Influence of the digestive/reproductive tract microbiota on chicken egg production beyond host genetics

Diyan Li

　Sichuan Agricultural University - Chengdu Campus　https://orcid.org/0000-0001-7490-3550

Shilin Tian

　Wuhan University

Wei Zhu

　Sichuan Agricultural University - Chengdu Campus

Yuan Su

　Sichuan Agricultural University - Chengdu Campus

Tao Wang

　Sichuan Agricultural University - Chengdu Campus

Tianyuan Ma

　Sichuan Agricultural University - Chengdu Campus

Shailendra Kumar Mishra

　Sichuan Agricultural University - Chengdu Campus

Lianzhe Shi

　Sichuan Agricultural University - Chengdu Campus

Ranlei Wei

　Sichuan University West China Hospital

Zhongxian Xu

　Sichuan Agricultural University - Chengdu Campus

Chunyou Ning

　Sichuan Agricultural University - Chengdu Campus

Dejing Zhang

　Novogene Bioinformatics institute

Xuxu Lin

　Sichuan Agricultural University - Chengdu Campus

Xiaoling Zhao

　Sichuan Agricultural University - Chengdu Campus

Yan Wang

　Sichuan Agricultural University - Chengdu Campus

Huadong Yin

Sichuan Agricultural University - Chengdu Campus

Yaodong Hu

Sichuan Agricultural University - Chengdu Campus

Xiaolan Fan

Sichuan Agricultural University - Chengdu Campus

Bo Zeng

Sichuan Agricultural University - Chengdu Campus

Mingyao Yang

Sichuan Agricultural University - Chengdu Campus

Deying Yang

Sichuan Agricultural University - Chengdu Campus

Qingyong Ni

Sichuan Agricultural University - Chengdu Campus

Yan Li

Sichuan Agricultural University - Chengdu Campus

Mingwang Zhang

Sichuan Agricultural University - Chengdu Campus

Yongfang Yao

Sichuan Agricultural University - Chengdu Campus

Huailiang Xu

Sichuan Agricultural University - Chengdu Campus

Mingzhou Li

Sichuan Agricultural University - Chengdu Campus

Qing Zhu

Sichuan Agricultural University - Chengdu Campus

Research

Keywords: Chicken digestive and reproductive microbiome, Composition, Host genetics, Egg production performance

---

Influence of the digestive/reproductive tract microbiota on chicken egg production beyond host genetics

Diyan Li[1†*], Shilin Tian[2,3†], Wei Zhu[1†], Yuan Su[1†], Tao Wang[1†], Tianyuan Ma[1†], Shailendra Kumar Mishra[1], Lianzhe Shi[1], Ranlei Wei[4], Zhongxian Xu[1], Chunyou Ning[1], Dejing Zhang[3], Xuxu Lin[1], Xiaoling Zhao[1], Yan Wang[1], Huadong Yin[1], Yaodong Hu[1], Xiaolan Fan[1], Bo Zeng[1], Mingyao Yang[1], Deying Yang[1], Qingyong Ni[1], Yan Li[1], Mingwang Zhang[1], Yongfang Yao[1], Huailiang Xu[1], Mingzhou Li & Qing Zhu[1*]

†Diyan Li, Shilin Tian, Wei Zhu, Yuan Su, Tao Wang and Tianyuan Ma contributed equally to this work.

*Correspondence: diyanli@sicau.edu.cn; zhuqing@sicau.edu.cn;

Abstract

Background: The microbiota of the digestive and reproductive systems has a prominent role in animal health and performance, but the extent of its contribution is difficult to determine. In chickens, the effect of host genetics on the reproductive and digestive tract microbiota is unclear, and the means by which digestive/reproductive microbiomes help improve egg production in chicken are unknown.

Results: To gain insight into this, we examined genomes from 128 chickens reared under identical conditions and described their digestive (crop, gizzard and small intestine) and reproductive tract (vagina, uterus and isthmus) microbiota. Although the diversity, composition and predicted function of the digestive and reproductive tract microbiota exhibited notable microbiota variation substantially between different parts, host genetics had limited effects on the reproductive and digestive tract microbial community. The digestive and reproductive tract microbiota had a significant effect on egg production (accounting for 52.31% - 98.86% of the variance), after correcting for host genetic effects; in particular, the uterus and isthmus microbiota accounted for an average of 93.59% and 98.86%, respectively, of variance in egg production. We further identified four reproductive tract microbial species which were related to immune system, *Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae* and *Clostridium leptum*, that were significantly positively correlated with egg production. Chickens with a lower abundance of these species had produced significantly fewer eggs at 300 days of age (37.13 vs. 113.75) than those with a higher abundance of these microorganisms. These taxa indicate potential roles play in promoting reproductive performance. Especially uterus and isthmus tract microbiota were major factors in regulating the chicken egg production.

Conclusions: Host genetics has limited effect on digestive/reproductive microbiome composition. The distinct site-associated chicken microbiome may be determined by the differences of their physical function. These findings may help design strategies for controlling and altering the digestive/reproductive tract microbiota in chickens to improve egg production.

Keywords: Chicken digestive and reproductive microbiome, Composition, Host genetics, Egg production performance

Background

The chicken (*Gallus gallus domesticus*) is a commonly raised livestock animal with important economic value, contributing to global food production for humans by providing meat and eggs, as these products have a very high nutritional value and are relatively inexpensive [1]. Improvement of egg laying performance, such as egg production up to 46 weeks of age, egg weight and age at first laying, is urgently needed in the poultry industry to meet the tremendous global demand [2]. With the modern revolution in advanced technologies, genome-wide association studies (GWASs) have been widely used to identify single nucleotide polymorphism (SNP) genotypes and candidate genes associated with quantitative traits in birds [3, 4]. To date, in egg production and egg quality traits of chickens, about 3,135 quantitative trait loci (QTLs) have been reported, among which 315 QTLs on 19 different chromosomes (1-11, 13, 17, 19-20, 22-24 autosomes and Z sex chromosome) were associated with egg number and are listed in the Animal QTLdb (https://www.animalgenome.org/cgi-bin/QTLdb/GG) [5]. There is plenty of GWAS evidence showing that promising mutants located within genes, such as *APOO, RGS1, PRDX4, and CLSPN*, are associated with reproductive traits [6-8].

Egg production is an important economic trait in the poultry industry [9], and it is a polygenic inheritance trait with low to moderate heritability values ranging from 0.05 to 0.44, depending on the period involved [10, 11]. Therefore, other effective approaches for modulating egg production in laying hens are urgently required. Among them, the role of the microbiota is a relatively new field of study that is gaining importance. Egg production may be regulated by the gut and reproductive tract microbiota. In humans, some studies have been conducted to explore the relationship between microbes and reproduction. For example, the dynamics of microbiome interactions with the host during pregnancy leading to preterm birth were investigated [12] and temporal changes in the vaginal microbiome associated with full-term pregnancies were identified [13]. Manipulation of the gut and reproductive tract microbiome can alter the outcome of hormone-related pathologies, including pregnancy and infertility [14-16]. An abnormal vaginal microbiota may predispose individuals to ascending colonization of the genital tract, microbial invasion of the amniotic cavity and fetal damage [17, 18].

The avian digestive and reproductive tracts house complex bacterial communities that are believed to play crucial roles in the overall health and egg production traits of birds [19-22]. Most of the studies have mainly focused on broilers to measure the composition of commensal and pathogenic bacteria in the chicken gut, likely due to the commercial importance of these birds in the meat industry [23, 24]. It has been identified that *Methanobrevibacter* and *Mucispirillum schaedleri* play crucial roles in chicken fat accumulation [25]. Typically, the chicken gut and reproductive tracts are colonized by bacterial phyla mainly Firmicutes, Bacteroidetes, Proteobacteria, Actinobacteria and Fusobacteria [26, 27], that are spatially organized within specific gut/reproductive compartments. In addition, *Lactobacillus* species were found to be as the keystone species residing in the chicken oviduct. The functions of these microorganisms in the reproductive system and their effects on egg production are worthy of study. More recently, a metagenomic association study showed that the gut microbial communities of chickens with high egg production performance differ significantly from those of chickens with low egg production performance [28].

Several synergistic factors, such as the environment and diet dominate over host genetics in determining gut microbiota composition [29-31]. A comparative study of gut microbial diversity among parrot species, indicated the potential role of host ancestry in shaping gut microbiomes [32]. In birds, genetics specialists applied GWASs to demonstrate the relationship between host genotype and

gut/reproductive microbiota population [33-38]. In a previous study, 14 QTLs were identified as exerting a strong effect on *C. leptum* and *Lactobacillus* and on related candidate genes involved in anti-inflammatory responses and in the motility of the digestive system [39]. On the other hand, recent studies have suggested that host genetics have a limited impact on the gut microbiota composition of humans [40] and chickens [25], and microorganisms play crucial roles in chicken fat accumulation [25], and microbes have important effects on host traits independent of host genetics. Therefore, we objectively realized that the colonization by some microorganisms is influenced by host genetics and that the microbiota of various sites can exert a great impact on different phenotypes of the host.

As suggested above, we speculated that the microbial complement of the digestive-reproductive system might be an important aspect of chicken egg production. To what extent host genetics affect the microbes of the digestive and reproductive system and the influence of these microbes on egg production are unknown. In this study, we performed whole-genome sequencing on 128 green-shell laying hens and 16S rRNA gene sequencing on 768 samples from digestive (crop, gizzard, small intestine) and reproductive (vagina, uterus, isthmus) sites (i) to characterize the digestive and reproductive tract microbiota of hens, (ii) to establish a correlation between host genetics and microbial diversity of the digestive/reproductive tract, and (iii) to identify the contribution of key microorganisms to egg production. The findings of this study will provide insights into the microbial communities in different digestive and reproductive tracts and their independent contributions to egg production.

Materials and Methods

Experimental Animals and Sample collection
The study was conducted on a flock of 128 green shell laying chickens reared on an experimental poultry farm at Sichuan Agricultural University in Ya'an, China. All birds were hatched on the same day and housed in individual pens. Feed intake was controlled daily according to standard farm husbandry practices and water was provided *ad libitum*. The number of eggs produced for the first 300 days of life was recorded daily for each bird. At the age of 300 days, 2 mL of whole blood was collected from the wing vein using venipuncture and stored at -20°C. Subsequently, each bird was culled by cervical dislocation followed by decapitation. After laparotomy, one centimeter of fresh tissues were collected from different reproductive tract (vagina, uterus and isthmus) and digestive tract sites (crop, gizzard and small intestine), as shown in Fig. 1. All 6 different tissue samples from each of 128 animals were snap-frozen in liquid nitrogen, transported to the laboratory and stored at -80°C until further processing.

DNA extraction
Total microbial genomic DNA (gDNA) was extracted from ~200 mg of the digestive and reproductive tissues of each individual chicken using a commercially available TIANamp Stool DNA Kit (Tiangen Biotech, Beijing, China) according to the manufacturer's instructions, with minor modifications. Host DNA was isolated from each blood sample using a TIANamp Genomic DNA Kit (Tiangen Biotech, Beijing, China) following the manufacturer's recommendation. The extracted DNA was quantified using a NanoDrop2000 spectrophotometer (Thermo Fisher Scientific, Chengdu), and the DNA integrity was determined by 1% agarose gel electrophoresis.

16S rRNA amplicon, sequencing and quality filtering
The V4 hypervariable regions of the bacterial 16S rRNA gene were amplified using a HOTSTAR Taq

Plus Master Mix Kit (Qiagen, Shanghai, China) and the universal primers 515 F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806 R (5′-GGACTACHVGGGTWTCTAAT -3′) [41]. Reactions were carried out using 15 μL of Phusion High-Fidelity PCR Master Mix (New England Biolabs), 3 μL of the forward and reverse primers, approximately 10 μL of template DNA, and 2 μL of ddH$_2$O. The PCR cycling conditions were as follows: an initial denaturation step at 98°C for 1 min and 30 cycles of 98°C for 10 sec, 50°C for 30 sec, and 72°C for 30 sec, followed by an extension step at 72°C for 5 min, using a Bio-Rad CFX96 thermal cycler (Bio-Rad, Hercules, USA). Amplicons were purified on agarose gel (1%) using a GeneJET Gel Extraction kit (Thermo Scientific, Schwerte, Germany). A DNA library was prepared using an Ion Plus Fragment Library Kit (Thermo Scientific, Schwerte, Germany) based on the manufacturer's instructions. Reads were barcoded per sample, combined for multiplexed sequencing with the Ion S5 TM XL platform (Thermo Fisher Scientific) to generate 400 bp single-end reads, and sequenced by Novogene Bioinformatics Technology Co. Ltd.

The raw data obtained by sequencing might contain a certain proportion of interference data, which were first filtered to obtain clean data. First, the adapter sequences in 61,051,578 raw 16S rRNA reads were trimmed using Cutadapt (ver 1.9.1) [42] when the overlap length between the read and the adapter was shorter than 10. Then, we trimmed barcode sequences and filtered reads that were too long (> 260) or too short (< 220) 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 [43] to detect chimera sequences, and then, the chimera sequences were removed [44]. Finally, low-quality reads (> 50% bases with phred quality < 5) were also removed. Consequently, we generated 57,609,211 high-quality reads for further analysis (Additional file 1: Table S1). The quality of clean data was assessed by the length distribution of reads, quality distribution, and error rate distribution. More than 93.85% of the high-quality reads had lengths of 250 - 260 nt (Additional file 2: Figure S1a). Data with a quality score greater than 20 accounted for 88.14% of all the effective bases (Additional file 2: Figures S1b, c). The error rate of the sequencing reads was relatively high in the ending position (Additional file 2: Figure S1d).

OTU cluster and species annotation

The remaining high-quality sequences were used to generate OTUs (operational taxonomic units) by Uparse software (ver 7.0.1001) [45] with an identity cutoff of 97%. Singleton OTUs that did not match the reference database were removed. Clustering across all samples from the 128 chickens produced 46,480 OTUs after singleton removal; however, most of those OTUs were present at low abundance and were found in very few samples. We then discarded OTUs that were not found in at least 20% of the chickens in each sampling site, yielding 6,776 OTUs. For each OTU, the SSUrRNA library in the Silva database (https://www.arb-silva.de/) [46] was used to annotate taxonomic information (i.e., kingdom, phylum, class, order, family, genus and species) based on the Mothur algorithm. Subsequently, we determined the phylogenetic relationship of different OTUs and difference of the dominant species in different samples (groups) after multiple sequence alignment using the MUSCLE software (Version 3.8.31) [47]. Additionally, OTU abundance information was normalized using a standard sequence number corresponding to the sample with the fewest sequences.

Alpha Diversity

Alpha diversity is used to analyze the complexity of species diversity for a sample based on the normalized OTUs through 6 indices (i.e., observed OTUs, Chao1, Shannon, Simpson, ACE, Good's

coverage), using the package QIIME2 (Quantitative Insights into Microbial Ecology). Of these indices, two were selected to identify community richness (i.e., Chao1 and ACE); two were used to identify community diversity (i.e., Shannon and Simpson); and was used to characterize sequencing depth (i.e., Good's coverage). The differences in alpha diversity indices among the six tracts were calculated with the Wilcoxon rank-sum test using R software (Ver 2.15.3).

Beta Diversity

Beta diversity was used to evaluate differences in samples. Beta diversity in Bray-Curtis (BC), weighted and unweighted UniFrac distances was calculated by QIIME software. The Bray-Curtis ordination provides position values along an ordination axis and distances from the axis for samples of communities [48].

Cluster analysis was preceded by principal component analysis (PCA), which was applied to reduce the dimension of the original variables using the FactoMineR package and ggplot2 package in R software.

Principal coordinate analysis (PCoA) was performed to obtain principal coordinates and visualize complex, multidimensional data. A distance matrix of weighted or unweighted UniFrac distances among samples obtained previously was transformed to a new set of orthogonal axes, by which the maximum variation factor was demonstrated by the first principal coordinate, the second maximum variation factor was demonstrated by the second principal coordinate, and so on. PCoA was performed using the WGCNA package, stat packages and ggplot2 package in R software.

Unweighted Pair-group Method with Arithmetic Means (UPGMA) Clustering was performed as a type of hierarchical clustering method to interpret the distance matrix using average linkage and was conducted by QIIME software.

Prediction of the functional profiles of microbial communities

The functions of the microorganisms present in the microbial communities detected at each of the six anatomical sites (uterus, isthmus, vagina, gizzard, crop and small intestine) were predicted using PICRUSt2 [49]. We used the Wilcoxon rank-sum test to investigate the differences in pathways among the six tracts. These p-values were adjusted using the Benjamini-Hochberg method by the false discovery rate (FDR) with the p.adjust function in R.

Community differences analysis

The different sites were statistically compared using ANOSIM (also named permutational MANOVA) [50] with 10,000 permutations based on Bray-Curtis (BC) ordination to evaluate the reasonability of the division of groups.

Between-group variation analysis

The high-dimensional biomarkers were discovered by LEfSe (linear discriminant analysis (LDA) Effect Size) with the parameter 'LDA score > 4' [51], for identifying characteristics of abundance and related classes (e.g., genes, metabolites or taxa).

Whole-genome sequencing and quality filtering

After qualified host DNA samples were tested, a Covaris ultrasonic fragmentation instrument was used to randomly fragment the DNA, 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. The DNA fragments were then end-repaired, dA-tailed and ligated with a full-length adapter for Illumina sequencing, PCR amplification and purification. Next, the isolated DNA libraries were constructed with an insert size of 350 bp. Finally, a total of 1.78 Tb (terabases) whole genomes of 128 host chickens were sequenced with the 150 bp paried-end based on the Illumina Novaseq platform (Novogene Bioinformatics Technology Co. Ltd., China).

To avoid analysis noise caused by sequencing errors, low-quality paired reads (reads with $\geqslant$ 10% unidentified nucleotides [N]; > 10 nt aligned to the adaptor, allowing $\leqslant$ 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 in house script [52]. Consequently, we obtained 1.30 Tb (~10.15-fold per individual) of high-quality paired-end reads, including 95.13% and 88.98% nucleotides with phred quality $\geq$ Q20 (with an accuracy of 99.0%) and $\geq$Q30 (with an accuracy of 99.9%), respectively (Additional file 1: Table S2).

Read mapping, genomic variant calling and annotation

The remaining high-quality reads of each birds were aligned to the chicken reference genome (Gallus_gallus-6.0 Ensembl release 94, http://asia.ensembl.org/Gallus_gallus/Info/Index/) using the Burrows-Wheeler Alignment tool (BWA ver 0.7.15) [53] with the command 'mem -t 10 -k 32'. BAM alignment files were then generated using SAMtools (ver 0.1.19). Additionally, we improved the alignment performance through filtering the alignment reads with mismatches $\leqslant$ 5 and mapping quality = 0. After sorting by SAMtools, the sorted bam file was marked in duplicate using the command "MarkDuplicates" in the package picard (ver 1.119).

Subsequently, we performed gVCF calling according to the best practices using the Genome Analysis Toolkit (GATK, ver v3.7) [54] with the HaplotypeCaller-based method, and then population SNP calling by merging all gVCFs with the commands "CombineGVCFs".

To obtain high-credibility SNPs, we applied hard filter command 'VariantFiltration' to exclude potential false-positive variant calls as follows: (a) quality by depth >10.0; (b) mapping quality score > 40.0; (c) FS < 60.0; (d) MQRank-Sum > −12.5; (e) ReadPosRankSum > −8.0; In addition, the filter variants were further filtered when more than three SNPs clustered within a 10-bp window were removed [55]. Finally, we used vcftools (ver 0.1.15) [56] to obtain biallelic variants with the following parameters: sample call rate > 90%, SNP call rate > 95%, minor allele frequencies > 1% and Hardy-Weinberg equilibrium p-value < 10^{-5} [57]. Ultimately, a total of 10,819,101 high-credibility SNPs in 128 chickens were retained (Additional file 1: Table S3). We further annotated SNPs using the package ANNOVAR [58], the SNPs were classified into different genomic regions (i.e., exonic, intronic, splice sites, upstream and downstream around gene regions and intergenic).

Construction of microbial relationship and host genetic relatedness matrices

The OTUs identified at each anatomical site were normalized to a zero mean and unit variance. We then constructed a microbial relationship matrix (MRM) [36] using an R script based on the following equation:

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

where r_{tij} represents the tested microbial relationship in tract t between chickens i and j; a_{tio} and a_{tjo} are the abundance of OTU o in tract t in chickens i and j, respectively; $\overline{t_{to}}$ is the average relative abundance of OTU o in tract t in the population; σ_{to}^2 is the variance in the abundance of OTU o in tract t; and N_T is the total number of OTUs in tract t used for the computation of relatedness. The high-quality SNPs were further used to detect independent markers using PLINK [59], with the following parameters: 50 kb window size, 10 SNPs per step and 0.2 as a squared Pearson's correlation coefficient (r^2). All 10,809,968 SNPs were used to compute the principal components (PCs) and GRM [60] using GCTA (ver 1.91.1) [61]:

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

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

Heritability analysis

To estimate the effects of host genetics on the digestive/reproductive tract microbiota at different anatomical sites, we first computed the correlation between the GRM and BC distances at each site using both Pearson's and Spearman's correlation-based Mantel tests with 10,000 permutations. The correlation between the GRM and the MRM was also computed. To estimate the correlation between the GRM and the reproductive/digestive tract microbiota community, we computed the heritability at the OTU, genus and species levels. OTU abundance information was normalized using a standard sequence number corresponding to the sample with the least sequences. The microorganisms that were present in < 60% but ≥ 20% of the samples were dichotomized as present or absent [62], and the microorganisms that were detected in < 20% of the samples from each site were excluded from the analysis.

Genetic and microbial parameters of egg production

As the chickens examined in this study had no pedigree information, we computed the SNP-based heritability of the egg production phenotype (EN300) instead, using the following model: [60]

$$y = K_c + g + e \quad [A]$$

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

The fraction of the EN300 variance explained by the microbial variance was calculated as $m^2 = \frac{\sigma_m^2}{\sigma_p^2}$; (called 'microbiability' (m^2) in animals [36] and 'microbiome-association index' in humans [30]), where σ_m^2 and σ_p^2 are the phenotypic variance and digestive/reproductive tract microbial variance, respectively. To adjust for host genetic effects, all valid individuals and SNPs were used in a GWAS with a univariate linear mixed model (LMM), which was performed using GEMMA [63]. The LMM was calculated as follows:

$$y = K_c + m_s + e \quad [B]$$

where the model parameters are the same as those described in model [A], except m_s, which is the random effect of the digestive/reproductive tract microbiota in location s following the multinomial distribution $m_s \sim N(0, M\sigma_m^2)$), and M is the MRM. We then used the MRM in GCTA to calculate m^2. The genome-wide significance threshold was 10e-6. We then extracted these SNPs with significant effects on EN300 and calculated the PCs using PLINK. The first two PCs and the top five host genetic PCs were then used as covariates in model [B] to account for host genetics.

Identification of microbiota constituents related to egg production

EN300 between two groups (the 20% lowest- and 20% highest-ranked chickens regarding EN300) was then compared using the Wilcoxon rank-sum test. Microorganisms with $P < 0.05$ and FDR < 0.05 were retained. Furthermore, we calculated the Spearman's and Pearson's correlations between EN300 and the abundance of each microbiota constituent at the genera, OTU and species levels. The presence of a microorganism was considered to correlate with EN300 if $P < 0.05$, as determined using the psych package in R with the p-value adjusted using the BH method. The overlapping microorganisms obtained from the Wilcoxon rank-sum test and Spearman's and Pearson's correlations were considered to have a potential relationship with EN300. We subsequently characterized the amount of each EN300-associated microbes at each site.

Results

Diversity and composition of the reproductive and digestive tract microbiota

To describe the microbial landscape of the chicken digestive and reproductive tracts, we profiled the taxonomic abundance in 768 samples from the crop, gizzard, small intestine, vagina, uterus and isthmus of 128 laying hens by sequencing the V4 region of the 16S rRNA gene. A total of 57,609,211 high-quality reads (with an average of 75,011 reads per sample, range 43,405-89,141) were obtained (Additional file 1: Table S1). *De novo* clustering after singleton removal produced 46,480 operational taxonomic units (OTUs) at an identity cutoff of 97%. We then discarded OTUs that were not found in at least 20% of these samples, resulting in a total of 6,776 OTUs.

For alpha diversity analysis, the Good's coverage ranged from 96.3% to 99.6% with an average of 98.7%, indicating that the measurement depth met the requirements (Additional file 2: Figure S2a). Analysis of the other five indices (observed OTUs, Chao1, Shannon, Simpson and ACE indices) indicated that all pairwise comparisons showed significant differences ($P < 0.001$) except for two comparisons (uterus vs. isthmus was not significant for all indices; and small intestine vs. vagina was not significant for observed OTUs, Shannon and Simpson indices) (Additional file 2: Figures S2b-g); the microbiota alpha diversity in the uterus and isthmus was higher than that in other sites, indicating that these sites contained more microbial taxa than the vagina and the digestive tracts (Fig. 1a, Additional file 2: Figure S2b-e). We then used principal coordinate analysis (PCoA) to visualize the differences in taxa composition among the six diverse sites. The first principal component (PC1), explaining most of the variance (34.96%) in weighted UniFrac distance matrices among the samples, separated the digestive tract samples from reproductive tract samples, similar results was observed using unweighted UniFrac distance matrices (Additional file 2: Figure S2f). The crop and gizzard samples were separated from samples from other sites, and an analysis of similarity (ANOSIM) based on Bray-Curtis distances confirmed this dissimilarity ($R = 0.266-0.790$, $P < 0.001$) (Additional file 1: Table S4). In addition, the results from the small intestine and vagina samples overlapped. The uterus

and isthmus showed essentially similar microbiota, and ANOSIM supported this (R = -0.003, P = 0.694).

Fig. 1 a An alpha diversity comparison based on the Shannon diversity index using the Wilcoxon rank-sum test to determine significant differences. **b** PCoA of the 768 samples based on weighted UniFrac distances. The first dimension and second dimension are shown. The fraction of the variance explained was 34.96% for component 1 (P < 0.05, Tracy-Widom test) and 24.23% for component 2 (P < 0.05, Tracy-Widom test). **c** Relative abundance of the microbiota from the six sites at the genus level. Only genera with an abundance of >1.0% in any of the sites are described (Additional file 1: Table S5).

Although similar phyla dominated the microbiota at all six sites (Firmicutes, Proteobacteria and Cyanobacteria, accounting for 71.45% to 97.86% of all OTUs), there were some noticeable differences (Additional file 2: Figure S2g). Firmicutes was the most abundant phylum at all sites except for the gizzard, accounting for 43.60% to 78.93% of the total abundance, while Cyanobacteria represented the dominant phylum in the gizzard (48.19%). We also found that the uterus and isthmus had similar dominant phyla, including Firmicutes (44.87% and 43.60%), Proteobacteria (26.25% and 23.77%) and Bacteroidetes (17.13% and 19.52%) (Additional file 2: Figure S2g). Interestingly, the vagina had the highest abundance of Fusobacteria (11.51%) among these six sites.

At the genus level, the crop (73.74%) and small intestine (30.70%) were dominated by *Lactobacillus*, and the crop exhibited the lowest alpha diversity (Fig. 1a). *Gallibacterium* (2.14%) was found only in the crop. *Lactobacillus* (24.76%) was found abundant in the gizzard. In the reproductive tract samples, *Lactobacillus* (7.24%-9.27%) were also dominant, and bacteria such as *Exiguobacterium* (2.79-4.68%), *Stenotrophomonas* (1.01-6.54%) and *Bacteroides* (3.40-4.86%) constituted a notable fraction of the microbiota (Fig. 1c, Additional file 1: Table S5); the uterus and isthmus also exhibited similar dominant genera. The proportion of bacteria *unidentified_Erysipelotrichaceae, unidentified_Chloroplast, Lactobacillus* and *Bacteroides* had an abundance of >1.0% in the vagina further increased in the uterus and isthmus (Additional file 1: Table S5). Furthermore, for the majority of genus, the relative abundance at one site was not associated with that at other sites. Among these 2,475 Spearman correlation coefficient, 2,113 (85.37%) of them were not significantly correlated (Additional file 2: Figure S3). A limited number of genera belonging to the Proteobacteria and Firmicutes phyla showed significant (P < 0.001) and positive correlations between the crop and gizzard, the gizzard and small intestine or the three reproductive tract sites.

Predicted functions of the microbiomes and site-associated microorganism identification

To determine whether the microbiomes identified as described above have specific functions within their distinct microenvironmental niches. We analyzed the functional capacity of the reproductive and digestive tract microbiota using PICRUSt2. Our results revealed that the top 50 KEGG pathways enriched at each site were mostly the same (shared 36 pathways), and twelve (33.33%) of the 36 shared pathways were associated primarily with metabolism (Additional file 2: Figure S4). Seven pathways were identified that were only enriched at a single site: 3 and 3 pathways for the crop and gizzard respectively, 'Riboflavin metabolism' for the vagina. The abundance of the OTUs involved in these pathways differed significantly (adjusted p-values < 0.001) among the six sites (Additional file 1: Table S6). For example, the crop microbiota was significantly enriched in 'D-Alanine metabolism' and

'Nucleotide excision repair', the gizzard microbial community had important roles in 'Sulfur relay system' and 'Starch and sucrose metabolism', and the small intestinal microbial community had an important role in 'Valine, leucine and isoleucine biosynthesis', as indicated by the moderate row z-scores for each pathway. Moreover, 'Propanoate metabolism' and 'Bacterial chemotaxis' were overrepresented in the vagina. Meanwhile, 'Bacterial secretion system' was overrepresented only in the uterus and isthmus (Fig. 2a).

We then used LEfSe to explore site-associated bacterial features (Additional file 2: Figure S5), and the relative abundances of these features was visualized on a heat map (Fig. 2b). LEfSe analysis confirmed most of the observations described above. For example, the isthmus and uterus had similar abundances of similar microorganisms; *Lactobacillus* and *Gallibacterium* were classified as crop-associated bacteria, whereas *unidentified_Chloroplast* was a gizzard-associated bacterium. Of note, *Helicobacter* and *unidentified Erysipelotrichaceae*, which were associated with the small intestine, showed the highest abundance among the six sites (Fig. 2b). *Unidentified_Erysipelotrichaceae* also showed high abundance in three reproductive sites (1.83%-2.40%, Additional file 1: Table S5). *Romboutsia*, *Fusobacterium*, *Turicibacter* and *Clostridium_sensu_stricto_1* were vagina-associated bacteria. The bacteria associated with isthmus and uterus both showed higher abundance at those sites than in other sites.

Fig. 2 Comparison of the predicted functional capacities and site-associated taxa of the six microbial communities at diverse anatomical sites. **a** Heatmap showing the predicted KEGG pathways and their different abundances at different reproductive and digestive tract locations (see also Additional file 1: Table S6). The heatmap is color coded based on row *z-scores* of means of KEGG pathway abundances. **b** Heat map showing the 65 site-associated bacterial taxa identified by LEfSe (LDA > 4) and taxa whose abundances differed significantly among the six sites. p, phylum; c, class; o, order; f, family; g, genus.

Weak association between host genetics and digestive/reproductive tract microbial communities
Next we sought to determine the association between the microbiota and genetics in the same cohort of laying hens. To do this, we first generated 1.76 tera bases (Tb) of high quality sequences from 128 chickens (Additional file 1: Table S2), of which the quality of 96.39% of the bases was more than Q20. Mapping to the chicken reference genome resulted in an average of 99.81% reads covering an average of 95.59% of the nongap genome with a ~10.15-fold average depth. A total of 10.82 Mb of highly credible single nucleotide polymorphisms (SNPs) with a density of 10.29 SNPs/kb were acquired from the 128 individuals. We then calculated the correlation between host genetics (using genetic relatedness matrix) and microbial beta diversity (based on Bray-Curtis (BC) distance) at all six sites. The results showed that there was no significant association ($|r|< 0.033$, $P>0.05$, Fig. 3a-f and Additional file 1: Table S7) between host genetics and microbiome composition. However, the correlation coefficients for adjacent anatomical sites of BC distances were larger than those for nonadjacent sites, indicating that microbiomes in adjacent sites were similar; for example, a positive correlation was detected between the uterus and isthmus ($r=0.4260$, $P<0.0001$, Spearman correlation). No significant association ($r=0.0186$, Spearman correlation) was found between the crop and isthmus microbial communities (Additional file 1: Table S7). We then estimated the association between the genetic relatedness matrix (GRM) and microbial relationship matrix (MRM), and obtained similar results: both

Pearson's and Spearman's correlation coefficients of the GRM with any of the six MRMs were close to zero (Additional file 1: Table S7), suggesting that there is no connection between genetics and the microbiota.

Then, we regarded each microorganism as a quantitative trait (abundance as a phenotype) and computed the heritability of each microorganism at the species, genus and OTU levels. The microorganisms in more than 20% but less than 60% of the samples were thus analyzed qualitatively as dichotomous traits; microorganisms present in less than 20% of samples were excluded from the analysis (Fig. 3g, Additional file 2: Figure S6a). At the species and genus level, there was no statistical correlation ($P < 0.05$) between the presence of a SNP and the presence of a specific microbe (Additional file 2: Figure S6b) in the crop. Three species in the small intestine (accounting for 0.21% of the tested crop microbiota), 6 in the vagina (0.39%), 14 in the uterus (0.92%) and 10 in the isthmus (0.66%) exhibited significant ($P < 0.05$) SNP-based heritability (Additional file 1: Table S8). Most of these heritable bacteria detected at all three levels belonged to the Firmicutes phylum (Additional file 2: Figure S7a, c, e). Reproductive segments had more heritable bacterial phyla than digestive segments (Additional file 2: Figure S7b, d, f). The cumulative abundances of these heritable bacteria were only 0.22%, 4.14%, 1.46% and 1.61% ($P < 0.05$) in the small intestine, vagina, uterus and isthmus, respectively (Fig. 3h). Similar results were obtained at the genus and OTU levels (Additional file 2: Figure S6c). This result was consistent with the results described above showing that host genetics have a limited effect on shaping the reproductive and digestive tract microbial composition.

Fig. 3 Host genetics have a limited effect on the reproductive/digestive tract microbiota. **a-f** Scatter plot of genetic kinship of pairs of individuals (x axis) and their microbiome dissimilarity (y axis) among all pairs (n= 16,256) of individuals, demonstrating that genetic kinship and the microbiome similarity of the crop, gizzard, small intestine, vagina, uterus and isthmus are uncorrelated ($P > 0.05$; Mantel test). NS, not significant. **g** Number of species identified in all chickens at the six different anatomical sites. **h** Cumulative relative abundance of heritable microbial species in each segment.

Heritability and microbiability of egg production
As there was no evidence of an association between host genetics and the digestive/reproductive tract microbiota, we next asked whether host genetics and the microbiota at different reproductive/digestive sites affect egg production. To do this, we determined the mean number of eggs that each hen had laid by 300 days of age (EN300). The mean EN300 was 75.32 (range: 24-129) (Additional file 2: Figure S8), and the trait fit a normal distribution (Shapiro-Wilk normality test, *p>0.05*). A genetic relationship matrix (GRM) of pairs of samples was used as input for the restricted maximum likelihood analysis to estimate the heritability of EN300 explained by the whole genome SNPs. As a result, we found that the EN300 exhibited relatively low heritability ($h^2 = 0.28$) (Additional file 1: Table S9). The heritability of EN300 was within the scope of previous reports (Additional file 1: Table S10). This suggests that host genetics have a limited effect on egg production in chickens.

Then, we performed a genome-wide association study (GWAS) and identified 11 EN300 related SNPs (Fig. 4a and Additional file 1: Table S11), five and four of which were located on chromosomes 18 and 3, respectively. All 11 SNPs were noncoding mutations: 7 were intergenic, and 4 were intronic. One of the intronic SNPs was located in *GRM5* (glutamate receptor, metabotropic 5), and three were located in *ROCK2* (rho-associated, coiled-coil containing protein kinase 2). The lower left corner of

Fig. 4b shows the rationality of the model, while the upper right corner shows the SNPs identified by GWAS.

The fraction of the EN300 variance explained by the microbial variance was calculated as microbiability (m^2). We used all 11 SNPs as additional covariates to correct for host genetic factors in the m^2 estimation. The estimated EN300 m^2 for the reproductive sites (0.923 for the vagina, 0.936 for the uterus and 0.989 for the isthmus) was higher than that for the digestive sites (0.867 for the crop, 0.873 for the gizzard and 0.523 for the small intestine) (Additional file 1: Table S12). Taken together, these results suggest that egg production chicken is primarily determined by the digestive and reproductive tract microbiota independent of host genetics.

Fig. 4 Genome-wide association study analysis of single-nucleotide polymorphisms associated with egg number at 300 days of age (EN300). **a** The horizontal dashed gray line indicates genome-wide significance thresholds (1.0×10^{-6}). Each point represents a single-nucleotide polymorphism. The association analysis was conducted with the genome-wide efficient mixed-model association (GEMMA) software package. **b** Quantile-quantile plot of the genome-wide efficient mixed model for EN300. A 45-degree reference line is shown in red. The greater the departure from this reference line, the greater the likelihood that the two data sets came from populations with different distributions.

The microorganisms significantly associated with egg production

We then analyzed how well egg production can be inferred from the digestive/reproductive tract microbial communities in comparison with that from host genetics. A double-divergent Wilcoxon rank-sum test was conducted to identify egg production–related microbial species, genera and OTUs ($p<0.05$). Most of the microorganisms detected at the three different levels that were significantly associated with egg production belonged to the Firmicutes phylum (Additional file 2: Figure S9). Only microorganisms that exhibited a significant correlation of egg production with the relative abundance as determined by both the Pearson correlation coefficient (PCC) and Spearman correlation coefficient (SCC) were considered causal ($p<0.05$). In total, 39 OTUs, 26 genera and 24 species fulfilled these criteria (Fig. 5a, Additional file 2: Figure S10a). In addition, we identified candidate genera related to egg production by LEfSe using downloaded data for commercial laying hens with high egg production rates and local chicken breeds with low egg production rates at similar ages (~50 weeks). The results showed that the microorganisms associated with egg production overlapped with those identified in this study (all five genera related to egg production in feces were detected in our study, and 4 of 7 genera related to egg production in the cecum were candidate genera in our study) (Fig S10b). Most OTUS, genus and species present at the three digestive tract sites were negatively correlated (using PCC) with egg production (negative/positive: 19/6, 8/8 and 16/5 respectively), whereas most taxa at the three reproductive tract sites were positively correlated with egg production (positive/negative: 13/9, 11/10 and 13/4 respectively) (Fig. 5b, Additional file 2: Figure S10c, d). The microorganisms in the uterus were most strongly correlated with each other (Additional file 2: Figure S11), which implied a strong symbiotic/competitive relationship.

Fig. 5 EN300-associated microorganisms. **a** The number of microbial species at the six anatomical sites associated with EN300 based on a significance threshold of $P < 0.05$, and the overlap between sites. **b** Pearson correlations between EN300 and the 24 EN300-associated microbial species;

significant *R*-values are filled numerically (P < 0.05). Red and blue tiles indicate positive and negative correlations, respectively. **c** Pearson correlations among microbial species in the uterus (UT).

At the genus level, we found that *Lactobacillus*, *Bacteroides* and *Desulfovibrio* were positively correlated with EN300 at all three reproductive sites; meanwhile, *Pseudomonas*, *Exiguobacterium* and *unidentified_Erysipelotrichaceae* were negatively correlated with EN300 at all three reproductive sites and positively correlated at all three digestive sites (Additional file 2: Figure S10b). At the species level, *Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae* and *Clostridium leptum* showed a significant positive correlation with EN300 in the three reproductive segments, a weak negative correlation in the small intestine, and no correlation in the crop and gizzard. The first three species belong to the genus *Bacteroides*, which showed a significant positive correlation with egg production at all three reproductive sites (r = 0.403-0.479). *Firmicutes bacterium ZOR0006* had a significant negative correlation with EN300 at all three reproductive sites and a significant positive correlation at the three digestive sites. EN300 was significantly lower in the 20% of birds (mean = 37.13) with the lowest *Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae* and *Clostridium leptum* abundances compared with that in the 20% of birds (mean = 113.75) with the highest abundances of those species (P < 0.05) at reproductive sites (Additional file 2: Figure S12). The 20% of birds with the highest *Firmicutes bacterium ZOR0006* abundance exhibited a significantly higher EN300 than the 20% of birds with the lowest abundance of this microorganism (Fig. 6a) (P < 0.05) at reproductive sites. Moreover, the 20% of birds with the highest *Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae* and *Clostridium leptum* abundances showed a significantly lower EN300 than the 20% of birds with the lowest abundance in reproductive segments, with the exception of *Clostridium leptum* abundance in the vagina (Fig. 6a).

We then characterized the spatial distribution of these five EN300-associated microorganisms (*Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae*, *Clostridium leptum* and *Firmicutes bacterium ZOR0006*) at multiple segments. *Bacteroides fragilis* was detected in almost all samples and accounted for 0.05% to 1.29% of the total abundance (Fig. 6b). *Bacteroides salanitronis* and *Bacteroides barnesiae* were detected at similar rates at the six sites and in most reproductive site samples; both accounted for the highest abundance in the vagina. *Firmicutes bacterium ZOR0006* was also detected in most reproductive segment samples (74.22%-89.8%) and in half of the digestive site samples (48.44%-58.59%), accounting for 0.61% to 2.40% of the total abundance. The detection rates (28.13%-64.84%) and relative abundance of *Clostridium leptum* were much lower than those of the other microorganisms at all six sites, but accounted for the highest abundance in the isthmus and uterus (Fig. 6b).

Fig. 6 Effect of microbial species associated with egg number at 300 days of age (EN300) on egg production. **a** EN300 for the 20% of birds with the highest and lowest abundances of *Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae*, *Clostridium leptum* and *Firmicutes bacterium ZOR0006* at the three reproductive sites. The box plots show the median and the 25% and 75% quantiles. The cross and horizontal line indicate the mean and median values in the corresponding group, respectively. Significance levels were calculated using a permutation test with 10,000 replicates. Significance was established at P < 0.05. **b** Relative abundance and detection rate of 5 species

(*Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae*, *Clostridium leptum* and *Firmicutes bacterium ZOR0006*) in the 6 segments. The values shown in the blue bars indicate the detection frequency of each species at each site.

Discussion

Alpha diversity

In order to characterize the digestive and reproductive tracts microbiota of hens, we demonstrated the spatial heterogeneity of the microbiota throughout contiguous segments of the digestive and reproductive tracts in chickens (including the crop, gizzard, small intestine, vagina, uterus and isthmus). Compared to the digestive system, the reproductive system, especially the upper reproductive tract (uterus and isthmus), exhibited higher alpha diversity (including observed OUT, ACE, Chao1, Simpson and Shannon indices), the exception was small intestine vs vagina, they showed similar alpha diversity. Similar to the case in humans, where the vaginal and upper reproductive tract microbiota are significantly different [64, 65], we also showed that the mature chicken reproductive tract had significant differences between the upper (isthmus and uterus) and lower (vagina) reproductive tract (vagina vs isthmus, $R = 0.4725$, $P = 0.001$; vagina vs uterus, $R = 0.4959$, $P = 0.001$)), whereas there were no differences between the isthmus and uterus. A recent study also showed no significant differences in the abundance of four microbes composition among different sections of the oviduct (infundibulum, magnum, isthmus, and uterus) [66]. Nonetheless, chicken digestive tract microbiome diversity was comparable with those of the reproductive tract, suggesting that the different function of chicken digestive and reproductive tract microbiome.

The chicken digestive and reproductive tract microbiome composition

This study provides a comprehensive view of the succession of the chicken digestive and reproductive tract microbiome. Such a study design allowed us to address the logical question regarding the chicken microbiome: How does the chicken microbiome change across different digestive and reproductive tracts?

Lactobacillus and *Exiguobacterium* were found at all sites due to their broad adaptability or beneficial functions. *Lactobacillus* are thought to play key protective roles by lowering the environmental pH through lactic acid production [67], so the high abundance of this genus in the digestive tracts corresponds to the low pH value of this organ system [68]. *Lactobacillus-* or other lactic acid bacteria–dominant communities were detected less abundant in the uterus and isthmus (~14.03%) and the relative abundance was lower than that of the mature human vaginal microbiota (~50%) [69]. Consistent with a recent study in chickens [66], this finding may be related to the alkaline pH value in the chicken reproductive tract that maintains sperm motility [70, 71], in contrast to the low pH in humans that prevents pathogen infection and growth [72-74]. *Exiguobacterium* is a genus of *bacilli* and a member of the low-GC phylum Firmicutes; it has been found in areas covering a wide range of temperatures, pH values, salinity and other conditions [75], and it also produces organic acids and hydrolytic enzymes [76], even having some potential function in the degradation of toxic substances [77]. Given that the reproductive tract and digestive tract share a common exit in the cloaca, we propose that microbial exchange occurs at this site, resulting in similar microbiotas at the distal end of both tracts. Thus, we found that the vagina as a reproductive organ that acquires microbe communities from the isthmus and uterus but also has diversity similar to that in the small intestine (Fig. 1b). A functional capacity analysis of the six sites microbial communities revealed that the top 50

pathways enriched in each segment were mostly the same (shared/total: 36/50) and primarily associated with metabolism. Specifically, 'bacterial secretion system' and 'bacterial chemotaxis' are over represented in the reproductive tract (vagina, uterus and isthmus). Previous study found that successful bacterial pathogens have evolved versatile protein secretion systems that they use to promote their survival and fitness in response to different environmental challenges, and also to modulate host immunity [78].

Our findings confirm that the chicken digestive and reproductive tract microbiota is primarily determined by the physiological function of each compartment within these systems.

Site-associated microorganisms in chicken digestive and reproductive tract

This study also enabled us to address some other important biological questions regarding the chicken digestive and reproductive tract microbiome, such as: What are the site-associated microorganisms in diverse sites? Site-associated bacterial features in this study were identified by using LEfSe, a total of 65 site-associated bacterial taxa were identified among six sites.

Six Lactobacillaceae bacterial taxa were detected as crop-associated. In the chicken digestive system, the crop acts as a reservoir for the storage of food prior to its digestion, where food mixes with many beneficial *Lactobacillus* bacteria (73.84% at the genus level) that produce lactic acid before moving on to the proventriculus [79, 80]. Next, a strong muscle known as the gizzard functions similar to teeth for chickens by grinding any remaining large food particles with the assistance of grit (upon kneading), releasing abundant *unidentified_Chloroplast* (2.94%) and mitochondria-like (2.37%) microbes from plant consumption. The small intestine had the most abundant microbes in the three digestive sites, which is the main area where further digestion occurs and fermentation begins. The overrepresented Paenibacillaceae species were found with an optimum growth at pH 6.0–7.0 [81]. *Helicobacter* specifically inhabits the chicken small intestine to maintain near neutral pH and a microaerophilic environment and may be involved in inflammation, metabolism and neutralization of gastric acid [82-84]; whether it is beneficial or harmful remains unknown.

In this study, we observed that some representative genera in the vagina, such as *Romboutsia*, *Fusobacterium* and *Clostridium_sensu_stricto_1*, accounted for more than 25% of the vaginal microbiota, which were in lower abundance in other sites, although the role of most of these dominant bacteria in vagina remains unknown. Previous study found that genus *Romboutsia* was negatively associated with the body weight, fasting serum glucose and insulin. *Fusobacterium* has been reckoned as a proinflammatory organism [85]. It has been observed that the genus *Clostridium* is capable of utilizing a broad range of organic substrates for efficient hydrogen production [86] and thus adjusting pH value of host. Gammaproteobacteria and Betaproteobacteria were detected as uterus-associated. Significant changes in these bacterial taxa (Gammaproteobacteria, Betaproteobacteria) were detected in immune-suppressed adults emerging from Varroa -infested colonies [87]. Six Bacteroidetes bacterial taxa were detected as isthmus-associated, *Bacteroides* species are thought to play a vital role in the breakdown of polysaccharides into simpler compounds that related to host immunity [88]. Acidic pH [68] and alkaline pH [70, 71] values were observed in the chicken digestive and reproductive sites respectively. Thus, we speculate that the reproductive microbiota plays a vital role in improving protective immunity, defending against invasion of the reproductive tract by harmful bacterial and adaptation to pH levels.

The predominant microbes identified at the different anatomical sites in this study largely correlated with the physiological functions carried out at those sites.

Host genetics has limited effect on chicken digestive/reproductive tract microbes

Recent studies have suggested that host genetics have an impact on the microbiota composition[25, 34, 89]. However, the extent to which host genetics shape the digestive/reproductive tract microbiota remains debatable. In order to explore the effect of host genetics on chicken microbes. We first performed a GWAS to obtain EN300-associated SNPs and introduced these SNPs as additional covariates in the model [B] to adjust the effects of host genetics. Our estimation of SNP-based heritability for EN300 was moderate (0.28), and the value was comparable to that obtained in previous estimates [10, 11]. We then identified SNPs in two genes (*GRM5* and *ROCK2*) that appear to be strongly associated with egg production. *GRM5* was found to be potentially involved in dermatological diseases/conditions in chickens [90], but the function of *ROCK2* is unknown. Further studies need to be conducted to explore the role of these two genes in regulating chicken egg production. Subsequently, after adjust the effects of host genetics, we found that the chicken digestive/reproductive tract microbiota is not significantly associated with host genetic factors, which is consistent with recent studies in chickens [25] and humans [30]. In chicken, the gut microbiota is largely independent of host genetics in regulating fat deposition [25].

Reproductive tract microbiota has a dominant effect on chicken egg production

'Microbiability' (m^2) can be used to evaluate the effect of microbial abundance on host phenotypes and represents a formal measure of predictability: larger m^2 values indicate that the microbiome is more informative with respect to a phenotype of interest [30]. This concept was proposed by Difford et al. for cattle [91] and has been used to predict several complex traits in pigs [36] and chickens [25]. A recent study in humans referred to this fraction as the 'microbiome-association index, b^2', and found that this index can be estimated for groups of hundreds of individuals [30]. Based on our m^2 estimates, we found that three sites (vagina, uterus and isthmus) from the reproductive tracts have a large effect on chicken egg production, with m^2 values for the vagina, uterus and isthmus of 0.9228, 0.9359 and 0.9886, respectively, after correcting for host genetic effects. The closer it was to the ovary, the larger m^2 value of the segment was observed, indicating that the microbiome of the isthmus was the most informative site with respect to egg production.

Commercial egg producers have special interest in hen oviducts because pathological changes or disrupted activity directly affect egg production efficiency and ultimately decrease economic profitability [92]. A previous study demonstrated that vaginal microbiota of sows could affect endometritis [93], with endometritis, the proportion of the smallest follicles was only 60% [94]. The isthmus in chicken forms both inner and outer shell membranes, while calcification of the eggshell and its subsequent pigmentation and cuticle deposition occur in the uterus, followed by expulsion of the egg through the vagina [95]. Functionally, it is likely that the microbiota in the uterus and isthmus may affect the health of the oviduct and thus influence chicken egg production, seen as an increased abundance of pathways for bacterial motility proteins, the bacterial secretion system, and membrane and intracellular structural molecules. Thus, we found that the microbiome could influence egg production in chickens, especially the microbiome in the reproductive tract.

Egg production associated microbial species found in the chicken reproductive tract were related to immune function

Our finding showed that the reproductive/digestive microbiota was significantly associated with

egg production led us to investigate which microorganisms play key roles in this association. At all three levels (OTU, species and genus), we found that the isthmus and uterus contained more microorganisms significantly associated with EN300 than other sites. We observed that the highly abundant species *Bacteroides fragilis*, *Bacteroides salanitronis*, *Bacteroides barnesiae* and *Clostridium leptum* were significantly associated with a higher EN300 in the three reproductive sites. Meanwhile, a high abundance of *Firmicutes bacterium ZOR0006* was significantly associated with a low EN300 in three reproductive sites and a high EN300 in three digestive sites. Among these five EN300-associated species, only one (*Bacteroides barnesiae*, h^2=0.43) was affected by host genetics.

In chickens, the genus *Bacteroides* is found in the caecum and represents one of the predominant anaerobic genera [96]. *Bacteroides* species are thought to play an important role in the breakdown of polysaccharides into simpler compounds that are used by the animal host as well as the microorganisms themselves [88]. Intestinal anaerobic bacteria such as *Bacteroides fragilis* and *Bacteroides salanitronis* have been suggested to possess metabolic pathways for N-glycan production [97]. The symbiont *Bacteroides fragilis* exists in a commensal relationship with the host as it expresses a relatively large number of genes involved in polysaccharide metabolism, which benefits the host [98]. The surface of *Bacteroides fragilis* can produce polysaccharides; in particular, capsule capsular polysaccharide A (CPSA) has been shown to be a key mediator of mammalian immune system development [99, 100]. Surprisingly, CPSA has also been shown to have protective effects on autoimmune disorder models, such as antibiotic-induced experimental encephalomyelitis [101]. It is thus likely that genus *Bacteroides* can regulate bird reproductive activity by mediating the avian immune system.

*Clostridium leptum*e a major member of the Firmicutes phylum, can alter the gut microbiota in rats, especially in obese individuals [102]. In human infants, fecal levels of *Clostridium leptum* were found to be negatively correlated with proinflammatory marker levels [103]. Colonic colonization of *Clostridium leptum* is associated with the accumulation of regulatory T cells, which inhibits the development of inflammatory lesions [104]. The proliferation and activation of regulatory T cells is crucial to establishing and maintaining an appropriate level of immune tolerance. In addition, our results demonstrated that *Clostridium leptum* had a large range of associations with other uterus or isthmus microbiota constituents (but limited association was observed with digestive microbiota constituents) and was not influenced by host genetic regulation. It is thus likely that this microorganism serves as a stimulator of regulatory T cell production and inhibitor of inflammatory lesions and then could regulate and maintain immunologic tolerance and the microbiota composition of the reproductive tract (especially the uterus and isthmus). The function of *Firmicutes bacterium ZOR0006* is unknown, the abundance of this species was negatively correlated with EN300. In summary, microbial species found in the chicken digestive/reproductive tract that are most strongly associated with egg production probably exert this effect by modulating immune function, and further studies need to be conducted to explore their function in detail.

Conclusions

In summary, our results indicated that the reproductive/digestive tract microbial community of chickens is largely independent of host genetics. However, the microbiota was highly adapted to the environment of a specific site. A small proportion of the variability in egg production was

associated with the reproductive/digestive tract microbiota in chickens. In particular, the genus *Bacteroides* and the species *Clostridium leptum* and *Firmicutes bacterium ZOR0006* were strongly associated with EN300. These findings provide new insight into the roles of the reproductive/digestive tract microbiota in complex traits, and may help contribute to the development of effective therapies for improving egg production.

Authors' contributions
D.L. and Q.Z. designed the study; T.W., S.K.M., Y.S., W.Z., M.Y., D.Y., Q.N., Y.L, and S.T. wrote the manuscript; W.Z., M.Y., X.L., D.Y., S.K.M. and Q.N. collected the sample and extracted the DNA. S.T., W.Z., T.M., L.S., Y.S., D.Z., C.N., X.Z., H.Y., X.F., B.Z, Y.H. and Y.W. and R.W. did bioinformatics analyses; M.Z., H.X., and Y.Y. wrote methods; D.L. and Q.Z. supervised the work.

Funding
This work was supported by the Sichuan Provincial Department of Science & Technology Program (2019JDTD0009), the Fok Ying-Tong Education Foundation for Young Teachers in the Higher Education Institutions of China (161026) and the China Agricultural Research System (CARS-41).

Availability of supporting data
The sequencing data for this project have been deposited in the National Genomics Data Center under accession number CRA002196 (microbiome) and CRA002195 (Genome resequencing). The authors declare that all other relevant data supporting the findings of the study are available in this article and its Supplementary Information files, or from the corresponding author upon request.

Ethics approval and consent to participate
Chickens were managed according to the Institutional Animal Care and Use Committee of Sichuan Agricultural University under permit number DKY- 2018102015.

Consent for publication
Not applicable

Competing interests
The authors declare no competing interests.

Author details
[1]Farm Animal Genetic Resources Exploration and Innovation Key Laboratory of Sichuan Province, Sichuan Agricultural University, Chengdu, China. [2]Department of Ecology, Hubei Key Laboratory of Cell Homeostasis, College of Life Sciences, Wuhan University, Wuhan, China. [3]Novogene Bioinformatics Institute, Beijing, China. [4]Center of Precision Medicine, West China Hospital, Sichuan University, Sichuan, China

References
1.   Jozefiak D, Jozefiak A, Kieronczyk B, Rawski M, Swiatkiewicz S, Dlugosz J, Engberg RM: **Insects - a Natural Nutrient Source for Poultry - a Review**. *Ann Anim Sci* 2016, **16**(2):297-313.
2.   Wolc A, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, Fernando R, Garrick DJ,

Dekkers JC: **Analysis of egg production in layer chickens using a random regression model with genomic relationships**. *Poult Sci* 2013, **92**(6):1486-1491.

3. Lamichhaney S, Fan G, Widemo F, Gunnarsson U, Thalmann DS, Hoeppner MP, Kerje S, Gustafson U, Shi C, Zhang H *et al*: **Structural genomic changes underlie alternative reproductive strategies in the ruff (Philomachus pugnax)**. *Nature Genetics* 2016, **48**(1):84-88.

4. Husby A, Kawakami T, Rönnegård L, Smeds L, Ellegren H, Qvarnström A: **Genome-wide association mapping in a wild avian population identifies a link between genetic and phenotypic variation in a life-history trait**. *Proc Biol Sci* 2015, **282**(1806):20150156.

5. Hu ZL, Park CA, 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.

6. Liu Z, Yang N, Yan Y, Li G, Liu A, Wu G, Sun C: **Genome-wide association analysis of egg production performance in chickens across the whole laying period**. *BMC Genet* 2019, **20**(1):67.

7. Yuan J, Sun C, Dou T, Yi G, Qu L, Qu L, Wang K, Yang N: **Identification of Promising Mutants Associated with Egg Production Traits Revealed by Genome-Wide Association Study**. *PLoS One* 2015, **10**(10):e0140615.

8. Zhang GX, Fan QC, Wang JY, Zhang T, Xue Q, Shi HQ: **Genome-wide association study on reproductive traits in Jinghai Yellow Chicken**. *Animal reproduction science* 2015, **163**:30-34.

9. Li DY, Zhang L, Yang MY, Xu HL, Yin HD, Li Y, Zhu Q: **Effect of luteinizing hormone/choriogonadotropin receptor (LHCGR) gene on chicken reproductive traits**. *Molecular biology reports* 2013, **40**(12):7111-7116.

10. Tongsiri S, Jeyaruban MG, Van Der Werf JH: **Genetic parameters for egg production traits in purebred and hybrid chicken in a tropical environment**. *British poultry science* 2015, **56**(6):613-620.

11. Savegnago RP, Caetano SL, Ramos SB, Nascimento GB, Schmidt GS, Ledur MC, Munari DP: **Estimates of genetic parameters, and cluster and principal components analyses of breeding values related to egg production traits in a White Leghorn population**. *Poultry science* 2011, **90**(10):2174-2188.

12. Fettweis JM, Serrano MG, Brooks JP, Edwards DJ, Girerd PH, Parikh HI, Huang B, Arodz TJ, Edupuganti L, Glascock AL *et al*: **The vaginal microbiome and preterm birth**. *Nat Med* 2019, **25**(6):1012-1021.

13. Serrano MG, Parikh HI, Brooks JP, Edwards DJ, Arodz TJ, Edupuganti L, Huang B, Girerd PH, Bokhari YA, Bradley SP *et al*: **Racioethnic diversity in the dynamics of the vaginal microbiome during pregnancy**. *Nat Med* 2019, **25**(6):1001-1011.

14. Baker JM, Al-Nakkash L, Herbst-Kralovetz MM: **Estrogen-gut microbiome axis: Physiological and clinical implications**. *Maturitas* 2017, **103**:45-53.

15. Franasiak JM, Scott RT, Jr.: **Introduction: Microbiome in human reproduction**. *Fertil Steril* 2015, **104**(6):1341-1343.

16. Franasiak JM, Scott RT, Jr.: **Reproductive tract microbiome in assisted reproductive technologies**. *Fertil Steril* 2015, **104**(6):1364-1371.

17. Gotsch F, Romero R, Kusanovic JP, Mazaki-Tovi S, Pineles BL, Erez O, Espinoza J, Hassan SS: **The fetal inflammatory response syndrome**. *Clinical obstetrics and gynecology* 2007,

50(3):652-683.

18.   Lamont RF, Sobel JD, Akins RA, Hassan SS, Chaiworapongsa T, Kusanovic JP, Romero R: **The vaginal microbiome: new information about genital tract flora using molecular based techniques**. *BJOG : an international journal of obstetrics and gynaecology* 2011, **118**(5):533-549.

19.   Gast RK, Guraya R, Guard-Bouldin J, Holt PS, Moore RW: **Colonization of specific regions of the reproductive tract and deposition at different locations inside eggs laid by hens infected with Salmonella enteritidis or Salmonella heidelberg**. *Avian diseases* 2007, **51**(1):40-44.

20.   Guard J, Gast RK, Guraya R: **Colonization of avian reproductive-tract tissues by variant subpopulations of Salmonella enteritidis**. *Avian diseases* 2010, **54**(2):857-861.

21.   Simon K, Verwoolde MB, Zhang J, Smidt H, de Vries Reilingh G, Kemp B, Lammers A: **Long-term effects of early life microbiota disturbance on adaptive immunity in laying hens**. *Poultry science* 2016, **95**(7):1543-1554.

22.   Waite DW, Taylor MW: **Characterizing the avian gut microbiota: membership, driving influences, and potential function**. *Front Microbiol* 2014, **5**:223.

23.   Johnson TJ, Youmans BP, Noll S, Cardona C, Evans NP, Karnezos TP, Ngunjiri JM, Abundo MC, Lee CW: **A Consistent and Predictable Commercial Broiler Chicken Bacterial Microbiota in Antibiotic-Free Production Displays Strong Correlations with Performance**. *Applied and environmental microbiology* 2018, **84**(12).

24.   Mohd Shaufi MA, Sieo CC, Chong CW, Gan HM, Ho YW: **Deciphering chicken gut microbial dynamics based on high-throughput 16S rRNA metagenomics analyses**. *Gut pathogens* 2015, **7**:4.

25.   Wen C, Yan W, Sun C, Ji C, Zhou Q, Zhang D, Zheng J, Yang N: **The gut microbiota is largely independent of host genetics in regulating fat deposition in chickens**. *ISME J* 2019, **13**(6):1422-1436.

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

27.   Lee S, La T-M, Lee H-J, Choi I-S, Song C-S, Park S-Y, Lee J-B, Lee S-W: **Characterization of microbial communities in the chicken oviduct and the origin of chicken embryo gut microbiota**. *Scientific Reports* 2019, **9**(1):6838.

28.   Elokil AA, Magdy M, Melak S, Ishfaq H, Bhuiyan A, Cui L, Jamil M, Zhao S, Li S: **Faecal microbiome sequences in relation to the egg-laying performance of hens using amplicon-based metagenomic association analysis**. *Animal* 2020, **14**(4):706-715.

29.   Carmody RN, Gerber GK, Luevano JM, Jr., Gatti DM, Somes L, Svenson KL, Turnbaugh PJ: **Diet dominates host genotype in shaping the murine gut microbiota**. *Cell host & microbe* 2015, **17**(1):72-84.

30.   Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, Costea PI, Godneva A, Kalka IN, Bar N *et al*: **Environment dominates over host genetics in shaping human gut microbiota**. *Nature* 2018, **555**(7695):210-215.

31.   Han Z, Willer T, Pielsticker C, Gerzova L, Rychlik I, Rautenschlein S: **Differences in host breed and diet influence colonization by Campylobacter jejuni and induction of local immune responses in chicken**. *Gut pathogens* 2016, **8**(1):56.

32.   Liu H, Chen Z, Gao G, Sun C, Li Y, Zhu Y: **Characterization and comparison of gut microbiomes**

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

33. Ding J, Dai R, Yang L, He C, Xu K, Liu S, Zhao W, Xiao L, Luo L, Zhang Y *et al*: **Inheritance and Establishment of Gut Microbiota in Chickens**. *Frontiers in microbiology* 2017, **8**:1967-1967.

34. Zhao L, Wang G, Siegel P, He C, Wang H, Zhao W, Zhai Z, Tian F, Zhao J, Zhang H *et al*: **Quantitative genetic background of the host influences gut microbiomes in chickens**. *Sci Rep* 2013, **3**:1163.

35. Kers JG, Velkers FC, Fischer EAJ, Hermes GDA, Stegeman JA, Smidt H: **Host and Environmental Factors Affecting the Intestinal Microbiota in Chickens**. *Frontiers in microbiology* 2018, **9**:235-235.

36. Camarinha-Silva A, Maushammer M, Wellmann R, Vital M, Preuss S, Bennewitz J: **Host Genome Influence on Gut Microbial Composition and Microbial Prediction of Complex Traits in Pigs**. *Genetics* 2017, **206**(3):1637-1644.

37. Fan P, Bian B, Teng L, Nelson CD, Driver J, Elzo MA, Jeong KC: **Host genetic effects upon the early gut microbiota in a bovine model with graduated spectrum of genetic variation**. *Isme j* 2020, **14**(1):302-317.

38. Ji J, Luo CL, Zou X, Lv XH, Xu YB, Shu DM, Qu H: **Association of host genetics with intestinal microbial relevant to body weight in a chicken F2 resource population**. *Poultry science* 2019, **98**(9):4084-4093.

39. Mignon-Grasteau S, Narcy A, Rideau N, Chantry-Darmon C, Boscher MY, Sellier N, Chabault M, Konsak-Ilievski B, Le Bihan-Duval E, Gabriel I: **Impact of Selection for Digestive Efficiency on Microbiota Composition in the Chicken**. *PloS one* 2015, **10**(8):e0135488.

40. Scepanovic P, Hodel F, Mondot S, Partula V, Byrd A, Hammer C, Alanio C, Bergstedt J, Patin E, Touvier M *et al*: **A comprehensive assessment of demographic, environmental, and host genetic associations with gut microbiome diversity in healthy individuals**. *Microbiome* 2019, **7**(1):130.

41. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities**. *Appl Environ Microbiol* 2009, **75**(23):7537-7541.

42. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads**. *EMBnet journal* 2011, **17**(1):10-12.

43. Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R: **UCHIME improves sensitivity and speed of chimera detection**. *Bioinformatics* 2011, **27**(16):2194-2200.

44. Haas BJ, Gevers D, Earl AM, Feldgarden M, Ward DV, Giannoukos G, Ciulla D, Tabbaa D, Highlander SK, Sodergren E *et al*: **Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons**. *Genome research* 2011, **21**(3):494-504.

45. Edgar RC: **UPARSE: highly accurate OTU sequences from microbial amplicon reads**. *Nature methods* 2013, **10**(10):996-998.

46. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO: **The SILVA ribosomal RNA gene database project: improved data processing and web-based tools**. *Nucleic acids research* 2013, **41**(Database issue):D590-596.

47. Edgar RC: **MUSCLE: multiple sequence alignment with high accuracy and high throughput**. *Nucleic acids research* 2004, **32**(5):1792-1797.

48. Bray JR, Curtis JT: **An ordination of the upland forest communities of southern Wisconsin**.

Ecol Monogr 1957, **27**(4):325-349.

49.   Douglas GM, Maffei VJ, Zaneveld J, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI: **PICRUSt2: An improved and extensible approach for metagenome inference**. *bioRxiv* 2019:672295.

50.   Chapman MG, Underwood AJ: **Ecological patterns in multivariate assemblages: information and interpretation of negative values in ANOSIM tests**. *Marine Ecology Progress Series* 1999, **180**:257-265.

51.   Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C: **Metagenomic biomarker discovery and explanation**. *Genome biology* 2011, **12**(6):R60.

52.   Li M, Tian S, Jin L, Zhou G, Li Y, Zhang Y, Wang T, Yeung CK, Chen L, Ma J *et al*: **Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars**. *Nature genetics* 2013, **45**(12):1431-1438.

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

54.   McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M: **The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data**. *Genome Res* 2010, **20**(9):1297-1303.

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

56.   Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST: **The variant call format and VCFtools**. *Bioinformatics* 2011, **27**(15):2156-2158.

57.   Browning SR, Browning BL: **Rapid and accurate haplotype phasing and missing-data inference for whole-genome association studies by use of localized haplotype clustering**. *The American Journal of Human Genetics* 2007, **81**(5):1084-1097.

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

59.   Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ *et al*: **PLINK: a tool set for whole-genome association and population-based linkage analyses**. *American journal of human genetics* 2007, **81**(3):559-575.

60.   Yang J, Bakshi A, Zhu Z, Hemani G, Vinkhuyzen AA, Lee SH, Robinson MR, Perry JR, Nolte IM, van Vliet-Ostaptchouk JV *et al*: **Genetic variance estimation with imputed variants finds negligible missing heritability for human height and body mass index**. *Nat Genet* 2015, **47**(10):1114-1120.

61.   Yang J, Lee SH, Goddard ME, Visscher PM: **GCTA: a tool for genome-wide complex trait analysis**. *Am J Hum Genet* 2011, **88**(1):76-82.

62.   Zierer J, Jackson MA, Kastenmuller G, Mangino M, Long T, Telenti A, Mohney RP, Small KS, Bell JT, Steves CJ *et al*: **The fecal metabolome as a functional readout of the gut microbiome**. *Nat Genet* 2018, **50**(6):790-795.

63.   Zhou X, Stephens M: **Genome-wide efficient mixed-model analysis for association studies**. *Nat Genet* 2012, **44**(7):821-824.

64.   Chen C, Song X, Wei W, Zhong H, Dai J, Lan Z, Li F, Yu X, Feng Q, Wang Z *et al*: **The microbiota continuum along the female reproductive tract and its relation to uterine-related diseases**.

Nat Commun 2017, **8**(1):875.

65. Verstraelen H, Vilchez-Vargas R, Desimpel F, Jauregui R, Vankeirsbilck N, Weyers S, Verhelst R, De Sutter P, Pieper DH, Van De Wiele T: **Characterisation of the human uterine microbiome in non-pregnant women through deep sequencing of the V1-2 region of the 16S rRNA gene**. *PeerJ* 2016, **4**:e1602.

66. Lee S, La TM, Lee HJ, Choi IS, Song CS, Park SY, Lee JB, Lee SW: **Characterization of microbial communities in the chicken oviduct and the origin of chicken embryo gut microbiota**. *Scientific reports* 2019, **9**(1):6838.

67. Boskey ER, Cone RA, Whaley KJ, Moench TR: **Origins of vaginal acidity: high D/L lactate ratio is consistent with bacteria being the primary source**. *Hum Reprod* 2001, **16**(9):1809-1813.

68. Mabelebele M, Alabi OJ, Ng`ambi JW, Norris D, Ginindza MM: **Comparison of Gastrointestinal Tracts and pH Values of Digestive Organs of Ross 308 Broiler and Indigenous Venda Chickens Fed the Same Diet**. *Asian J Anim Vet Adv* 2014, **9**(1):71-76.

69. Ravel J, Gajer P, Abdo Z, Schneider GM, Koenig SS, McCulle SL, Karlebach S, Gorle R, Russell J, Tacket CO *et al*: **Vaginal microbiome of reproductive-age women**. *Proceedings of the National Academy of Sciences of the United States of America* 2011, **108 Suppl 1**:4680-4687.

70. Mishra AK, Kumar A, Swain DK, Yadav S, Nigam R: **Insights into pH regulatory mechanisms in mediating spermatozoa functions**. *Vet World* 2018, **11**(6):852-858.

71. Fiser PS, Macpherson JW: **pH values in the oviduct of the hen during egg formation**. *Poult Sci* 1974, **53**(2):827-829.

72. Hickey RJ, Zhou X, Settles ML, Erb J, Malone K, Hansmann MA, Shew ML, Van Der Pol B, Fortenberry JD, Forney LJ: **Vaginal microbiota of adolescent girls prior to the onset of menarche resemble those of reproductive-age women**. *MBio* 2015, **6**(2).

73. Aagaard K, Riehle K, Ma J, Segata N, Mistretta TA, Coarfa C, Raza S, Rosenbaum S, Van den Veyver I, Milosavljevic A *et al*: **A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy**. *PLoS One* 2012, **7**(6):e36466.

74. Romero R, Hassan SS, Gajer P, Tarca AL, Fadrosh DW, Nikita L, Galuppi M, Lamont RF, Chaemsaithong P, Miranda J *et al*: **The composition and stability of the vaginal microbiota of normal pregnant women is different from that of non-pregnant women**. *Microbiome* 2014, **2**(1):4.

75. White RA, 3rd, Soles SA, Gavelis G, Gosselin E, Slater GF, Lim DSS, Leander B, Suttle CA: **The Complete Genome and Physiological Analysis of the Eurythermal Firmicute Exiguobacterium chiriqhucha Strain RW2 Isolated From a Freshwater Microbialite, Widely Adaptable to Broad Thermal, pH, and Salinity Ranges**. *Front Microbiol* 2018, **9**:3189.

76. Kulshreshtha NM, Kumar A, Bisht G, Pasha S, Kumar R: **Usefulness of organic acid produced by Exiguobacterium sp. 12/1 on neutralization of alkaline wastewater**. *ScientificWorldJournal* 2012, **2012**:345101.

77. Kasana RC, Pandey CB: **Exiguobacterium: an overview of a versatile genus with potential in industry and agriculture**. *Crit Rev Biotechnol* 2018, **38**(1):141-156.

78. Yu MD, Lai EM: **Warfare between Host Immunity and Bacterial Weapons**. *Cell Host & Microbe* 2017, **21**(1):3-4.

79. Pritchard PJ: **Digestion of sugars in the crop**. *Comp Biochem Physiol A Comp Physiol* 1972, **43**(1):195-205.

80. Moore RW, Park SY, Kubena LF, Byrd JA, McReynolds JL, Burnham MR, Hume ME, Birkhold SG,

Nisbet DJ, Ricke SC: **Comparison of zinc acetate and propionate addition on gastrointestinal tract fermentation and susceptibility of laying hens to Salmonella enteritidis during forced molt**. *Poult Sci* 2004, **83**(8):1276-1286.

81.   Mayilraj S, Stackebrandt E: **The Family Paenibacillaceae**; 2014.

82.   Kalisperati P, Spanou E, Pateras IS, Korkolopoulou P, Varvarigou A, Karavokyros I, Gorgoulis VG, Vlachoyiannopoulos PG, Sougioultzis S: **Inflammation, DNA Damage, Helicobacter pylori and Gastric Tumorigenesis**. *Front Genet* 2017, **8**:20.

83.   Franceschi F, Annalisa T, Teresa DR, Giovanna D, Ianiro G, Franco S, Viviana G, Valentina T, Riccardo LL, Antonio G: **Role of Helicobacter pylori infection on nutrition and metabolism**. *World J Gastroenterol* 2014, **20**(36):12809-12817.

84.   Ansari S, Yamaoka Y: **Survival of Helicobacter pylori in gastric acidic territory**. *Helicobacter* 2017, **22**(4).

85.   Moreira G, Sobrinho APR, Bambirra BHS, Bambirra FHS, Carvalho MAR, Farias LM, Nicoli JR, Moreira ES: **Synergistic Growth Effect among Bacteria Recovered from Root Canal Infections**. *Braz J Microbiol* 2011, **42**(3):973-979.

86.   Jeong DY, Cho SK, Shin HS, Jung KW: **Application of an electric field for pretreatment of a seeding source for dark fermentative hydrogen production**. *Bioresource technology* 2013, **139**:393-396.

87.   Marche MG, Satta A, Floris I, Pusceddu M, Ruiu L: **Quantitative variation in the core bacterial community associated with honey bees from Varroa- infested colonies**. *Journal of Apicultural Research* 2019, **58**(3):444-454.

88.   Reeves AR, Wang GR, Salyers AA: **Characterization of four outer membrane proteins that play a role in utilization of starch by Bacteroides thetaiotaomicron**. *Journal of bacteriology* 1997, **179**(3):643-649.

89.   Kers JG, Velkers FC, Fischer EAJ, Hermes GDA, Stegeman JA, Smidt H: **Host and Environmental Factors Affecting the Intestinal Microbiota in Chickens**. *Frontiers in microbiology* 2018, **9**:235.

90.   Jang HM, Erf GF, Rowland KC, Kong BW: **Genome resequencing and bioinformatic analysis of SNP containing candidate genes in the autoimmune vitiligo Smyth line chicken model**. *BMC genomics*, **15**(1).

91.   Difford GF, Plichta DR, Lovendahl P, Lassen J, Noel SJ, Hojberg O, Wright AG, Zhu Z, Kristensen L, Nielsen HB *et al*: **Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows**. *PLoS genetics* 2018, **14**(10):e1007580.

92.   Chousalkar KK, Roberts JR: **Ultrastructural changes in the oviduct of the laying hen during the laying cycle**. *Cell and tissue research* 2008, **332**(2):349-358.

93.   Wang J, Li C, Nesengani LT, Gong Y, Zhang S, Lu W: **Characterization of vaginal microbiota of endometritis and healthy sows using high-throughput pyrosequencing of 16S rRNA gene**. *Microb Pathog* 2017, **111**:325-330.

94.   Zralý Z, Kummer V, Canderle J: **Activity of the ovarian follicle system in cows with various puerperal courses**. 1989, **34**(10):593-602.

95.   Hrabia A, Lesniak-Walentyn A, Sechman A, Gertler A: **Chicken oviduct-the target tissue for growth hormone action: effect on cell proliferation and apoptosis and on the gene expression of some oviduct-specific proteins**. *Cell and tissue research* 2014, **357**(1):363-372.

96.   Salanitro JP, Blake IG, Muirhead PA: **Studies on the cecal microflora of commercial broiler**

chickens. *Appl Microbiol* 1974, **28**(3):439-447.

97. Nihira T, Suzuki E, Kitaoka M, Nishimoto M, Ohtsubo K, Nakai H: **Discovery of beta-1,4-D-mannosyl-N-acetyl-D-glucosamine phosphorylase involved in the metabolism of N-glycans**. *The Journal of biological chemistry* 2013, **288**(38):27366-27374.

98. Chung H, Kasper DL: **Microbiota-stimulated immune mechanisms to maintain gut homeostasis**. *Curr Opin Immunol* 2010, **22**(4):455-460.

99. Mazmanian SK, Liu CH, Tzianabos AO, Kasper DL: **An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system**. *Cell* 2005, **122**(1):107-118.

100. Mostafavi AZ, Troutman JM: **Biosynthetic assembly of the Bacteroides fragilis capsular polysaccharide A precursor bactoprenyl diphosphate-linked acetamido-4-amino-6-deoxygalactopyranose**. *Biochemistry* 2013, **52**(11):1939-1949.

101. Ochoa-Reparaz J, Mielcarz DW, Wang Y, Begum-Haque S, Dasgupta S, Kasper DL, Kasper LH: **A polysaccharide from the human commensal Bacteroides fragilis protects against CNS demyelinating disease**. *Mucosal Immunol* 2010, **3**(5):487-495.

102. Eslinger AJ, Eller LK, Reimer RA: **Yellow pea fiber improves glycemia and reduces Clostridium leptum in diet-induced obese rats**. *Nutr Res* 2014, **34**(8):714-722.

103. Partty A, Kalliomaki M, Salminen S, Isolauri E: **Infantile Colic Is Associated With Low-grade Systemic Inflammation**. *J Pediatr Gastr Nutr* 2017, **64**(5):691-695.

104. Li YN, Huang F, Cheng HJ, Li SY, Liu L, Wang LY: **Intestine-derived Clostridium leptum induces murine tolerogenic dendritic cells and regulatory T cells in vitro**. *Hum Immunol* 2014, **75**(12):1232-1238.

Figures
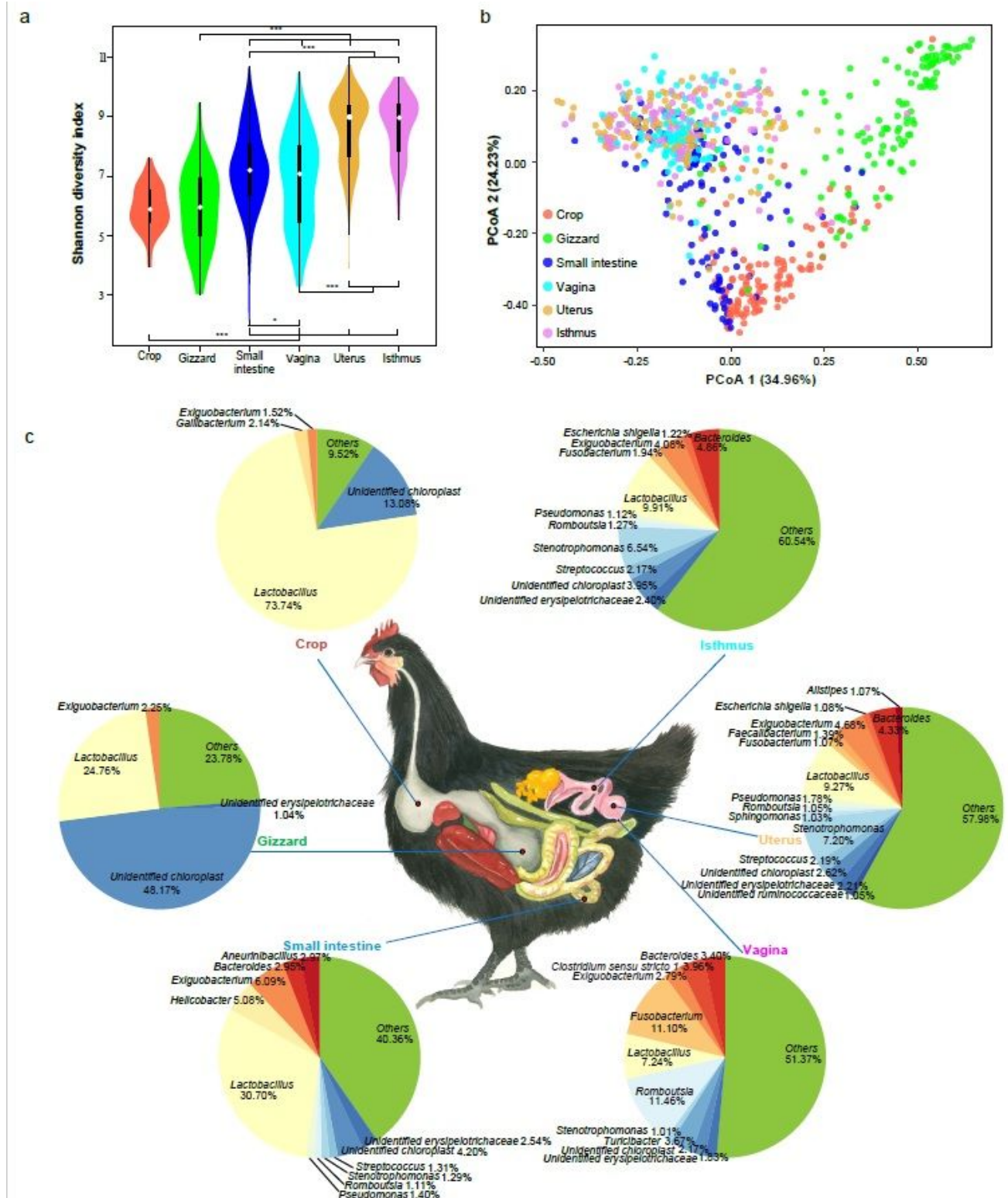
Figure 1

a An alpha diversity comparison based on the Shannon diversity index using the Wilcoxon rank-sum test to determine significant differences. b PCoA of the 768 samples based on weighted UniFrac distances. The first dimension and second dimension are shown. The fraction of the variance explained was 34.96%

for component 1 (P < 0.05, Tracy-Widom test) and 24.23% for component 2 (P < 0.05, Tracy-Widom test). c Relative abundance of the microbiota from the six sites at the genus level. Only genera with an abundance of >1.0% in any of the sites are described (Additional file 1: Table S5).
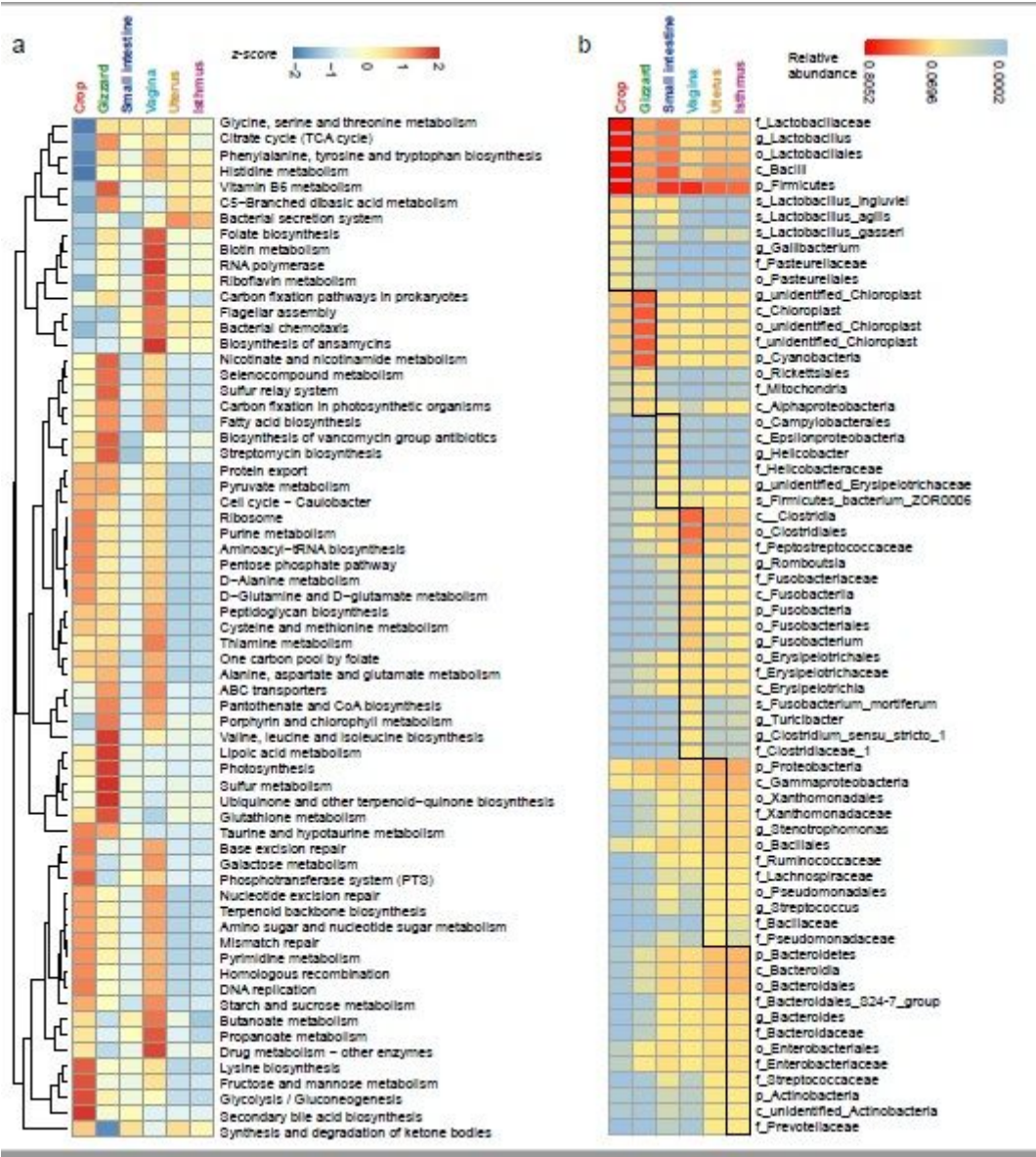
Figure 2

Comparison of the predicted functional capacities and site-associated taxa of the six microbial communities at diverse anatomical sites. a Heatmap showing the predicted KEGG pathways and their different abundances at different reproductive and digestive tract locations (see also Additional file 1: Table S6). The heatmap is color coded based on row z-scores of means of KEGG pathway abundances. b Heat map showing the 65 site-associated bacterial taxa identified by LEfSe (LDA > 4) and taxa whose abundances differed significantly among the six sites. p, phylum; c, class; o, order; f, family; g, genus.
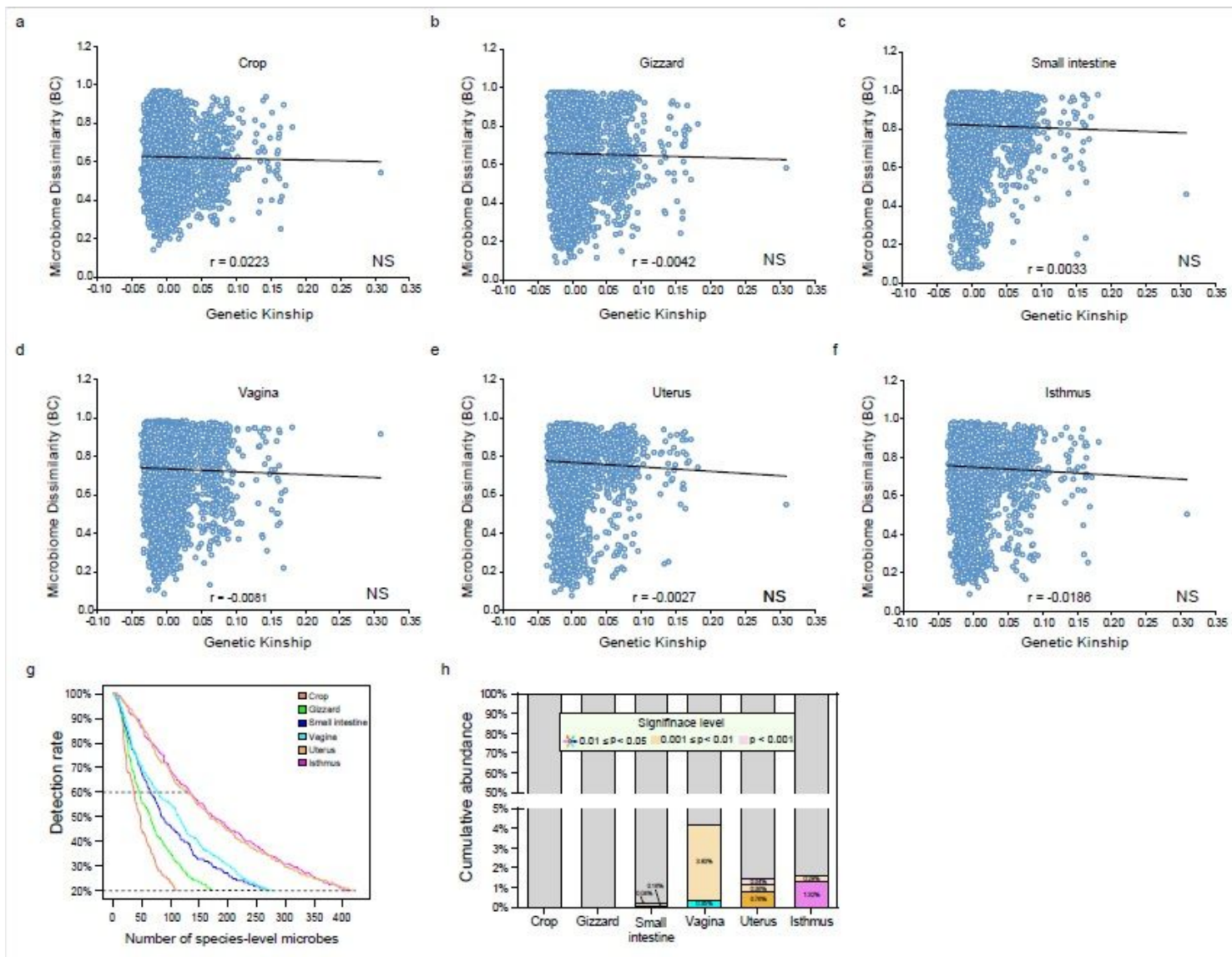
Figure 3

Host genetics have a limited effect on the reproductive/digestive tract microbiota. a-f Scatter plot of genetic kinship of pairs of individuals (x axis) and their microbiome dissimilarity (y axis) among all pairs (n= 16,256) of individuals, demonstrating that genetic kinship and the microbiome similarity of the crop, gizzard, small intestine, vagina, uterus and isthmus are uncorrelated (P > 0.05; Mantel test). NS, not significant. g Number of species identified in all chickens at the six different anatomical sites. h Cumulative relative abundance of heritable microbial species in each segment.
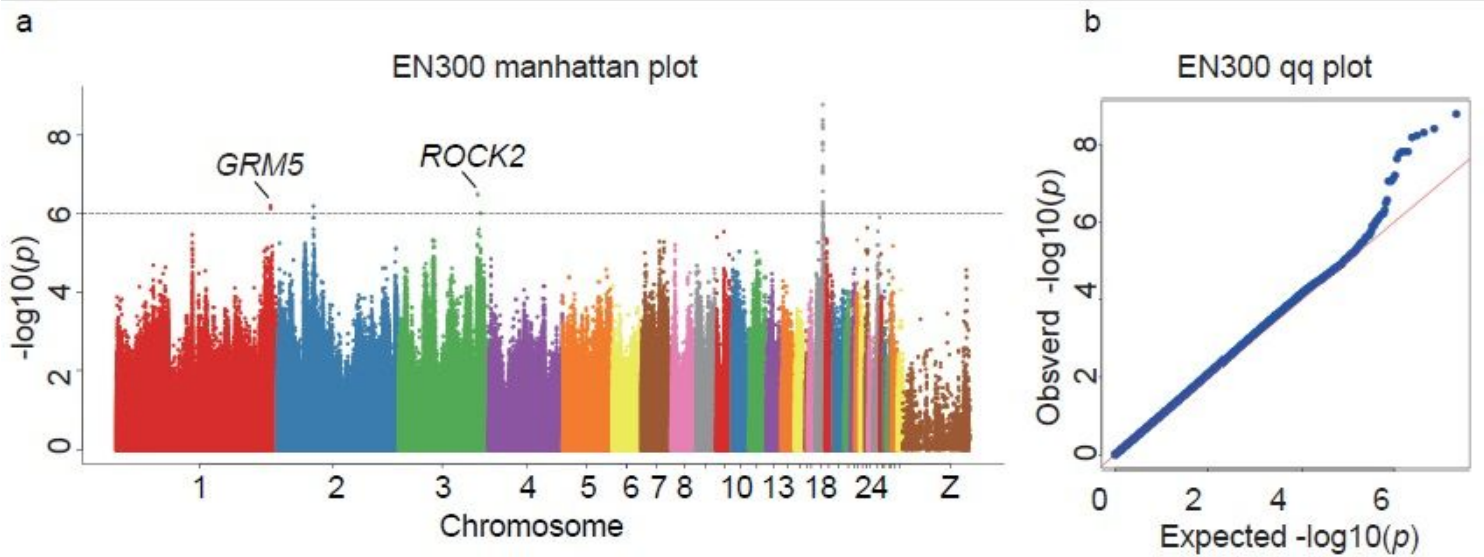
Figure 4

Genome-wide association study analysis of single-nucleotide polymorphisms associated with egg number at 300 days of age (EN300). a The horizontal dashed gray line indicates genome-wide significance thresholds (1.0 × 10-6). Each point represents a single-nucleotide polymorphism. The association analysis was conducted with the genome-wide efficient mixed-model association (GEMMA) software package. b Quantile-quantile plot of the genome-wide efficient mixed model for EN300. A 45-degree reference line is shown in red. The greater the departure from this reference line, the greater the likelihood that the two data sets came from populations with different distributions.

Figure 5

EN300-associated microorganisms. a The number of microbial species at the six anatomical sites associated with EN300 based on a significance threshold of P < 0.05, and the overlap between sites. b Pearson correlations between EN300 and the 24 EN300-associated microbial species; significant R-values are filled numerically (P < 0.05). Red and blue tiles indicate positive and negative correlations, respectively. c Pearson correlations among microbial species in the uterus (UT).

Figure 6

Effect of microbial species associated with egg number at 300 days of age (EN300) on egg production. a EN300 for the 20% of birds with the highest and lowest abundances of Bacteroides fragilis, Bacteroides salanitronis, Bacteroides barnesiae, Clostridium leptum and Firmicutes bacterium ZOR0006 at the three reproductive sites. The box plots show the median and the 25% and 75% quantiles. The cross and horizontal line indicate the mean and median values in the corresponding group, respectively.

Significance levels were calculated using a permutation test with 10,000 replicates. Significance was established at P < 0.05. b Relative abundance and detection rate of 5 species (Bacteroides fragilis, Bacteroides salanitronis, Bacteroides barnesiae, Clostridium leptum and Firmicutes bacterium ZOR0006) in the 6 segments. The values shown in the blue bars indicate the detection frequency of each species at each site.