Coursework Declaration and Feedback Form

The Student should complete and sign this part

Student Number: 2698734C	Student Name: Youqi Cheng
Programme of Study: MSc in Computer Systems Engineering	
Course Code: ENG5059P	Course Name: MSc Project
Name of **First** Supervisor: Umer Zeeshan Ijaz	Name of **Second** Supervisor: Lukasz Kaczmarczyk
Title of Project: Longitudinal microbiome analysis of healthy chicken ceca	

Declaration of Originality and Submission Information

I affirm that this submission is all my own work in accordance with the University of Glasgow Regulations and the School of Engineering requirements Signed (Student) : *Youqi Cheng*	
Date of Submission : August 19th 2022	

Feedback from Lecturer to Student – to be completed by Lecturer or Demonstrator	
Grade Awarded: Feedback (as appropriate to the coursework which was assessed):	
Lecturer/Demonstrator:	Date returned to the Teaching Office:

University of Glasgow | School of Engineering

Longitudinal microbiome analysis of healthy chicken ceca

Youqi Cheng

2698734C

Supervised by Dr. Umer Zeeshan Ijaz and

Dr. Lukasz Kaczmarczyk

August 18th 2022

Abstract

Chickens offer a substantial amount of protein, and the diverse microbiome of chickens, which plays a crucial part in the digestion of nutrients as well as the creation of vitamins and amino acids, is crucial for gut health [1]. Chicken metabolism and immunological capability are influenced by the growth of the cecum microbiota. Understanding the regular succession inside the cecum microbiota can help guide the selection of probiotics as well as other treatments to enhance the cecum microbiome. Little is known about the microbiome-related factors that affect host-pathogen ecology.

This study examines 16S rRNA sample data that was gathered from day 3 to day 35 of the experiment. The major objectives of this research are to comprehend the evolution of the cecum microbiota and the likely ecological factors that shape the gut population.

The results show that the diversity within samples increases as the birds growth and the microbial community in the chicken's cecum is changing from being driven by competing forces to environmental ones from day08-14 to day15-25 as the chickens mature. The studies indicate that *Lachnospiraceae*, *Ruminococcaceae*, *Oscillospiraceae*, and *Clostridiales vadin BB60 group* are the major Firmicutes in the chicken cecum microbiome. The bird's age and diet have significant influence on microbial diversity.

Early interventions are expected to be greater successful in changing the mature microbiome as well as maximizing the microbiome's impact on the development of the immune and metabolic processes. The goal of interventions should be to encourage early maturity, especially in *Oscillospiraceae*, *Lachnospiraceae*, and *Ruminococcaceae*.

Keywords: chicken, microbiome, diversity, environmental filtering, competitive exclusion

Acknowledgements

I owe my deep gratitude and privilege to my friendly and inspirational supervisor supervisor Dr. Umer Zeeshan Ijaz. His assistance, direction, helpful advice, and conversations made it possible for me to finish this project. I am also thankful to Dr. Uzma Khan, Aqsa Ameer and Dr. Farrukh Saleem.

List of Abbreviations

Abbreviation	Explanation
DNA	Deoxyribonucleic acid
BW_mean	Mean body weight
PCR	Polymerase chain reaction
OTU	Operational Taxonomic Unit
NRI	Nearest relatedness index
MNTD	Mean nearest taxon distance
ANOVA	Analysis of variance
PERMANOVA	Permutational multivariate ANOVA
FI	Feed intake
QPE	Quantitative Process Estimate
βMNTD	β mean nearest taxon distance
NTI	Nearest taxa index
NST	Normalized stochasticity ratio
FCR	Feed conversion ratio
MPD	Mean phylogenetic distance

List of Figures

Figure		Page
1	The alpha diversity for each of group samples	25
2	The results of beta diversity distance measure	25
3	The result of NTI/NRI and NST	28
4	The result of QPE and incidence-based (Raup-Crick) beta-diversity	28
5	The differential analysis in three comparisons in genera level	30
6	The mean expression value of different genera between different comparisons	30
7	Core microbiome that persists in greater than 85 percent of samples for comparing between four groups.	32
8	The top-25 most abundant genera for each group samples	33
9	Heatmap showing major external factors that affect various microbiome traits	33

Contents

Abstract 2

Acknowledgements 3

List of Abbreviations 4

List of Figures 5

1 Introduction 8

1.1 Background 8

1.2 Related Research 10

1.3 Aims and Objectives 12

2 Methods 13

2.1 Data Description 13

2.2 DNA Extraction 14

2.3 Bioinformatics 15

2.4 Statistical Analysis 17

2.4.1 Alpha Diversity 17

2.4.2 Beta Diversity 18

2.4.3 NRI and NTI 20

2.4.4 Null Modelling Approaches: QPE and Incidence-based (Raup-Crick) Beta-diversity 21

2.4.5 Normalized Stochasticity Ratio (NST) Analysis 22

2.4.6 Differential Analysis: DeSeq 22

2.4.7 Core Microbiome Analysis and The Top 25 Most Abundant Genera 23

2.4.8 Direction of Effects for External Factors Determining Microbiome's Key Metrics (Subset Regression Analysis) 23

3 Results 25

3.1 Succession in the Cecum Microbiota 25

3.1.1 Alpha Diversity 25

3.1.2 Beta Diversity 25
3.1.3 NTI and NRI 27
3.1.4 Null Modelling Approaches: QPE and Incidence-based (Raup-Crick) Beta-diversity 27
3.1.5 Normalized Stochasticity Ratio (NST) Analysis 28
3.2 Main Species that Represent the Majority of Community Dynamics Change (Differential Analysis) 30
3.3 Analysis of Prevalent Species Over Time 32
3.4 Direction of Effects for External Factors Determining Microbiome's Key Metrics (Subset Regression Analysis) 35
4 Discussion 36
4.1 Diversity Settle into a Stable Status as Time goes on 36
4.2 A Shift in Microbial Community Drivers from Competitive to Environmental 36
4.3 The Primary Causes of Microbial Community Pattern Variations 36
4.4 Dominant Genera Change over Time 37
5 Conclusions 38
Reference 39
Appendix 46
Appendix I 46
Appendix II 47
Appendix III 48

1 Introduction

1.1 Background

Chickens are one of the most important food resources for humans. Chicken meat is rich in nutrients, with good concentrations of all the essential vitamins (mainly vitamin B) [2]. Chickens are most commonly raised and slaughtered than all other poultry species.

The crop is the first part of the digestive system of a chicken, and it is where food is kept [3]. The proventriculous stomach, also known as the true stomach, comes next, and it ultimately connects to the muscular gizzard [4]. Digesta are mashed up in the muscular gizzard before going into the small and large intestines. [5]. The ceca, a set of sacs that emerge from the chicken large intestine, contain the most densely populated microbial community of any part of the chicken's digestive system [6]. Long villi that intertwine with one another can be found at the opening of each cecum. These villi filter out large particles of material, enabling only soluble or finely divided digesta to pass through into the lumen [6]. Digesta remain in the ceca for 12–20 hours, whereas the time it takes for food to traverse the intestine's top portion is only 2.5 hours [7]. Theoretically, this might give the body more time to digest the food and absorb the nutrients. There have been several theories on the potential functions of wild birds' cecal bacteria. [8]. However, its function in today's commercial broilers is not well understood, and in some trials, the caeca have been removed from chickens without causing any negative effects on the chicken's growth. On the other hand, numerous other processes have been connected to this microbiome, including the recycling of nitrogen through the breakdown of uric acid [9], provide their hosts with B vitamins and create critical amino acids that are required by their hosts respectively.

Chickens carry many pathogenic microbes including *Campylobacter*, causing bacterial foodborne gastroenteritis, and in later stages Guillain-Barre syndrome, development retardation in children, particularly in countries with little resources [10-11]. Despite an evident *Campylobacter* impact and economic burden, it is unclear that how and when it appears in chickens [12-14]. Our knowledge of *Campylobacter* colonization in chicken gut has been

delayed by a paucity of repeatable daily microbiome data, accompanying metadata, and natural infection studies.

Between 200 and 350 different bacterial species may make up the intestinal microbiota of a single chicken [15], and the chicken digestive system currently contains about 640 different bacterial species [16]. Recent research has shown that this diverse range of bacteria performs a number of functions in the host, from metabolic to immunological development [17]. The desire to find helpful bacteria and control their abundance to enhance their effects has arisen as a result of this awareness. However, a deeper comprehension of the typical evolution of the gut microbiota is necessary before effective therapies can be put into place. As a result of their significance in both metabolic and immunological maturation, the caeca are frequently selected as a focus for microbiota research. About 10 percent of the total of a chicken's digestible energy is derived through caecal digestion process, like microbial action to produce short chains of fatty acids [18].The cecum is a crucial location for the colonization of zoonotic pathogens like *Campylobacter* jejuni and is also continually threatened by ambient microorganisms introduced via urodeum and cloaca reflux [19-20].

So it is important to study the succession of the microbiome in cecum of healthy chickens. This study analyses 16S rRNA sample data, which is collected from the experimental period of day 3 to day 35. Understanding the evolution of the cecum microbiota and the probable ecological processes that shape the microbial community are the main goals of this research.

1.2 Related Research

The research community has been closely examining the tripartite interaction the relationship between intestinal microbiota, birds health, and performances in an effort to increase the effectiveness of current microbiome manipulation techniques [21, 22].

For instance, the overexpression of a xylanase gene from in chicken cecum may lead to the development of novel feed ingredients for industrial usage. [23]. More research is required since our understanding of nutrition and how it impacts the intestinal bacteria is still developing.

Another classic example is *Campylobacter* jejuni, which is mostly responsible for human infections, colonizes chicken cecum at rather high densities (> 109 CFU/g), and which, depending on the host's genetics and the strain of the infection, may be dangerous to chickens [24-27]. Although there are numerous intervention studies, we still don't fully understand why *Campylobacter* normally appears at around 2 or 3 weeks into the chicken lifetime on an industrial farm [28–30]. The gut microbiome's normal expansion and flux may be important [31]. The fact that there has been relatively little research done on how the diet of chickens affects the existence of *Campylobacter* in the microbiota of chicken cecum particularly in industrial farming environments. In the study of Ijaz UZ et. al (2018), An in-depth study conducted on the microbiome of the chicken cecum on a daily basis revealed a dramatic rise in microbial diversity up to day 12, and microbial diversity have be observed in the genera and quantity of the different types of microbes. 20th day later, significant stability of the chicken cecum microbiota indicates that the relative variety and species richness of the microbial community have become uniform. Environmental factors, in this case the chicken that is serving as the host, impose a stronger impact on any alterations to the microbial diversity. There is a shift from competing drivers of the microbial community to environmental forces between days 12 and 20, which creates an opening for *Campylobacter* to make its appearance. Day 16, the day following the most notable alterations in metabolic profiles seen over the entire period, was when *Campylobacter* was discovered. Despite the fact that *Campylobacter* was discovered in 25 percent of the cages on day 16th, the study would reasonably expect that it will expand towards other chicks and cages and be discovered on following days. Significant changes in the

samples' beta dispersion and OTU relative abundance, which are often linked to feed changes, suggesting that the food of the chickens is the most likely factor influencing the progression of microbial variety through time. Notably, a sizable number of OTUs between the day 9th and 11th and days 26th and 28th were log2-fold different. Additionally, the beta distribution on days 11–13 was adjusted to reflect variations in the feed given to the finishers by the producers. These results are noteworthy. To determine the full extent of the impact on the cecum microbiome of chickens, additional research studying the effects of varying meal amounts is required.

In terms of the dominated microbiome in the chicken. From days 1-3, the microbiome of chickens develops quickly, with *Enterobacteriaceae* as dominating family. Day 7 onwards, *Firmicutes* become more abundant and diverse taxonomically [32-34]. The environment in which the chickens are raised, their food and water influence the bacterial populations in the chickens' guts [35].

1.3 Aims and Objectives

The aims and objectives of study are:

- To analyze the sample data (OTUs and meta data) of chicken gut microbial communities during the experimental period from day 3 to day 35.
- To investigate the successive development of cecum microbiome.
- To highlight the factors that affect the assembly of microbial communities.
- To comprehend the underlying ecological phenomena that structure the chicken gut microbial communities.

2 Methods

2.1 Data Description

A number of 396 Ross-308 male broiler chickens were used in this investigation. There were 12 pens with 33 chickens in each pen. From day 0 to day 35, birds were reared on a three stage diet. The birds were given starter meals between day 0 and 10, grower diets between days 11 and 25, and finisher diets between days 26 and 35. One chicken from each of 12 pens was randomly selected every 24 hours and put to death in accordance with the ASPA schedule 1 rules. In a nutshell, birds under 250 g were put to death by dislocating their necks, while birds over 250 g and up to 1 kilogram were euthanized after receiving isoflurane anesthesia. An overdose of the anesthetic isoflurane was used to euthanize birds weighing more than 1 kg, followed by the rupture in their neck. The bird's head was covered with an anaesthetic mask during anaesthesia administration so that oxygen and vapourized isoflurane could be given, which was served as conclusive evidence that death had occurred in all of the birds. After this step was completed, genomic DNA, also known as gDNA, was obtained from the chicken cecum. Because of the low quality of their gDNA, a total of 17 of the 396 samples were eliminated from the final analysis, bringing the total number of samples to 379. The following measures of performance were investigated: feed intake (FI), mean body weight (BW mean), body weight increase (Gain), and feed conversion ratio (FCR) [31]. At time points ranging from 3 to 7 days, 8 to 14 days, 15 to 24 days, and 25 to 35 days, measurements were taken. After that, a number of statistical studies were performed, each of which involved correlating these characteristics with the make-up of the microbial community.

2.2 DNA Extraction

The gDNA from the cecum was collected using the QIAamp DNA Stool Mini Kit in accordance with the instructions provided by the manufacturer and then frozen at 20 degrees Celsius. The preparation of the 16S metagenomic sequencing library was carried out in accordance with Illumina's standards (Illumina, U.S.A). The 16S ribosomal primers used were V3 (tcgtcggcagcgtcagatgtgtataagag acagcctacgggnggcwgcag) and V4 (gtctcgtgggctcggagatgtgtataaga gacaggactachvgggtatctaatcc) [36-37]. With the use of the Nextera XT Index kit, a second round of PCR was performed to attach dual indices as well as Illumina sequencing adapters. A v3 300 bp paired-end kit was utilized during the sequencing process that took place at LSHTM using an Illumina MiSeq.

2.3 Bioinformatics

The following construction of OTUs, which acts as a proxy for species, was used to create abundance tables. Sickle v1.200 was used to trim and filter paired-end readings [38]. This was achieved by applying a sliding window methodology and removing locations with an estimated average value under 20. Following this, a length restriction of 10 base pairs was set, and any reads shorter than this threshold were eliminated. After that, the paired-end reads were error-corrected using BayesHammer [39] from of the Spades v2.5.0 assembler. After that, We assembled the forward and reverse reads together into single sequence including the full V3-V4 region using pandaseq (v2.4) with a minimum overlap of 20 bp. [31]. The decision to use the software described above was made as a direct result of the most recent research conducted by the author [37,40], in which it was demonstrated that the strategy described above significantly lowers the substitution rates, which are the most common type of error. After obtaining the consensus sequences for each sample, then constructed OTUs with the help of the VSEARCH (v2.3.4) pipeline (all of these stages are published in https://github.com/torognes/vsearch/wiki/VSEARCH-pipeline) [31]. The following is the method used: first, combine the reads obtained from the various samples, then we append barcodes so that we can record the samples that each read came from. Following deduplication and sorting in decreasing order on abundance, singletons are removed from the dataset. The last stage is clustering the reads based on how similar they are to one another by 97%, and then using the -uchime denovo option in the vsearch software, removing cluster that contains chimeric models created from more plentiful reads. There is a possibility that some chimaeras will be overlooked, particularly if their parents do not appear in the reads or appear in very low abundance if they do appear. Therefore, in the following stage, employ a reference-based chimaera filtering step (the –uchime ref option in vsearch), and we do so with a gold database (https://www.mothur.org/w/images/f/f1/Silva.gold.bacteria.zip) [31]. In this step, the researchers filter the chimaeras based on the references that they have. In order to build an OTU table for n = 379 samples, the initial barcoded readings were compared to clean OTUs with a 97% similarity (which serves as a substitute for the species level separation). This resulted in a

total of 18,588 unique sequences. After that, the Qiime [41] workflow's assign taxonomy.py script was used to taxonomically classify the typical OTUs. This categorization was performed using the release v123 of the SILVA SSU Ref NR database. We first used Kalign v2.0.4 to multisequence align the OTUs with one another [42] (using the options -gpo 11 -gpe 0.85) and then used FastTree v2.1.7 [43] to generate the phylogenetic tree in NEWICK format. This allowed us to determine the phylogenetic distances between OTUs. In the final step, the otu make table. In order to create a biome file for OTUs, the abundance table and the taxonomy data were combined using the script from the Qiime procedure. Based on 16S rRNA information, Tax4Fun was used to forecast the functional capacities of microbial communities (For SILVA v123 and KEGG dataset version 64.0, Tax4Fun has access to all prokaryotic KEGG species) [31], after the data was corrected for 16S rRNA gene copy counts, metabolic functional profiles were generated using the ultrafast protein classification (UProC) algorithm [44].When developing Tax4Fun, we made advantage of the MoP-Pro technique [45] to provide pre-calculated reference profiles for 274 KEGG Pathways. Despite the fact that the number of taxa that can be predicted using Tax4Fun's metabolic model is constrained by the data set, it provides a statistic known as fraction-of-taxonomic-units-unexplained (FTU). The number of sequences that are attributed to a taxonomy unit but cannot be transferred to KEGG reference organisms is represented by this statistic [31]. This may serve as a gauge to determine how much confidence one has in the accuracy of the predictions.

2.4 Statistical Analysis

The analysis of the statistics was done in R using the tables and data that were prepared in the manner described above, together with the study's accompanying meta data.

2.4.1 Alpha Diversity

Alpha diversity quantifies the diversity in a single sample, normally accounting for the variety of species observed. Indicators of alpha diversity can be weighted according to the observed abundances of the various microbes. There are several widely used metrics to represent alpha diversity such as: richness, Shannon entropy and Simpson diversity, which we can calculate by using the vegan package [46].

Richness is the simplest alpha diversity metric, which simply adds up the variety of taxa observed in a sample at a particular taxonomic level, and in this research we used OTU level. Both species richness and evenness can be estimated using Shannon entropy, but richness is given more weight.

Formula for Shannon entropy:

$$\mathrm{H(X)} = -\sum_{x\in\chi} p(x) log_b\, p(x) \qquad (1)$$

Where p(x) is probability, and log() is a typical log function.

Another indicator of evenness within communities is Simpson diversity, and the resulting scores range from 0 to 1. Where scores around 1 means high diversity and scores near 0 imply poor diversity [47]. The calculation of the Simpson diversity as follows:

$$\mathrm{D} = 1 - \frac{\sum n(n-1)}{N(N-1)} \qquad (2)$$

Where: n Indicates number of specimens of each species and N means overall amount of specimens of all species [47].

2.4.2 Beta Diversity

Beta diversity was assessed using three different distance metrics in order to study the dietary disparities that exist between communities. The ratio of the regional species diversity to the local species diversity is referred to as beta diversity. In other words, it determines how far apart each sample pair is from one another and measures the distance. To begin, Bray-Curtis distances [48] were computed. These distances determine whether or not populations are largely different from one another in the aspects of the abundance counts of OTUs. After that is Unifrac distances, which were determined by Lozupone and Knight (2005), were used to establish whether or not on the basis of phylogeny alone, populations differ greatly. The Unifrac approach involves determining the phylogenetic distance between any two OTUs that appear together in a phylogeny. A branch is considered to be shared if it leads to an OTU from both of the samples, whereas a branch is considered to be unshared if it leads to OTUs that are unique to a single samples [49]. We made use of the phyloseq [50] package in order to do the calculation of Unifrac distances, which take into consideration phylogenetic proximity. The following nonmetric distance measures were utilized in the Nonmetric Distance Scaling (NMDS) visualization of the community data (OTUs) generated by Vegan's metamds() function: Unweighted Unifrac considers the phylogeny distance between both the branching lengths of OTUs observed in various samples, but disregards the abundances, and Weighted Unifrac is an unweighted unifrac distance that takes OTU abundance into account [31]. Bray-Curtis takes into account the species richness counts. The phylogenetic distance between branching lengths of OTUs seen in various samples is taken into account by Weighted Unifrac. The samples have been categorized according to the various treatments, and the mean ordination value and distribution of points have also been determined (the ellipses representing the 95% confidence interval of the standard errors of the groups were produced using Vegan's ordiellipse() function) [31]. After this, the phylogenetic distance between the samples is determined by dividing the sum of all the branch lengths that are shared by the sum of all the branch lengths that are not shared. According to Lozupone and Knight (2005) The Weighted unifrac distances [51] were determined. These distances take into account both the abundance counts of OTUs as well as the phylogenetic distances between the species [51]. The phyloseq software was utilized in

order to perform the calculations for unifrac and weighted unifrac [50].

2.4.3 NRI and NTI

To find if phylogenetic overdispersion or clumping within each population, environmental filtering was performed for the samples taken from chicken cecum. The nearest taxa index (NTI) and nearest relatedness index (NRI) were used to characterise phylogenetic distances . The phylogenetic tree's various branches have varied environmental filtering processes, which are reflected by the NTI and NRI [52]. The mean nearest taxon distance (MNTD) and its closest relative in the phylogenetic tree serve as the foundation for the NTI. NTI depicts overdispersion or clustering at the phylogenetic tree's tip [52]. The mean phylogenetic distance (MPD) is the average distance between each pair of OTUs within a sample and indicates clustering or overdispersion over the whole phylogeny, serves as the foundation for the NRI [52].We calculated mean-nearest-taxon-distance (MNTD) and the nearest taxon index (NTI) by using 'ses.mntd' and 'mntd' in package 'picante'. and using ses.mpd() and mpd() function from the picante [53] package to calculate mean phylogenetic diversity (MPD) and nearest relativeindex (NRI). NTI has the opposite relationship of what "ses.mntd" produces. MNTD can be used to calculate the phylogenetic separation among each OTU inside a sample and its closest relative. After that, these phylogenetic distances are averaged using the abundance weighted mean.

$$\text{MNTD} = \sum_{i_k=1}^{n_k} f_{i_k} \min(\Delta i_k j_k) \qquad (3)$$

In this equation, f_{ik} is relative abundance of OTU i in community k and n_k is order of OTUs in k [54].

Environmental filtering results in large phylogenetic clustering when NRI or NTI values recovered as >0. The environment has little influence on the OTUs that are found when NRI or NTI values are less than 0, which indicates phylogenetic over dispersion. When environmental competition affects the OTUs in a sample, this is referred to as phylogenetic overdispersion, and the species found within samples will have closer genetic ties than would be expected by chance. [55]. NTI greater than +2 for a single community signifies that coexist taxa are more strongly linked than would be predicted by chance (phylogenetic clustering). Coexisting taxa are more weakly correlated than would be expected by chance if NTI was less than -2. (phylogenetic overdispersion). If the average NTI across all communities is statistically

different from 0, there is either overdispersion or clustering.

In summary, the value of NTI/NRI can help to determine if the microbial community structure is deterministic (environmental filtering) or stochastic (competitive exclusion). If the value >0, there is environmental filtering factor in community assembly while value <0 means there is competitive exclusion factor in community assembly. Value greater than 2 or less than 2 means strong environmental filtering or strong competitive exclusion [56].

2.4.4 Null Modelling Approaches: QPE and Incidence-based (Raup-Crick) Beta-diversity

The null modelling approach was conducted to investigate the ecological factors influencing the microbiome community in each group's chicken cecum. Here conducted null modelling approaches in two ways. One method is to compute Quantitative process estimates to measure the relative importance of possible species grouping, dispersal limitation, drift, and mass effects. The other measure is to calculate incidence-based (Raup-Crick) dissimilarity indices (βRC) to test whether communities are assembled stochastically or deterministically.

We firstly calculate how far the βMNTD (β mean nearest taxon distance) diverged from the mean of the null distribution and analysed significance by using the β-Nearest Taxon Index [57]. If the observed MNTD value is substantially higher (βNTI > 2) or lower (βNTI < −2) than the null prediction, then variable or homogenous selection can be used to assemble the community. The observed changes in phylogenetic community structure should have the result in dispersion limitation, homogenizing dispersal (mass effect), or random drift if there is no significant departure from the null expectation. In the second stage, pairwise Bray-Curtis dissimilarity (βRCbray) is used to determine the abundance-based (Raup-Crick) beta-diversity in order to estimate the relative relevance of these processes [58]. We can infer from the calculated RCbray that communities that were not assembled via selection and were not chosen in the first phase were shaped by: (i) If RCbray>+0.95, dispersal limitation and drift are combined. (ii) If RCbray +0.95, homogenising dispersal may either represent the "true" impacts of dispersal limitation. (iii) If RCbray is between 0.95 and +0.95, random processes operating on their own.

The second measure investigated incidence-based (Raup-Crick) beta-diversity (βRC), which does not take into account phylogeny but instead examines for stochastic and deterministic processes in samples by examining the existence and not existence of OTUs in an abundance table [59]. The community is regarded as stochastically constructed when βRC is not considerably different from 0. βRC values close close to +1 suggest that deterministic processes prefer diverse communities, βRC values close to -1 suggest that communities are deterministically formed and more similar to one another than expected by chance [60].

2.4.5 Normalized Stochasticity Ratio (NST) Analysis

Similar to QPE, the analysis of normalized stochasticity ratio (NST) was conducted to reveal the ecological factors influencing the microbiome community, which supports any sort of incidence-based (presence-absence) or abundance-based beta-diversity measure and provides a percentage of stochasticity of samples falling within a single category (Exploration of marine bacterioplankton community assembly mechanisms during chemical dispersant and surfactant-assisted oil biodegradation (wiley.com)). The NST for each treatment was determined to assess the real contribution of determinism in relation to stochasticity using incidence-based Jaccard metrics and null model algorithms of Taxa-Richness constraints of proportional-fixed (PF) [61]. To perform this, use the function of 'tNST' from NST package in R [62][66]. When NST is 0%, the microbial community assembly is entirely stochastic, and when NST is 100%, it is entirely deterministic. To see if the amount of stochasticity changed over time, we calculated the NST of the four groups: day03-07, day08-14, day15-24, day25-3.

2.4.6 Differential Analysis: DeSeq

There were three pairs of comparisons: Day03-07 + Day08-14, Day08-14 + Day15-24, Day15-24+ Day25-35. To identify the main difference within the three pairs, the DeSeq analysis was conducted by using the function of DeSeqDataSetFromMatrix() in the DeSeq2 package (Love et al., 2014) [63]. In order to enable a more quantitative analysis of comparison data utilizing

shrinkage estimators for dispersion and fold change, DESeq2 integrates methodological advancements with a number of unique features (Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2). In order to obtain the maximum likelihood estimates for the OTUs log fold change across two groups, a negative binomial GLM is applied to the data in this analysis [64]. The log fold change was set at 2, and the significance threshold for the adjusted p-value was set at 0.005. If the |log2 fold change| >2, then that microbe between two condition is significantly different.

2.4.7 Core Microbiome Analysis and The Top 25 Most Abundant Genera

To make a visual comparison of how microbial taxonomic abundance within the samples varied among the four groups, the 25 most abundant genera were computed for day03-07,day08-14, day15-24, day25-35. The core microbiome analysis revealed the genera of core microbiome which persist in 85% of samples for various groups. The abundances of genera are ordered, with low abundant genera at the top and high abundant prevalent genera at the bottom.

2.4.8 Direction of Effects for External Factors Determining Microbiome's Key Metrics (Subset Regression Analysis)

We performed subset regressions to find out whether an increase or reduction in these factors caused an increase or decrease in the microbiome's characteristics (Influence of chicken microbiomes on the production system characteristics of industrial chicken farms: pathways for enhanced performance) [65, 67]. These subset regressions discretized all of the available subsets of the independent variable (external forces that were taken into consideration in this study) by ranking them according to their statistical fit following cross-validation (Appendix I, II, III). Take note of the orange and blue backgrounds; the orange ones indicate a positive association with the significances in the regression model, whereas the blue ones indicate a negative relationship with the significances in the model. All of the categorical variables were given a tag that indicated their presence or absence and were then included in the regression model. This was done so that it could be determined whether the inclusion of category variables or their exclusion had an effect

on the final model.

3 Results

3.1 Succession in the Cecum Microbiota

3.1.1 Alpha Diversity

As the samples was divided in to four groups, the alpha diversity analysis showed how diversity varies among day03-07, day08-14, day15-24, day25-35 (Figure 1). The higher the value, the more diverse the community is. The diversity within samples rapidly increase in day03-07 to day08-14, and grow slowly over the next time, which means that as the chicken growth, the OTU abundance are upregulated and settle to stable community when the birds have been mature.

3.1.2 Beta Diversity

In terms of the differences between periods that are driven by abundance count, the sequences in the first three periods stay far away from one other (Figure 2). This indicates that the diversity between samples in day03-07, day08-14, and day15-24 is larger than it is in the most recent period (day25-35). During the time span of day 15 to day 24, the sequences that were present clustered relatively closely together and had less diversity overall. In conclusion, the level of diversity between samples continues to rise with increasing age up until the third week, after which point the microbial communities appear to become more stable.

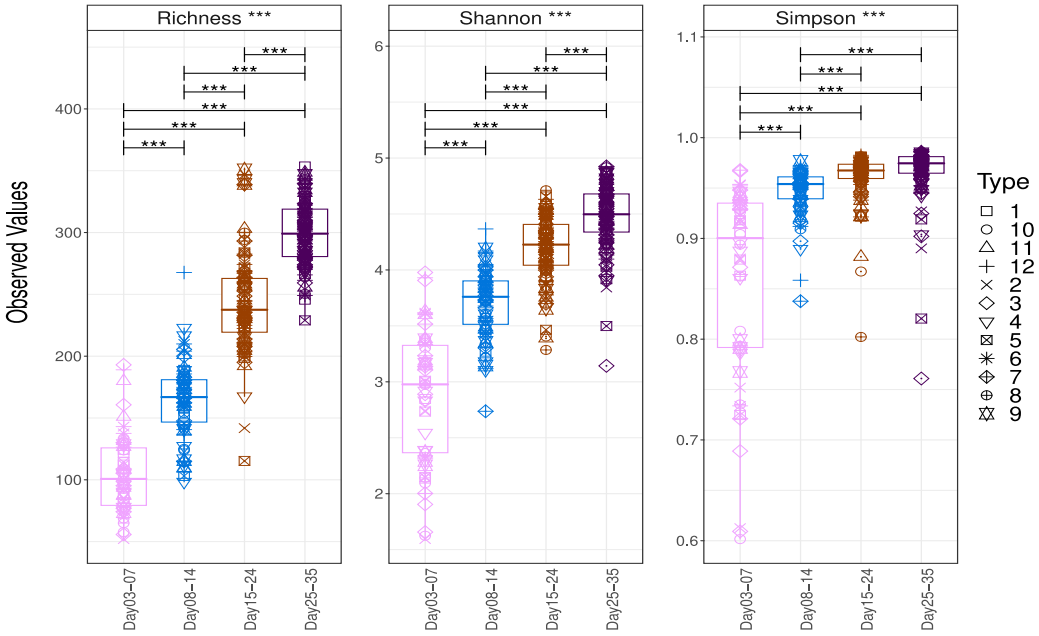


Figure 1: *The alpha diversity for each of group samples*

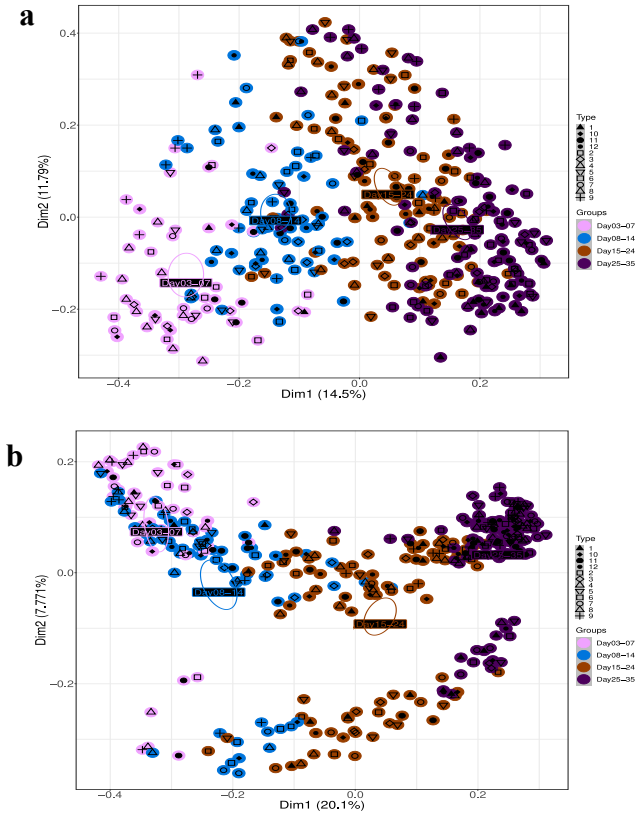


Figure 2: *a) The beta diversity using Bray-curtis distance measure*
b) The beta diversity using Unweighted UniFrac distance measure

3.1.3 NTI and NRI

The samples were divided into 4 groups: day03-07, day08-14, day15-24, day25-35. For all groups, the NRI and NTI were present as >0, which means there is strong deterministic factors, driven by environmental filtering was present in microbial community assembly (Figure 3a). It is possible to rule out the presence of OTUs based on stochastic procedures in all groups because the OTUs within each group are more closely related to one another than would be predicted by chance. But in the process of day15-24 to day25-35, the value of NTI and NRI became decreasing, which means as the birds age, the effect of environmental filerting is decreasing, allowing random processes to come and play.

3.1.4 Null Modelling Approaches: QPE and Incidence-based (Raup-Crick) Beta-diversity

On Figure 4b, the value of group 'day25-35' is −1 indicate that this community is deterministically assembled and more similar to each other than expected by chance. The values of last 3 groups are closer to -1 than group 'day03-07' revealing that the communities within last 3 groups were more similar to each other than of day03-07.

Figure 4a shows that in day03 to day07, the undominated was the most dominated assembly processes (88.24%) while dispersal limitation was the second most dominant assembly processes (11.43%). For the remaining groups, the dispersal limitation was the most dominated assembly processes while undominated was the second most dominant assembly processes. As the birds growth, the role of dispersal limitation is increasing while the role of undominated is decreasing.

3.1.5 Normalized Stochasticity Ratio (NST) Analysis

As seen in Figure 3b, the NST generally showed that the microbial community assembly in the samples from the four groups was neither entirely deterministic nor purely stochastic. The NST value of group ‘day03-07’ to ‘day08-14’ was increasing from 68% to 88% suggesting that in the early days, there were more chances of random processes. Early colonizers may inhibit or compete subsequent microbes, playing role in community assembly. The in the remaining days, the NST value present in decreasing trend from 59% to 41% which revealed that there were deterministic factors. In summary, overall stochasticity is decreasing as the birds age. There is more role of competitive exclusion in initial weeks, later deterministic or environmental factors started playing their role.

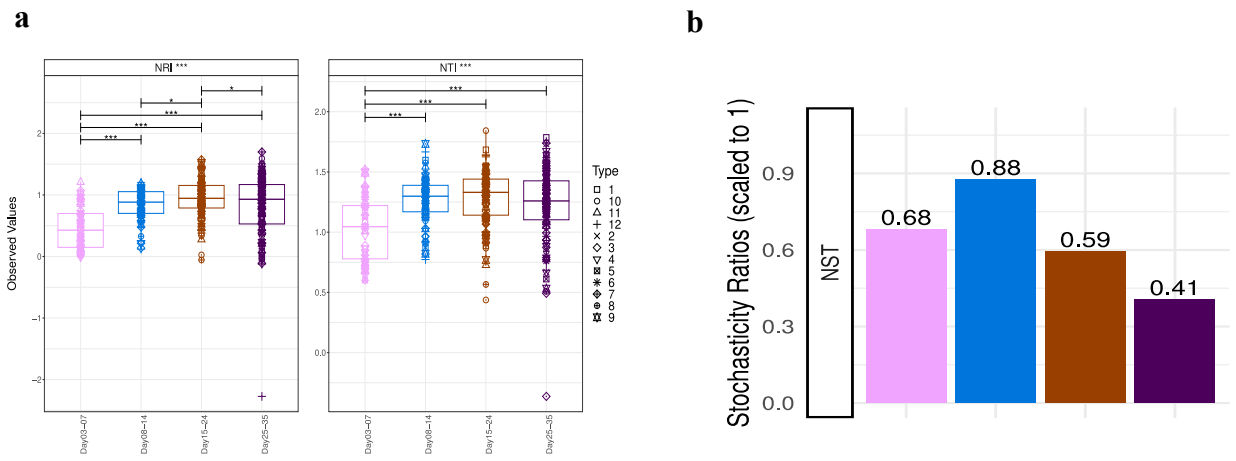


Figure 3: *a) The NTI/NRI measure for each group samples*
b) The Normalized stochasticity ratio (NST) analysis calculated as Jaccard with PF being the null modelling regime used.

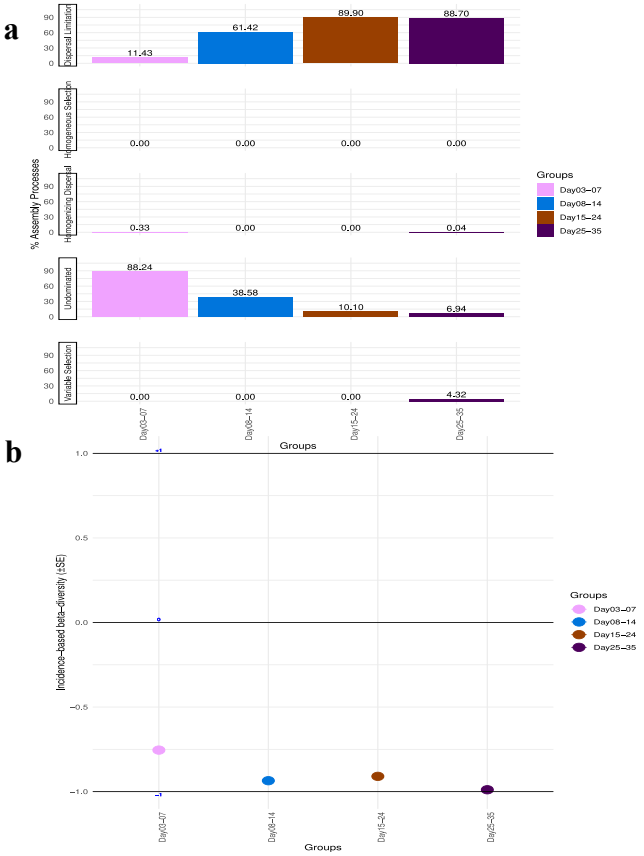


Figure 4: *a) The Quantitative Process Estimate (QPE) result*
b) The incidence-based (Raup-Crick) beta-diversity result

3.2 Main Species that Represent the Majority of Community Dynamics Change (Differential Analysis)

Figure 5 demonstrates that as the birds mature, there is a significant shift in the microbiome from days 8 to 14 and 15 to 24 with regard to the genus levels. The mean value of the majority of microbial species rises throughout time, with the exception of several genera, including *Enterococcus*, *Streptococcus*, and *Paenibacillus* from Day03-07 to Day08-14, *Candidatus Arthromitus* from Day08-14 to Day15-24, and *Anaerotruncus* from Day15-24 to Day25-35. They are all exhibiting a decreasing trend as birds aged.

Oscillospiraceae is the major genus that has changed from Day03-07 to Day08-14, while *Bifidobacterium* and *Angelakisella* are the major genuses that have changed from Day08-14 to Day15-24. *Butyricicoccaceae* and *Anaerostignum* are the two primary genera that have undergone significant shifts from Day08-14 to Day15-24 (Figure 6).

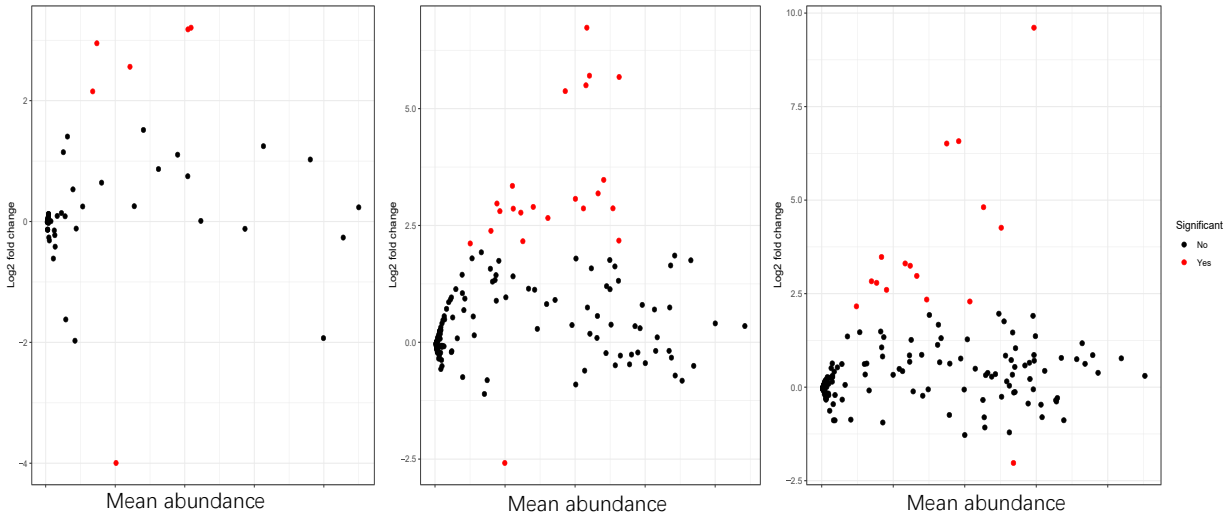


***Figure 5**: The differential analysis in three comparisons in genera level (From left to right: day03-07 vs day08-14, day08-14 vs day15-24, day15-24 vs day25-35)*

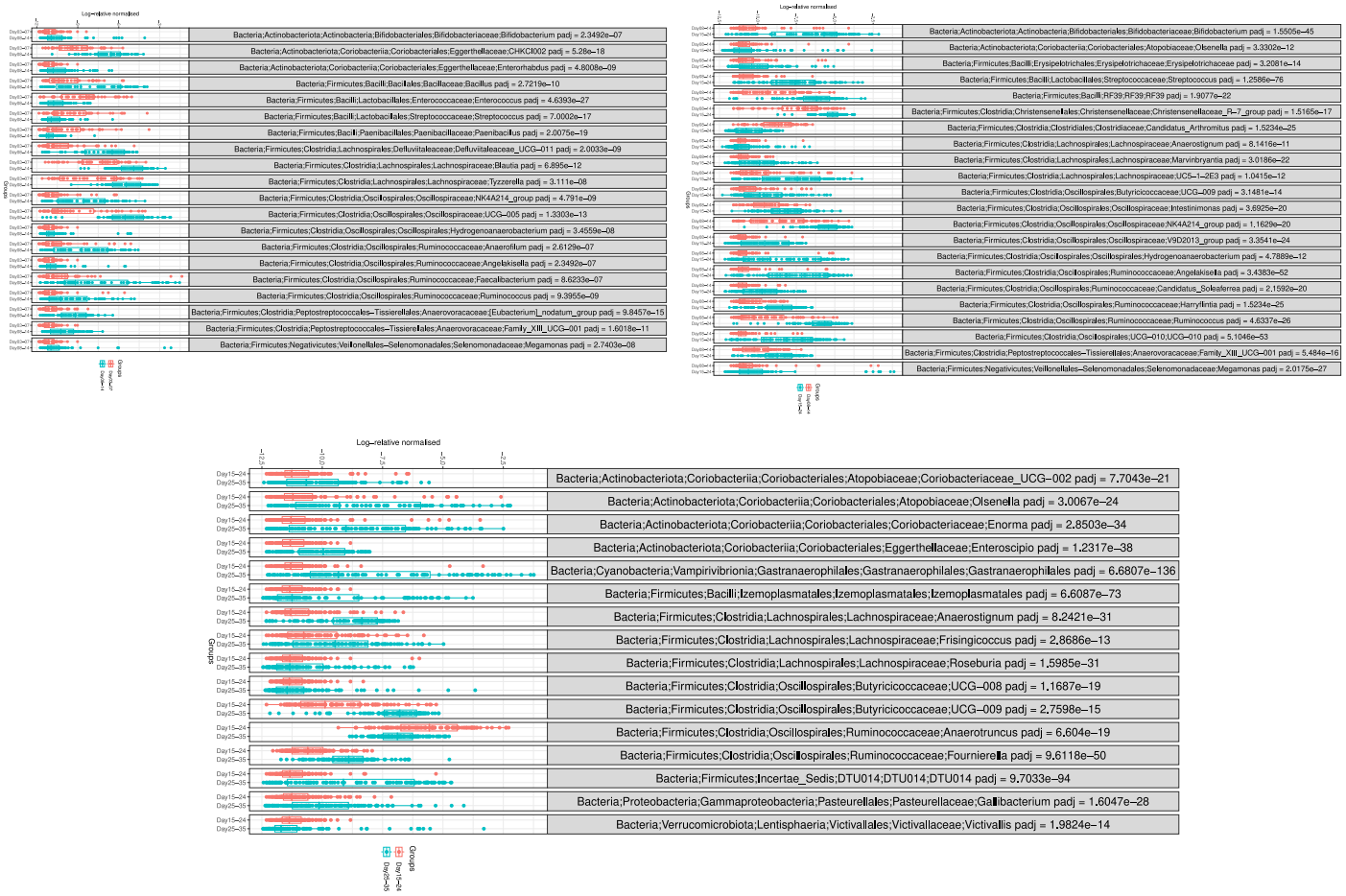


***Figure 6**: The mean expression value of different genera between different comparisons (From left to right: day03-07 vs day08-14, day08-14 vs day15-24, day15-24 vs day25-35)*

3.3 Analysis of Prevalent Species Over Time

By observing the Figure 6, we can observe changes in following groups:

1) *Escherichia Shigella*: In Day03-07, it was the highest abundant core microbiome. But its abundance in the samples continued to decrease over the following dates. There are four pathogenic species of Shigella that are known to exist: S. flexneri, S. dysenteriae, S. sonnei, and S. boydii. An infection with S. flexneri typically develops into the most severe type of dysentery, which can lead to potentially life-threatening consequences [68]. The development of the immune system and host regulation of the microbiota may be the cause.
2) *Oscillibacter*: The major genus that have changed in Day03-07 to Day08-14, which shows upwards trend and was the highest abundant core microbiome in Day08-14. *Oscillibacter*, a member of *Clostridium cluster IV*, has been found to produce valerate as an anaerobe and has been linked to diet-induced obesity (i.e. Birds growth).
3) *Clostridia vadinBB60 group*: The highest abundant core microbiome in Day15-24 and Day25-35, which present in increasing trend in the whole period. But very little is documented about their metabolism or function in the microbiome, and they are poorly categorised.
4) *Ruminococcaceae* and *Lachnospiraceae*: Besides *Oscillospiraceae* (*Oscillibacter*), *Lachnospiraceae* and *Ruminococcaceae* were substantially more numerous. There is little information available about these two families' potential roles in the microbiota, and their classification is poor.

Our results demonstrated that the Firmicutes is predominant in the chicken cecum microbiome, largely consisting of *Ruminococcaceae*, *Lachnospiraceae* , *Oscillospiraceae* and *Clostridiales vadin BB60 group* (Figure 8). In regards of modifying the mature microbiota and maximising the microbiome's influence on immune system maturation and metabolic function, earlier interventions in the microbiota are probably more successful.

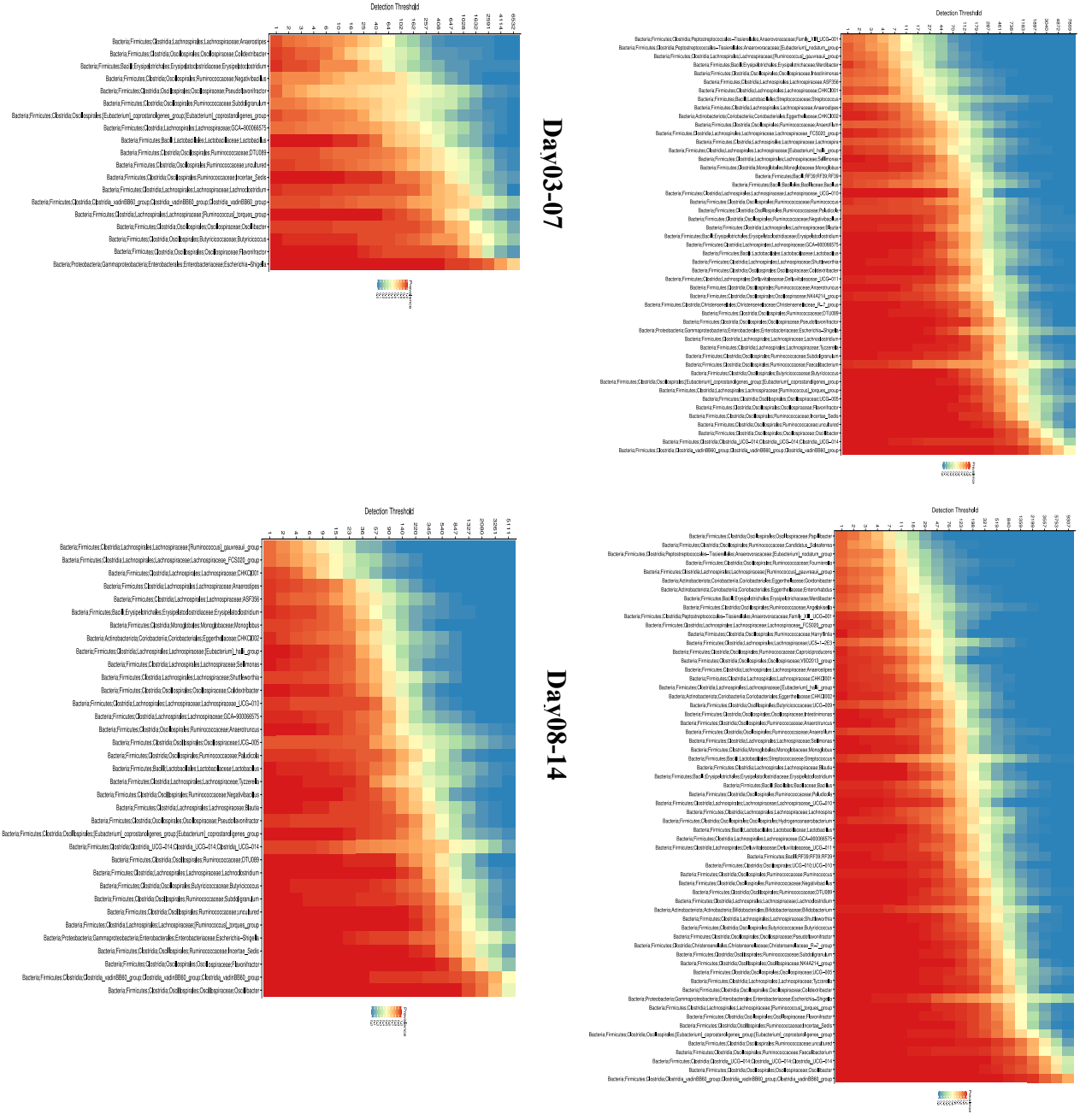


***Figure 7**: Core microbiome that persists in greater than 85 percent of samples for comparing between four groups.*

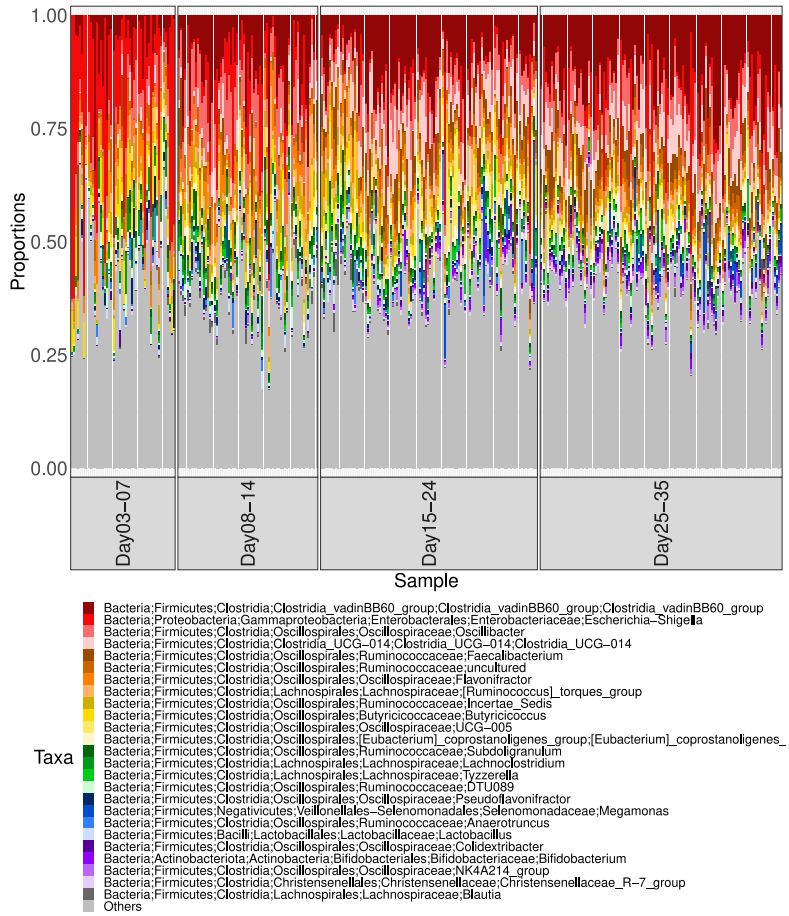


***Figure 8**: The top-25 most abundant genera for each group samples*

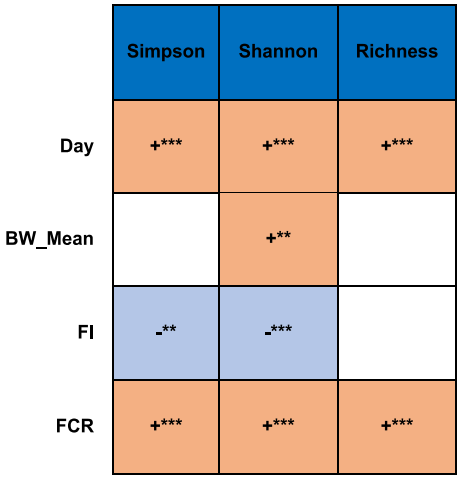


***Figure 9**: Heatmap showing major external factors that affect various microbiome traits, The figure is based on subset regressions (Appendix 1), With orange and blue stand for the relevant positive and negative beta coefficients which were regularly chosen from different regression models.*

3.4 Direction of Effects for External Factors Determining Microbiome's Key Metrics (Subset Regression Analysis)

The Figure 9 is the heatmap of key extrinsic parameters that influence different attributes of microbiome, The figure is based on subset regressions (Appendix I, II, III), where orange and blue represent the significant positive and negative beta coefficients that were consistently selected indifferent regression models. An increase in day led to a positive effect on Simpson, Shannon and Richness diversity, while FI leads to opposite influence on Simpson and Shannon diversity. As for the FCR, it has the same impact as day on the alpha diversity of the samples, mainly because that the day and FCR both have close ties with the growth of the birds. In terms of the mean body weight of the chickens, it plays the roles mainly by influencing the Shannon diversity of the samples.

4 Discussion

4.1 Diversity Settle into a Stable Status as Time goes on

Considering the alpha diversity and NTI figures, diversity within samples is increasing with bird's age, more role of environmental filtering there in causing diversity respectively. Diversity between samples is also increasing with age up to 3rd week, after that microbial communities seems to get consistent. The GI tract of chickens is exposed to foreign bacteria as soon as they hatch, and as the hosts grow, this microbiome becomes increasingly varied before reaching a relatively stable but dynamic state, according to earlier research [69].

4.2 A Shift in Microbial Community Drivers from Competitive to Environmental

According to NST results, there is more role of stochasticity at first, in shaping the microbial community, mainly dispersal limitation as observed in QPE. It is believed that the early variation is caused by competitive elements influenced by available food resources and spatial constraints. The relative microbial diversity are standardized in the chicken cecal microbiome after day 25, with external influences (in this example, the host chicken) having a higher impact on the changes in the microbial diversity.

4.3 The Primary Causes of Microbial Community Pattern Variations

The diet and age are the two key features influencing the microbial community's ecology, as shown by the various correlation analysis between changes in FCR, day and OTU abundances. To fully understand their impact on the chicken cecal microbiota, additional research examining various feed contents is needed.

4.4 Dominant Genera Change over Time

Our results demonstrated that the *Firmicutes* is predominant in the chicken cecum microbiome, largely consisting of *Ruminococcaceae, Lachnospiraceae* , *Oscillospiraceae* and *Clostridiales vadin BB60 group* .

Chicks that are produced for commercial purposes are extremely vulnerable to enteric infections; however, their resistance can be improved through the injection of complex adult microbiota. On the other hand, it is unknown which adult members of the microbiota are able to colonize the caecum of newly hatched chicks [70]. The results may answer this question in some extent. The *Escherichia Shigella* present in decreasing trend mostly in early days (first week), mainly because of the host regulation of the microbiota and maturation of the immune system. The *Oscillibacter*, which was the highest abundant core microbiome in Day08-14, has been identified as an anaerobe producer of valerate and associated with diet-induced obesity (i.e. Birds growth). In the next times, the *Ruminococcaceae* and *Lachnospiraceae* began to dominate the community.

There's a study about the microbiome changes in chicken's fecal [71]. During the time period that was investigated, the fecal microbiomes of the developing chickens demonstrated a linear-log increase in the community richness as well as continuous alterations in the community composition. There were found to be three distinct stages of succession: *Streptococcus* and *Escherichia/Shigella* were among the vertically transmitted or quickly colonizing taxa that dominated the first stage. These taxa were then replaced by *Lachnospiraceae* and *Ruminococcus*-like species variants in the second stage, which began on day 4, and slow-growing, specialized taxa like *Candidatus Arthrobacter* and *Romboutsia* were discovered in the third stage, which began on day 10. Our findings may also be supported by alterations in the microbiota found in the feces of chickens.

5 Conclusions

The host age and diet are the key features in shaping the gut microbial community. Early microbiota interventions are expected to be more effective, both in terms of changing the mature microbiota and maximizing the microbiome's influence on metabolic and immune system development. The goal of interventions should be to encourage early maturity, especially in *Oscillospiraceae*, *Lachnospiraceae*, and *Ruminococcaceae*.

Reference

1. Zhang, X., Akhtar, M., Chen, Y. et al. Chicken jejunal microbiota improves growth performance by mitigating intestinal inflammation. Microbiome 10, 107. https://doi.org/10.1186/s40168-022-01299-8 (2022).
2. Xiao, Sha-Sha, Jian-Dui Mi, Liang Mei, Juanboo Liang, Kun-Xian Feng, Yin-Bao Wu, Xin-Di Liao, and Yan Wang. Microbial Diversity and Community Variation in the Intestines of Layer Chickens. Animals 11, no. 3: 840. https://doi.org/10.3390/ani11030840 (2021).
3. Zhou W, Wang Y, Lin J. Functional Cloning and Characterization of Antibiotic Resistance Genes from the Chicken Gut Microbiome. Appl Environ Microb 78: 3028–3032 (2012).
4. Huyghebaert G, Ducatelle R, Van Immerseel F. An update on alternatives to antimicrobial growth promoters for broilers. Vet J 187: 182–188 (2011).
5. Sergeant MJ, Constantinidou C, Cogan TA, Bedford MR, Penn CW, Pallen MJ. Extensive Microbial and Functional Diversity within the Chicken Cecal Microbiome. PLoS ONE 9(3): e91941. https://doi.org/10.1371/journal.pone.0091941 (2014).
6. Clench MH, Mathias JR. The Avian Cecum – a Review. Wilson Bull 107: 93–121 (1995)
7. Karasawa Y. Significant role of the nitrogen recycling system through the ceca occurs in protein-depleted chickens. J Exp Zool 283: 418–425 (1999).
8. Obst BS, Diamond JM. Interspecific Variation in Sugar and Amino-Acid-Transport by the Avian Cecum. J Exp Zool: 117–126 (1989).
9. Karasawa, Y.. Significant role of the nitrogen recycling system through the ceca occurs in protein‐depleted chickens. Journal of Experimental Zoology Part A: Ecological Genetics and Physiology, 283(4‐5), 418-425 (1999).
10. Ijaz, U. Z., Sivaloganathan, L., McKenna, A., Richmond, A., Kelly, C., Linton, M., ... & Gundogdu, O.. Comprehensive longitudinal microbiome analysis of the chicken cecum reveals a shift from competitive to environmental drivers and a window of opportunity for Campylobacter. Frontiers in microbiology, 9, 2452 (2018).
11. Amour, C., Gratz, J., Mduma, E., Svensen, E., Rogawski, E. T., Mcgrath, M., et al..

Epidemiology and impact of campylobacter infection in children in 8 low-resource settings: results from the MAL-ED study. Clin. Infect. Dis. 63, 1171–1179. doi: 10.1093/cid/ciw542 (2016).

12. Neill, S. D., Campbell, J. N., and Greene, J. A.. Campylobacter species in broiler chickens. Avian Pathol. 13, 777–785. doi: 10.1080/03079458408418574 (1984).
13. Klaus, B., and Strimmer, K.. Signal identification for rare and weak features: higher criticism or false discovery rates? Biostatistics 14, 129–143. doi: 10.1093/biostatistics/kxs030 (2013).
14. Thibodeau, A., Fravalo, P., Yergeau, E., Arsenault, J., Lahaye, L., and Letellier, A.. Chicken caecal microbiome modifications induced by Campylobacter jejuni colonization and by a non-antibiotic feed additive. PLoS ONE 10:e0131978. doi: 10.1371/journal.pone.0131978 (2015).
15. Sergeant MJ, Constantinidou C, Cogan TA, Bedford MR, Penn CW, Pallen MJ. Extensive microbial and functional diversity within the chicken cecal microbiome. PLoS ONE. 9:e91941. doi: 10.1371/journal.pone. 0091941 (2014).
16. Apajalahti J, Kettunen A, Graham H. Characteristics of the gastrointestinal microbial communities, with special reference to the chicken. World's Poult Sci J. 60:223–32. doi: 10.1079/WPS20040017 (2004).
17. Oakley BB, Lillehoj HS, KogutMH, KimWK, Maurer JJ, Pedroso A, et al. The chicken gastrointestinal microbiome. FEMS Microbiol Lett.10:100–12. doi: 10.1111/1574-6968.12608 (2014).
18. Jozefiak D, Rutkowski A, Martin SA. Carbohydrate fermentation in the avian ceca: a review. Anim Feed Sci Technol.133:1–15. doi: 10.1016/j.anifeedsci.2003.09.007 (2004).
19. Crhanova M, Hradecka H, Faldynova M, Matulova M, Havlickova H, Sisak F, et al. Immune response of chicken gut to natural colonization by gut microflora and to Salmonella enterica serovar Enteritidis infection. Infect Immun. 79:2755–63. doi: 10.1128/IAI.01375-10 (2011).
20. Paper Braun EJ, Campbell CE. Uric acid decomposition in the lower gastrointestinal tract. J Exp Zool. 252:70–4. doi: 10.1002/jez.1402520512 (1989).
21. Connerton PL, Richards PJ, Lafontaine GM, O'Kane PM, Ghaffar N, Cummings NJ, Smith

DL, Fish NM, Connerton IF. The effect of the timing of exposure to Campylobacter jejuni on the gut microbiome and inflammatory responses of broiler chickens. Microbiome.6:88 (2018).

22. Huyghebaert G, Ducatelle R, Van Immerseel F. An update on alternatives to antimicrobial growth promoters for broilers. Vet J. 187:182–8 (2011).
23. AL-Darkazali H, Meevootisom V, Isarangkul D, Wiyakrutta S. Gene expression and molecular characterization of a xylanase from chicken cecum metagenome. International journal of microbiology. 2017 (2017).
24. Hermans D, Pasmans F, Heyndrickx M, Van Immerseel F, Martel A, Van Deun K, Haesebrouck F. A tolerogenic mucosal immune response leads to persistent Campylobacter jejuni colonization in the chicken gut. Crit Rev Microbiol.38:17–29 (2012).
25. Humphrey S, Chaloner G, Kemmett K, Davidson N, Williams N, Kipar A, Humphrey T, Wigley P. Campylobacter jejuni is not merely a commensal in commercial broiler chickens and affects bird welfare. MBio.5:e01364–14 (2014).
26. Van Deun K, Pasmans F, Ducatelle R, Flahou B, Vissenberg K, Martel A, Van den Broeck W, Van Immerseel F, Haesebrouck F. Colonization strategy of Campylobacter jejuni results in persistent infection of the chicken gut. Vet Microbiol.130:285–97 (2008).
27. Wigley P. Blurred lines: pathogens, commensals, and the healthy gut. Front Vet Sci.2:40 (2015).
28. Kalupahana RS, Kottawatta KS, Kanankege KS, van Bergen MA, Abeynayake P, Wagenaar JA. Colonization of Campylobacter spp. in broiler chickens and laying hens reared in tropical climates with low-biosecurity housing. Appl Environ Microbiol.79:393–5 (2013).
29. Neill SD, Campbell JN, Greene JA. Campylobacter species in broiler chickens. Avian Pathology.13:777–85 (1984).
30. Thibodeau A, Fravalo P, Yergeau E, Arsenault J, Lahaye L, Letellier A. Chicken caecal microbiome modifications induced by Campylobacter jejuni colonization and by a non-antibiotic feed additive. PLoS One.10: e0131978 (2015).
31. Ijaz UZ, Sivaloganathan L, McKenna A, Richmond A, Kelly C, Linton M, Stratakos AC, Lavery U, Elmi A, Wren BW, et al. Comprehensive longitudinal microbiome analysis of the

chicken cecum reveals a shift from competitive to environmental drivers and a window of opportunity for Campylobacter. Front Microbiol.9:2452 (2018).

32. Danzeisen, J. L., Kim, H. B., Isaacson, R. E., Tu, Z. J., and Johnson, T. J.. Modulations of the chicken cecal microbiome and metagenome in response to anticoccidial and growth promoter treatment. PLoS ONE 6:e27949. doi: 10.1371/journal.pone.0027949 (2011).
33. Ballou, A. L., Ali, R. A., Mendoza, M. A., Ellis, J. C., Hassan, H. M., Croom, W. J., et al.. Development of the chick microbiome: how early exposure influences future microbial diversity. Front. Vet. Sci. 3:2. doi: 10.3389/fvets.2016.00002 (2016).
34. Mancabelli, L., Ferrario, C., Milani, C., Mangifesta, M., Turroni, F., Duranti, S., et al.. Insights into the biodiversity of the gut microbiota of broiler chickens. Environ. Microbiol. 18, 4727–4738. doi: 10.1111/1462-2920.13363 (2016).
35. Connerton, P. L., Richards, P. J., Lafontaine, G. M., O'kane, P. M., Ghaffar, N., Cummings, N. J., et al..The effect of the timing of exposure to Campylobacter jejuni on the gut microbiome and inflammatory responses of broiler chickens. Microbiome 6:88. doi: 10.1186/s40168-018-0477-5 (2018).
36. Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al.. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res. 41:e1. doi: 10.1093/nar/gks808 (2013).
37. D'Amore, R., Ijaz, U. Z., Schirmer, M., Kenny, J. G., Gregory, R., Darby, A. C., et al.. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. BMC Genomics 17:55. doi: 10.1186/s12864-015-2194-9 (2016).
38. Joshi, N., and Fass, J.. Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for Fastq Files (Version 1.33). 1.33 ed. San Francisco, CA. Available online at: https://github.com/najoshi/sickle (2011).
39. Nikolenko, S. I., Korobeynikov, A. I., and Alekseyev, M. A.. BayesHammer: bayesian clustering for error correction in single-cell sequencing. BMC Genomics 14(Suppl. 1):S7. doi: 10.1186/1471-2164-14-S1-S7 (2013).
40. Schirmer, M., Ijaz, U. Z., D'amore, R., Hall, N., Sloan, W. T., and Quince, C.. Insight into

biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic Acids Res. 43:e37. doi: 10.1093/nar/gku1341 (2015).

41. Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al.. QIIME allows analysis of highthroughput community sequencing data. Nat. Methods 7, 335–336. doi: 10.1038/nmeth.f.303 (2010).
42. Lassmann, T., and Sonnhammer, E. L.. Kalign–an accurate and fast multiple sequence alignment algorithm. BMC Bioinformatics 6:298. doi: 10.1186/1471-2105-6-298 (2005).
43. Price, M. N., Dehal, P. S., and Arkin, A. P.. FastTree 2–approximately maximum-likelihood trees for large alignments. PLoS ONE 5:e9490. doi: 10.1371/journal.pone.0009490 (2010).
44. Meinicke, P.. UProC: tools for ultra-fast protein domain classification. Bioinformatics 31, 1382–1388. doi: 10.1093/bioinformatics/btu843 (2015).
45. Asshauer, K. P., and Meinicke, P.. “On the estimation of metabolic profiles in metagenomics,” in German Conference on Bioinformatics 2013, eds T. Beissbarth, M. Kollmar, A. Leha, B. Morgenstern, A. Schultz, S. Waack and E. Wingender (Dagstuhl: Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik), 1–13 (2013).
46. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O’Hara RB, Simpson GL, Solymos P, Stevens MHH, Szoecs E, Wagner H vegan: Community Ecology Package, R package version 2.5-6 (2019).
47. Stephanie Glen. "Simpson’s Diversity Index: Definition, Formula, Calculation" From StatisticsHowTo.com: Elementary Statistics for the rest of us! https://www.statisticshowto.com/simpsons-diversity-index/ (viewed 2022 July 29).
48. Bray JR, Curtis JT. An ordination of the upland forest communities of southern wisconsin. Ecological Monographs 27:325–349 (1957).
49. Lozupone C, Knight R. UniFrac: A new phylogenetic method for comparing microbial communities. Applied and Environmental Microbiology 71:8228–82 (2005).
50. McMurdie PJ, Holmes S. phyloseq: An r package for reproducible interactive analysis and graphics of microbiome census data. PLoS ONE 8:e61217 (2013).
51. Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities.

Applied and environmental microbiology 73:1576-85 (2007).

52. Cooper N, Rodr´ıguez J, Purvis A. A common tendency for phylogenetic overdispersion in mammalian assemblages. Proc. R. Soc. B.2752031–2 (2008).
53. Kembel S, Cowan P, Helmus M, Cornwell W, Morlon H, Ackerly D, Blomberg S, Webb C. Picante:R tools for integrating phylogenies and ecology. Bioinformatics 26:1463–1464 (2010).
54. Stegen, J., Lin, X., Konopka, A. et al. Stochastic and deterministic assembly processes in subsurface microbial communities. ISME J 6, 1653–1664. https://doi.org/10.1038/ismej.2012.22 (2012).
55. Campbell O Webb MAM David D Ackerly, Donoghue MJ. Phylogenies and community ecology. Annual Review of Ecology and Systematics 33:475-505 (2002).
56. Stegen, J. C., Lin, X., Fredrickson, J. K., & Konopka, A. E..Estimating and mapping ecological processes influencing microbial community assembly. Frontiers in microbiology, 6, 370 (2015).
57. Vass, M., Székely, A. J., Lindström, E. S., & Langenheder, S.. Using null models to compare bacterial and microeukaryotic metacommunity assembly under shifting environmental conditions. Scientific Reports, 10(1), 1-13 (2020).
58. Stegen, J. C. et al. Quantifying community assembly processes and identifying features that impose them. ISME J. 7, 2069–2079 (2013).
59. Vass M, Szekely AJ, Lindstr ´ om ES, Langenheder S. Using null models to compare bacterial and microeukaryotic metacommunity assembly under shifting environmental conditions. Scientific Reports 10:2455 (2020).
60. Chase, J. M., Kraft, N. J. B., Smith, K. G., Vellend, M. & Inouye, B. D. Using null models to disentangle variation in community dissimilarity from variation in α-diversity. Ecosphere 2, art24 (2011).
61. W. T. Sloan et al., Quantifying the roles of immigration and chance in shaping prokaryote community structure. Environ. Microbiol. 8, 732–740 (2006).
62. A. R. Burns et al., Contribution of neutral processes to the assembly of gut microbial communities in the zebrafish over host development. ISME J. 10, 655–664 (2016).

63. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for rna-seq data with deseq2. Genome Biology 15:550 (2014).
64. Anders S, Huber W: Differential expression analysis for sequence count data. Genome Biol, 11:106 (2010).
65. McKenna, A., Ijaz, U. Z., Kelly, C., Linton, M., Sloan, W. T., Green, B. D., ... & Gundogdu, O. Impact of industrial production system parameters on chicken microbiomes: mechanisms to improve performance and reduce Campylobacter. Microbiome, 8(1), 1-13 (2020).
66. Ning, D., Deng, Y., Tiedje, J. M., & Zhou, J.. A general framework for quantitatively assessing ecological stochasticity. Proceedings of the National Academy of Sciences, 116, 16892–16898 (2019).
67. Pan D, Yu Z. Intestinal microbiome of poultry and its interaction with host and diet. Gut Microbes.5:108–19 (2014).
68. Webb CO, Ackerly DD, McPeek MA, Donoghue MJ. .Phylogenies and community ecology. Annu Rev Ecol Syst 33: 475–505 (2002).
69. Pan, D., and Yu, Z.. Intestinal microbiome of poultry and its interaction with host and diet. Gut Microbes 5, 108–119. doi: 10.4161/gmic.26945 (2014).
70. Kubasova, T., Kollarcikova, M., Crhanova, M., Karasova, D., Cejkova, D., Sebkova, A., ... & Rychlik, I.. Gut anaerobes capable of chicken caecum colonisation. Microorganisms, 7(12), 597 (2019).
71. Rubio, L. A., Peinado, M. J., Ruiz, R., Suárez‐Pereira, E., Ortiz Mellet, C., & García Fernández, J. M.. Correlations between changes in intestinal microbiota composition and performance parameters in broiler chickens. Journal of animal physiology and animal nutrition, 99(3), 418-423 (2015).

Appendix

Appendix I

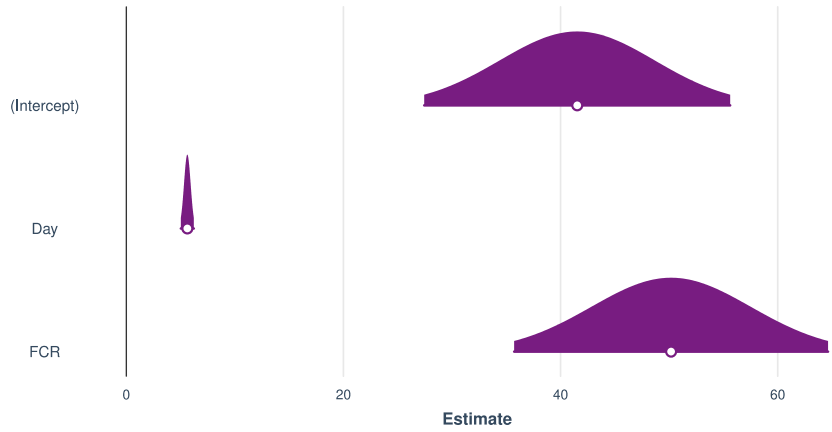
The model selection of richness.

Top models (Model parameters given with significant positive influencers highlighted in orange and negative in blue

	Model	Cross-validation Errors
2	Richness ~ Day + FCR	29.43793
5	Richness ~ Day + BW_Mean + Gain + FI + FCR	29.47557
4	Richness ~ Day + BW_Mean + FI + FCR	29.48940
3	Richness ~ Day + FI + FCR	29.71066
1	Richness ~ Day	31.46717

	Richness								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	41.53559 ***	7.16245	-0.00000	0.02115	27.44397 – 55.62720	-0.04161 – 0.04161	5.79907	**1.604e-08**	319.00000
Day	5.62083 ***	0.30450	0.69964	0.03790	5.02174 – 6.21992	0.62507 – 0.77421	18.45897	**2.866e-52**	319.00000
FCR	50.18258 ***	7.34778	0.25886	0.03790	35.72635 – 64.63881	0.18429 – 0.33343	6.82963	**4.314e-11**	319.00000
Observations	322								
R^2 / R^2 adjusted	0.857 / 0.856								

** p<0.05 ** p<0.01 *** p<0.001*



Appendix II

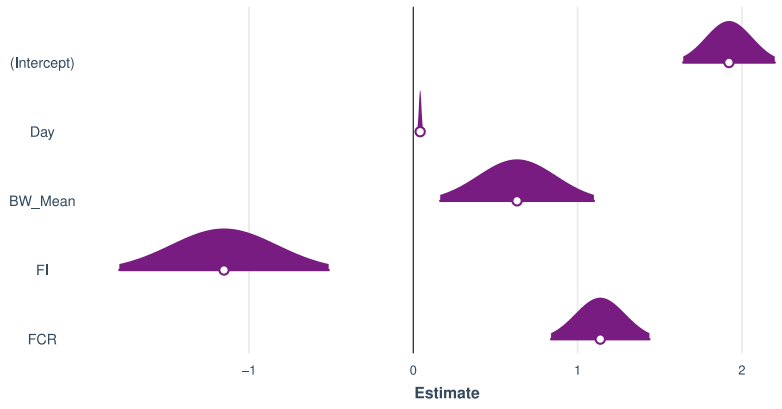
The model selection of Shannon.

Top models (Model parameters given with significant positive influencers highlighted in orange and negative in blue

	Model	Cross-validation Errors
4	Shannon ~ Day + BW_Mean + FI + FCR	0.37331
3	Shannon ~ Day + FI + FCR	0.37333
5	Shannon ~ Day + BW_Mean + Gain + FI + FCR	0.37506
2	Shannon ~ Day + FCR	0.37726
1	Shannon ~ Day	0.40477

	Shannon								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	1.92150 ***	0.14145	-0.00000	0.03088	1.64319 – 2.19980	-0.06076 – 0.06076	13.58390	**1.940e-33**	317.00000
Day	0.04180 ***	0.00775	0.60504	0.11213	0.02656 – 0.05704	0.38442 – 0.82566	5.39566	**1.339e-07**	317.00000
BW Mean	0.63146 **	0.23779	0.86097	0.32422	0.16361 – 1.09932	0.22307 – 1.49888	2.65550	**8.319e-03**	317.00000
FI	-1.15211 ***	0.32375	-1.26198	0.35462	-1.78907 – -0.51515	-1.95968 – -0.56427	-3.55869	**4.297e-04**	317.00000
FCR	1.13814 ***	0.15219	0.68278	0.09130	0.83871 – 1.43757	0.50315 – 0.86241	7.47842	**7.434e-13**	317.00000
Observations	322								
R^2 / R^2 adjusted	0.697 / 0.693								

** p<0.05 ** p<0.01 *** p<0.001*



Appendix III

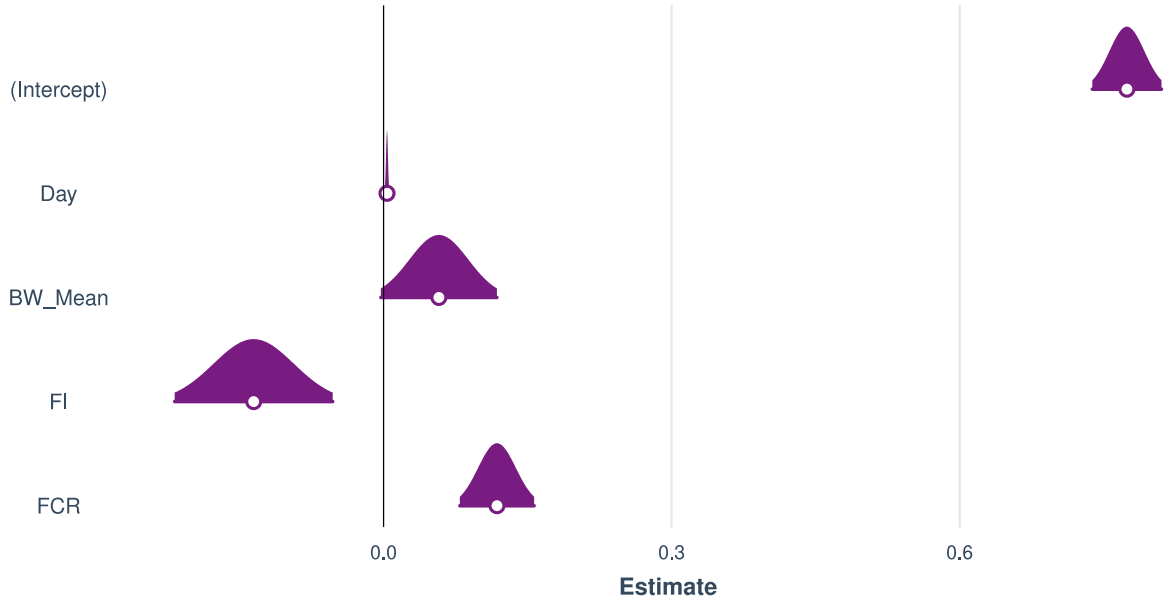
The model selection of Simpson.

Top models (Model parameters given with significant positive influencers highlighted in orange and negative in blue

	Model	Cross-validation Errors
4	Simpson ~ Day + BW_Mean + FI + FCR	0.04694
5	Simpson ~ Day + BW_Mean + Gain + FI + FCR	0.04696
3	Simpson ~ Day + FI + FCR	0.04724
2	Simpson ~ Day + FCR	0.04831
1	Simpson ~ FCR	0.04857

	Simpson								
Predictors	*Estimates*	*std. Error*	*std. Beta*	*standardized std. Error*	*CI*	*standardized CI*	*Statistic*	*p*	*df*
(Intercept)	0.77421 ***	0.01827	-0.00000	0.04504	0.73826 – 0.81016	-0.08862 – 0.08862	42.37047	**1.354e-132**	317.00000
Day	0.00360 ***	0.00100	0.58772	0.16356	0.00163 – 0.00556	0.26593 – 0.90952	3.59339	**3.783e-04**	317.00000
BW Mean	0.05762	0.03072	0.88716	0.47290	-0.00281 – 0.11806	-0.04327 – 1.81759	1.87597	6.158e-02	317.00000
FI	-0.13525 **	0.04182	-1.67282	0.51724	-0.21753 – -0.05297	-2.69048 – -0.65516	-3.23413	**1.349e-03**	317.00000
FCR	0.11813 ***	0.01966	0.80021	0.13317	0.07945 – 0.15681	0.53820 – 1.06221	6.00901	**5.123e-09**	317.00000
Observations	322								
R^2 / R^2 adjusted	0.355 / 0.347								

** p<0.05 ** p<0.01 *** p<0.001*


(Intercept)
Day
BW_Mean
FI
FCR
0.0
0.3
0.6
Estimate