

Bacterial meta-analysis of chicken cecal microbiota

Luis Alberto Chica Cardenas^{1,2}, Viviana Clavijo³, Martha Vives³ and Alejandro Reyes^{1,2,4}

¹ Research Group on Computational Biology and Microbial Ecology, Department of Biological Sciences, Universidad de Los Andes, Bogotá, Colombia

² Max Planck Tandem Group in Computational Biology, Universidad de Los Andes, Bogotá, Colombia

³ Centro de Investigaciones Microbiológicas, Department of Biological Sciences, Universidad de Los Andes, Bogotá, Colombia

⁴ The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of Medicine, St. Louis, MO, USA

ABSTRACT

Poultry production is an industry that generates 90,000 metric tons of chicken meat worldwide. Thus, optimizing chicken growth and sustainable production is of great importance. A central factor determining not only production parameters, but also stability of the immune system and chicken health, is the diversity and variability of the microbiota present throughout the gastrointestinal tract. To date, several studies have investigated the relationship between bacterial communities and the gut microbiome, with limited data to compare. This study aims to create a bacterial meta-analysis based on studies using amplicon sequencing with Illumina sequencing technologies in order to build a baseline for comparison in future analyses of the cecal bacterial composition in chicken. A systematic literature review was performed (SYRF ID: e84f0468-e418-4eec-9da4-b517f1b4809d. Full project URL: <https://app.syrf.org.uk/projects/e84f0468-e418-4eec-9da4-b517f1b4809d/detail>). From all the available and analyzed manuscripts only nine contained full raw-sequence data available and the corresponding metadata. A total of 324 samples, comprising three different regions within the *16S rRNA* gene, were analyzed. Due to the heterogeneity of the data, each region was analyzed independently and an effort for a joint analysis was performed as well. Taxonomic profiling revealed 11 phyla, with *Firmicutes* as the most prevalent phylum, followed by *Bacteroidetes* and *Proteobacteria*. At genus level, 109 genera were found. Shannon metric for alpha diversity showed that factors like type of chickens (Commercial or experimental) and *16S rRNA* gene subregion have negligible effect on diversity. Despite the large number of parameters that were taken into account, the identification of common bacteria showed five genera to be common for all sets in at least 50% of the samples. These genera are highly associated to cellulose degradation and short chain fatty acids synthesis. In general, it was possible to identify some commonalities in the bacterial cecal microbial community despite the extensive variability and factors differing from one study to another.

Submitted 11 August 2020
Accepted 23 November 2020
Published 5 January 2021

Corresponding author
Alejandro Reyes,
a.reyes@uniandes.edu.co

Academic editor
Mikhail Gelfand

Additional Information and
Declarations can be found on
page 13

DOI [10.7717/peerj.10571](https://doi.org/10.7717/peerj.10571)

© Copyright
2021 Chica Cardenas et al.

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

Subjects Agricultural Science, Bioinformatics, Microbiology, Veterinary Medicine, Zoology
Keywords Meta-analysis, Chicken microbiota, 16S rRNA sequencing, Bacterial diversity

INTRODUCTION

Chickens are considered to be one of the main sources of food production worldwide (Godfray *et al.*, 2010). Chicken meat production accounts for more than 128 million tons per year worldwide (USDA Foreign Agricultural Service, 2019). It is also considered as the most efficient production of animal protein, due to its property of doubling the weight of food ingested into meat weight at the end of the production cycle. Furthermore, chicken meat has been reported as a source of highly digestible proteins, minerals and vitamins with low levels of saturated fatty acids, which accompanied with a balanced diet, might help reduce the risk of having cardiovascular and endocrine problems (Marangoni *et al.*, 2015). Likewise, the efficiency of poultry meat production has had an impact on the price of the product, which make it more accessible for a broader number of social classes than other sources of meat protein (Scanes, 2007).

The microbial community (microbiota) present in the gastrointestinal tract (gut) has been widely associated with factors involving the health of chickens such as the immune system, the physiology of the digestive system and exclusion of pathogens, as well as the performance in production (Clavijo & Flórez, 2017). Commensal bacteria have been positively associated with the generation and regulation of the mucus layer, which is involved in the protection of epithelial cells against pathogenic bacteria (Clavijo & Flórez, 2017). For instance, chicken gut microbiota might have a direct effect on controlling the prevalence of pathogenic bacteria by competitive exclusion (La Ragione & Woodward, 2003). Additionally, commensal bacterial metabolites, such as short chain fatty acids (SCFA), are responsible for the expansion of the absorption surface in the gut by increasing the number of its proliferating cells (Kien *et al.*, 2007).

Among the different segments that comprise the chicken's gut, the cecum is the place where the food is retained the longest time. It presents the highest rate of water absorption and bacterial diversity (Xiao *et al.*, 2016; Shaufi *et al.*, 2015). Of all phyla that colonize the cecum, *Firmicutes* and *Bacteroidetes* are reported to be the most abundant. Their prevalence has been associated to their capacity for digesting cellulose and non-starch polysaccharides, which cannot be digested in the small intestine and leads to SCFA production (Clavijo & Flórez, 2017; Sergeant *et al.*, 2014). Moreover, the *Firmicutes* that colonize the cecal region are also responsible for nitrogen cycling, which is highly associated with the efficiency of chickens to extract energy from food (Mancabelli *et al.*, 2016; Oakley *et al.*, 2014). Hence, most of the information available on chicken gut microbiota has been focused on this organ.

For microbial diversity analyses based on *16S rRNA* gene data it has been proven that the use of different sequencing technologies may alter species richness and estimates of microbial diversity for the same sample (Allali *et al.*, 2017). Therefore, comparisons between samples sequenced by different methods or bacterial consensus using data mixtures might result in significant biases for further analysis. However, a bacterial consensus or meta-analysis might provide useful information in order to evaluate the effect of different factors in the normal composition of cecal microbiota. To the best of our knowledge, three meta-analysis of chicken cecal composition has been performed

([Wei, Morrison & Yu, 2013](#); [Waite & Taylor, 2014](#); [Zou, Sharif & Parkinson, 2018](#)).

Even though these studies were proposed as a good model for new projects, the data used was retrieved from samples sequenced by Sanger technology ([Wei, Morrison & Yu, 2013](#)), 454 pyrosequencing ([Waite & Taylor, 2014](#)) or by a mixture of data from different methods ([Zou, Sharif & Parkinson, 2018](#)). Nowadays, Illumina Sequencing platforms are the predominant methods to obtain biological sequence information. In consideration of the biases that a comparison of reads obtained by different sequencing strategies ([Plummer et al., 2015](#)), we performed a bacterial meta-analysis based on sequences generated with Illumina technology with the aim of identifying bacteria, that despite the variations associated with each study, are prevalent in the cecum and therefore, could be crucial in chicken gut modulation.

MATERIALS AND METHODS

Data collection and quality filters

In order to retrieve the data, an extensive literature search using “cecal microbiota composition in chickens” as search term was performed. Moreover, keywords such as “cecum”, “broiler chicken”, “microbiota composition”, “16S rRNA” and “Illumina” were also taken into account. Different search engines (PubMed, Google Scholar, Elsevier) were used to conduct the search. Literature search was performed exclusively for english written papers and conducted until February 25, 2020. For the systematic search developed in this study, authors Luis Alberto Chica and Alejandro Reyes serve as screeners of the evaluated studies, whereas Martha Vives took the place of referee. However, there were no disagreements between screeners. A reverse search strategy, applying the same search terms described above, was made in SRA database in order to find data that might have been excluded in the initial search. Review articles were also checked for additional studies.

With the aim of choosing the most relevant studies, the following filters were used as inclusion parameters: (i) data must be available in public databases, (ii) metadata must be provided within the manuscript or in the database where the data was published and (iii) the number of samples reported in the article must coincide with the data submitted to the database. Those three steps are encompassed within a two-step exclusion criteria according to PRISMA guidelines ([Moher et al., 2009](#)), in which the screen criteria encompasses all studies that did not have available data and the eligibility criteria states for all studies in which the metadata was absent or the number of samples reported was different to the data submitted in public databases ([Fig. 1](#)). The entire set of 16S rRNA gene sequences used for this analysis was retrieved from the Sequence Read Archive (SRA) database ([Leinonen, Sugawara & Shumway, 2010](#)) and downloaded using the SRA Toolkit software ([Leinonen, Sugawara & Shumway, 2010](#)). Trimmomatic software ([Bolger, Lohse & Usadel, 2014](#)) was used to filter low quality reads using a sliding window of 4 and a minimum Phred score of 20. Due to the lack of reported information regarding barcodes, primers and adapters for the different samples, a headcrop of 15 bp was performed for all samples. In order to validate trimming procedures, we conducted a primer search in raw and clean reads by using the specific primers reported in each study and the software Seqkit v0.14 ([Shen et al., 2016](#)) ([Table S1](#)).

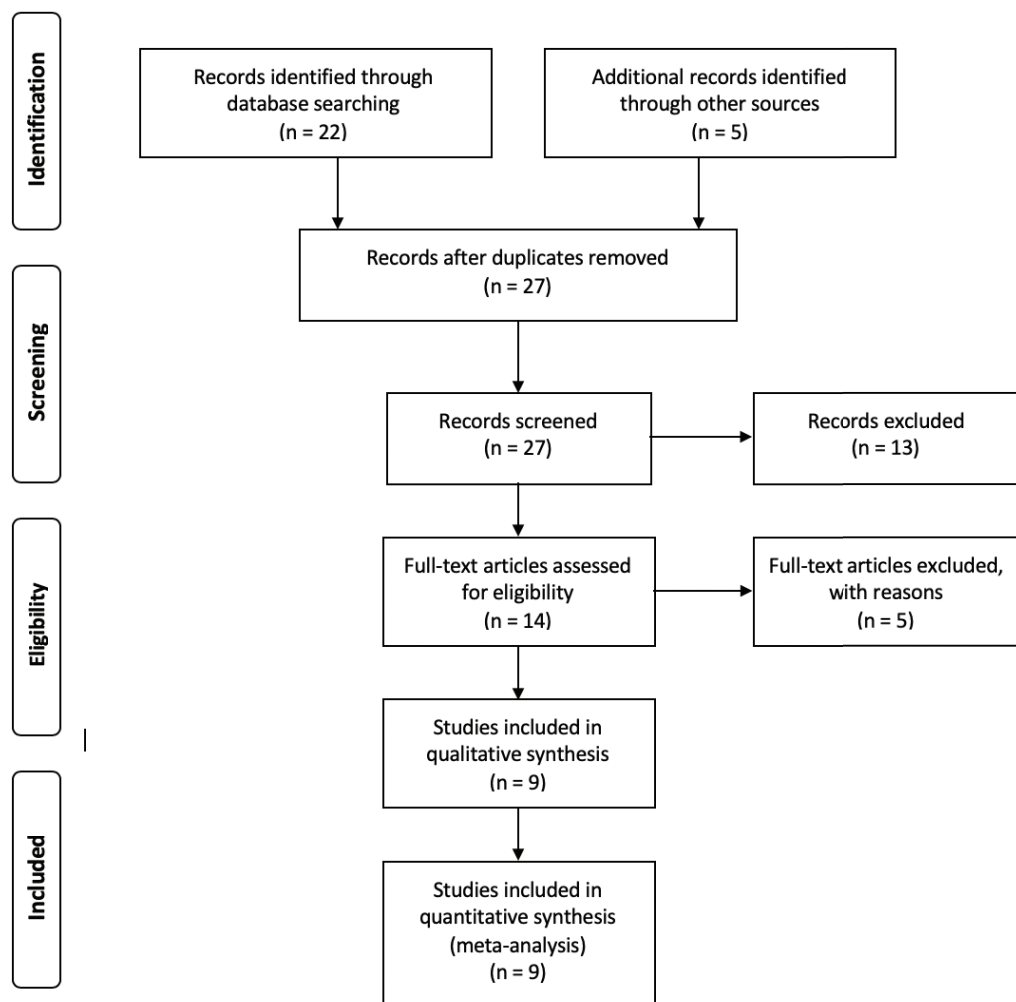


Figure 1 Flow chart showing the filtering steps, resulting in the nine articles that were used for this meta-analysis. This flow chart was made according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) (Moher et al., 2009). Four steps are considered to be crucial for the correct evaluation of different records, in order to include them into the analysis. Whereas identification step only involves the finding of records and the elimination of redundant papers, in the screening step, all records in which data was not available are removed. After the screening process, the eligibility step involves the exclusion of all records that, even with available data, were removed for multiple reasons. For the purpose of this article, those reasons are considered as filters ii and iii (see Methods: “Data Collection and Quality Filters”). Finally, the inclusion step accounts for the final number of records used for the analysis. [Full-size !\[\]\(5f471a71b78d7676bc356df190b88ab4_img.jpg\) DOI: 10.7717/peerj.10571/fig-1](https://doi.org/10.7717/peerj.10571/fig-1)

ASVs assignment

Due to the diversity of the hypervariable regions sequenced by the different studies, we split the entire dataset in three subsets. Each subset comprises samples generated with varying primers and corresponding to different hypervariable regions (V3, V4, V3V4). Deblur software (Amir et al., 2017) implemented in Qiime2 (Bolyen et al., 2019) was used for performing denoising and dereplication of the reads by applying a truncation length of 130 bp. ASV tables generated by deblur were filtered by a minimum frequency of 10, where singletons and doubletons were also excluded. In order to compare the taxa

identified by the different amplicon regions, ASV tables were summarized at the genus level using sklearn classifier for taxonomic classification. GreenGenes database v13.18 (DeSantis *et al.*, 2006) was used for training the classifier. Subsequently, all taxa that did not reach genus level were removed from collapsed tables. Finally, filtered tables resulting from the three datasets were merged together.

Core microbiome generation, taxonomic evaluation and diversity estimates

Independent core microbiome analysis was performed for all ASVs tables that were collapsed at genus level. Two approaches to define the core microbiome were assessed. (i) a strict approach closer to the general definition of core, where a genus should be present in at least 80% of the samples and (ii) a more relaxed search requiring a genus to be present in more than 50% of the samples. By using an in-house script, Venn diagrams were constructed with the aim of relating common genera present on the three hypervariable regions. Moreover, a core microbiome analysis per study was performed, selecting genera that were present in at least 80% of the samples of each given study. A comparison of the study specific core microbiomes was then performed to identify taxa that were present in the nine cores microbiome. The merged ASV table was transformed into relative frequency for computing taxonomic heatmaps at phylum and genus levels. For diversity estimates, rarefaction was performed at a sequence depth of 500. Alpha diversity was estimated using Shannon index and Observed species metrics. Bray-Curtis metric was employed to calculate Beta diversity. The significance of alpha diversity differences was evaluated by a Kruskal Wallis test. Meanwhile, the influence of different hypervariable regions on beta diversity estimates, was evaluated by a permutational analysis of variance (PERMANOVA), both tests as implemented in Qiime2.

RESULTS

The literature search resulted in 27 studies that contained information fulfilling the search criteria (Table 1). After applying the filters described previously (see “Materials and Methods”), nine studies were selected to perform this meta-analysis (Fig. 1). Those studies contained a total of 324 samples and 36,743,185 of single end sequences, with a number of sequences per sample varying between 15,181 and 1,548,296. After applying sequencing quality filters, the number of sequences was reduced to 8,843,573. Among the different hypervariable regions, 4,178,126 sequences covered the V4 16S rRNA gene region, 1,897,781 the V3 gene and 3,280,185 covered both the V3 and V4 region (V3V4) (Table S2).

After denoising the reads, filtering the resulting ASV tables, collapsing all taxa at genus level and merging the ASVs, 109 genera and 11 phyla were identified. *Firmicutes* was the most prevalent phylum, followed by *Bacteroidetes* and *Proteobacteria*. In some cases, a given taxonomic profile was highly related to a single study, however, a similar abundance profile of *Firmicutes* and *Proteobacteria* was seen in more than 95% of the samples (Fig. 2A). At genus level *Oscilospira*, *Bacteroides*, *Helicobacter* and *Lactobacillus* were found as the most representative genera (Fig. 2B).

Table 1 Characteristics of the manuscripts reviewed to perform the analyses.

Study	Available in public databases	Complete metadata	Information in databases according to article	Used for this study	Index
<i>Mancabelli et al. (2016)</i>	x	x	x	x	1
<i>Shaufi et al. (2015)</i>	x	x	x	x	2
<i>Ballou et al. (2016)</i>	x	x	x	x	3
<i>Costa et al. (2017)</i>	x	x	x	x	4
<i>Xu et al. (2016)</i>	x	x	x	x	5
<i>Xia et al. (2019)</i>	x	x	x	x	6
<i>Varmuzova et al. (2015)</i>	x	x	x	x	7
<i>Biasato et al. (2018)</i>	x	x	x	x	8
<i>Zhou et al. (2016)</i>	x	x	x	x	9
<i>Awad et al. (2016)</i>	x				NA
<i>Han et al. (2016)</i>	x				NA
<i>Kim et al. (2018)</i>					NA
<i>Kollarcikova et al. (2019)</i>					NA
<i>Krueger et al. (2017)</i>	x				NA
<i>Li et al. (2016)</i>		x			NA
<i>Ma et al. (2017)</i>					NA
<i>Marimuthu et al. (2019)</i>					NA
<i>Mon et al. (2015)</i>		x			NA
<i>Park, Lee & Ricke (2016)</i>					NA
<i>Park et al. (2017)</i>					NA
<i>Polansky et al. (2016)</i>					NA
<i>Saxena et al. (2016)</i>					NA
<i>Varmuzova et al. (2016)</i>					NA
<i>Wang et al. (2018)</i>	x				NA
<i>Wu et al. (2019)</i>					NA
<i>Xiao et al. (2016)</i>					NA
<i>Yan et al. (2017)</i>					NA

Note:

Each column represents one of the search parameters that were taken into account in order to select the most suitable studies. Articles that fulfilled each criterion were labeled with an X. The column titled “Used for this study” lists the studies selected for the analysis, whereas the last column contains the index assigned to each study, which is used as reference in different figures. NA refers to Not Applicable.

For the evaluation of common bacterial taxa, core microbiome of the samples was assessed by two different approaches. Common bacteria in (i) at least 50% and (ii) at least 80% of the samples were evaluated. The traditional cut-off of 80% abundance for core definition showed *Oscillospira* as the only genus that was identified in all 3 regions (Fig. 3A). After relaxing the threshold and evaluating the 50% core microbiome at genus level, 5 genera were found to be common for all sets (*Oscillospira*, *Lactobacillus*, *Faecalibacterium*, *Clostridium* and *Ruminococcus*). In addition, 6 genera were shared by V3 and V4 regions, 3 were exclusively shared between V3 and V3V4 and 2 genera were common for V4 and V3V4 regions (Fig. 3B). It is important to highlight that all genera that were exclusively found in a particular region were present in all studies

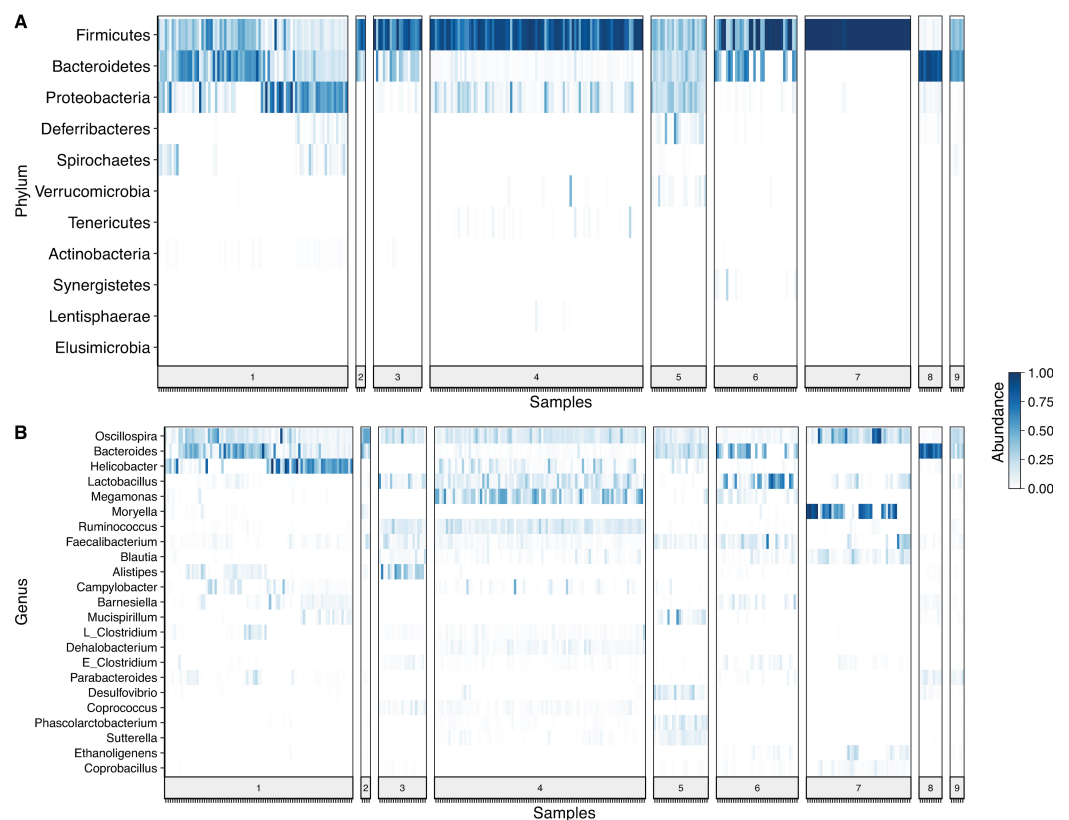


Figure 2 Heat-map showing relative abundance of each sample from each study at different taxonomic assignments. (A) Phylum and (B) genus level, for the merged ASV table. The most prevalent genera are shown (accounting for up to 90% of cumulative abundance). Index numbers at the bottom represent each study that was analyzed (Table 1). Relation of index numbers and metadata for each study is also shown on Table S1. In the y-axis, “L_Clostridium” refers to Clostridium genus assigned to Lachnospiraceae family and “E_Clostridium” refers to Clostridium genus assigned to Erysipelotrichaceae family. [Full-size !\[\]\(ba1b80118482ccef74a5d718ca4d7242_img.jpg\) DOI: 10.7717/peerj.10571/fig-2](https://doi.org/10.7717/peerj.10571/fig-2)

concerning that specific region, but not necessarily in all samples (Table S3). Finally, we evaluated the core microbiome per study, here, *Oscillospira* remained as the only genus shared by all the studies. Interestingly, multiple taxa were part of the core microbiome of several but one or two studies. Overall, *Oscillospira*, *Faecalibacterium*, *Lactobacillus*, *Clostridium*, *Bacteroides*, *Blautia*, *Ruminococcus* and *Coprobacillus* were present in the core of 6 studies or more (Fig. 3C; Table S4), which includes all genera identified at the 50% overall core-microbiome.

Due to the high number of metadata categories considered in this analysis, alpha diversity variation was evaluated based on four categories: (i) *16S rRNA* gene hypervariable region (ii) study, (iii) type of chickens (chickens under commercial rearing conditions or chickens rearing under controlled environment “Experimental”) and (iv) chicken breed. Variables such as age, due to the number of studies with missing data make it impossible to consider it as a variable for comparison. Alpha rarefaction plot for all samples using a depth of 500 sequences per sample showed that all samples reached a saturation asymptote, which confirms that the samples can be used for further analyses and

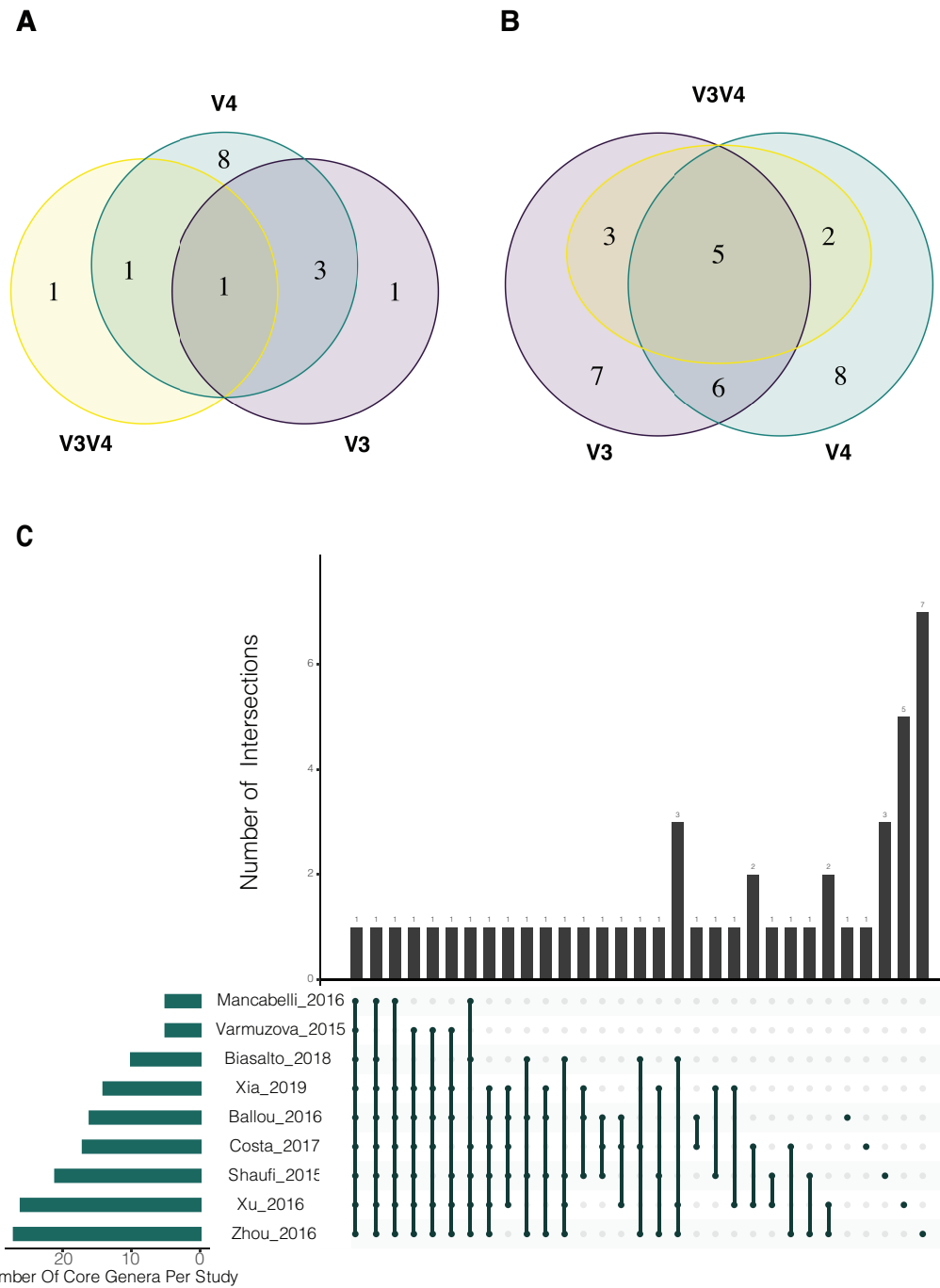


Figure 3 Analysis of shared taxons in the three different hypervariable regions considered in the study. (A) Analysis at genus level, using a 80% threshold, which means that bacteria must be present in at least the 80% of the samples. (B) Genus level at 50% threshold. (C) Core microbiome per study using an 80% threshold. Number of Core genera per study are shown on the horizontal bar plot, while the number of genera shared in the core of one or more studies are indicated in the vertical bar plots, the shared studies for each specific columns are shown by the connected dots.

Full-size  DOI: [10.7717/peerj.10571/fig-3](https://doi.org/10.7717/peerj.10571/fig-3)

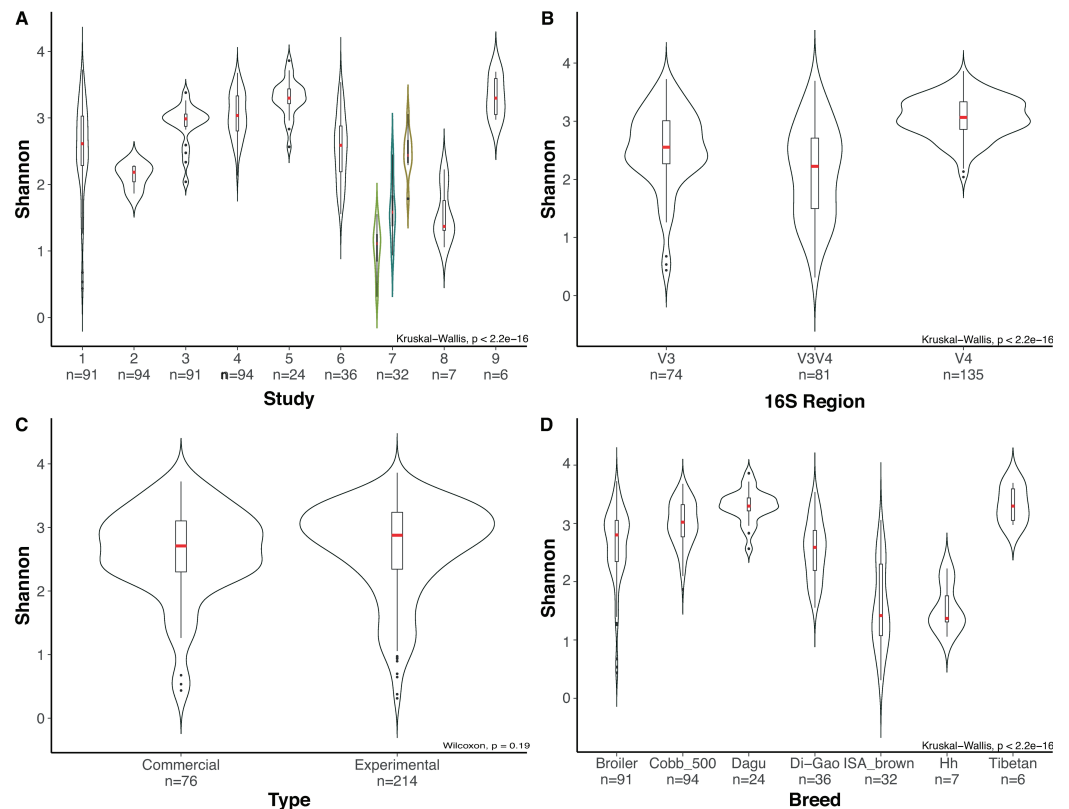


Figure 4 Evaluation of alpha diversity of the entire set of samples by using Shannon metric and Kruskal-Wallis statistics for significance. Violin plots represent different groups for each category. (A) Study. (B) Hypervariable region amplified from the *16S rRNA* gene. (C) Type of chickens (commercial, experimental). (D) Breed of the chickens, where Hh corresponds to Hubbard Hybrid breed and Broiler to all fattening breeds that are not specified. Boxplots represent the interquartile range and in red the median for each group. The number of samples in each group is shown right below the name. For Panel A, Indexes of metadata were used for visualization purposes and are related with the specific studies in Table S1. The colored violin plots correspond to *Varmuzova et al. (2015)* study, which is highlighted due to the inclusion of *Salmonella enteritidis* samples (green) that showed lower diversity than treated samples with plant extract (blue) and control samples (yellow) from the same study.

Full-size [DOI: 10.7717/peerj.10571/fig-4](https://doi.org/10.7717/peerj.10571/fig-4)

confirming a relatively low diversity in the chicken cecum (Fig. S1). Rarefaction process discarded 33 samples that did not fulfill the minimum sampling depth, leaving 290 samples for the diversity analyses.

Most variation for alpha diversity was observed when comparing among the different studies. Interestingly, one of the groups of samples that presented low diversity values was a subset of the study of *Varmuzova et al. (2015)*, which contains chickens infected with *Salmonella enteritidis*, showing a significant reduction (Wilcoxon, p -value = 0.009), compared with healthy chickens of the same study and healthy chickens from other studies (Fig. 4A). Likewise, when analyzing the effect of the selected hypervariable region, we noticed a significant difference on Shannon values for all groups (Kruskal Wallis, p -value = $1.9e-19$), being V4 the region with the higher values (Fig. 4B). In order to prove if the difference presented in the V3V4 group was a consequence of the infected chickens

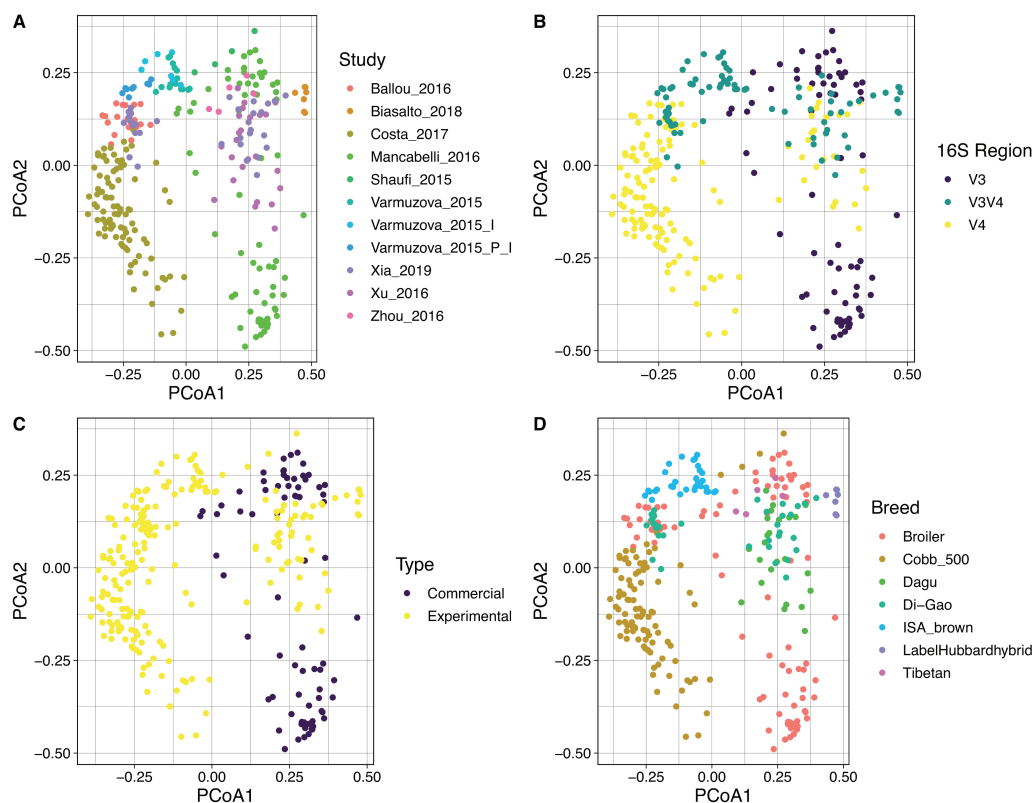


Figure 5 PCoA showing the spatial dispersion of the samples. Distance matrix was constructed based on Bray–Curtis metric. Coordinate plots represent (A) Study; the study of *Varmuzova et al. (2015)* was split due to the high intra-study variation obtained. Varmuzova_2015 = Control Chickens and uninfected Chickens treated with the plant extracts; Varmuzova_2015_I = Infected chickens; Varmuzova_2015_P_I = Infected chickens, treated with the plant extracts. (B) Hypervariable region of the 16S rRNA gene. (C) Type of chickens (commercial, experimental), when commercial chickens refer to animals raised on normal production environments and experimental refers to animals that were kept into environments with controlled conditions. (D) Breed of the animals, where Broiler corresponds to all fattening breeds that are not specified. [Full-size !\[\]\(1679558f37f6db0dd8360a2a7e913e90_img.jpg\) DOI: 10.7717/peerj.10571/fig-5](https://doi.org/10.7717/peerj.10571/fig-5)

that was previously described, we removed those samples from the dataset and performed again the analysis obtaining similar results. The comparison of filtered V3V4 group against V4 (Wilcoxon, p -value = $1.2e-13$) and V3 (Wilcoxon, p -value = 0.0067) groups independently, still showed significant differences on both cases. On the other hand, comparisons between commercial and experimental chickens reveal no significant variation (Wilcoxon, p -value = 0.188) (Fig. 4C). When using the breed of the chickens as variable, even though significant differences were observed, we cannot be sure of its relevance due to the lack of repeating breeds over different studies (Fig. 4D).

By using Bray Curtis metric, distance matrices were built, and beta diversity was estimated for the same categories evaluated above. No strong clustering was observed by any of the variables analyzed (Fig. 5), however, some recognizable clustering was observed for the samples belonging to *Mancabelli et al. (2016)* and *Costa et al. (2017)*. Further statistical analysis using Permanova pairwise method was applied on the variable Study and hypervariable region, both analyses showed significant differences. Whereas all

pairwise comparisons performed to the variable study reported q -values ranging between 0.00125 and 0.423, being the comparison of two groups of the same study (control and infected chickens) the only with no significant differences, the comparisons applied to the hypervariable region variable report all q -values of 0.001.

DISCUSSION

The aim of this study was to create a baseline for future research in order to give a higher insight into the general composition of the microbial communities in the cecum of chickens. This analysis aimed to collect the largest possible dataset of studies describing bacterial cecal content using Illumina technologies and amplicon sequencing.

Unfortunately, the study was limited since many of the published studies did not have their data available in public databases or did not share the metadata related with each project, making impossible to use the data or evaluate and replicate the results. Problems such as the lack of data available and the high number of sequences that were excluded after the quality filters were applied can be avoided by implementing common quality standards in the experimental design, data generation and data deposition. Only with serious commitment to high standards it is possible to achieve reproducible and comparable results among different studies. Another important point to be taken into account is the lack of a consensus on the specific hypervariable region of the *16S rRNA* gene to focus. In order to compare the different results obtained it was needed to collapse the ASVs tables at genus level and merge them.

After taxonomic classification of the three sets of samples (different hypervariable regions) at phylum and genus level, the results showed the predominance of *Firmicutes* and *Bacteroidetes* in most of the samples. These results are concordant with previously reported consensus based on sequences generated by different sequencing platforms (Wei, Morrison & Yu, 2013). When looking at the genus level, the number of genera reported were much higher in this meta-analysis than what was reported for individual studies (Xu et al., 2016; Zhou et al., 2016; Biasato et al., 2018) and comparable with the meta-analysis of Sanger generated sequences, performed by Wei, Morrison & Yu (2013), in which 117 genera were found to be present in the chicken gastrointestinal tract. Thus, the increase in sensitivity, achieved by the higher number of sequences generated using Illumina sequencing technology, might lead to comparable results, even when fewer samples were used.

The behavior of diversity estimates for *16S rRNA* gene regions showed that the V4 region presents the highest diversity values at ASV level when comparing with V3 and V3V4. This behavior is supported by experimental results (Sperling et al., 2017). On the other hand, when the V3V4 dataset is separated in groups of control chickens and chickens infected with *Salmonella enteritidis*, control chickens show the same diversity behavior than animals belonging to the other studies, however differences when comparing V3V4 region against V3 and V4 regions were still present. Those differences are likely due to factors such as biases on the primer's specificity, although the V3V4 primers used on the different studies did not have the same sequence, or the resolution obtained by the

inter-specific variability of the amplified region. When comparing diversity estimates between commercial and experimental chickens no difference was found, leading to assume that variables like temperature, food sources, water supply, housing, hygiene, and the number of animals per cage may not have as significant effects as other variables on bacterial diversity, although it is important to highlight that only two studies encompass commercial samples, from which a single study groups 92% of the samples, leading to a potential bias. Despite the large number of variables inherent to each study, the collapse of ASVs tables to genus level allowed to diminish the influence of those variables in beta diversity analysis and led to the clusterization of several samples corresponding to different studies. Although, neither of the metadata categories considered in this metaanalysis could group the samples, leading to assume that beta diversity grouping is driven by multiple categories.

The evaluation of common bacteria, in different percentages of the samples at the genus level, revealed the presence of several bacterial genera with an important role in food conversion. The genera *Faecalibacterium*, *Oscillospira*, *Lactobacillus*, *Clostridium* and *Ruminococcus* were found to be present in more than 50% of the samples. The importance of *Faecalibacterium* has been attributed to their capacity to express acetyl-CoA acetyltransferase and several enzymes involved in the production of butyrate (Polansky *et al.*, 2016). Production of butyrate could also be accomplished by *Clostridium* and *Oscillospira* species and due to the role of butyrate on anti-inflammatory responses and grow performance, the remarkable prevalence of this genera could be explained (Biasato *et al.*, 2018). On the other hand, *Ruminococcus* and *Lactobacillus* have been extensively described as major colonizers of cecum and ileum, and as part of *Ruminococcaceae* family, their role in SCFA production has been proposed (Wang, Lilburn & Yu, 2016; Rinttilä & Apajalahti, 2013; Wei, Morrison & Yu, 2013; Gophna, Konikoff & Nielsen, 2017). Moreover, the ability of *Ruminococcus* to degrade cellulose, can explain its major presence and abundance in the cecum (Deveillard *et al.*, 2004).

Beta diversity results using different studies as variable showed no clear differentiation of samples from the same study, only two studies showed this behavior. These findings differ with the analysis performed by Zou, Sharif & Parkinson (2018), in which the clustering of samples from the same study is strong. Is not clear the potential source of the variations in the results observed, but they might come from the different efforts in cleaning and normalizing the datasets in order to make them as comparable as possible thus reducing the biases inherent to experimental procedures. However, differences in the software used for the analysis and other technical factors could also have an influence in this variation.

Even though this study may have potential biases related to the data availability and different experimental procedures performed by different groups in different countries, it presents a useful model of comparison and baseline for future studies when cecal microbiome composition wants to be assessed (Table S5; Additional File 1). Despite the heterogeneity of the samples, several bacterial genera, with implications in food conversions rate, were reported to be present in an important proportion of the sample.

CONCLUSIONS

Microbial communities' meta-analysis is a very valuable tool to identify commonalities as well as sources of variation that structures the communities in different environmental settings. Our study showed that the first and most important limitation is the availability of the data and thorough and complete description of the metadata associated. Once those factors were taken into consideration, the source of variation can be both technical and experimental, such as the choice of primers to use or the presence of certain pathogens as part of the experimental design; however, no single factor was the major driver of inter-sample variation, although in some cases samples from a given study clearly clustered together.

Finding commonalities among the different datasets depends on the taxonomical level of resolution desired, at phylum level we were able to detect 11 different phyla and a general structure with dominance of Firmicutes, Bacteroidetes and Proteobacteria. At higher taxonomical levels such as genus, although some genera were identified in a majority (>50%) of the samples overall or within studies, few of them were consistently present suggesting a functional redundancy among closely related bacteria. Most of these common bacteria have been associated with growth and health of different hosts. As we improve our methods for generating and reporting metagenomic and metabarcoding studies, more insights into the ecological and functional role of the different members of bacterial communities associated to a given host will be elucidated.

ACKNOWLEDGEMENTS

The authors would like to acknowledge doctors Maryam Chaib, Angela Viviana Peña and Juan Manuel Anzola for their help in the revision of the manuscript. Likewise, we thank the IT Services Department and ExaCore - IT Core-facility of the Vice Presidency for Research & Creation at the Universidad de Los Andes that allow us to perform the computational analysis.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The authors received no funding for this work.

Competing Interests

Martha Vives is an Academic Editor for PeerJ.

Author Contributions

- Luis Alberto Chica Cardenas conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Viviana Clavijo conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the paper, and approved the final draft.

- Martha Vives conceived and designed the experiments, authored or reviewed drafts of the paper, and approved the final draft.
- Alejandro Reyes conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

We didn't develop or generated any raw data or code for the current study, all raw data used was previously published and is referenced and the software used is third party and has already been published.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.10571#supplemental-information>.

REFERENCES

- Allali I, Arnold JW, Roach J, Cadenas MB, Butz N, Hassan HM, Koci M, Ballou A, Mendoza M, Ali R, Azcarate-Peril MA. 2017. A comparison of sequencing platforms and bioinformatics pipelines for compositional analysis of the gut microbiome. *BMC Microbiology* 17(1):194 DOI 10.1186/s12866-017-1101-8.
- Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP, Thompson LR, Hyde ER, Gonzalez A, Knight R. 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. *MSystems* 2(2):e00191-16 DOI 10.1128/mSystems.00191-16.
- Awad WA, Mann E, Dzieciol M, Hess C, Schmitz-Esser S, Wagner M, Hess M. 2016. Age-related differences in the luminal and mucosa-associated gut microbiome of broiler chickens and shifts associated with *Campylobacter jejuni* infection. *Frontiers in Cellular and Infection Microbiology* 6:154 DOI 10.3389/fcimb.2016.00154.
- Biasato I, Ferrocino I, Biasibetti E, Grego E, Dabbou S, Sereno A, Gai F, Gasco L, Schiavone A, Coccolin L, Capucchio MT. 2018. Modulation of intestinal microbiota, morphology and mucin composition by dietary insect meal inclusion in free-range chickens. *BMC Veterinary Research* 14(1):383 DOI 10.1186/s12917-018-1690-y.
- Ballou AL, Ali RA, Mendoza MA, Ellis JC, Hassan HM, Croom WJ, Koci MD. 2016. Development of the chick microbiome: How early exposure influences future microbial diversity. *Frontiers in Veterinary Science* 3(2):1–12 DOI 10.3389/fvets.2016.00002.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120 DOI 10.1093/bioinformatics/btu17.
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolk T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Loftfield E, Lozupone C, Maher M,

- Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Priesse E, Rasmussen LB, Rivers A, Robeson MS II, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, Van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, Von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* 37(8):852–857 DOI 10.1038/s41587-019-0209-9.
- Costa MC, Bessegatto JA, Alfieri AA, Weese JS, Filho JAB, Oba A. 2017. Different antibiotic growth promoters induce specific changes in the cecal microbiota membership of broiler chicken. *PLOS ONE* 12(2):e0171642 DOI 10.1371/journal.pone.0171642.
- Clavijo V, Flórez MJV. 2017. The gastrointestinal microbiome and its association with the control of pathogens in broiler chicken production: a review. *Poultry Science* 97(3):1006–1021 DOI 10.3382/ps/pex359.
- DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi D, Hu P, Andersen GL. 2006. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Applied and Environmental Microbiology* 72(7):5069–5072 DOI 10.1128/AEM.03006-05.
- Devillard E, Goodheart DB, Karnati SKR, Bayer EA, Lamed R, Miron J, Nelson KE, Morrison M. 2004. Ruminococcus albus 8 mutants defective in cellulose degradation are deficient in two processive endocellulases, Cel48A and Cel9B, both of which possess a novel modular architecture. *Journal of Bacteriology* 186(1):136–145 DOI 10.1128/JB.186.1.136-145.2004.
- Godfray HCJ, Beddington JR, Crute IR, Haddad L, Lawrence D, Muir JF, Pretty J, Robinson S, Thomas SM, Toulmin C. 2010. Food security: the challenge of feeding 9 billion people. *Science* 327(5967):812–818 DOI 10.1126/science.1185383.
- Gophna U, Konikoff T, Nielsen HB. 2017. Oscillospira and related bacteria: from metagenomic species to metabolic features. *Environmental Microbiology* 19(3):835–841 DOI 10.1111/1462-2920.13658.
- Han GG, Kim EB, Lee J, Lee JY, Jin G, Park J, Huh CS, Kwon IK, Kil DY, Choi YJ, Kong C. 2016. Relationship between the microbiota in different sections of the gastrointestinal tract, and the body weight of broiler chickens. *SpringerPlus* 5(1):911 DOI 10.1186/s40064-016-2604-8.
- Kien CL, Blauwiel R, Bunn JY, Jetton TL, Frankel WL, Holst J. 2007. Cecal infusion of butyrate increases intestinal cell proliferation in piglets. *Journal of Nutrition* 137(4):916–922 DOI 10.1093/jn/137.4.916.
- Kim S, Rubinelli PM, Park SH, Ricke SC. 2018. Ability of Arkansas LaKast and LaKast hybrid rice bran to reduce Salmonella Typhimurium in chicken cecal incubations and effects on cecal microbiota. *Frontiers in Microbiology* 9:134 DOI 10.3389/fmicb.2018.00134.
- Kollarcikova M, Kubasova T, Karasova D, Crhanova M, Cejkova D, Sisak F, Rychlik I. 2019. Use of 16S rRNA gene sequencing for prediction of new opportunistic pathogens in chicken ileal and cecal microbiota. *Poultry Science* 98(6):2347–2353 DOI 10.3382/ps/pey594.
- Krueger LA, Spangler DA, Vandermyde DR, Sims MD, Ayangbile GA. 2017. Avi-Lution[®] supplemented at 1.0 or 2.0 g/kg in feed improves the growth performance of broiler chickens during challenge with bacitracin-resistant Clostridium perfringens. *Poultry Science* 96(8):2595–2600 DOI 10.3382/ps/pex074.

- La Ragione RM, Woodward MJ. 2003.** Competitive exclusion by *Bacillus subtilis* spores of *Salmonella enterica* serotype Enteritidis and *Clostridium perfringens* in young chickens. *Veterinary Microbiology* **94**(3):245–256 DOI [10.1016/S0378-1135\(03\)00077-4](https://doi.org/10.1016/S0378-1135(03)00077-4).
- Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database Collaboration. 2010.** The sequence read archive. *Nucleic Acids Research* **39**(Suppl. 1):D19–D21.
- Li Y, Xu Q, Huang Z, Lv L, Liu X, Yin C, Yan H, Yuan J. 2016.** Effect of *Bacillus subtilis* CGMCC 1.1086 on the growth performance and intestinal microbiota of broilers. *Journal of Applied Microbiology* **120**(1):195–204 DOI [10.1111/jam.12972](https://doi.org/10.1111/jam.12972).
- Ma X, Wang Q, Li H, Xu C, Cui N, Zhao X. 2017.** 16S rRNA genes Illumina sequencing revealed differential cecal microbiome in specific pathogen free chickens infected with different subgroup of avian leukosis viruses. *Veterinary Microbiology* **207**:195–204 DOI [10.1016/j.vetmic.2017.05.016](https://doi.org/10.1016/j.vetmic.2017.05.016).
- Mancabelli L, Ferrario C, Milani C, Mangifesta M, Turrone F, Duranti S, Lugli GA, Viappiani A, Ossiprandi MC, Van Sinderen D, Ventura M. 2016.** Insights into the biodiversity of the gut microbiota of broiler chickens. *Environmental Microbiology* **18**(12):4727–4738 DOI [10.1111/1462-2920.13363](https://doi.org/10.1111/1462-2920.13363).
- Marangoni F, Corsello G, Cricelli C, Ferrara N, Ghiselli A, Lucchin L, Poli A. 2015.** Role of poultry meat in a balanced diet aimed at maintaining health and wellbeing: an Italian consensus document. *Food & Nutrition Research* **59**(1):27606 DOI [10.3402/fnr.v59.27606](https://doi.org/10.3402/fnr.v59.27606).
- Marimuthu SK, Balasubramanian B, Selvam R, D’Souza P. 2019.** Modulation of chicken cecal microbiota by a phyto-genic feed additive, Stodi®: a metagenomic analysis. *Pharmacognosy Research* **11**(3):201.
- Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma Group. 2009.** Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLOS Medicine* **6**(7):e1000097 DOI [10.1371/journal.pmed.1000097](https://doi.org/10.1371/journal.pmed.1000097).
- Mon KKZ, Saelao P, Halstead MM, Chanthavixay G, Chang HC, Garas L, Maga EA, Zhou H. 2015.** *Salmonella enterica* serovars enteritidis infection alters the indigenous microbiota diversity in young layer chicks. *Frontiers in Veterinary Science* **2**:61 DOI [10.3389/fvets.2015.00061](https://doi.org/10.3389/fvets.2015.00061).
- Oakley BB, Lillehoj HS, Kogut MH, Kim WK, Maurer JJ, Pedroso A, Lee MD, Collett SR, Johnson TJ, Cox NA. 2014.** The chicken gastrointestinal microbiome. *FEMS Microbiology Letters* **360**(2):100–112 DOI [10.1111/1574-6968.12608](https://doi.org/10.1111/1574-6968.12608).
- Park SH, Lee SI, Ricke SC. 2016.** Microbial populations in naked neck chicken ceca raised on pasture flock fed with commercial yeast cell wall prebiotics via an Illumina MiSeq platform. *PLOS ONE* **11**(3):e0151944 DOI [10.1371/journal.pone.0151944](https://doi.org/10.1371/journal.pone.0151944).
- Park SH, Roto S, Pavlidis H, McIntyre D, Striplin K, Brammer L, Ricke SC. 2017.** Effects of feeding Original XPCTM to broilers with a live coccidiosis vaccine under industrial conditions: Part 2. Cecal microbiota analysis. *Poultry Science* **96**(7):2400–2411 DOI [10.3382/ps/pex014](https://doi.org/10.3382/ps/pex014).
- Plummer E, Twin J, Bulach DM, Garland SM, Tabrizi SN. 2015.** A comparison of three bioinformatics pipelines for the analysis of preterm gut microbiota using 16S rRNA gene sequencing data. *Journal of Proteomics & Bioinformatics* **8**(12):283.
- Polansky O, Sekelova Z, Faldynova M, Sebkova A, Sisak F, Rychlik I. 2016.** Important metabolic pathways and biological processes expressed by chicken cecal microbiota. *Applied and Environmental Microbiology* **82**(5):1569–1576 DOI [10.1128/AEM.03473-15](https://doi.org/10.1128/AEM.03473-15).
- Rinttilä T, Apajalahti J. 2013.** Intestinal microbiota and metabolites—implications for broiler chicken health and performance1. *Journal of Applied Poultry Research* **22**(3):647–658 DOI [10.3382/japr.2013-00742](https://doi.org/10.3382/japr.2013-00742).

- Saxena S, Saxena VK, Tomar S, Sapkota D, Gonmei G. 2016. Characterisation of caecum and crop microbiota of Indian indigenous chicken targeting multiple hypervariable regions within 16S rRNA gene. *British Poultry Science* 57(3):381–389 DOI 10.1080/00071668.2016.1161728.
- Scanes CG. 2007. The global importance of poultry. *Poultry Science* 86(6):1057–1058 DOI 10.1093/ps/86.6.1057.
- Sergeant MJ, Constantinidou C, Cogan TA, Bedford MR, Penn CW, Pallen MJ. 2014. Extensive microbial and functional diversity within the chicken cecal microbiome. *PLOS ONE* 9(3):e91941 DOI 10.1371/journal.pone.0091941.
- Shaafi MAM, Siew CC, Chong CW, Gan HM, Ho YW. 2015. Deciphering chicken gut microbial dynamics based on high-throughput 16S rRNA metagenomics analyses. *Gut Pathogens* 7(1):4 DOI 10.1186/s13099-015-0051-7.
- Shen W, Le S, Li Y, Hu F. 2016. SeqKit: a cross-platform and ultrafast toolkit for FASTA/Q file manipulation. *PLOS ONE* 11(10):e0163962 DOI 10.1371/journal.pone.0163962.
- Sperling JL, Silva-Brandão KL, Brandão MM, Lloyd VK, Dang S, Davis CS, Sperling FAH, Magor KE. 2017. Comparison of bacterial 16S rRNA variable regions for microbiome surveys of ticks. *Ticks and Tick-Borne Diseases* 8(4):453–461 DOI 10.1016/j.ttbdis.2017.02.002.
- Varmuzova K, Matulova ME, Gerzova L, Cejkova D, Gardan-Salmon D, Panhéleux M, Robert F, Sisak F, Havlickova H, Rychlik I. 2015. Curcuma and Scutellaria plant extracts protect chickens against inflammation and Salmonella Enteritidis infection. *Poultry Science* 94(9):2049–2058 DOI 10.3382/ps/pev190.
- Varmuzova K, Kubasova T, Davidova-Gerzova L, Sisak F, Havlickova H, Sebkova A, Faldynova M, Rychlik I. 2016. Composition of gut microbiota influences resistance of newly hatched chickens to Salmonella enteritidis infection. *Frontiers in Microbiology* 7:957 DOI 10.3389/fmicb.2016.00957.
- USDA Foreign Agricultural Service. 2019. Livestock and Poultry: World Markets and Trade. Available at https://downloads.usda.library.cornell.edu/usda-esmis/files/73666448x/g445ct12h/ff365k146/Livestock_poultry.pdf (accessed 10 October 2019).
- Waite DW, Taylor MW. 2014. Characterizing the avian gut microbiota: membership, driving influences, and potential function. *Frontiers in Microbiology* 5:223 DOI 10.3389/fmicb.2014.00223.
- Wang L, Lilburn M, Yu Z. 2016. Intestinal microbiota of broiler chickens as affected by litter management regimens. *Frontiers in Microbiology* 7:593.
- Wang J, Nesengani LT, Gong Y, Yang Y, Lu W. 2018. 16S rRNA gene sequencing reveals effects of photoperiod on cecal microbiota of broiler roosters. *PeerJ* 6(11):e4390 DOI 10.7717/peerj.4390.
- Wei S, Morrison M, Yu Z. 2013. Bacterial census of poultry intestinal microbiome. *Poultry Science* 92(3):671–683 DOI 10.3382/ps.2012-02822.
- Wu S, Li T, Niu H, Zhu Y, Liu Y, Duan Y, Sun Q, Yang X. 2019. Effects of glucose oxidase on growth performance, gut function, and cecal microbiota of broiler chickens. *Poultry Science* 98(2):828–841 DOI 10.3382/ps/pey393.
- Xia Y, Kong J, Zhang G, Zhang X, Seviour R, Kong Y. 2019. Effects of dietary supplementation with lysozyme on the structure and function of the cecal microbiota in broiler chickens. *PLOS ONE* 14(6):e0216748 DOI 10.1371/journal.pone.0216748.
- Xiao Y, Xiang Y, Zhou W, Chen J, Li K, Yang H. 2016. Microbial community mapping in intestinal tract of broiler chicken. *Poultry Science* 96(5):1387–1393 DOI 10.3382/ps/pew372.
- Xu Y, Yang H, Zhang L, Su Y, Shi D, Xiao H, Tian Y. 2016. High-throughput sequencing technology to reveal the composition and function of cecal microbiota in Dagu chicken. *BMC Microbiology* 16(1):1–9 DOI 10.1186/s12866-015-0617-z.

- Yan W, Sun C, Yuan J, Yang N. 2017.** Gut metagenomic analysis reveals prominent roles of *Lactobacillus* and cecal microbiota in chicken feed efficiency. *Scientific Reports* 7:45308 DOI [10.1038/srep45308](https://doi.org/10.1038/srep45308).
- Zhou X, Jiang X, Yang C, Ma B, Lei C, Xu C, Zhang A, Yang X, Xiong Q, Zhang P, Men S, Xiang R, Wang H. 2016.** Cecal microbiota of Tibetan Chickens from five geographic regions were determined by 16S rRNA sequencing. *MicrobiologyOpen* 5(5):753–762 DOI [10.1002/mbo3.367](https://doi.org/10.1002/mbo3.367).
- Zou A, Sharif S, Parkinson J. 2018.** *Lactobacillus* elicits a ‘Marmite effect’ on the chicken cecal microbiome. *NPJ Biofilms and Microbiomes* 4(1):1–5 DOI [10.1038/s41522-018-0070-5](https://doi.org/10.1038/s41522-018-0070-5).