dominant material in the gizzard (48.19% of the total abundance)." l.462,463, I would suggest alternative explanation in this case and this is living on expense of host mucus secreted polysaccharides, Sonnenburg, J. L. et al. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science 307, 1955-1959 (2005) Response: As suggested, we have cited the reference and revised the sentence in the manuscript (Lines 192-194). "Bacteroides species live on host mucus-secreted polysaccharides and this flexible foraging behavior contributes to diversity and stability [38]." l.559, check that B. fragilis might be a consequence of intensive human care, Kollarcikova et al. Different Bacteroides Species Colonise Human and Chicken Intestinal Tract. Microorganisms 2020, 8, 1483. Response: We are grateful for the suggestion. We have cited the useful and interesting reference in the revised manuscript (Lines 301-303). "Interestingly, a recent study reported that the human-adapted Bacteroides species are likely introduced to chicken flocks by contact with humans and can temporarily persist in chickens [49]." l.615, how can you know this? What if all of this the other way round, and I indeed believe that this is the other way, i.e. hens becomes of compromised performance, due to whatever factor, within but possibly also outside of those which you have monitored. This naturally results in decrease in egg lay but also in increased inflammatory response. Locally changed conditions due to inflammatory signaling change, infiltrating heterophils and macrophages produce antimicrobial peptides and reactive oxygen species and strict anaerobes will be the firs bacterial species to decrease in a response of increasing oxygen concentration. What is cause and what is consequence. I do not know, I think that you do not know either, though you blame bacteria that these are responsible for the response. Response: Thank you for the comments. We deleted the sentence that caused misunderstanding. l.631, similar to previous comment, I do not think that there is any downregulation. In high egg producers, there is basal, background expression of inflammatory marker genes. And these are induced in the hens with compromised performance. Be also careful, whether this induction since this could also be a cause of infiltration of macrophage with their specific expression profile, and you then mistakenly conclude on induction when purifying mRNA from a total complex tissue. Response: Thanks for your comments. "downregulated" has been changed to "different expressed" in the revised manuscript (Line 383). l.656, the same as above, be careful what is cause and what is consequence. Increase in reactive oxygen species may affect the most strict anaerobes. When these are present, there is no inflammation. When these are eliminated by increase in oxygen concentration and all other inflammatory responses, this is explained that these bacteria are anti-inflammatory. These are not, these only dislike inflammation and oxygen. Response: Thanks for your helpful suggestion. We have deleted the sentence in the revised manuscript.
Additional Information:	

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Question	Response
Are you submitting this manuscript to a special series or article collection?	No
Experimental design and statistics Full details of the experimental design and statistical methods used should be given in the Methods section, as detailed in our Minimum Standards Reporting Checklist. Information essential to interpreting the data presented should be made available in the figure legends. Have you included all the information requested in your manuscript?	Yes
Resources A description of all resources used, including antibodies, cell lines, animals and software tools, with enough information to allow them to be uniquely identified, should be included in the Methods section. Authors are strongly encouraged to cite Research Resource Identifiers (RRIDs) for antibodies, model organisms and tools, where possible. Have you included the information requested as detailed in our Minimum Standards Reporting Checklist?	Yes
Availability of data and materials All datasets and code on which the conclusions of the paper rely must be either included in your submission or deposited in publicly available repositories (where available and ethically appropriate), referencing such data using a unique identifier in the references and in the "Availability of Data and Materials" section of your manuscript.	Yes

Have you have met the above requirement as detailed in our Minimum Standards Reporting Checklist?	

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Manuscript w. DOI
Click here to access/download;Manuscript;Revised Manuscript DOI.docx

1 **Female reproductive tract microbiota influence egg production in layer chickens**

2 Yuan Su[1,†] , Shilin Tian[2,3,†], Diyan Li[1,*,†], Wei Zhu[1,†], Tao Wang[1,†], Shailendra Kumar Mishra[1], Ranlei

3 Wei[4], Zhongxian Xu[1], Mengnan He[1], Xiaoling Zhao[1], Huadong Yin[1], Xiaolan Fan[1], Bo Zeng[1], Mingyao

4 Yang[1], Deying Yang[1], Qingyong Ni[1], Yan Li[1], Mingwang Zhang[1], Qing Zhu[1,*], Mingzhou Li[1,*]

5

6 [1]Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province,

7 Sichuan Agricultural University, Chengdu 611130, China.

8 [2]Department of Ecology, Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan

9 University, Wuhan 430072, China.

10 [3]Novogene Bioinformatics Institute, Beijing 100000, China.

11 [4]Center of Precision Medicine, West China Hospital, Sichuan University, Chengdu 610065, China.

12 [†]These authors contributed equally to this paper.

13 [*]Corresponding author. e-mail: Diyan Li: diyanli@sicau.edu.cn; Qing Zhu: zhuqing@sicau.edu.cn;

14 Mingzhou Li: mingzhou.li@sicau.edu.cn

15 **Abstract**

16 **Background:** The microbiota of the female reproductive tract is increasingly
17 recognized as having fundamental roles in animal reproduction. To explore the relative
18 contribution of reproductive tract microbiomes on egg production in chickens, we
19 investigated the microbiota in multiple reproductive and digestive tract sites from 128
20 female layer (egg-producing) chickens in comparable environments.

21 **Results:** We identified substantial differences between the diversity, composition, and
22 predicted function of site-associated microbiota. Reproductive tract microbiota were
23 more profoundly affected egg production than that in the digestive tract. We identified
24 four reproductive tract microbial species, *Bacteroides fragilis*, *B. salanitronis*, *B.*
25 *barnesiae,* and *Clostridium leptum*, which were related to immune function and
26 potentially contribute to enhanced egg production.

27 **Conclusions:** These findings provide insights into the diverse microbiota characteristics
28 of reproductive and digestive tracts, and may help design strategies for controlling and
29 manipulating chicken reproductive tract microbiota to improve egg production.

30 **Keywords:** microbiota, reproductive tract, egg production, chicken

31

32 **Background**

33 The domestic chicken (*Gallus gallus domesticus*) is of enormous agricultural
34 significance, comprising broiler (meat-producing) and layer (egg-producing) chickens.
35 Specialized commercial layer breeds were established during the twentieth century with
36 greatly improved reproductive traits [1]. Currently, thousands of quantitative trait loci
37 (QTLs) [2] and many gene mutations [3, 4] are reportedly associated with chicken

38 reproductive traits. Nonetheless, egg production, as a polygenic inheritance trait,
39 exhibits low to moderate heritability (*h^2*, ranging from 0.05 to 0.44, depending on the
40 period involved) [5, 6]. Alternative effective approaches for modulating egg production
41 in laying hens are urgently required for the poultry industry to meet consumer demand.

42 Distinct bacterial communities throughout the female reproductive tract form a
43 microbiota continuum from the vagina to the isthmus, which has a prominent role in
44 animal reproduction [7, 8]. In humans, microbiome interactions with the host during
45 pregnancy leading to preterm birth were investigated [9], and temporal changes in the
46 vaginal microbiome associated with full-term pregnancies were identified [10]. An
47 abnormal vaginal microbiota may predispose individuals to increased microbial
48 invasion of the amniotic cavity and fetal damage [11, 12]. The avian reproductive tract
49 houses complex bacterial communities that are believed to play crucial roles in egg
50 production [13]. Chicken digestive and reproductive tracts are mainly colonized by
51 *Firmicutes*, *Bacteroidetes*, *Proteobacteria*, *Actinobacteria*, and *Fusobacteria*, which
52 are spatially organized within specific digestive and reproductive compartments [14,
53 15]. Additionally, *Lactobacillus* species were found to be keystone species residing in
54 the chicken oviduct [16].

55 Several synergistic factors, such as environment and diet, dominate over host
56 genetics in determining gut microbiota composition [17, 18]. A comparative study of
57 gut microbial diversity among parrot species indicated the potential role of host
58 ancestry in shaping the gut microbiome [19]. A genome-wide association study (GWAS)
59 in chickens demonstrated the genetic loci rs15142709 and rs15142674 which are
60 located in the pleiomorphic adenoma gene 1 (*PLAG1*) and lck/yes-related novel
61 tyrosine kinase (*LYN*) genes were significantly associated with microbial
62 *Methanobacterium* abundance [20]. In a previous study, 14 identified QTLs strongly

63 influenced *Clostridium leptum* and *Lactobacillus* abundance, as well as related
64 candidate genes involved in anti-inflammatory responses and the motility of the
65 digestive tract [21]. On the other hand, recent studies have suggested that host genetics
66 have limited impact on gut microbiota composition in humans [22].

67 We speculated that the microbial component of the reproductive tract might be an
68 important aspect of egg production in chicken. Here, we performed 16S rDNA
69 sequencing on 768 samples from three reproductive (vagina, uterus, and isthmus) and
70 three digestive (crop, gizzard, and small intestine) tract sites and whole-genome
71 sequencing of 128 laying hens. We characterized the reproductive tract microbiota and
72 its features compared with those of the digestive tract microbiota of hens. We identified
73 the contribution of key microorganisms to egg production, and established a correlation
74 between host genetics and the microbial diversity of six tract sites. These findings
75 provide insights into the microbial communities in the reproductive tract of highly
76 specialized layer populations, which may help develop strategies to enhance
77 commercial egg production.

78

79 **Results and Discussion**

80 **Discriminative characteristics of microbiota in reproductive and digestive tract**
81 **sites**

82 The 16S rDNA sequencing in 768 samples generated a total of ~57.61 M high-quality
83 reads (~75.01 K reads per sample). *De novo* clustering after singleton removal produced
84 46,480 operational taxonomic units (OTUs) at an identity cutoff of 97%, among which
85 6,776 OTUs found in > 20% of samples were used for subsequent analysis
86 (**Supplementary Fig. S1a, b**). We performed alpha diversity analysis based on

87 qualified sequencing depth with an average Good's coverage of 98.69% (ranging from
88 96.30% - 99.60%) (**Supplementary Fig. S2a**). Analysis of five indices (i.e., observed
89 OTUs, ACE, Chao1, Simpson, and Shannon) (**Supplementary Fig. S2b-f**) indicated
90 that the vast majority of pairwise comparisons between sites (10 of 90 pairwise
91 comparisons, or 11.11%) showed significant differences ($P < 0.001$, Wilcoxon rank-
92 sum test), with the exception of two comparisons (uterus vs. isthmus for all indices; and
93 small intestine vs. vagina for the observed OTUs, Simpson, and Shannon indices)
94 (**Supplementary Fig. S2g, h**).

95 Compared to the digestive system, the reproductive system exhibited higher alpha
96 diversity (all five indices) and thus contained more microbial taxa, especially in the
97 upper reproductive tract (i.e., uterus and isthmus) (**Fig. 1a, b**). Similar to significant
98 microbiota differences between the vaginal and upper reproductive tracts in humans [7,
99 23], we found highly discriminative microbial communities in chickens between the
100 upper (isthmus and uterus) and lower (vagina) reproductive tracts (isthmus vs. vagina,
101 $R = 0.473$, $P < 0.001$, Analysis of Similarity [ANOSIM]; uterus vs. vagina, $R = 0.496$,
102 $P < 0.001$, ANOSIM), but indistinguishable microbiota between the isthmus and uterus
103 (isthmus vs. uterus, $R = -0.003$, $P = 0.694$, ANOSIM) (**Supplementary Table S1**).
104 These results demonstrated microbiota heterogeneity throughout contiguous sites of the
105 digestive and reproductive tracts in hens.

106 We used principal coordinates analysis (PCoA) to visualize differences in taxa
107 composition between microbiota in the reproductive and digestive tracts. The first
108 principal component, explaining 34.96% of the variance in weighted UniFrac distance
109 matrices among the samples, separated reproductive and digestive tract samples (**Fig.
110 1c**). Given that reproductive and digestive tracts share a common exit in the cloaca, the
111 frequently exchanged microbiome likely resulted in similar microbiota at the distal end

112 of both tracts. Consequently, we found that the vagina acquired microbe communities
113 from the isthmus and uterus, which all belong to the reproductive tract. Nonetheless,
114 the vagina microbiota was partially indistinguishable from that of the small intestine
115 (**Fig. 1c**). The unweighted UniFrac distance matrices (**Supplementary Fig. S2i**),
116 weighted UniFrac distance (**Supplementary Fig. S2j**) and an ANOSIM based on Bray-
117 Curtis (BC) distances (**Supplementary Table S1**) recapitulated these findings.

118 Similar phyla dominated the microbiota in the six sites; *Firmicutes*, *Proteobacteria*,
119 and *Cyanobacteria* (likely an artifact of feed-derived chloroplast DNA) accounted for
120 71.45% - 97.86% of all OTUs. Nonetheless, some differences were observed among
121 the sites. *Cyanobacteria* (likely an artifact of feed-derived chloroplast DNA) was the
122 dominant material in the gizzard (48.19% of the total abundance); however, *Firmicutes*
123 was the most abundant phylum (43.60% to 78.93%) in the other five sites. We also
124 found that the uterus and isthmus had similar dominant phyla, including *Firmicutes*
125 (44.87% and 43.60%), *Proteobacteria* (26.25% and 23.77%), and *Bacteroidetes* (17.13%
126 and 19.52%) (**Supplementary Fig. S2k**). Strikingly, the vagina had the highest
127 abundance of *Fusobacteria* (11.51%) among the six sites.

128 At the genus level, *Lactobacillus* (7.24% - 73.74% in the six sites), *Exiguobacterium*
129 (2.79% - 4.68%), *Stenotrophomonas* (1.01% - 6.54%), and *Bacteroides* (3.40% -
130 4.86%) were ubiquitously found across all sites with higher abundances than other
131 bacteria due to their broad adaptability and beneficial functions (**Fig. 1a**). We found
132 *Lactobacillus* to be more dominant in the digestive tract (73.74% in crop, 24.76% in
133 gizzard, 30.70% in small intestine) compared with the reproductive tract (7.24% in
134 vagina, 9.27% in uterus, 9.91% in isthmus). *Lactobacillus* is thought to inhibit
135 pathogenic bacteria by lowering the environmental pH through lactic acid and hydrogen
136 peroxide production [24]. This genus was highly abundant in the digestive tracts which

137 were characterized by low pH values which strongly limits the growth of most
138 pathogens [25, 26]. In contrast, *Lactobacillus* was less abundant in the reproductive
139 tract where an alkaline pH is needed to maintain sperm motility [27, 28]. *Unidentified*
140 *Erysipelotrichaceae, Unidentified Chloroplast, Lactobacillus,* and *Bacteroides* had
141 abundances of >1.0% in the vagina, which was further increased in the uterus and
142 isthmus (**Fig. 1a**).

143 Furthermore, 14.63% of genera (362 of 2,475) demonstrated associations between
144 sites ($P < 0.05$ of Spearman' s r, Z-test) (**Supplementary Fig. S2l**). Typically, genera
145 belonging to *Proteobacteria* and *Firmicutes* showed significantly positive correlations
146 ($P < 0.001$, Z-test) between the crop and gizzard, the gizzard and small intestine, or the
147 three reproductive tract sites. These results imply there is a connection of microbiome
148 communities possibly caused because by the flow of material from different sites.

149

150 **Site-associated microorganisms in reproductive and digestive tracts**

151 We analyzed the functional capacity of the microbiota in each reproductive and
152 digestive tract site using PICRUSt2, and found that 72.00% of the representative
153 pathways (36 of the top 50 Kyoto Encyclopedia of Genes and Genomes [KEGG]
154 pathways) were shared across the six sites, one third of which (12 of 36) were primarily
155 involved in metabolism (**Supplementary Fig. S3a, b**). Specifically, 'bacterial secretion
156 system' and 'bacterial chemotaxis' were enriched in the reproductive tract. Previous
157 studies found that successful bacterial pathogens evolved versatile protein secretion
158 systems to promote their survival and fitness in response to different environmental
159 challenges, and to modulate host immunity [29-32]. Seven pathways were specifically
160 enriched at a site (three of six were site-specific to crop and gizzard, and 'riboflavin

161 metabolism' was specific to vagina). Abundances of the OTUs involved in these
162 pathways differed among the six sites ($P < 0.001$, Wilcoxon rank-sum test)
163 (**Supplementary Table S2**). For example, the microbial community of the small
164 intestine had important roles in 'valine, leucine, and isoleucine biosynthesis', as
165 indicated by the moderate row *Z* scores (-0.66) for each pathway. Moreover,
166 'propanoate metabolism' (*Z* score = 1.72) and 'bacterial chemotaxis' (*Z* score = 1.53)
167 were overrepresented in the vagina. Meanwhile, 'bacterial secretion system' was
168 overrepresented only in the uterus (*Z* score = 1.29) and isthmus (*Z* score = 0.94)
169 compared with the vagina and the three digestive tract sites (**Fig. 2a**).

170 We next identified 65 site-associated bacterial taxa among the six sites using Linear
171 discriminant analysis (LDA) Effect Size (LEfSe); [33] (**Fig. 2b and Supplementary**
172 **Fig. S3c**), which confirmed most of the observations described above (i.e., the uterus
173 and isthmus showed essentially similar microbiota). Of note, *Helicobacter* and
174 *Unidentified Erysipelotrichaceae*, which were associated with the small intestine,
175 showed the highest abundance among the six sites (**Fig. 2b**). Six genera from
176 *Lactobacillaceae* were crop-associated bacteria. In chicken, the crop acts as a reservoir
177 for the storage of food prior to its digestion, where food mixes with many beneficial
178 *Lactobacillus* bacteria (73.84% at the genus level) that produce lactic acid before
179 moving on to the proventriculus [34]. Next, the gizzard grinds any remaining large food
180 particles with the assistance of grit, releasing abundant *Unidentified Chloroplast*
181 (2.94%) and mitochondria-like (2.37%) materials from plant consumption. The small
182 intestine exhibited the most abundant microbes of the three digestive tract sites, which
183 is mainly where further digestion occurs, and fermentation begins. *Paenibacillaceae*
184 species, with optimum growth at pH 6.0-7.0, were also overrepresented. As a possible
185 pathogen, *Helicobacter* specifically inhabits the small intestine in chickens, and may

186 be involved in inflammation, metabolism, and neutralization of gastric acid [35-37].

187 *Unidentified Erysipelotrichaceae* showed higher abundance in the three reproductive
188 tract sites (1.83% - 2.40%). Bacteria associated with the isthmus and uterus both
189 showed higher abundances than in the other sites. Several genera (typically, *Romboutsia*,
190 *Fusobacterium,* and *Clostridium sensu stricto 1*) were dominant in the vagina (> 25%
191 of the microbiota) but had lower abundances in the other sites (**Fig. 2b**). Among these,
192 vaginal *Romboutsia* could be employed as a predictor for egg number in laying hens
193 [8]. Six *Bacteroidetes* bacterial taxa were isthmus-associated; *Bacteroides* species live
194 on host mucus-secreted polysaccharides and this flexible foraging behavior contributes
195 to diversity and stability [38].

196 These findings confirm that the digestive and reproductive tract microbiotas in
197 chicken are primarily determined by the physiological function of each compartment
198 within these systems.

199

200 **Weak association between host genetics and microbial communities**

201 In order to explore the relationship between host genome and microbiome of six sites,
202 we generated a total of 1.76 tera bases of high-quality genome sequences from 128
203 chickens with ~10.15-fold average depth for each individual (**Supplementary Table**
204 **S3**), and identified a total of 10.82 M single nucleotide polymorphisms (SNPs) with a
205 density of ~10.29 SNPs per kb.

206 The correlation between host genetics (using genetic relatedness matrix [GRM] and
207 microbial beta diversity based on BC distance) at the six sites in the same cohort of
208 laying hens was not statistically significant ($|r| < 0.033$, $P > 0.05$, Mantel test, **Fig. 3a-**
209 **f and Supplementary Table S4**). Nonetheless, the microbiomes of anatomically

210 neighboring sites were similar. Typically, the microbial communities of the isthmus
211 were positively correlated with those of the neighboring uterus (Spearman's $r = 0.426$,
212 $P < 0.0001$, Mantel test), but not significantly associated with the relatively distant crop
213 (Spearman's $r = 0.019$, $P = 0.335$, Mantel test, **Supplementary Table S4**). We also
214 estimated the association between GRM and microbial relationship matrix (MRM), and
215 obtained similar results: both Pearson's and Spearman's correlations suggesting that
216 host genetics and the microbiota composition are weakly associated (**Supplementary**
217 **Table S4**).

218 We next regarded the abundance of each microorganism as a quantitative trait to
219 estimate the h^2 of each microorganism at the species, genus, and OTU level.
220 Microorganisms in > 20% but < 60% of samples were analyzed qualitatively as
221 dichotomous traits (**Fig. 3g-i**). At the species and genus levels, no significant correlation
222 ($P > 0.05$, Wilcoxon rank-sum test) was found between the presence of a SNP and the
223 presence of a specific microbe (**Fig. 3j-l**) in the crop. Three species in the small intestine
224 (accounting for 0.21% of the microbiota species tested in the small intestine), six in the
225 vagina (0.39%), 14 in the uterus (0.92%), and ten in the isthmus (0.66%) exhibited
226 significant SNP-based heritability ($P < 0.05$, Wilcoxon rank-sum test) (**Fig. 3k and**
227 **Supplementary Table S5**). Most of these heritable bacteria belonged to the *Firmicutes*
228 phylum (**Supplementary Fig. S4a, c, e**). Reproductive tract sites had more heritable
229 bacterial phyla than digestive tract sites (**Supplementary Fig. S4b, d, f**). The
230 cumulative abundances of these heritable bacteria were only 0.22%, 4.14%, 1.46%, and
231 1.61% ($P < 0.05$, Wilcoxon rank-sum test) in the small intestine, vagina, uterus, and
232 isthmus, respectively (**Fig. 3n and Supplementary Table S5**). Similar results were
233 observed at the genus and OTU levels (**Fig. 3m, o**). These results supported that host
234 genetics have limited effect on shaping the microbial composition of the reproductive

235 and digestive tracts.

236

237 **Heritability (h^2) and microbiability (m^2) of EN300**

238 To further explore the effect of genome and microbiome on egg number at 300 days of
239 age (EN300), we used a GRM of sample pairs to estimate the h^2 value of EN300
240 explained by whole genome SNPs using the restricted maximum likelihood method.
241 We found that EN300 exhibited relatively low to medium heritability ($h^2 = 0.282$, $P =$
242 0.048, likelihood ratio test), which was comparable to previous estimations
243 (**Supplementary Table S6**) [5, 6]. The fraction of EN300 variance explained by
244 microbial variance was measured by microbiability (m^2) [18]. After correcting for host
245 genetic factors using EN300-related SNPs as additional covariates, we found that the
246 estimated EN300 m^2 values for digestive tract sites (0.523 for small intestine, 0.869 for
247 crop, and 0.873 for gizzard) were lower than those for reproductive tract sites (0.923
248 for vagina, 0.936 for uterus, and 0.989 for isthmus) (**Table 1**). Generally, higher EN300
249 m^2 values were observed for sites neighboring the ovaries; the isthmus was the most
250 pertinent site with respect to egg production. Commercial egg producers are acutely
251 interested in hen oviducts because pathological changes or disrupted activity directly
252 affect egg production efficiency and ultimately decrease economic profitability [39]. In
253 chickens, the inner and outer shell membranes form in the isthmus, while calcification
254 of the eggshell, subsequent pigmentation, and cuticle deposition occur in the uterus and
255 is followed by expulsion of the egg through the vagina [40]. These results suggest that
256 EN300 in layer chickens is determined more by the microbiota in the reproductive tract
257 than in the digestive tract.

258 **Table 1.** Estimated microbiability (m^2) of EN300

	Site	m^2	Standard error	*P* value
Digestive tract	Crop	0.869	0.049	$< 10^{-16}$
	Gizzard	0.873	0.045	$< 10^{-16}$
	Small intestine	0.523	0.111	2.56×10^{-11}
Reproductive tract	Isthmus	0.989	0.011	$< 10^{-16}$
	Uterus	0.936	0.030	$< 10^{-16}$
	Vagina	0.923	0.028	$< 10^{-16}$
Host genetics		0.282	0.231	0.049

259

260 **Microorganisms in the reproductive tract are significantly associated with EN300**

261 We next focused on the microorganisms which are highly associated with EN300. The
262 results showed that most of the microorganisms detected at the microbial species,
263 genus, and OTU levels that significantly associated with EN300 belonged to the
264 *Firmicutes* phylum ($P < 0.05$, Wilcoxon rank-sum test) (**Supplementary Fig. S5**). Only
265 microorganisms that exhibited a significant correlation between egg production and
266 relative abundance as determined by both Pearson's *r* and Spearman's *r* were
267 considered a causal relationship ($P < 0.05$, Wilcoxon rank-sum test) Consequently, 39
268 OTUs, 26 genera, and 24 species fulfilled these criteria (**Fig. 4a and Supplementary**
269 **Fig. S6a**).

270 Most OTUs, genera, and species present in the three digestive tract sites were
271 negatively correlated (Pearson's *r*) with egg production (negative/positive: 19/6, 8/8
272 and 16/5, respectively), whereas they were positively correlated with egg production in
273 the reproductive tract sites (positive/negative: 13/9, 11/10 and 13/4, respectively) (**Fig.**
274 **4b and Supplementary Fig. S6b, c**). Microorganisms in the uterus were most strongly
275 correlated with each other (**Fig. 4c and Supplementary Fig. S7**), which implied a
276 strong symbiotic/competitive relationship.

277 At the genus level, *Lactobacillus*, *Bacteroides,* and *Desulfovibrio* were positively

278 correlated with EN300 in the three reproductive tract sites. *Pseudomonas*,
279 *Exiguobacterium,* and *Unidentified Erysipelotrichaceae* were negatively correlated
280 with EN300 in the three reproductive tract sites, but were positively correlated in the
281 three digestive tract sites (**Supplementary Fig. S6b**). In accordance with previous
282 findings, there is clear evidence of the role of fecal microbiomes in low and high egg-
283 laying performance in hens; Elokil et al [16] demonstrated a significantly positive
284 association between the microbial genus *Lactobacillus* and egg-laying performance (P
285 < 0.05). Likewise, Wang et al [41] reported that *Lactobacillus* was also abundant in the
286 feces of high-yield hens. The genus *Lactobacillus* produces growth promoters and
287 exhibits antimicrobial activity against pathogenic microbes [42, 43] which may explain
288 why the increasing abundance of *Lactobacillus* in the high-yield group is beneficial to
289 egg-laying performance. The cecum has distinct microbial community profiles [44-46]
290 that were not explored in this study. Microbial community analysis of the cecum
291 microbiota in chickens exhibiting different egg production performances requires
292 further investigation.

293 At the species level, *Bacteroides fragilis*, *B. salanitronis*, *B. barnesiae,* and *C. leptum*
294 were positively correlated with EN300 in the three reproductive tract sites, but were
295 weakly negatively correlated with EN300 in the small intestine, while no correlation
296 was found in the crop and gizzard. The first three species belong to the genus
297 *Bacteroides*, which had a significantly positive correlation with egg production in the
298 three reproductive sites (Pearson's r = 0.403 - 0.479). *Bacteroides* species have been
299 identified as the predominant anaerobic genera in chicken cecum [47], which were
300 thought to play an important role in the breakdown of polysaccharides into simpler
301 compounds used by the animal host as well as the microorganisms themselves [48].
302 Interestingly, a recent study reported that the human-adapted *Bacteroides* species are

303 likely introduced to chicken flocks by contact with humans and can temporarily persist
304 in chickens [49]. Intestinal anaerobic bacteria such as *B. fragilis* and *B. salanitronis*
305 have been suggested to possess metabolic pathways for N-glycan production [50]. The
306 symbiont *B. fragilis* exists in a commensal relationship with the host as it expresses a
307 relatively large number of genes involved in polysaccharide metabolism, which benefits
308 the host. The surface of *B. fragilis* can produce polysaccharides; in particular, capsular
309 polysaccharide A (CPSA) is a key mediator of mammalian immune system
310 development [51]. Surprisingly, CPSA has also been shown to exert protective effects
311 in autoimmune disorder models, such as antibiotic-induced experimental
312 encephalomyelitis. It is thus suggested that the genus *Bacteroides* could regulate
313 reproductive activity by mediating the avian immune system.

314 *Firmicutes* bacterium ZOR0006 had a significantly negative correlation with EN300
315 in the three reproductive tract sites and a significantly positive correlation in the three
316 digestive tract sites. The 20% of chickens with the lowest EN300 values (mean = 37.13)
317 had significantly lower *B. fragilis*, *B. salanitronis*, *B. barnesiae,* and *C. leptum*
318 abundances (P < 0.05, Wilcoxon rank-sum test) (**Supplementary Fig. S8a, b**)
319 compared with the highest EN300 values (mean = 113.75) of the 20% of chickens in
320 the reproductive tract sites. Although its function is unknown, 20% of chickens with
321 the highest abundance of *Firmicutes* bacterium ZOR0006 exhibited significantly lower
322 EN300 values than the 20% of chickens with the lowest abundance of this
323 microorganism (**Fig. 5a**) (P < 0.05, Wilcoxon rank-sum test) in the reproductive tract
324 sites.

325 Moreover, the 20% of chickens with the highest *B. fragilis*, *B. salanitronis*, *B.*
326 *barnesiae,* and *C. leptum* abundances exhibited significantly higher EN300 values than
327 the 20% of chickens with the lowest abundances of these microorganisms in the

328 reproductive tract sites, with the exception of *C. leptum* abundance in the vagina (**Fig.
329 5a**). *C. leptum,* a major member of the *Firmicutes* phylum, can alter the gut microbiota
330 in rats, especially in obese individuals. In human infants, fecal levels of *C. leptum* were
331 found to be negatively correlated with proinflammatory marker levels [52]. Colonic
332 colonization of *C. leptum* was associated with accumulation of regulatory T cells, which
333 inhibited the development of inflammatory lesions. The proliferation and activation of
334 regulatory T cells is crucial to establishing and maintaining an appropriate level of
335 immune tolerance. In addition, our results demonstrated that *C. leptum* was associated
336 with a large range of other uterus or isthmus microbiota constituents (but limited
337 association was observed with digestive microbiota constituents) and was not
338 influenced by host genetics. Thus, this microorganism might serves as a stimulator of
339 regulatory T cell production and inhibitor of inflammatory lesions, then regulating and
340 maintaining immunologic tolerance and microbiota composition of the reproductive
341 tract (especially the uterus and isthmus). These results suggest that the microbial species
342 contribute to the enhanced egg production are modulated by influencing the immune
343 processes.

344 We then characterized the spatial distribution of these five EN300-associated
345 microorganisms (*B. fragilis*, *B. salanitronis*, *B. barnesiae*, *C. leptum,* and *Firmicutes*
346 bacterium ZOR0006). *B. fragilis* was detected in almost all samples and accounted for
347 0.05% - 1.29% of the total abundance (**Fig. 5b**). *B. salanitronis* and *B. barnesiae* were
348 detected at similar ratios in the six sites and in most samples from the reproductive tract
349 sites; both accounted for the highest abundance in the vagina. *Firmicutes* bacterium
350 ZOR0006 was also detected in most samples from the reproductive tract sites (74.22%
351 - 89.84%) and in half of the samples from the digestive tract sites (48.44% - 58.59%),
352 accounting for 0.61% - 2.40% of the total abundance. Although the detection ratio

353 (28.13% - 64.84%) and relative abundance of *C. leptum* were much lower than those of
354 other microorganisms in all six sites, they accounted for the highest abundance in the
355 isthmus and uterus (**Fig. 5b**).

356 Multiple factors, especially host species, potential pathogens, and immune status of
357 the host, all play a major role in the female reproductive organs adversely interfering
358 with the egg industry in laying flocks [53]. Additionally, the digestive tract environment
359 of low-egg producing hens is fragile and susceptible to the influence of exogenous
360 microorganisms [41]. Pathogenic infection, room temperature fluctuations,
361 management systems, and other sudden changes to various factors can alter the
362 composition of microbiota [54, 55]. These alterations may cause a significant
363 degradation in production performance. Here, our results indicate that the reproductive
364 tract microbiota play an important role in egg production.

365

366 **Transcriptomic divergence in the uterus between hens with high and low egg**
367 **production**

368 Furthermore, we compared the transcriptional profiles in the uterus between the two
369 groups composed of hens with either the 20% lowest or 20% highest EN300 values (six
370 hens for each group). As expected, the correlation rates between the high- and low-egg
371 production groups (mean Pearson's $r = 0.93$) were relatively lower than those between
372 biological replicates (mean Pearson's $r = 0.95$ and 0.96 for groups with high and low
373 egg production, respectively) (**Fig. 6a, b**), indicating significant biological differences
374 between groups. We identified 1,051 genes that exhibited significant expression
375 changes (false discovery rate [FDR] ≤ 0.01 and $|\log2 \text{ (fold change)}| \geq 1$) between groups
376 with high and low egg production (**Fig. 6c**), which are mainly involved in immune-

377 related categories, including the 'NF−κB signaling pathway' and 'chronic inflammatory
378 response' (**Supplementary Fig. S9**). Of these, 739 genes that were significantly
379 downregulated in the high-egg production group were overrepresented in the categories
380 related to the inflammatory response, including 'T cell costimulation', 'B cell receptor
381 signaling pathway' and 'lymphocyte activation' (**Fig. 6e**). Notably, we observed eight
382 well-documented inflammatory markers (two Toll-like receptors [*TLR15* and *TLR1A*]
383 and six interleukins [*IL21R*, *IL18RAP*, *IL22RA2*, *IL4I1*, *IL17REL* and *IL8*]) (**Fig. 6c, d**)
384 that were significantly different expressed in the uterus of the high-egg production
385 group and the low-egg production group. Functionally, the microbiota of the uterus
386 affects the health of the oviduct and thus influence chicken egg production, which is
387 manifested as increased pathway abundance for bacterial motility proteins, the bacterial
388 secretion system, and membrane and intracellular structural molecules.

389

390 **Conclusion**

391 Our study provides a comprehensive view of the microbial community in the digestive
392 and reproductive tracts of layer chickens. The diversity, composition, and predicted
393 function of the microbiota varied considerably according to location within the
394 reproductive and digestive tracts. Our results indicate that the reproductive tract
395 microbiota in the hen influences egg production more than the digestive tract
396 microbiota, and host genome has limited effect on their microbial composition. A small
397 proportion of the variability in egg production was associated with the microbiota in
398 the reproductive and digestive tracts of chickens. Remarkably, the genus *Bacteroides*
399 and the species *C. leptum* and *Firmicutes* bacterium ZOR0006 were strongly associated
400 with egg production, indicating their potential role in promoting reproductive
401 performance. These findings provide new insight into the roles of reproductive and

402 digestive tract microbiota for complex traits, and may help contribute to the
403 development of effective therapies for improving commercial egg production in
404 chickens.

405

406 **Methods**

407 **Chickens**

408 The study was conducted on a common flock of 128 Dongxiang green shell laying hens
409 reared on an experimental poultry farm at Sichuan Agricultural University in Ya'an,
410 Sichuan, China. All chicks were hatched on the same day and housed in individual pens.
411 Feed intake was controlled daily according to standard farm husbandry practices and
412 water was provided *ad libitum*. The number of eggs produced for the first 300 days of
413 life was recorded daily for each individual. We determined that the mean number of
414 eggs (~75.32; ranging from 24-129) that each hen laid by 300 days of age (EN300) fit
415 a normal distribution pattern (P = 0.725, Kolmogorov-Smirnov test) (**Supplementary**
416 **Fig. S10**). At the age of 300 days, 2 mL of whole blood was collected from the wing
417 vein using venipuncture and stored at -20 °C. Subsequently, each individual was culled
418 by cervical dislocation followed by decapitation. After the abdomen was opened, fresh
419 tissue was collected from three sites in the reproductive tract (vagina, uterus and
420 isthmus) and three sites in the digestive tract (crop, gizzard and small intestine) (**Fig.**
421 **1a**). A 12-cm-long fixed mid-region of the small intestine (jejunum) was collected from
422 each bird. All samples were snap-frozen in liquid nitrogen, transported to the laboratory,
423 and stored at -80 °C until further analysis. Chickens were managed according to the
424 Institutional Animal Care and Use Committee of Sichuan Agricultural University under
425 permit number DKY- 2018102015.

426

427 **Host and microbial genomic DNA extraction**

428 Host DNA was isolated from blood using a TIANamp Genomic DNA Kit (Tiangen
429 Biotech) following the manufacturer's instructions. Total microbial genomic DNA
430 from lumen of digestive tracts and mucus of reproductive tracts were extracted from
431 ~200 mg tissues using a TIANamp Stool DNA Kit (Tiangen Biotech, Beijing, China)
432 according to the manufacturer's instructions. The extracted DNA was quantified using
433 a NanoDrop2000 spectrophotometer (Thermo Fisher Scientific, Chengdu, China), and
434 DNA integrity was determined by 1% agarose gel electrophoresis.

435

436 **16S rDNA amplicon and sequencing**

437 The V4 hypervariable region of the bacterial 16S rDNA was amplified using Phusion®
438 High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, Ma, USA) and the
439 universal primers 515 F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806 R (5′-
440 GGACTACHVGGGTWTCTAAT -3′) [56]. Reactions were carried out using 15 μL of
441 Phusion® High-Fidelity PCR Master Mix, 3 μL of the forward and reverse primers, 10
442 μL of template DNA, and 2 μL of ddH_2O, no template control was also performed. The
443 PCR cycling conditions were as follows: an initial denaturation step at 98 °C for 1 min
444 and 30 cycles of 98 °C for 10 sec, 50 °C for 30 sec, and 72 °C for 30 sec, followed by
445 an extension step at 72 °C for 5 min, using a Bio-Rad CFX96 thermal cycler (Bio-Rad
446 Laboratories, Hercules, CA, USA). Amplicons were purified on agarose gel (1%) using
447 a GeneJET Gel Extraction kit (Thermo Fisher Scientific, Schwerte, Germany). A DNA
448 library was prepared using an Ion Plus Fragment Library Kit (Thermo Fisher Scientific)
449 based on the manufacturer's instructions. Reads were barcoded per sample, combined

for multiplexed sequencing with the Ion S5™ XL platform (Thermo Fisher Scientific) to generate 400-bp single-end reads, and sequenced by Novogene Bioinformatics Technology Co. Ltd of China.

Whole-genome sequencing

After qualified host DNA samples were tested, the DNA was randomly fragmented using an ultrasonicator (Covaris Inc., Woburn, MA, USA), and then a sequencing library was prepared using a TruSeq Nano DNA HT Sample Preparation Kit (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Index codes were added to tag each sample. DNA fragments were then end-repaired, dA-tailed, and ligated with a full-length adapter for Illumina sequencing, PCR amplification, and purification. Next, isolated DNA libraries were constructed with an insert size of 350 bp. Finally, genomes of the 128 individuals were separately sequenced with 150 bp paired-end reads using the Illumina Novaseq platform by Novogene Bioinformatics Technology Co. Ltd.

16S rDNA-seq data processing

The raw data obtained by sequencing were first filtered to obtain high-quality data. First, the adapter sequences in 61.05 million (M) raw reads (**Supplementary Table S7**) were trimmed using Cutadapt (version 1.9.1) [57] when the overlap length between the read and the adapter was shorter than 10 bp. Then, barcode sequences were trimmed and reads that were too long (> 260 bp) or too short (< 220 bp) were filtered using Cutadapt with the parameters "-e 0 -q 17 -m 200 -M 2600". Next, the remaining reads were compared with the ChimeraSlayer reference database using the UCHIME algorithm [58]

474 to detect and then remove chimera sequences. Finally, low-quality reads (i.e., more than
475 50% of bases with a phred quality lower than five) were also removed. Consequently,
476 57.61 M high-quality reads were generated for subsequent analysis (**Supplementary**
477 **Table S8**). Greater than 93.85% of the high-quality reads had lengths of 250 - 260 nt
478 (**Supplementary Fig. S11a**). Data with a quality score > 20 accounted for 88.14% of
479 all the effective bases (**Supplementary Fig. S11b, c**). The error ratio of the sequencing
480 reads was relatively high in the ending position (**Supplementary Fig. S11d**).

481

482 **OTU cluster and species annotation**

483 The remaining high-quality sequences were used to generate OTUs by Uparse software
484 (version 7.0.1001) [59] with an identity cutoff of 97%. Singleton OTUs (OTUs found
485 in only one sample) that did not match the reference database were removed. Clustering
486 across all samples from the 128 chickens produced 46,480 OTUs after singleton
487 removal. Nonetheless, most of those OTUs were present in low abundance and were
488 found in very few samples. We then discarded OTUs that were not found in at least 20%
489 of the chickens in each sampling site, yielding 6,776 OTUs (**Supplementary Fig. S1a,**
490 **b**). For each OTU, the SSUrRNA library in the Silva (https://www.arb-silva.de/) [60]
491 was used to annotate taxonomic information (i.e., kingdom, phylum, class, order, family,
492 genus and species) based on the Mothur algorithm [61]. Subsequently, we determined
493 the phylogenetic relationship of different OTUs and dominant species differences in
494 samples (groups) after multiple sequence alignment using MUSCLE software (version
495 3.8.31) [62]. Additionally, OTU abundance information was normalized using a
496 standard sequence number corresponding to the sample with the fewest sequences.

497

498 **Alpha diversity**

499 Alpha diversity was used to analyze the complexity of species diversity for a sample
500 based on normalized OTUs through six indices (i.e., observed OTUs, ACE, Chao1,
501 Simpson, Shannon, and Good's coverage), using the QIIME2 software [63]. Among
502 these, Chao1 and ACE were selected to identify community richness, Shannon and
503 Simpson were used to identify community diversity, and Good's coverage was used to
504 characterize sequencing depth. Differences in alpha diversity indices among the six
505 sites were calculated with the Wilcoxon rank-sum test using R software (version 2.15.3).

506

507 **Beta diversity**

508 Beta diversity was used to evaluate differences in samples. Beta diversity in BC and
509 weighted/unweighted UniFrac distances were calculated using QIIME2 software [63].
510 The BC ordination provided position values along an ordination axis and distances from
511 the axis for samples of communities.

512

513 **Principal coordinates analysis (PCoA)**

514 PCoA was performed to obtain principal coordinates and to visualize complex,
515 multidimensional data. A distance matrix of previously obtained weighted/unweighted
516 UniFrac distances among samples was transformed to a new set of orthogonal axes, by
517 which the maximum variation factor was demonstrated by the first principal coordinate,
518 the second maximum variation factor was demonstrated by the second principal
519 coordinate, and so on. PCoA was performed using the WGCNA package [64], stat
520 packages, and ggplot2 package in R software.

521

522 **Prediction of the functional profiles of microbial communities**

523 The functions of the microorganisms present in the microbial communities detected in
524 the six sites were predicted using PICRUSt2 [65]. We used the Wilcoxon rank-sum test
525 to investigate differences in pathways among sites. *P*-values were adjusted using the
526 Benjamini-Hochberg method by the FDR with the p.adjust function in R.

527

528 **Community difference analysis**

529 Pairwise comparisons between different sites were statistically compared using
530 ANOSIM (also named permutational MANOVA) with 10,000 permutations based on
531 BC ordination to evaluate the reasonability of the division of groups.

532

533 **Between-group variation analysis**

534 High-dimensional biomarkers were discovered by LEfSe using the parameter 'LDA
535 score > 4' [33] to identify characteristics of abundance and related classes (e.g., genes,
536 metabolites, or taxa).

537

538 **Identification of microbiota constituents related to egg production**

539 EN300 values between two groups (the lowest- and highest-ranked 20% of chickens
540 with respect to their EN300 value) were then compared using the Wilcoxon rank-sum
541 test. Microorganisms with $P < 0.05$ and FDR < 0.05 were retained. Furthermore, we
542 calculated the Spearman's *r* and Pearson's *r* between EN300 and the abundance of each

microbiota constituent at genus, OTU, and species levels. A significant correlation between the presence of a microorganism and the EN300 value was considered if $P < 0.05$, as determined using the psych package in R with the P value adjusted using the Benjamini-Hochberg method. Overlapping microorganisms obtained from the Wilcoxon rank-sum test and Spearman's r and Pearson's r were considered to have a potential relationship with EN300. We subsequently characterized EN300-associated microbes in the six sites.

Whole-genome sequencing data processing

To avoid analysis noise caused by sequencing errors, low-quality paired reads (reads with ≥ 10% unidentified nucleotides [N]; > 10 nt aligned to the adaptor, allowing ≤ 10% mismatches; > 50% bases having phred quality < 5; and putative PCR duplicates generated in the library construction process), which mainly resulted from base-calling duplicates and adaptor contamination, were removed using an in-house script [66]. Consequently, 1.30 terabases (~10.15-fold per individual) of high-quality paired-end reads were obtained, including 95.13% and 88.98% nucleotides with phred quality ≥ Q20 (with an accuracy of 99.00%) and ≥Q30 (with an accuracy of 99.90%), respectively (**Supplementary Table S3**).

Read mapping, and genomic variant calling and annotation

The remaining high-quality reads of each individual were aligned to the reference chicken genome (Gallus_gallus-6.0 Ensembl release 98, http://asia.ensembl.org/Gallus_gallus/Info/Index/) using the Burrows-Wheeler Alignment tool (BWA) (version 0.7.15) [67] with the command 'mem -t 10 -k 32'.

567 BAM alignment files were then generated using SAMtools (version 0.1.19) [67].
568 Additionally, we improved alignment performance through filtering the alignment
569 reads with mismatches ≤ 5 and mapping quality = 0. After sorting by SAMtools, the
570 sorted BAM file was marked in duplicate using the command "MarkDuplicates" in the
571 package Picard (version 1.119).

572 Subsequently, we performed gVCF calling in accordance with the Genome Analysis
573 Toolkit (GATK) best practices pipeline (version v3.7) [68] using the HaplotypeCaller-
574 based method, and then population SNP calling by merging all gVCFs with the
575 commands "CombineGVCFs".

576 To obtain high-credibility SNPs, we applied the hard filter command
577 'VariantFiltration' to exclude potential false-positive variant calls as follows: (a) quality
578 by depth >10.0; (b) mapping quality score > 40.0; (c) FS < 60.0; (d) MQRank-Sum >
579 −12.5; (e) ReadPosRankSum > −8.0. In addition, In addition, we further removed the
580 SNPs with adjacent distances ≤ 5 [69]. Finally, we used vcftools (version 0.1.15) to
581 obtain biallelic variants with the following parameters: sample call rate > 90%, SNP
582 call ratio > 95%, minor allele frequencies > 1%, and Hardy-Weinberg equilibrium *P*
583 value < 10^{-5}. Ultimately, a total of 10.82 M high-credibility SNPs in 128 individuals
584 were retained (**Supplementary Table S9**). SNPs were classified into different genomic
585 regions (i.e., exonic, intronic, splice sites, upstream and downstream around gene
586 regions, and intergenic) using the ANNOVAR package [70].

587

588 **Construction of microbial relationship and host genetic relatedness matrices**

589 OTUs identified in each site were normalized to a zero mean and unit variance. We then
590 constructed a MRM [71] using an R script based on the following equation:

591

$$r_{tij} = \frac{1}{N_T}\sum_{o=1}^{N_T}\frac{(a_{tio} - \overline{t_{to}})(a_{tjo} - \overline{a_{to}})}{\sigma_{to}^2}$$

592 where r_{tij} represents the tested microbial relationship in tract t between chickens i and
593 j; a_{tio} and a_{tjo} are the abundance of OTU o in tract t in chickens i and j,
594 respectively; $\overline{t_{to}}$ is the average relative abundance of OTU o in tract t in the
595 population; σ_{to}^2 is the variance in the abundance of OTU o in tract t; and N_T is the total
596 number of OTUs in tract t used for the computation of relatedness. High-quality SNPs
597 were further used to detect independent markers using PLINK [72], with the following
598 parameters: 50 kb window size, 10 SNPs per step, and 0.2 as a squared Pearson's *r* (r^2).
599 All 10,809,968 SNPs were used to compute the principal components (PCs) and GRM
600 [73] using GCTA version 1.91.1 [74]:

601

$$h_{ij} = \frac{1}{N}\sum_{a=1}^{N}\frac{(r_{ia} - 2\overline{f_a})(r_{ja} - 2\overline{f_a})}{2\overline{f_a}(1 - \overline{f_a})}$$

602 where h_{ij} is the tested genetic relationship between chickens i and j; r_{ia} and r_{ja}
603 represent the number of reference alleles in chickens i and j, respectively; $\overline{f_a}$ is the
604 frequency of the reference allele in the population; and N is the number of variants.

605

606 **Heritability (h^2) analysis**

607 To estimate the effects of host genetics on the microbiota at different sites, we computed
608 the correlation between GRM and BC distances at each site using both Pearson's *r* and
609 Spearman's *r,* based on Mantel tests with 10,000 permutations. The correlation between
610 GRM and MRM was also computed. To estimate the correlation between GRM and the
611 microbiota community, we computed heritability at OTU, genus, and species levels.
612 OTU abundance information was normalized using a standard sequence number

613 corresponding to the sample with the least number of sequences. Microorganisms that
614 were present in < 60% but ≥ 20% of the samples were dichotomized as present or absent
615 [75], and the microorganisms that were detected in < 20% of the samples from each site
616 were excluded from the analysis.

617

618 **Genetic and microbial parameters of egg production**

619 As the individuals examined in this study had no pedigree information, we computed
620 the SNP-based heritability of the egg production phenotype (i.e., EN300) instead, using
621 the following model [73]:

622 $$y = K_c + g + e[A]$$

623 where y is an observed value (EN300); c is a vector of fixed covariates with the
624 corresponding design matrix K; e is the residual effect; and g is a vector of aggregate
625 effects of all SNPs with an ~N(0,$G_{\sigma_A^2}$), where G and σ_A^2 are the GRM and polygenetic
626 variance (overall SNP effects), respectively. The top five host genetic PCs were
627 considered covariates in the model to account for the calculated population stratification,
628 as described above. The likelihood ratio test P value was calculated to examine the
629 significance of the association between SNPs and EN300.

630 The fraction of EN300 variance explained by microbial variance was calculated as
631 $m^2 = \frac{\sigma_m^2}{\sigma_p^2}$; (called 'microbiability' [m^2] in animals [71] and 'microbiome-association
632 index' in humans [18]), where σ_m^2 and σ_p^2 are the phenotypic variance and microbial
633 variance, respectively. To adjust for host genetic effects, all valid individuals and SNPs
634 were used in a GWAS with a univariate linear mixed model (LMM), which was
635 performed using GEMMA [76]. The LMM was calculated as follows:

636 $$y=K_c+m_s+e \text{ [B]}$$

637 where the model parameters are the same as those described in model [A], except m_s,
638 which is the random effect of the microbiota in locations following the multinomial
639 distribution m_s~N(0,Mσ_m^2), and M is the MRM. We then used the MRM in GCTA to
640 calculate m^2. The genome-wide significance threshold was 10^{-6}. We then extracted
641 these SNPs with significant effects on EN300 and calculated the PCs using PLINK.
642 The first two PCs and the top five host genetic PCs were then used as covariates in
643 model [B] to account for host genetics.

644

645 **RNA sequencing (RNA-seq) analysis**

646 For RNA-seq, total RNA was extracted from uterine tissue of twelve hens (six for each
647 of the groups with either the 20% lowest or 20% highest EN300 values) using the
648 RNeasy Mini Kit (Qiagen). We used an rRNA depletion protocol (Ribo-Zero kit,
649 Epicenter) coupled with the Illumina TruSeq stranded RNA-seq library protocol to
650 construct the RNA-seq libraries. A total of 12 libraries were quantified using the Qubit
651 dsDNA High Sensitivity Assay Kit (Invitrogen) and separately sequenced on the
652 NovaSeq 6000 platform (Illumina) to produce an average of ~31.86 M 150-bp paired-
653 end raw reads and ~30.52 M high-quality reads for each library. Sequence reads were
654 aligned to the chicken reference genome (Gallus_gallus-6.0 Ensembl release 98) by the
655 STAR alignment tool (version 2.5.3a). On average, ~96% of reads of individual
656 libraries were aligned to the chicken reference genome, generating an average of 29.30
657 M aligned reads for each sample. The gene expression level was then estimated as

658 transcripts per million (TPM) using the high-speed transcript quantification tool

659 Kallisto (V0.43.0) [77].

660 We used the edgeR package [78] to identify differentially expressed genes (FDR <

661 0.01 and |log2(fold change)| ≥ 1) between the two groups with either the 20% lowest or

662 20% highest EN300 values. Functional enrichment analysis of differentially expressed

663 genes was performed using the Metascape tool [79]. Only Gene Ontology (GO) terms

664 and KEGG pathways with a P value < 0.05 were considered significant and are listed.

665 The expression levels of eight genes were verified using a quantitative PCR (q-PCR)

666 approach. The β-actin gene of chicken was used as an endogenous control gene.

667 Relative expression levels of objective mRNAs were calculated using the ΔΔCt method.

668 The primer sequences used for q-PCR are shown in **Supplementary Table S10**. All

669 measurements contained a negative control (no cDNA template), and each RNA sample

670 was analyzed in triplicate.

671

672 **Figure Legends**

673 **Figure 1.** Diversity and composition of the reproductive and digestive tract microbiota

674 in chickens. (**a**) Relative abundance of the microbiota from six sites at the genus level.

675 Only genera with an abundance > 1% in a site are shown. (**b**) Alpha diversity

676 comparison based on the Shannon diversity index (*P < 0.05; **P < 0.01, Wilcoxon

677 rank-sum test). **c.** Principal coordinates analysis of the 768 samples based on weighted

678 UniFrac distances. 34.96% of variance was explained for component 1 (P < 0.05,

679 Tracy-Widom test) and 24.23% for component 2 (P < 0.05, Tracy-Widom test).

680 **Figure 2.** Comparison of predicted functional capacities and site-associated taxa of
681 microbial communities. (**a**) Heatmap showing the predicted KEGG pathways and their
682 abundances at reproductive and digestive tract sites (Supplementary Table S2). Z scores
683 indicate the means of KEGG pathway abundances. (**b**) Heatmap showing the 65 site-
684 associated bacterial taxa identified by LEfSe (LDA > 4). Z scores indicate the relative
685 abundances of site-associated bacterial taxa. Black frames represent site-associated
686 bacterial taxa whose *Z* scores of relative abundances differed significantly among the
687 six sites. p, phylum; c, class; o, order; f, family; g, genus.

688 **Figure 3.** Effect of host genetics on microbiota of the reproductive and digestive tracts.
689 (**a-f**) Density scatter plots of genetic kinship of pairs of individuals (x axis) and their
690 microbiome dissimilarity (y axis) among all pairs of individuals (n = 16,256). NS, not
691 significant ($P > 0.05$; Mantel test). **(g-i)** Distribution of species, genera, and OTUs
692 identified in six sites for all chickens. Microorganisms present in < 20% of samples
693 were excluded. **(j-l)** Proportion of heritable microbial genera, species, and OTUs in
694 each site. **(m-o)** Cumulative relative abundances of heritable microbial genera, species,
695 and OTUs in each site.

696 **Figure 4.** Microorganisms associated with egg number at 300 days of age (EN300). (**a**)
697 Microbial species associated with EN300 ($P < 0.05$) among the six sites. **(b)** Pearson's
698 r values between EN300 and 24 EN300-associated microbial species; only significant
699 r values are given numerically ($P < 0.05$). **c** Pearson's r (lower diagonal) and
700 Spearman's r (upper diagonal) values among microbial species in the uterus (UT);
701 significant r values are given numerically ($P < 0.05$).

702 **Figure 5.** Effect of microbial species associated with EN300. (**a**) EN300 values for the
703 20% of chickens with the highest and lowest abundances of *B. fragilis*, *B. salanitronis*,

704 *B. barnesiae*, *C. leptum,* and *Firmicutes* bacterium ZOR0006 in the three reproductive
705 tract sites. The plots show the median, as well as the 25% and 75% quantiles. The cross
706 and horizontal lines indicate the mean and median values in the corresponding group,
707 respectively. Significance levels were calculated using a permutation test with 10,000
708 replicates. *$P < 0.05$; NS, not significant. (**b**) Relative abundance and detected ratio of
709 five species (*B. fragilis*, *B. salanitronis*, *B. barnesiae*, *C. leptum* and *Firmicutes*
710 bacterium ZOR0006) in the six sites. Blue bars indicate the detection ratio of each
711 species at each site.

712 **Figure 6.** Microorganisms influence the expression of immune-related genes in the
713 uterus. **(a)** Hierarchical clustering and (**b)** pairwise Pearson's correlations of 12 samples
714 using transcriptional profiles. **(c)** Differentially expressed genes between groups with
715 high and low egg production. Eight well-documented inflammatory markers are labeled.
716 (**d)** Comparison of expression changes of eight inflammatory markers in the uterus
717 between groups with high and low egg production by a q-PCR approach. **(e)** Top 20
718 functional categories enriched for 739 genes that were significantly downregulated in
719 the high-egg production group. The enrichment analysis was performed using the
720 Metascape tool (see Method). GO-BP: biological process (blue) and GO-MF:
721 molecular function (yellow).

722

723 **Data Availability**

724 The sequencing data for this project have been deposited in the NCBI and can be
725 accessed with BioProject Nos. PRJNA730194 (Microbiome), PRJNA731001 (Whole
726 genome resequencing) and PRJNA730355 (RNA-seq). Other data further supporting
727 this work are openly available in the GigaScience repository, GigaDB [80].

728

729 **Additional Files**

730 **Supplementary Figure S1.** The number distribution of OTUs with different existing
731 ratio of samples. (**a**) The number plot of OTUs with different existing ratio of samples.
732 (**b**) The relationship of existing ratio and the slope of the curve in **a**. The dotted line
733 indicated the threshold utilized to remove existing ratio distribution trend due to
734 fluctuate greatly.

735 **Supplementary Figure S2.** Alpha diversity, relative abundance, and Spearman's *r*
736 values of specific microbiota among the six sites. (**a-e**) Alpha diversity comparison
737 based on Good's coverage, observed OTU , ACE, Chao1, and Simpson indices, using
738 Wilcoxon rank-sum test to determine significant differences. (**f**) Rarefaction curves of
739 observed OTU. (**g**) Alpha diversity values of the six sites. Values are represented as
740 median ± SD. (**h**) *P* values of Wilcoxon rank-sum test of each comparison for six alpha
741 diversity indices. (**i**) PCoA of the 768 samples based on unweighted UniFrac distances.
742 (**j**) Beta diversity comparison based on the weighted UniFrac distances among the six
743 sites. The values are filled with weighted UniFrac distances (mean ± SD) in the
744 corresponding comparisons. All comparisons were significantly different using
745 Wilcoxon rank-sum test ($P < 0.05$). (**k**) Relative abundance of the top ten dominant
746 microbial phyla in the six sites. (**l**) Only microbial genera that were present in at least
747 461 samples (60% of the total) were plotted. Each row represents a microorganism.
748 Among 2,475 Spearman's *r* values, only 362 (14.62%) were significantly correlated (P
749 < 0.05).

750 **Supplementary Figure S3.** Comparison of the functional capacities of the reproductive
751 and digestive microbial communities among the six sites. (**a**) Overlap of the top 50
752 predictions among the six sites. (**b**) Heatmap showing the 36 overlapped predictions
753 with different abundances among the six sites. The heatmap is color-coded based on
754 row *Z* scores. (**c**) Map showing 65 site-associated bacterial taxa identified by LEfSe
755 (LDA score > 4) in the test trial.
756 **Supplementary Figure S4.** Significantly heritable microorganisms. The number of
757 significantly heritable microorganism OTUs, genera, and species ($P < 0.05$) grouped by
758 sampling phyla (**a, c, e**) and site (**b, d, f**).
759 **Supplementary Figure S5.** Significantly EN300-associated microorganisms. The
760 number of microorganisms significantly associated with EN300 detected at OTU, genus,
761 and species ($P < 0.05$) levels grouped by sampling phyla (**a, c, e**) and site (**b, d, f**).
762 **Supplementary Figure S6.** EN300-associated microorganisms. (**a**) The number of
763 microbial genera (left) and OTUs (right) associated with EN300 at $P < 0.05$ of three
764 test methods of six sites and their overlap. (**b-c**) Pearson's and Spearman's *r* values
765 between EN300 and EN300-associated 26 genera and 39 OTUs. Red and blue tiles
766 indicate positive and negative correlations, respectively. Significant *r* values are filled
767 in numerically ($P < 0.05$).
768 **Supplementary Figure S7.** Pearson correlations between EN300 and EN300-
769 associated microorganisms. (**a**) Pearson' s *r* values of candidate microbial species in the
770 six sites. (**b**) Pearson' s *r* values among microbial species in each site. CP: Crop, GZ:
771 Gizzard, SI: small intestine, UT: uterus, IS: Isthmus, VA: vagina. Red and blue tiles

772 indicate positive and negative correlations, respectively. The ratios on the right side of

773 each site represents the number of significant correlations. *$P < 0.05$.

774 **Supplementary Figure S8.** Differences in the relative abundance of five species

775 between the 20% of chickens with the highest and lowest egg production (EN300). (**a**)

776 EN300 values for the 20% of individuals with the highest and lowest egg production.

777 (**b**) EN300 values for the 20% of individuals with the highest and lowest abundances

778 of *Clostridium leptum*, *Bacteroides salanitronis*, *Firmicutes* bacterium ZOR0006, *B.*

779 *barnesiae,* and *B. fragilis* in the three reproductive tract sites. All comparisons were

780 significantly different, established at $P < 0.05$.

781 **Supplementary Figure S9.** Top 20 functional categories enriched by 1,051 genes

782 exhibited significant expression changes between groups with the high- and low- egg

783 production. The enrichment analysis was performed using the Metascape tool (See

784 Method). GO-BP: biological process (blue), GO-MF: molecular function (yellow) and

785 KEGG (red).

786 **Supplementary Figure S10.** Distribution of egg number at 300 days of age (EN300).

787 Compared to the 20% of hens with the highest EN300 values, the 20% of hens with the

788 lowest EN300 values exhibited a later start laying age, an earlier stop aging day and

789 irregular lay performance.

790 **Supplementary Figure S11.** Quality assessment of sequencing data. (**a**) Length

791 distribution of reads. (**b**) Quality score of each base. (**c**) Quality score distribution of

792 sequencing data. (**d**) Error rate distribution of reads.

793 **Supplementary Table S1.** Analysis of Bray-Curtis distance similarities.

794 **Supplementary Table S2.** Statistical test for the 65 functional capacities among the six
795 sites.
796 **Supplementary Table S3.** Summary of host whole genome sequencing.
797 **Supplementary Table S4.** Correlation between genetic relatedness matrix (GRM) and
798 each Bray-Curtis (BC) distance or microbial relationship matrix (MRM) by Mantel test.
799 **Supplementary Table S5.** Heritability (h^2) of the microbiota and cumulative
800 abundance of heritable microbiota.
801 **Supplementary Table S6.** Heritability (h^2) of reproductive traits from previous reports.
802 **Supplementary Table S7.** Summary of 16S rDNA sequencing.
803 **Supplementary Table S8.** Summary statistics of 16S rDNA sequencing.
804 **Supplementary Table S9.** Summary statistics of host whole genome sequencing.
805 **Supplementary Table S10.** Primer sequences for q-PCR.
806
807 **Abbreviations**

808 QTLs: quantitative trait loci; h^2: heritability; GWAS: genome-wide association study;
809 *PLAG1*: pleiomorphic adenoma gene 1; *LYN*: lck/yes-related novel tyrosine kinase gene;
810 EN300: egg number at 300 days of age; M: million; OTUs: operational taxonomic units;
811 BC: Bray-Curtis; PCoA: Principal coordinates analysis; FDR: false discovery rate;
812 ANOSIM: Analysis of Similarity; LDA: Linear discriminant analysis; LEfSe: Linear
813 discriminant analysis Effect Size; BWA: Burrows-Wheeler Alignment tool; GATK:
814 Genome Analysis Toolkit; SNP: single nucleotide polymorphism; MRM: microbial
815 relationship matrix; PCs: principal components; GRM: genetic relatedness matrix; m^2:

microbiability; LMM: linear mixed model; RNA-seq: RNA sequencing; TPM: transcripts per million; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; q-PCR: quantitative PCR; CPSA: capsular polysaccharide A.

Ethics Statement

All animal experiments were approved and reviewed by Animal Care and Use Committee Institutional of Sichuan Agricultural University (Approval No. DKY-2018102015).

Competing Interests

The authors declare that they have no competing interests.

Funding

This work was supported by the Sichuan Science and Technology Program (2019JDTD0009, 2020YFH0138 and 2021YFYZ0009), the Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China (161026).

Authors' Contributions

D.L., M.L. and Q.Z. designed the study; T.W., S.K.M., Y.S., W.Z., and S.T. wrote the manuscript; Z.X., M.H., X.Z., H.Y., X.F., and Q.N. collected the sample and extracted the DNA. M.Y., D.Y., Y.L., B.Z, and M.Z., did bioinformatics analyses; D.L. wrote methods; D.L., M.L. and Q.Z. supervised the work.

Acknowledgements

We thank the High-Performance Computing Platform of Sichuan Agricultural University and Ya'an Big Data Industrial Park for providing computing resources and support that have contributed to these research results.

References

1. Wolc A, Arango J, Settar P, et al. Analysis of egg production in layer chickens using a random regression model with genomic relationships. Poult Sci 2013;**92**(6):1486-1491.
2. Hu ZL, Park CA and Reecy JM. Building a livestock genetic and genomic information knowledgebase through integrative developments of Animal QTLdb and CorrDB. Nucleic Acids Res 2019;**47**(D1):D701-D710.
3. Yuan J, Sun C, Dou T, et al. Identification of Promising Mutants Associated with Egg Production Traits Revealed by Genome-Wide Association Study. PLoS One 2015;**10**(10):e0140615.
4. Zhang GX, Fan QC, Wang JY, et al. Genome-wide association study on reproductive traits in Jinghai Yellow Chicken. Anim Reprod Sci 2015;**163**:30-34.
5. Tongsiri S, Jeyaruban MG and Van Der Werf JH. Genetic parameters for egg production traits in purebred and hybrid chicken in a tropical environment. Br Poult Sci 2015;**56**(6):613-620.
6. Savegnago RP, Caetano SL, Ramos SB, et al. Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population. Poult Sci 2011;**90**(10):2174-2188.
7. Chen C, Song X, Wei W, et al. The microbiota continuum along the female reproductive tract

860 and its relation to uterine-related diseases. Nat Commun 2017;**8**(1):875.

861 8. Wen C, Li Q, Lan F, et al. Microbiota continuum along the chicken oviduct and its association
862 with host genetics and egg formation. Poult Sci 2021;**100**(7):101104.

863 9. Fettweis JM, Serrano MG, Brooks JP, et al. The vaginal microbiome and preterm birth. Nat Med
864 2019;**25**(6):1012-1021.

865 10. Serrano MG, Parikh HI, Brooks JP, et al. Racioethnic diversity in the dynamics of the vaginal
866 microbiome during pregnancy. Nat Med 2019;**25**(6):1001-1011.

867 11. Gotsch F, Romero R, Kusanovic JP, et al. The fetal inflammatory response syndrome. Clin
868 Obstet Gynecol 2007;**50**(3):652-683.

869 12. Lamont RF, Sobel JD, Akins RA, et al. The vaginal microbiome: new information about genital
870 tract flora using molecular based techniques. BJOG 2011;**118**(5):533-549.

871 13. Den Hartog G, De Vries-Reilingh G, Wehrmaker AM, et al. Intestinal immune maturation is
872 accompanied by temporal changes in the composition of the microbiota. Benef Microbes
873 2016;**7**(5):677-685.

874 14. Choi JH, Kim GB and Cha CJ. Spatial heterogeneity and stability of bacterial community in the
875 gastrointestinal tracts of broiler chickens. Poult Sci 2014;**93**(8):1942-1950.

876 15. Lee S, La T-M, Lee H-J, et al. Characterization of microbial communities in the chicken oviduct
877 and the origin of chicken embryo gut microbiota. Sci Rep 2019;**9**(1):6838.

878 16. Elokil AA, Magdy M, Melak S, et al. Faecal microbiome sequences in relation to the egg-laying
879 performance of hens using amplicon-based metagenomic association analysis. Animal
880 2020;**14**(4):706-715.

881 17. Carmody RN, Gerber GK, Luevano JM, Jr., et al. Diet dominates host genotype in shaping the

882 murine gut microbiota. Cell Host Microbe 2015;**17**(1):72-84.

883 18. Rothschild D, Weissbrod O, Barkan E, et al. Environment dominates over host genetics in

884 shaping human gut microbiota. Nature 2018;**555**(7695):210-215.

885 19. Liu H, Chen Z, Gao G, et al. Characterization and comparison of gut microbiomes in nine

886 species of parrots in captivity. Symbiosis 2019;**78**(3):241-250.

887 20. Ji J, Luo CL, Zou X, et al. Association of host genetics with intestinal microbial relevant to body

888 weight in a chicken F2 resource population. Poult Sci 2019;**98**(9):4084-4093.

889 21. Mignon-Grasteau S, Narcy A, Rideau N, et al. Impact of Selection for Digestive Efficiency on

890 Microbiota Composition in the Chicken. PLoS One 2015;**10**(8):e0135488.

891 22. Scepanovic P, Hodel F, Mondot S, et al. A comprehensive assessment of demographic,

892 environmental, and host genetic associations with gut microbiome diversity in healthy

893 individuals. Microbiome 2019;**7**(1):130.

894 23. Verstraelen H, Vilchez-Vargas R, Desimpel F, et al. Characterisation of the human uterine

895 microbiome in non-pregnant women through deep sequencing of the V1-2 region of the 16S

896 rRNA gene. PeerJ 2016;**4**:e1602.

897 24. Anahtar MN, Gootenberg DB, Mitchell CM, et al. Cervicovaginal Microbiota and Reproductive

898 Health: The Virtue of Simplicity. Cell Host Microbe 2018;**23**(2):159-168.

899 25. Thompson JL and Hinton M. Antibacterial activity of formic and propionic acids in the diet of

900 hens on Salmonellas in the crop. Br Poult Sci 1997;**38**(1):59-65.

901 26. Khan S, Moore RJ, Stanley D, et al. Gut microbiota of laying hens and its manipulation with

902 prebiotics and probiotics to enhance gut health and food safety. Appl Environ Microbiol

903 2020;**86**(13):e00600-00620.

904 27. Mishra AK, Kumar A, Swain DK, et al. Insights into pH regulatory mechanisms in mediating
905 spermatozoa functions. Vet World 2018;**11**(6):852-858.
906 28. Fiser PS and Macpherson JW. pH values in the oviduct of the hen during egg formation. Poult
907 Sci 1974;**53**(2):827-829.
908 29. Storek KM and Monack DM. Bacterial recognition pathways that lead to inflammasome
909 activation. Immunol Rev 2015;**265**(1):112-129.
910 30. Fang S, Zhang L, Lou Y, et al. Intracellular translocation and localization of Edwardsiella tarda
911 type III secretion system effector EseG in host cells. Microb Pathog 2016;**97**:166-171.
912 31. Yu MD and Lai EM. Warfare between Host Immunity and Bacterial Weapons. Cell Host
913 Microbe 2017;**21**(1):3-4.
914 32. Chen H, Yang D, Han F, et al. The bacterial T6SS effector EvpP prevents NLRP3 inflammasome
915 activation by inhibiting the Ca2+-dependent MAPK-Jnk pathway. Cell Host Microbe
916 2017;**21**(1):47-58.
917 33. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. Genome
918 Biol 2011;**12**(6):R60.
919 34. Moore RW, Park SY, Kubena LF, et al. Comparison of zinc acetate and propionate addition on
920 gastrointestinal tract fermentation and susceptibility of laying hens to Salmonella enteritidis
921 during forced molt. Poult Sci 2004;**83**(8):1276-1286.
922 35. Kalisperati P, Spanou E, Pateras IS, et al. Inflammation, DNA Damage, Helicobacter pylori and
923 Gastric Tumorigenesis. Front Genet 2017;**8**:20.
924 36. Franceschi F, Annalisa T, Teresa DR, et al. Role of Helicobacter pylori infection on nutrition
925 and metabolism. World J Gastroenterol 2014;**20**(36):12809-12817.

926 37. Ansari S and Yamaoka Y. Survival of Helicobacter pylori in gastric acidic territory. Helicobacter
927 2017;**22**(4):e12386.
928 38. Sonnenburg JL, Xu J, Leip DD, et al. Glycan foraging in vivo by an intestine-adapted bacterial
929 symbiont. Science 2005;**307**(5717):1955-1959.
930 39. Chousalkar KK and Roberts JR. Ultrastructural changes in the oviduct of the laying hen during
931 the laying cycle. Cell Tissue Res 2008;**332**(2):349-358.
932 40. Hrabia A, Lesniak-Walentyn A, Sechman A, et al. Chicken oviduct-the target tissue for growth
933 hormone action: effect on cell proliferation and apoptosis and on the gene expression of some
934 oviduct-specific proteins. Cell Tissue Res 2014;**357**(1):363-372.
935 41. Wang Y, Xu L, Sun X, et al. Characteristics of the fecal microbiota of high-and low-yield hens
936 and effects of fecal microbiota transplantation on egg production performance. Res Vet Sci
937 2020;**129**:164-173.
938 42. Choe D, Loh T, Foo H, et al. Egg production, faecal pH and microbial population, small intestine
939 morphology, and plasma and yolk cholesterol in laying hens given liquid metabolites produced
940 by Lactobacillus plantarum strains. Br Poult Sci 2012;**53**(1):106-115.
941 43. Lee W-J and Hase K. Gut microbiota–generated metabolites in animal health and disease. Nat
942 Chem Biol 2014;**10**(6):416-424.
943 44. Pandit RJ, Hinsu AT, Patel NV, et al. Microbial diversity and community composition of caecal
944 microbiota in commercial and indigenous Indian chickens determined using 16s rDNA amplicon
945 sequencing. Microbiome 2018;**6**(1):115.
946 45. Cui L, Zhang X, Cheng R, et al. Sex differences in growth performance are related to cecal
947 microbiota in chicken. Microb Pathog 2021;**150**:104710.

948 46. Xiang H, Gan J, Zeng D, et al. Specific Microbial Taxa and Functional Capacity Contribute to
949 Chicken Abdominal Fat Deposition. Front Microbiol 2021;**12**:643025.
950 47. Salanitro JP, Blake IG and Muirhead PA. Studies on the cecal microflora of commercial broiler
951 chickens. Appl Microbiol 1974;**28**(3):439-447.
952 48. Reeves AR, Wang GR and Salyers AA. Characterization of four outer membrane proteins that
953 play a role in utilization of starch by Bacteroides thetaiotaomicron. J Bacteriol 1997;**179**(3):643-
954 649.
955 49. Kollarcikova M, Faldynova M, Matiasovicova J, et al. Different Bacteroides species colonise
956 human and chicken intestinal tract. Microorganisms 2020;**8**(10):1483.
957 50. Nihira T, Suzuki E, Kitaoka M, et al. Discovery of beta-1,4-D-mannosyl-N-acetyl-D-
958 glucosamine phosphorylase involved in the metabolism of N-glycans. J Biol Chem
959 2013;**288**(38):27366-27374.
960 51. Mazmanian SK, Liu CH, Tzianabos AO, et al. An immunomodulatory molecule of symbiotic
961 bacteria directs maturation of the host immune system. Cell 2005;**122**(1):107-118.
962 52. Partty A, Kalliomaki M, Salminen S, et al. Infantile Colic Is Associated With Low-grade
963 Systemic Inflammation. J Pediatr Gastr Nutr 2017;**64**(5):691-695.
964 53. Hassan MS and Abdul-Careem MF. Avian Viruses that Impact Table Egg Production. Animals
965 2020;**10**(10):1747.
966 54. Reid G, Younes JA, Van der Mei HC, et al. Microbiota restoration: natural and supplemented
967 recovery of human microbial communities. Nat Rev Microbiol 2011;**9**(1):27-38.
968 55. Lee KN and Lee OY. Intestinal microbiota in pathophysiology and management of irritable
969 bowel syndrome. World J Gastroenterol 2014;**20**(27):8886-8897.

970 56. Caporaso JG, Lauber CL, Walters WA, et al. Global patterns of 16S rRNA diversity at a depth
971 of millions of sequences per sample. Proc Natl Acad Sci U S A 2011;**108**:4516-4522.
972 57. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
973 EMBnet J 2011;**17**(1):10-12.
974 58. Edgar RC, Haas BJ, Clemente JC, et al. UCHIME improves sensitivity and speed of chimera
975 detection. Bioinformatics 2011;**27**(16):2194-2200.
976 59. Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature
977 Methods 2013;**10**(10):996-998.
978 60. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved
979 data processing and web-based tools. Nucleic Acids Res 2013;**41**:D590-D596.
980 61. Kozich J, Westcott S, Baxter N, et al. Development of a dual-index sequencing strategy and
981 curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing
982 platform. Appl Environ Microbiol 2013;**79**(17):5112-5120.
983 62. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput.
984 Nucleic Acids Res 2004;**32**(5):1792-1797.
985 63. Bolyen E, Rideout J, Dillon M, et al. Reproducible, interactive, scalable and extensible
986 microbiome data science using QIIME 2. Nat Biotechnol 2019;**37**(8):852-857.
987 64. Langfelder P and Horvath S. WGCNA: an R package for weighted correlation network analysis.
988 BMC Bioinformatics 2008;**9**:559.
989 65. Douglas GM, Maffei VJ, Zaneveld J, et al. PICRUSt2: An improved and extensible approach
990 for metagenome inference. BioRxiv 2019:672295.
991 66. Li M, Tian S, Jin L, et al. Genomic analyses identify distinct patterns of selection in

992 domesticated pigs and Tibetan wild boars. Nat Genet 2013;**45**(12):1431-1438.

993 67. Li H and Durbin R. Fast and accurate long-read alignment with Burrows–Wheeler transform.
994 Bioinformatics 2010;**26**(5):589-595.

995 68. McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework
996 for analyzing next-generation DNA sequencing data. Genome Res 2010;**20**(9):1297-1303.

997 69. Li H, Ruan J and Durbin R. Mapping short DNA sequencing reads and calling variants using
998 mapping quality scores. Genome Res 2008;**18**(11):1851-1858.

999 70. Wang K, Li M and Hakonarson H. ANNOVAR: functional annotation of genetic variants from
1000 high-throughput sequencing data. Nucleic Acids Res 2010;**38**(16):e164.

1001 71. Camarinha-Silva A, Maushammer M, Wellmann R, et al. Host Genome Influence on Gut
1002 Microbial Composition and Microbial Prediction of Complex Traits in Pigs. Genetics
1003 2017;**206**(3):1637-1644.

1004 72. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and
1005 population-based linkage analyses. Am J Hum Genet 2007;**81**(3):559-575.

1006 73. Yang J, Bakshi A, Zhu Z, et al. Genetic variance estimation with imputed variants finds
1007 negligible missing heritability for human height and body mass index. Nat Genet
1008 2015;**47**(10):1114-1120.

1009 74. Yang J, Lee SH, Goddard ME, et al. GCTA: a tool for genome-wide complex trait analysis. Am
1010 J Hum Genet 2011;**88**(1):76-82.

1011 75. Zierer J, Jackson MA, Kastenmuller G, et al. The fecal metabolome as a functional readout of
1012 the gut microbiome. Nat Genet 2018;**50**(6):790-795.

1013 76. Zhou X and Stephens M. Genome-wide efficient mixed-model analysis for association studies.

1014 Nat Genet 2012;**44**(7):821-824.

1015 77. Bray NL, Pimentel H, Melsted P, et al. Near-optimal probabilistic RNA-seq quantification. Nat
1016 Biotechnol 2016;**34**(5):525-527.

1017 78. Robinson MD, McCarthy DJ and Smyth GK. edgeR: a Bioconductor package for differential
1018 expression analysis of digital gene expression data. Bioinformatics 2010;**26**(1):139-140.

1019 79. Zhou Y, Zhou B, Pache L, et al. Metascape provides a biologist-oriented resource for the analysis
1020 of systems-level datasets. Nat Commun 2019;**10**(1):1523.

1021 80. Su Y; Tian S; Li D; Zhu W; Wang T; Mishra SK; Wei R; Xu Z; He M; Zhao X; Yin H; Fan X;
1022 Zeng B; Yang M; Yang D; Ni Q; Li Y; Zhang M; Zhu Q; Li M: Supporting data for "Female
1023 reproductive tract microbiota influence egg production in layer chickens" GigaScience Database.
1024 2021. http://dx.doi.org/10.5524/100928.

Figure 1

Click here to access/download;Figure;Figure 1.tif

a

Unidentified Chloroplast 13.08%
Others 9.52%
Exiguobacterium 1.52%
Gallibacterium 2.14%
Lactobacillus 73.74%

Unidentified Erysipelotrichaceae 2.40%
Unidentified Chloroplast 3.95%
Streptococcus 2.17%
Stenotrophomonas 6.54%
Romboutsia 1.27%
Pseudomonas 1.12%
Lactobacillus 9.91%
Fusobacterium 1.94%
Exiguobacterium 4.08%
Escherichia shigella 1.22%
Bacteroides 4.86%
Others 60.54%

Crop
Isthmus

Unidentified Erysipelotrichaceae 1.04%
Others 23.78%
Exiguobacterium 2.25%
Lactobacillus 24.76%
Unidentified Chloroplast 48.17%

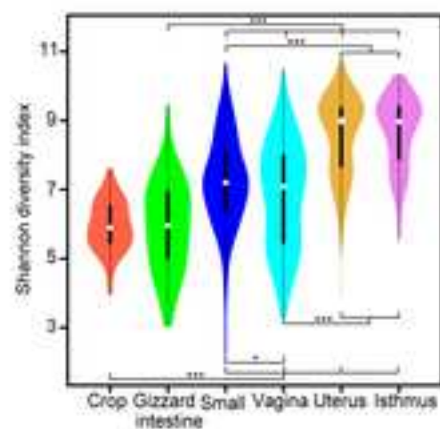
Unidentified Ruminococcaceae 1.05%
Unidentified Erysipelotrichaceae 2.21%
Unidentified Chloroplast 2.62%
Streptococcus 2.19%
Stenotrophomonas 7.20%
Sphingomonas 1.03%
Romboutsia 1.05%
Pseudomonas 1.78%
Lactobacillus 9.27%
Fusobacterium 1.07%
Faecalibacterium 1.39%
Exiguobacterium 4.68%
Escherichia shigella 1.08%
Bacteroides 4.33%
Alistipes 1.07%
Others 57.99%

Gizzard
Uterus
Small intestine
Vagina

Aneurinibacillus 2.97%
Bacteroides 2.95%
Exiguobacterium 6.09%
Helicobacter 5.08%
Lactobacillus 30.70%
Others 40.36%
Pseudomonas 1.40%
Romboutsia 1.11%
Stenotrophomonas 1.29%
Streptococcus 1.31%
Unidentified Chloroplast 4.20%
Unidentified Erysipelotrichaceae 2.54%

Unidentified Erysipelotrichaceae 1.83%
Unidentified Chloroplast 2.17%
Turicibacter 3.67%
Stenotrophomonas 1.01%
Romboutsia 11.46%
Lactobacillus 7.24%
Fusobacterium 11.10%
Exiguobacterium 2.79%
Clostridium sensu stricto 1 3.96%
Bacteroides 3.40%
Others 51.37%

b



c

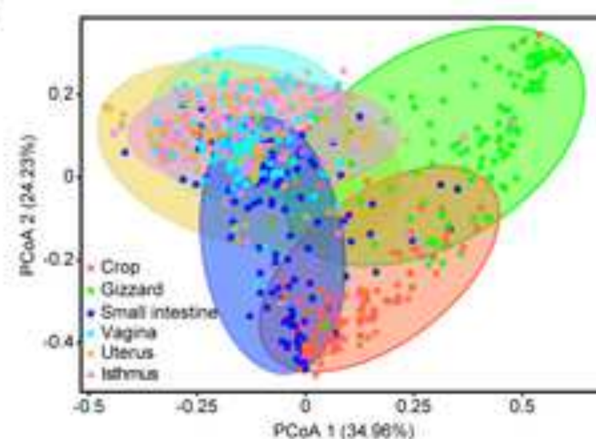


Figure 2

Click here to access/download;Figure;Figure 2.tif

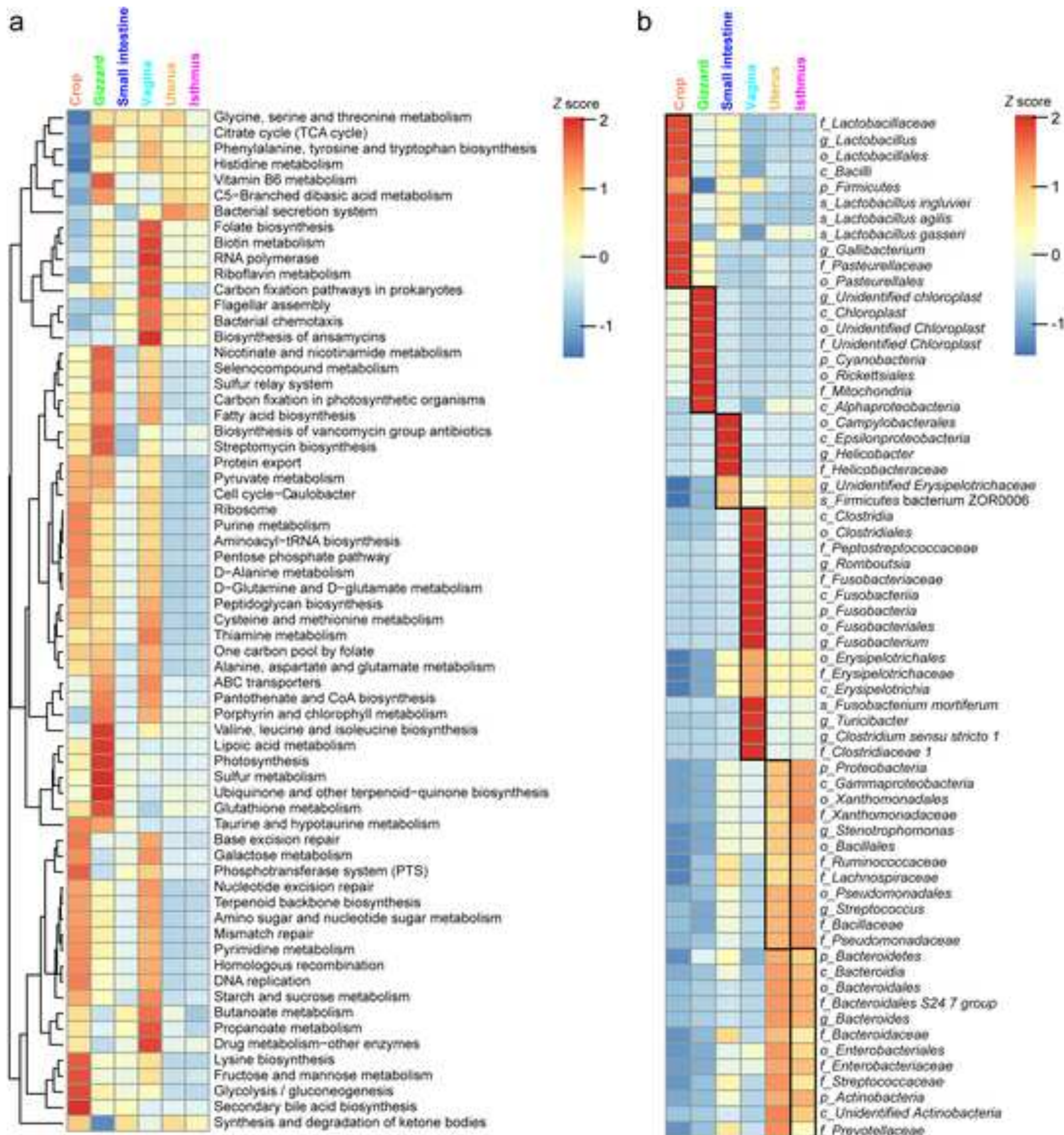
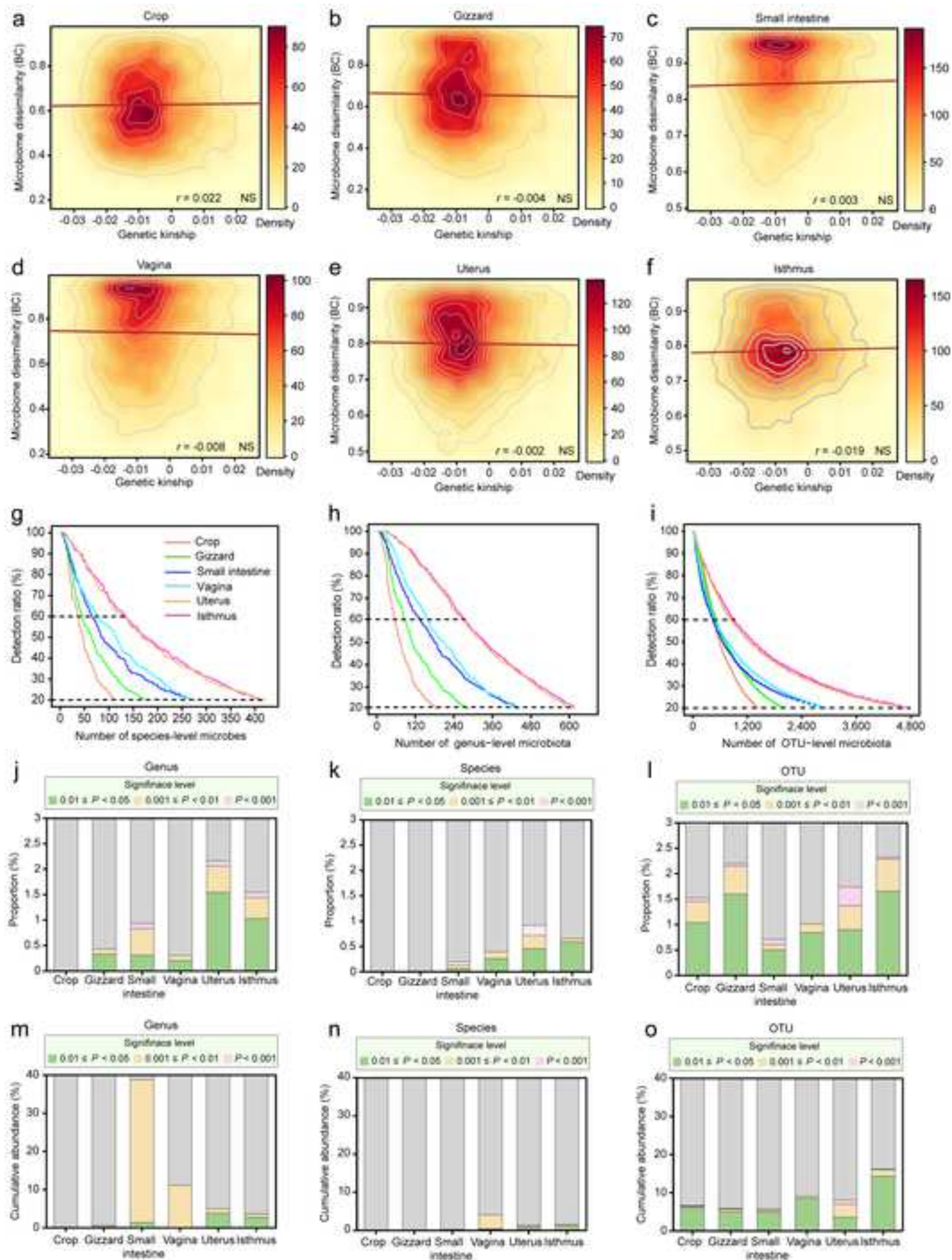


Figure 3 Click here to access/download;Figure;Figure 3.tif



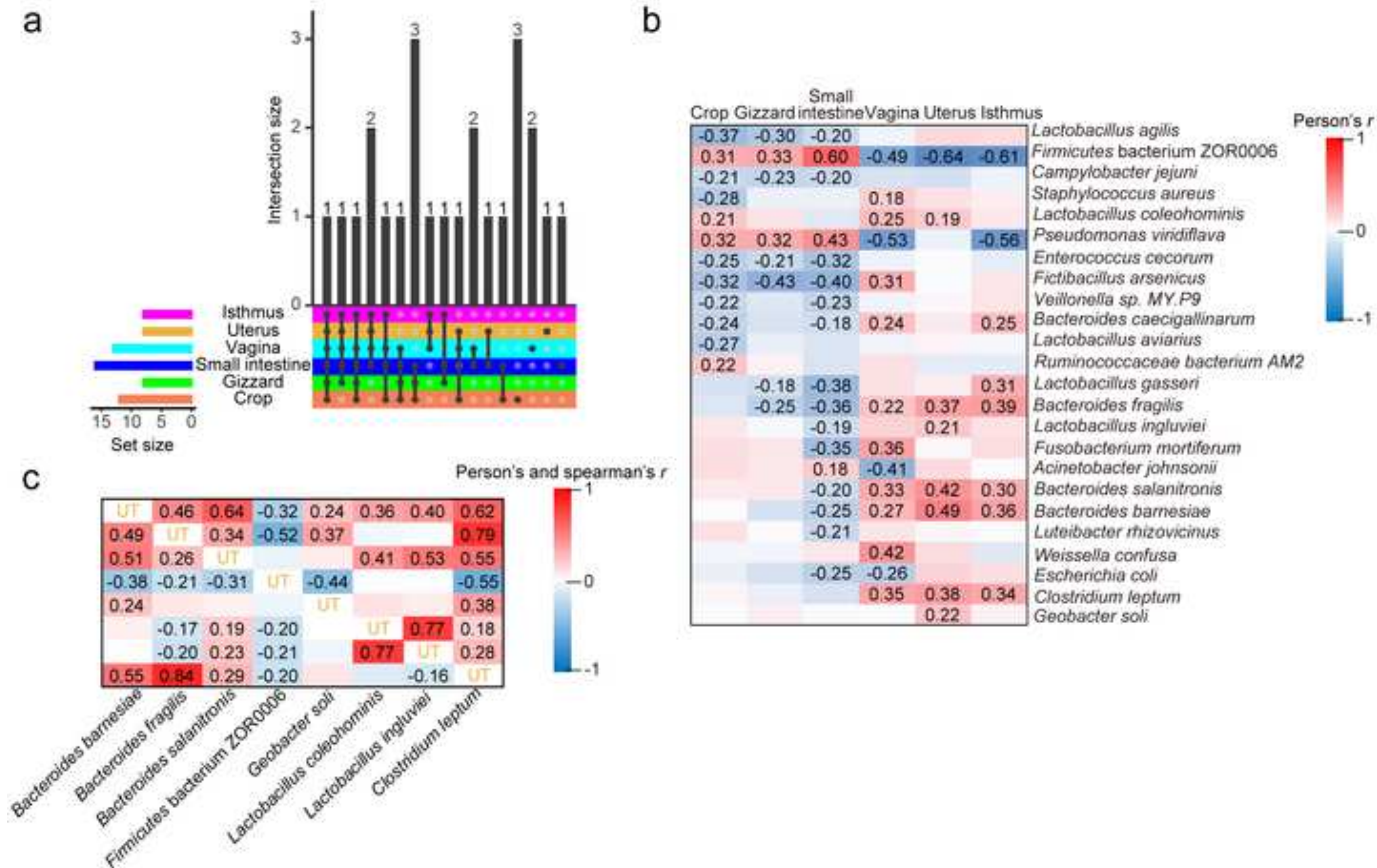

a
Intersection size
3
2
1
0
1 1 1 2 1 1 3 1 1 1 2 1 1 3 2 1 1
Isthmus
Uterus
Vagina
Small intestine
Gizzard
Crop
15 10 5 0
Set size
b
Crop Gizzard Small intestine Vagina Uterus Isthmus
Person's r
1
0
-1
Lactobacillus agilis -0.37 -0.30 -0.20
Firmicutes bacterium ZOR0006 0.31 0.33 0.60 -0.49 -0.64 -0.61
Campylobacter jejuni -0.21 -0.23 -0.20
Staphylococcus aureus -0.28 0.18
Lactobacillus coleohominis 0.21 0.25 0.19
Pseudomonas viridiflava 0.32 0.32 0.43 -0.53 -0.56
Enterococcus cecorum -0.25 -0.21 -0.32
Fictibacillus arsenicus -0.32 -0.43 -0.40 0.31
Veillonella sp. MY.P9 -0.22 -0.23
Bacteroides caecigallinarum -0.24 -0.18 0.24 0.25
Lactobacillus aviarius -0.27
Ruminococcaceae bacterium AM2 0.22
Lactobacillus gasseri -0.18 -0.38 0.31
Bacteroides fragilis -0.25 -0.36 0.22 0.37 0.39
Lactobacillus ingluviei -0.19 0.21
Fusobacterium mortiferum -0.35 0.36
Acinetobacter johnsonii 0.18 -0.41
Bacteroides salanitronis -0.20 0.33 0.42 0.30
Bacteroides barnesiae -0.25 0.27 0.49 0.36
Luteibacter rhizovicinus -0.21
Weissella confusa 0.42
Escherichia coli -0.25 -0.26
Clostridium leptum 0.35 0.38 0.34
Geobacter soli 0.22
c
Person's and spearman's r
1
0
-1
UT 0.46 0.64 -0.32 0.24 0.36 0.40 0.62
0.49 UT 0.34 -0.52 0.37 0.79
0.51 0.26 UT 0.41 0.53 0.55
-0.38 -0.21 -0.31 UT -0.44 -0.55
0.24 UT 0.38
-0.17 0.19 -0.20 UT 0.77 0.18
-0.20 0.23 -0.21 0.77 UT 0.28
0.55 0.84 0.29 -0.20 -0.16 UT
Bacteroides barnesiae
Bacteroides fragilis
Bacteroides salanitronis
Firmicutes bacterium ZOR0006
Geobacter soli
Lactobacillus coleohominis
Lactobacillus ingluviei
Clostridium leptum

Figure 5

Click here to access/download;Figure;Figure 5.tif

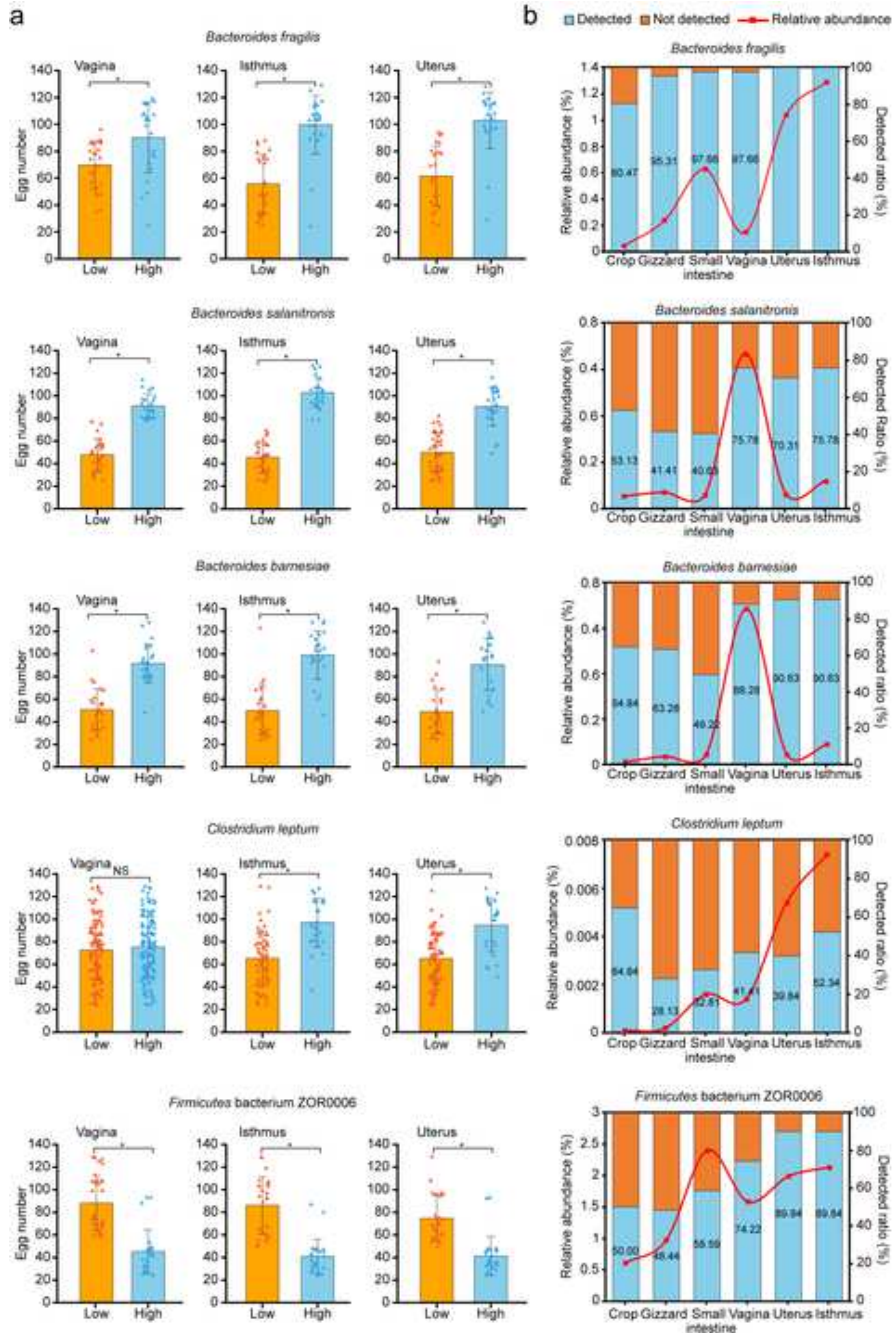
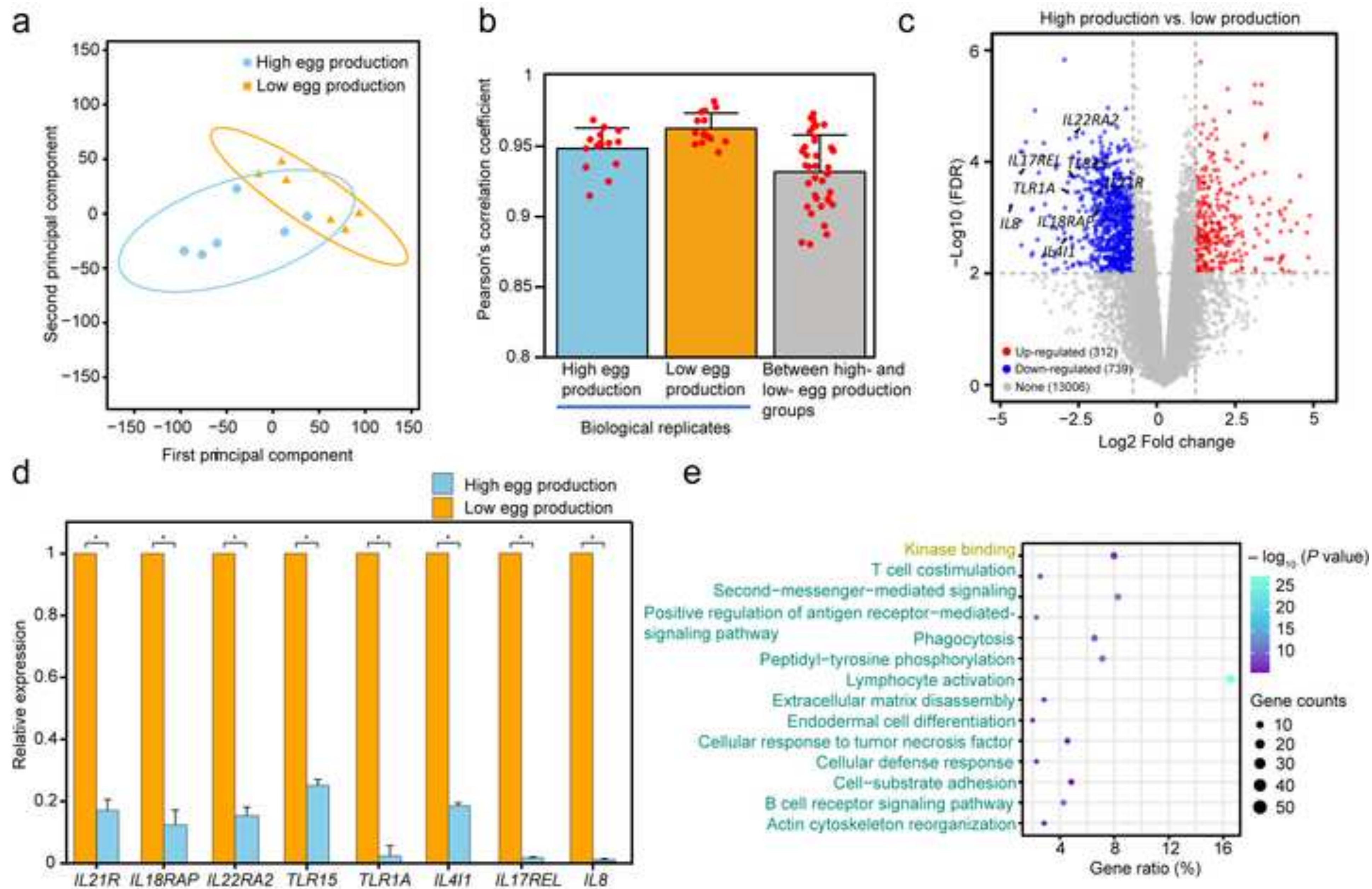


Figure 6

Click here to access/download;Figure;Figure 6.tif



Supplementary Figures

Click here to access/download
Supplementary Material
Supplementary Tables.xlsx

 Click here to access/download;Personal Cover;Respone letter.docx

Dear editor and reviewers,

Thank you very much for reviewing our manuscript "*Female reproductive tract microbiota influence egg production in layer chickens*" (ID: GIGA-D-21-00132) for possible publication in *GigaScience*. We sincerely thank the editor and two reviewers for their valuable feedback that we have used to improve the quality of our manuscript. According to the reviewers' comments, we have made the relevant modifications to our manuscript. All modifications are marked in red in the revised manuscript and a point-by-point response to the reviewers' comments follows. We hope these revisions meet your satisfaction and make our manuscript acceptable for publication in your journal.

We look forward to hearing a positive response from you.

Best regards,

Diyan Li, Professor, Ph.D.
Address: Institute of Animal Genetics and Breeding, College of Animal Science and Technology, Sichuan Agricultural University, Chengdu 611130, China.
E-mail: diyanli@sicau.edu.cn

Detailed responses to reviewers:

All comments provided by reviewers are in black, and our responses are in red. This following text is also included in the uploaded file "Response letter":

Reviewer #1: Dear authors, the work described in the manuscript is very extensive! I have not yet seen an analysis of host genetics, microbiota composition and host transcriptomics coupled with egg production data. Also while not perfect the manuscript is written well.

Response: Thank you for your positive comments.

Main problems:

1. It is clear you found an association between specific bacteria and egg production. You have also shown that some form of inflammation is associated with the changes in microbiota and egg production. However, you assume that the bacteria affect inflammation which affect egg production. But other options also exist. For example, it is possible that waning egg production, from other causes, changes the conditions in the oviduct so that the microbiota would change. Or inflammation, for example because of a pathogen, might modulate both egg production as well as microbiota composition. Please change the text so that the readers understand you are speculating and briefly mention the other options.

Response: As suggested, we have added discussion of the other options affecting egg production in the "Results and Discussion" sections (Lines 355-363).

"Multiple factors, especially host species, potential pathogens, and immune status of the host, all play a major role in the female reproductive organs adversely interfering with the egg industry in laying flocks [53]. Additionally, the digestive tract environment

of low-egg producing hens is fragile and susceptible to the influence of exogenous microorganisms [41]. Pathogenic infection, room temperature fluctuations, management systems, and other sudden changes to various factors can alter the composition of microbiota [54, 55]. These alterations may cause a significant degradation in production performance. Here, our results indicate that the reproductive tract microbiota play an important role in egg production."

2. While it is too late to change, it is a pity that you did not characterize the cecum community. Of all of the intestinal communities the cecum is by far the biggest and the most likely to affect the nutrition of the hen, thereby possibly affecting egg production.
Response: Thank you for your comments and understanding. We have included some information regarding this in the Discussion (Lines 288-291).
"The cecum has distinct microbial community profiles [44-46] that were not explored in this study. Microbial community analysis of the cecum microbiota in chickens exhibiting different egg production performances requires further investigation."

3. Please make sure to reference and discuss relevant literature. I quickly identified Elokil A. A. Animal 2020 which performed a limited but similar analysis. Please perform a literature search and make sure to reference and discuss relevant work.
Response: Thank you for your helpful suggestions. As suggested, we have added and discussed additionally relevant literature references (Lines 280-288).
"In accordance with previous findings, there is clear evidence of the role of fecal microbiomes in low and high egg-laying performance in hens; Elokil et al [16] demonstrated a significantly positive association between the microbial genus *Lactobacillus* and egg-laying performance ($P < 0.05$). Likewise, Wang et al [41] reported that *Lactobacillus* was also abundant in the feces of high-yield hens. The genus *Lactobacillus* produces growth promoters and exhibits antimicrobial activity against pathogenic microbes [42, 43] which may explain why the increasing abundance of *Lactobacillus* in the high-yield group is beneficial to egg-laying performance."

Minor problems:
Lines 48-49 - something is wrong with this sentence.
Response: Thank you for the pointing this out. We have revised the sentence in the manuscript (Lines 46-48).
"An abnormal vaginal microbiota may predispose individuals to increased microbial invasion of the amniotic cavity and fetal damage [11, 12]."

Lines 67-69 - it is not clear what you are trying to say here.
Response: As commented, the sentence was not clear and had limited relevance to the manuscript so we deleted the sentence.

Line 71 - 16S rDNA sequencing and not whole-genome sequencing.
Response: In this study, we conducted whole-genome sequencing on 128 laying hens and 16S rDNA sequencing on 768 samples from six sites. We have revised the sentence

in the manuscript (Lines 68-71).
"Here, we performed 16S rDNA sequencing on 768 samples from three reproductive (vagina, uterus, and isthmus) and three digestive (crop, gizzard, and small intestine) tract sites and whole-genome sequencing of 128 laying hens."

Line 96 - please be more specific regarding which part of the small intestine was sampled.
Response: The jejunum of the small intestine was sampled. As suggested, we provide a more specific explanation of which part of the small intestine was sampled in the revised manuscript (Lines 420-421).
"A 12-cm-long fixed mid-region of the small intestine (jejunum) was collected from each bird."

Lines 100-108 - description of host DNA extraction is not appropriate another the heading "microbial genomic DNA extraction".
Response: "Microbial genomic DNA extraction" has been changed to "Host and microbial genomic DNA extraction" in the revised manuscript (Line 426).

Line 109 and elsewhere - 16S rDNA and not 16S rRNA.
Response: "16S rRNA" was changed to "16S rDNA" throughout the whole manuscript.

Line 154 - why are you using OTUs with 97% identity and not 100% identical amplicon sequence variants (ASVs)?
Response: As an alternative to OTUs, ASVs have been proposed as a way to adapt the thresholds suggested by genome sequencing to microbial community analysis using 16S rDNA sequences. Meanwhile, the OTU approach is still one of the primarily used methods for analyzing 16S rDNA-seq data (Dvergedal et al., 2020; Pinna et al., 2021; Wen et al., 2021). Schloss PD recently evaluated the clustering risk among ASV and OTU methods, and reached a conclusion that ASVs and the use of overly narrow thresholds to identify OTUs increase the risk of splitting a single genome into separate clusters (Schloss, 2021).

Since there is no consensus for a biological definition of a bacterial species (Sanford et al., 2021), microbiologists should accept that how bacterial species are named is biased and that taxonomic rules are not applied in a consistent manner. This makes it impossible to fit a distance threshold that matches a set of species names (Konstantinidis and Tiedje, 2005). Furthermore, the 16S rDNA sequence does not evolve at the same rate across all bacterial lineages (Schloss and Westcott, 2011), which limits the biological interpretation of a common OTU definition. A distance-based definition of a taxonomic unit based on the 16S rDNA or full-genome sequences is operational and not necessarily grounded in biological theory (Yarza et al., 2014; Barco et al., 2020). One benefit of a distance-based OTU definition is the ability to mask residual sequencing errors. The sequences generated in microbiome studies can harbor PCR and sequencing errors. These errors would only exacerbate the inflated number of ASVs.

Although there are multiple reasons why proponents favor ASVs, we feel the significant risk of artificially splitting genomes into separate clusters is too high to warrant their use.

References

Barco, R., G. Garrity, J. Scott, J. Amend, K. Nealson, and D. Emerson. 2020. A genus definition for bacteria and archaea based on a standard genome relatedness index. mBio 11:e02475-02419.

Dvergedal, H., S. R. Sandve, I. Angell, G. Klemetsdal, and K. Rudi. 2020. Association of gut microbiota with metabolism in juvenile Atlantic salmon. Microbiome 8:160.

Konstantinidis, K. T., and J. M. Tiedje. 2005. Towards a genome-based taxonomy for prokaryotes. Journal of Bacteriology 187:6258-6264.

Pinna, N. K., R. M. Anjana, S. Saxena, A. Dutta, V. Gnanaprakash, G. Rameshkumar, S. Aswath, S. Raghavan, C. S. S. Rani, and V. Radha. 2021. Trans-ethnic gut microbial signatures of prediabetic subjects from India and Denmark. Genome Medicine 13:36.

Sanford, R. A., K. G. Lloyd, K. T. Konstantinidis, and F. E. Löffler. 2021. Microbial taxonomy run amok. Trends in Microbiology 29:394-404.

Schloss, P. D. 2021. Amplicon sequence variants artificially split bacterial genomes into separate clusters. mSphere Jul 21:e0019121. doi: 10.1128/mSphere.00191-21.

Schloss, P. D., and S. L. Westcott. 2011. Assessing and improving methods used in operational taxonomic unit-based approaches for 16S rRNA gene sequence analysis. Applied and Environmental Microbiology 77:3219-3226.

Wen, C., W. Yan, C. Mai, Z. Duan, J. Zheng, C. Sun, and N. Yang. 2021. Joint contributions of the gut microbiota and host genetics to feed efficiency in chickens. Microbiome 9:126.

Yarza, P., P. Yilmaz, E. Pruesse, F. O. Glöckner, W. Ludwig, K.-H. Schleifer, W. B. Whitman, J. Euzéby, R. Amann, and R. Rosselló-Móra. 2014. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. Nature Reviews Microbiology 12:635-645.

Line 157 - what are "singleton OTUs"? OTUs with only one read? OTUs found in only one sample?

Response: "singleton OTUs" represent OTUs found in only one sample (Lines 483-484).

"Singleton OTUs (OTUs found in only one sample) that did not match the reference database were removed."

Lines 159-161 - why were low abundance OTUs discarded? Does it affect your analysis? In general, it is better not to modify the data base extensively. Specifically, some phylogenetic groups are represented by multiple low abundant strains with a high total abundance whereas other phylogenetic groups are represented by a single highly abundant strain. By getting rid of such lower abundance OTUs you might be creating an artifact.

Response: Thank you for the helpful comments. Errors introduced by next-generation amplicon sequencing tend to induce spurious OTUs and spurious counts in OTU tables,

both of which are especially prevalent at low abundances. Despite the power of NGS and the progress achieved, generated data is imperfect, being subject to different types of errors, including those inherent to PCR amplification (substitutions and chimeric sequence formation) and sequencing-specific biases which are characteristic of each sequencing technology. Sequencing errors are predominantly caused by base substitutions, although base deletions, low-quality reads, variable read lengths and non-target amplification are also known error sources which may result in false species affiliation. Undetected chimeric sequences, caused by the hybridization of DNA fragments from different species also reduce the reliability of the 16S rDNA sequence-based phylogenetic composition of microbial communities. Together, these different errors generate a high number of lower-abundance sequences, which lead to overestimations of actual community diversity and the creation of many false taxa. Those spurious sequences are hard to filter out using current mainstream pipelines implementing error correction, denoising, and stringent filtration of chimeric sequences, contaminants and non-bacterial contents. Previous study have reported that although the overall abundance of these pseudo sequences was low, introducing them into analysis increased the total number of features to 10 times higher than expected and enlarged the divergence of the alpha and beta diversity analyses among the different methods (Wang et al., 2018). Lower-abundance and lower-quality sequences were observed to surround higher-abundance, biologically real sequences, forming error clouds (Bokulich et al., 2013; Edgar, 2013). Various researchers have developed different approaches to remove these pseudo sequences.

In our work, we described the numerical distribution of OTUs with different existing ratios in samples, after which the slope fluctuation is estimated. We considered there are a lot of false positives for these microbiotas when fluctuations are great. Therefore, we selected OTUs that existed in more than 20% of samples according to the slope distribution curve (**Supplementary Fig. S1a, b**). This filtering rule will improve the stability and accuracy of further analyses. Similar filtering rules have also been reported in previous analyses (Zierer et al., 2018). In addition, we have cited **Supplementary Fig. S1a, b** in the revised manuscript (Lines 488-489).

References

Bokulich, N. A., S. Subramanian, J. J. Faith, D. Gevers, J. I. Gordon, R. Knight, D. A. Mills, and J. G. Caporaso. 2013. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. Nature Methods 10:57-59.

Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature Methods 10:996-998.

Wang, J., Q. Zhang, G. Wu, C. Zhang, M. Zhang, and L. Zhao. 2018. Minimizing spurious features in 16S rRNA gene amplicon sequencing. PeerJ Preprints 6:e26872v1 https://doi.org/10.7287/peerj.preprints.26872v1

Zierer, J., M. A. Jackson, G. Kastenmüller, M. Mangino, T. Long, A. Telenti, R. P. Mohney, K. S. Small, J. T. Bell, and C. J. Steves. 2018. The fecal metabolome as a functional readout of the gut microbiome. Nature Genetics 50:790-795.

Lines 378-382 - this is an interesting idea that you did not establish well. Could you determine unifrac (or any other metric) distances of samples from different organs. i.e. compare all vagina samples to all small intestine samples and determine average Unifrac distances and stDev, and do so for each pair of organ sites. Perhaps you will gain a new figure showing how similar or different are the communities in each site.
Response: We are grateful for the suggestion. As suggested by the reviewer, we added a new figure (**Supplementary Fig. S2j**) showing the weighted UniFrac distances among the six sites.

Line 387 - cyanobacteria are not likely gut or reproductive tract inhabitants. This is more likely an artifact of identifying feed derived chloroplast DNA. You might want to mention this.
Response: Thanks for your helpful suggestions. The chickens used in our study are fed using corn-soybean-based diets, so the *Cyanobacteria* (likely an artifact of feed-derived chloroplast DNA) were misidentified as microbes. Additionally, we have revised the sentence in the revised manuscript (Lines 118-120).
"*Firmicutes*, *Proteobacteria*, and *Cyanobacteria* (likely an artifact of feed-derived chloroplast DNA) accounted for 71.45% - 97.86% of all OTUs."

Lines 398-402 - what is the connection between the "well-documented protective role of lactobacillus…" and differences in the levels of lactobacillus between gut and reproductive tract? Does the reproductive tract not deserve protection? Please stick to the data. Do not attach unsupported conclusions. Instead connect the levels of Lactobacillus and pH.
Response: As suggested, we have connected the levels of *Lactobacillus* and pH in the revised manuscript (Lines 134-139).
"*Lactobacillus* is thought to inhibit pathogenic bacteria by lowering the environmental pH through lactic acid and hydrogen peroxide production [24]. This genus was highly abundant in the digestive tracts which were characterized by low pH values which strongly limits the growth of most pathogens [25, 26]. In contrast, *Lactobacillus* was less abundant in the reproductive tract where an alkaline pH is needed to maintain sperm motility [27, 28]."

Lines 407-411 - if I understand correctly, you are implying a connection possibly caused because of the flow of material from different sites. You may want to note that.
Response: Thank you for the helpful comments. Yes, we want to note that, and we have added a sentence in our revised manuscript (Lines 146-147).
"These results imply there is a connection of microbiome communities possibly caused because by the flow of material from different sites."

Line 419-421 - you have reduced a whole field of study to just one "previous study". And in any case I am not sure what you wanted to write here.
Response: We have now included additional references in the revised manuscript (Line 158).

Line 442 - not chloroplast and mitochondria-like microbes but rather true chloroplasts and mitochondria which you misidentified as microbes.

Response: Thanks for your helpful suggestion. We have revised the sentence in the manuscript (Lines 179-180).

"*Unidentified Chloroplast* (2.94%) and mitochondria-like (2.37%) materials from plant consumption."

Line 447 - does Helicobactor maintain near natural pH and a microaerophilic environment? Is that its role in the gut? Please do not bring up bits and pieces of information if this is not really relevant to your results. What is important is that a possible pathogen was found in your chickens.

Response: Thanks for the comments. In order to infect the gastric mucosa, *Helicobacter pylori* has to survive in the gastric acidic pH, and *Helicobacter pylori* has well developed mechanisms to neutralize the effects of acidic pH (Ansari, S. and Y. Yamaoka. 2017. Survival of *Helicobacter pylori* in gastric acidic territory, helicobacter 22:e12386), that is its role in the gastric tract. In addition, we have reorganized the sentence in the revised manuscript (Lines 183-185).

"As a possible pathogen, *Helicobacter* specifically inhabits the small intestine in chickens, and may be involved in inflammation, metabolism, and neutralization of gastric acid [35-37]."

Lines 450-463 - this whole paragraph contains a lot of disjoined bits and pieces of information. Does it really matter that some bacteria were changed in immune suppressed honey bees? What are you trying to say? Consider taking this whole paragraph off.

Response: Thanks for the comment. This whole paragraph mainly displays the results of the site-associated bacterial taxa in the three reproductive tract sites identified by LEfSe and discussed association with the previous literature. Additionally, we have deleted the disjoined bits and pieces of information and reorganized the sentence in the paragraph in our revised manuscript (Lines 186-194).

"*Unidentified Erysipelotrichaceae* showed higher abundance in the three reproductive tract sites (1.83% - 2.40%). Bacteria associated with the isthmus and uterus both showed higher abundances than in the other sites. Several genera (typically, *Romboutsia*, *Fusobacterium*, and *Clostridium sensu stricto 1*) were dominant in the vagina (> 25% of the microbiota) but had lower abundances in the other sites (**Fig. 2b**). Among these, vaginal *Romboutsia* could be employed as a predictor for egg number in laying hens [8]. Six *Bacteroidetes* bacterial taxa were isthmus-associated; *Bacteroides* species live on host mucus-secreted polysaccharides and this flexible foraging behavior contributes to diversity and stability [38]."

Lines 464-466 - do your results support this speculation? If yes please expend, if not please remove.

Response: Thank you for the suggestion. Lines 464-466 have been removed from the

revised manuscript.

Line 492 - 0.21% of total abundance? Of total species?
Response: Yes, it is "of total species". We have reorganized the sentence and clarified in the manuscript (Line 223).
"accounting for 0.21% of the microbiota species tested in the small intestine".

Lines 560-568 - this is a good example of discussion! Here you brought relevant information and created a data based speculation.
Response: Thank you for your positive remarks about this discussion.

Line 646 - what do you mean in "microbiome succession"?
Response: "microbiome succession" has been changed to "microbial community" in the revised manuscript (Line 390).

Supplementary figure 2 contains two separate topics - functional sequencing data and bacterial community analysis. Please divide into separate figures.
Response: As suggested, we have divided Supplementary figure 2 into two separate figures (**Supplementary Fig. S2 and S11**).

Reviewer #2: This is a very good study to which I have only a few comments, mostly to microbiota part in which I am stronger than in chicken genetics. The only weaker part, but not weak, is the fact that the study was performed with only a single flock and was not repeated with different hens at all. I understand that it is impossible to repeat the study in another flock but for example you could have set up Cl. leptum PCR and check for its presence in reproductive tract in completely different birds.
Response: We appreciate the positive feedback and sincerely thank the reviewer for the thoughtful and supportive recommendations, which are of great help in improving the quality of this manuscript.

Please consider the following points.
line 94, indeed laparotomy?
Response: "After laparotomy" has been changed to "After the abdomen was opened" in the revised manuscript (Line 417).

l.111 and 114, I do not understand wha PCR kit you used. Please, reword.
Response: The PCR amplification was performed using Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, Ma, USA). We have revised the sentence in the revised manuscript (Lines 436-439).
"The V4 hypervariable region of the bacterial 16S rDNA was amplified using Phusion® High-Fidelity PCR Master Mix (New England Biolabs, Ipswich, Ma, USA) and the universal primers 515 F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806 R (5′-GGACTACHVGGGTWTCTAAT -3′) [56]."

l.253, 254, this sentence is somewhat compromised. Please check and reword.

Response: Thank you for the comments, and we have revised the sentence in the revised manuscript (Lines 578-579).

"In addition, we further removed the SNPs with adjacent distances ≤ 5."

l.342, rather verified, or confirmed, then detected. By the way, I am not sure whether you used data from the qPCR in the rest of the manuscript

Response: We have changed "detected" to "verified" in the revised manuscript (Line 664). We only used qPCR data in the "RNA sequencing (RNA-seq) analysis" section.

l.388, these are not Cyanobacteria, this is chloroplast DNA from plants in the feed, you may check very recent paper Volf et al. Eggshell and Feed Microbiota Do Not Represent Major Sources of Gut Anaerobes for Chickens in Commercial Production. Microorganisms 2021, 9, 1480.

Response: Thanks for the valuable suggestions, and we have read the above reference carefully. The chickens used in our study are fed using corn-soybean-based diets, so the *Cyanobacteria* (likely an artifact of feed derived-chloroplast DNA) was misidentified as microbes. Additionally, we have revised the sentence in the manuscript (Lines 121-122).

"*Cyanobacteria* (likely an artifact of feed-derived chloroplast DNA) was the dominant material in the gizzard (48.19% of the total abundance)."

l.462,463, I would suggest alternative explanation in this case and this is living on expense of host mucus secreted polysaccharides, Sonnenburg, J. L. et al. Glycan foraging in vivo by an intestine-adapted bacterial symbiont. Science 307, 1955-1959 (2005)

Response: As suggested, we have cited the reference and revised the sentence in the manuscript (Lines 192-194).

"*Bacteroides* species live on host mucus-secreted polysaccharides and this flexible foraging behavior contributes to diversity and stability [38]."

l.559, check that B. fragilis might be a consequence of intensive human care, Kollarcikova et al. Different Bacteroides Species Colonise Human and Chicken Intestinal Tract. Microorganisms 2020, 8, 1483.

Response: We are grateful for the suggestion. We have cited the useful and interesting reference in the revised manuscript (Lines 301-303).

"Interestingly, a recent study reported that the human-adapted *Bacteroides* species are likely introduced to chicken flocks by contact with humans and can temporarily persist in chickens [49]."

l.615, how can you know this? What if all of this the other way round, and I indeed believe that this is the other way, i.e. hens becomes of compromised performance, due to whatever factor, within but possibly also outside of those which you have monitored. This naturally results in decrease in egg lay but also in increased inflammatory response.

Locally changed conditions due to inflammatory signaling change, infiltrating heterophils and macrophages produce antimicrobial peptides and reactive oxygen species and strict anaerobes will be the firs bacterial species to decrease in a response of increasing oxygen concentration. What is cause and what is consequence. I do not know, I think that you do not know either, though you blame bacteria that these are responsible for the response.
Response: Thank you for the comments. We deleted the sentence that caused misunderstanding.

l.631, similar to previous comment, I do not think that there is any downregulation. In high egg producers, there is basal, background expression of inflammatory marker genes. And these are induced in the hens with compromised performance. Be also careful, whether this induction since this could also be a cause of infiltration of macrophage with their specific expression profile, and you then mistakenly conclude on induction when purifying mRNA from a total complex tissue.
Response: Thanks for your comments. "downregulated" has been changed to "different expressed" in the revised manuscript (Line 383).

l.656, the same as above, be careful what is cause and what is consequence. Increase in reactive oxygen species may affect the most strict anaerobes. When these are present, there is no inflammation. When these are eliminated by increase in oxygen concentration and all other inflammatory responses, this is explained that these bacteria are anti-inflammatory. These are not, these only dislike inflammation and oxygen.
Response: Thanks for your helpful suggestion. We have deleted the sentence in the revised manuscript.