

Article

# Comparison of Bacterial Populations in the Ceca of Swine at Two Different Stages and Their Functional Annotations

Himansu Kumar <sup>1</sup>, Woncheol Park <sup>1</sup>, Krishnamoorthy Srikanth <sup>1</sup>, Bong-Hwan Choi <sup>1</sup>, Eun-Seok Cho <sup>2</sup>, Kyung-Tai Lee <sup>3</sup>, Jun-Mo Kim <sup>4</sup>, Kwangmin Kim <sup>5</sup>, Junhyung Park <sup>5</sup>, Dajeong Lim <sup>1</sup> and Jong-Eun Park <sup>1,\*</sup>

<sup>1</sup> Division of Animal Genomics and Bioinformatics, National Institute of Animal Science, RDA, Wanju 55365, Korea; himanshu.genetics@gmail.com (H.K.); wcpark1982@korea.kr (W.P.); kris87@korea.kr (K.S.); bhchoi@korea.kr (B.-H.C.); djlim@korea.kr (D.L.)

<sup>2</sup> Swine Science Division, National Institute of Animal Science, RDA, Cheonan 31000, Korea; segi0486@korea.kr

<sup>3</sup> Animal Genetics and Breeding Division, National Institute of Animal Science, RDA, Cheonan 31000, Korea; leekt@korea.kr

<sup>4</sup> Department of Animal Science and Technology, Chung-Ang University, Anseong 17546, Korea; junmokim@cau.ac.kr

<sup>5</sup> 3 BIGS CO. LTD., Hwaseong 18454, Korea; kmkim@3big.com (K.K.); jhpark@3big.com (J.P.)

\* Correspondence: jepark0105@korea.kr; Tel.: +82-63-238-7309; Fax: +82-63-238-7347

Received: 18 March 2019; Accepted: 16 May 2019; Published: 20 May 2019



**Abstract:** The microbial composition in the cecum of pig influences host health, immunity, nutrient digestion, and feeding requirements significantly. Advancements in metagenome sequencing technologies such as 16S rRNAs have made it possible to explore cecum microbial population. In this study, we performed a comparative analysis of cecum microbiota of crossbred Korean native pigs at two different growth stages (stage L = 10 weeks, and stage LD = 26 weeks) using 16S rRNA sequencing technology. Our results revealed remarkable differences in microbial composition,  $\alpha$  and  $\beta$  diversity, and differential abundance between the two stages. Phylum composition analysis with respect to SILVA132 database showed *Firmicutes* to be present at 51.87% and 48.76% in stages L and LD, respectively. Similarly, *Bacteroidetes* were present at 37.28% and 45.98% in L and LD, respectively. The genera *Prevotella*, *Anaerovibrio*, *Succinivibrio*, *Megasphaera* were differentially enriched in stage L, whereas *Clostridium*, *Terrisporobacter*, *Rikenellaceae* were enriched in stage LD. Functional annotation of microbiome by level-three KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis revealed that glycine, serine, threonine, valine, leucine, isoleucine arginine, proline, and tryptophan metabolism were differentially enriched in stage L, whereas alanine, aspartate, glutamate, cysteine, methionine, phenylalanine, tyrosine, and tryptophan biosynthesis metabolism were differentially enriched in stage LD. Through machine-learning approaches such as LefSe (linear discriminant analysis effect size), random forest, and Pearson's correlation, we found pathways such as amino acid metabolism, transport systems, and genetic regulation of metabolism are commonly enriched in both stages. Our findings suggest that the bacterial compositions in cecum content of pigs are heavily involved in their nutrient digestion process. This study may help to meet the demand of human food and can play significant roles in medicinal application.

**Keywords:** cecum microbiota; 16S rRNA sequencing; SILVA database

## 1. Introduction

Pigs are principally farmed to meet the demand of human food but also useful in clothing, cosmetics, processed food ingredients, and medicinal purposes [1]. Numerous studies have shown that the microbial community plays a significant role in multiple functions of the host, for example food digestion, absorption, and the immune system [2]. It was reported that around 1014 bacterial populations encompassing 500–1000 bacterial species are present in the mammalian gastrointestinal tract (GIT) [3,4]. Previously, the gut microbiota of mammals or pigs was studied by culture-dependent method [5]. However, due to unknown growth requirements of various species of bacteria, the method was not supposed to be sufficient to explore the microbial population [1]. Multiple advanced techniques such as denaturing gradient gel electrophoresis (DGGE) [6,7], terminal-restriction fragment length polymorphism (T-RFLP) [8], temporal temperature gradient gel electrophoresis (TTGE) [9], and Sanger sequencing [10] have been introduced to decipher the gut microbiota. Although those techniques are more robust than microbial culture-dependent ones, they still lack the coverage of the diverse microbial community [11]. Recently, a 16s rRNA high-throughput next-generation sequencing technique with high coverage diversity was introduced for microbial population study [12]. Since then, 16s rRNA genes are well known and used to classify and identify microbes, because of their ubiquitous presence in microbes. Apart from the NCBI (National Center for Biotechnology Information) database, secondary databases specifically for 16s rRNA genes, such as SILVA and Greengenes are being developed [13,14]. SILVA consists of quality-checked, updated, and comprehensive datasets of aligned small 16S rRNA sequences for *Bacteria*, *Archaea*, and *Eukarya* [13], whereas Greengenes is a taxonomy database based on de novo phylogeny that provides comprehensive 16s reference datasets [14].

The GIT microbiome of pig plays essential roles, such as a pathogen barrier, and the microbial population in cecum reveals the correlation between microbiomes and metabolites [15,16]. It has been proven that the microbial communities of GIT are changing from childhood to adulthood, and dominated by bacteria. Kim et al. in 2011 reported that the microbiome population changes from neonates to adults as aerobes to anaerobes, respectively [17]. Multiple machine-learning algorithms have been proposed for taxonomical classifications as well as functional annotation [18–21]. Most recently, Fiannaca et al., [22] has introduced deep-learning models for microbial taxonomic classification from 16 s shotgun and amplicon metagenomic sequencing data. Furthermore, Maltecca, Christian et al., has also predicted the growth and carcass traits of swine through Bayesian, random forest, gradient boosting etc. [23].

In this study, we have compared the bacterial composition in the ceca of pigs at different ages (10 weeks, L; 26 weeks, LD) through 16s rRNAs sequencing [24,25]. Pigs from weaning to 10 weeks are termed as weaners and they experience numerous changes in their life such as changes in diet, stress of being separated from the sow, environmental changes, etc. It has been noticed that microbial populations of pigs of <10 weeks of age are dynamically changing and, in contrast to pigs between 10 and 26 weeks, the cecum microbiota is relatively stable. We have used 16 animals to collect the sample, 8 from each stage. Each sequence assigned an operational taxonomy unit (OTU) based on alignment with the databases SILVA. Furthermore, functional annotation of microbial sequences was also done with the help of PICRUST (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) [26] and through machine-learning approaches such as LEfSe, random forest, and Pearson's correlation. Multivariate analysis was also done through principal component analysis (PCA) [27], redundancy analysis (RDA) [28], canonical correspondence analysis (CCA) [29], and detrended correspondence analysis (DCA) [30] to explore the ecological association. We performed a Pearson's correlation network analysis to establish the relationship among identified functions. Comprehensive study of microbial interactions at different time points can help to implement the strategy to improve age-related pig farming as well as the well-being of swine [31,32].

## 2. Materials and Methods

### 2.1. Ethics Statement, Animal Rearing, and Feeding

A total of 16 F1 crossbred pigs between Korean native and Yorkshire breed were used in this study. The experiment was reviewed and approved by the Institutional Animal Care and Use Committee, National Institute of Animal Science, South Korea (IACUC no. NIAS2016-848). Feed and water were supplied ad-libitum during the experiment. These piglets were randomly allocated into two groups, raised for 10 weeks and 26 weeks. Out of 16 piglets, 8 of them were male and 8 were female. Feeding components as commercial formula diet for each stage is shown in Table 1.

**Table 1.** Feed composition for stage 1 (10 weeks) and stage 2 (26 weeks), Ca: Calcium, P: Phosphorus, DE: Digestible energy, DCP: Digestible crude protein.

Samples	Protein	Crude Fat	Ca	P	Crude Fiber	Crude Ash	Lysine	DE	DCP
Stage 1 (L)	19.0	6.7	0.4	1.5	4.0	8.0	1.40	3.6	16.0
Stage 2 (LD)	14.5	5.5	0.4	1.2	7.0	8.0	0.80	3.5	12.0

### 2.2. Sample Collection, Nucleic Acid Extraction, and 16s rRNA High-Throughput Sequencing

The animals were sacrificed at the end of the growth stage and samples were collected from cecum. The experimental area was cleaned and disinfected before every sample collection to avoid cross-contamination among samples. All collected samples were transferred to liquid nitrogen temporarily and sent to lab for storage at  $-80^{\circ}\text{C}$ . The total microbial DNA of all the samples were following the standard protocol of the DNA isolation kit (Qiagen, Germany) and quality of the extracts were checked using NanoDrop (Thermo scientific, Waltham, MA, USA). Good-quality samples (according to the ratio of absorbance at 260/280 nm) were used for library preparation. During all experimental procedure, pigs were not treated with any antibiotics. V4–V5 hyper-variable regions have been amplified by using universal primers of bacterial 16S rRNA gene. For the use of PCR, samples were diluted as 1:10, subsequently, the universal primer was used as 515F (5'-GTGCCAGCMGCCGCGGTAA) and 806R (5'-GGACTACHVGGGTWTCTAAT). PCR products were purified and used to construct the library and sequenced on MiSeq platform (Illumina, San Diego, CA, USA) [33].

### 2.3. Sequence Quality, Assembly, Preprocessing, and Clustering

FLASH 1.2.11 (Fast Length Adjustment of Short Reads) assembly program was used for assembly [34]. It is a fast and accurate tool to merge paired end reads from next-generation sequencing experiments [35,36]. Short reads were filtered out and extra-long tails were trimmed. Filtered reads were clustered at 100% identity using CD-HIT-DUP [37] which provided an OTU. It uses three-step clustering methods, first as raw read filtering and trimming [38], second error-free reads picking, and at the last step, it clusters at different distance cutoffs (0.03). The rDnaTools is a Python package from the MOTHRUR suit of utilities for working with rDNA sequences data generated by PacBio SMRT sequencing technology [39]. Currently, rDnaTools implements a single pipeline for the export, filtering, and cluster of 16S sequences [40].

By using CD-HIT-DUP and rDnaTools, chimeric reads were identified [41]. Secondary clusters are recruited into primary clusters. Noise sequences such as ambiguous base detection, miss-matched primer, extra-long read, etc. were removed from the cluster, and the remaining representative reads were clustered using a greedy algorithm and assigned OTU at a user-specific OTU cutoff (e.g., 97% ID at species level) [42].

### 2.4. Taxonomic Assignment and Diversity Analysis

QIIME (v1.91) is known bioinformatics pipeline to analyze the microbial samples, was used to assemble the paired end reads into tags according to their sequence overlap relationship [43]. Preprocessing consisted of removal of primers, demultiplexing and quality filtering (Phred  $\geq 20$ ),

chimera removal, etc. by using usearch61 module [44]. We used Silva132 database (April 2018 release) to pick up the OTU using the open-reference analysis method, allocated the taxonomy by using the uclust, created the phylogenetic tree by using QIIME FastTree, and checked the OTU identifier [13,41]. After filtering, an OTU table was generated in biom format [45]. The resulting sequences were clustered into OTUs based on a similarity threshold of  $\geq 97\%$  by PyNAST [46]. We performed comparative OTU assignment against the database at phylum, class, order, family, genus, and species levels. We used  *$\alpha$  diversity.py* script in QIIME to analyze the  $\alpha$ -diversity to understand the local population of the microbiome. Multiple algorithms such as ACE (abundance-based coverage estimator), Chao1 (richness estimator), Observed\_otus (distinct OTUs estimator), Shannon (entropy estimator), and Simpson (Simpson's index calculator) were used to estimate the  $\alpha$  diversity. The rarefaction curves of  $\alpha$  diversity were visualized with the help of Phyloseq v1.8.2 of R Package [47].

*$\beta$  diversity.py* of QIIME was used to calculate  $\beta$  diversity for estimating the correlation among other factors and microbes. Microbial distribution was analyzed by the principal coordinate analysis (PCoA) and 2D plots were produced with the help of *make\_2d\_plots.py* [48]. Phylogenies among identified OTUs were calculated by unweighted pair-group method using arithmetic averages (UPGMA) clustering techniques. To calculate the robustness of individual UPGMA clusters and PCoA plot cluster, jackknife  $\beta$  diversity program was also used [49].

### 2.5. Bioinformatics Analysis for Functional Annotations

Multiple bioinformatics approaches such as multivariate statistical analysis, machine learning, PICRUST, STAMP, and Pearson's correlation were used to annotate the functions of bacterial compositions present in the cecum of pigs. Differential abundance of microbial functions has also been analyzed.

#### 2.5.1. Multivariate Statistical Analysis to Explore Complex Ecological Associations

In this study, we have used multivariate statistical tools such as PCA [27], redundancy analysis (RDA) [28], canonical correspondence analysis (CCA) [29], detrended correspondence analysis (DCA) [30], and non-metric multidimensional scaling (NMDS) [50] to explore the microbial community composition and other explanatory variables. Furthermore, the mixMC (mixOmics microbial community) has also been used. It provides access to multivariate methods and is implemented in mixMC R package. It is a multivariate framework that takes into account the sparsity and compositionality of microbiome data. mixMC aims to identify specific associations between microbial communities and explanatory variables, such as habitat [51]. It builds on the hypothesis that multivariate methods can help identify microbial communities that modulate and influence biological systems as a whole.

#### 2.5.2. Comparative Functional Annotation of Cecum Content

Functional studies have been done on the basis of phylogenetic investigations of communities by reconstructing the unobserved states using PICRUST [26]. Relative abundance of phyla, class, order, genus, and species were estimated with the help of PICRUST and STAMP program [52,53]. Multivariate analysis between two growth stages (L vs. LD) was done with the help of PCA, whereas univariate analysis was done with the help of an ANOVA (analysis of variance) test [54]. Random forest [55], a robust machine-learning algorithm was used to describe the metagenome and their involvement in metabolic processes.

#### 2.5.3. Network Analysis through Pearson's Correlation

Network analysis has been done to identify the co-occurring and mutually exclusive microbial functions through pathway analysis. Pathways are represented as nodes, and its abundance as node size, and edges represent positive and negative associations [56]. Nodes can be colored based on their association with selected environmental variables (Pearson's correlation) [57].

### 3. Results

#### 3.1. Sequence Analysis of Cecum Content

All 16 pigs (L = 8, LD = 8) were raised separately and their body weight and feed intake were recorded and regulated daily. All sequences were generated by 16S rRNA sequencing technology of F1. Crossbreed Korean native x Yorkshire pigs are shown in Table 2, as total bases, read count, GC percentage, Q20, Q30 values, etc.

**Table 2.** Result summary of Assembly (by FLASH), Q20 (%): The percentage of bases in which the phred score is above 20; Q30 (%): The percentage of bases in which the phred score is above 30.

Sample Name	Sex	Total Bases	Read	GC (%)	Q20 (%)	Q30 (%)
K4.4L	F	86,303,564	189,866	53.08	98.54	95.24
K4.10L	M	100,659,045	223,531	53.1	98.56	95.22
Y5.2L	F	78,466,731	172,363	52.8	98.51	95.14
Y5.3L	M	79,260,896	175,057	53.03	98.55	95.26
Y5.4L	F	74,925,854	165,186	53.18	98.58	95.33
Y5.6L	M	85,490,768	189,176	53.11	98.52	95.06
Y5.9L	M	93,195,851	206,722	53.3	98.55	95.25
Y5.12L	F	92,420,536	204,127	53.21	98.63	95.43
K4.5L.D	M	70,633,514	155,507	53.32	98.62	95.35
K4.8L.D	F	77,651,573	170,983	53.21	98.62	95.47
Y5.1L.D	F	74,716,039	164,985	53.61	98.78	95.75
Y5.5L.D	F	81,054,520	178,286	53.31	98.78	95.8
Y5.7L.D	F	76,811,356	168,951	53.58	98.56	95.29
Y5.8L.D	M	85,996,517	188,902	53.44	98.54	95.29
Y5.10L.D	M	70,647,269	155,842	53.62	98.52	95.27
Y5.11L.D	M	62,782,136	138,382	53.68	98.65	95.46

FATSQC-approved sequencing reads of each sample are assembled and shown in Table 2. The DNA sequences of all samples were pooled to calculate the microbial diversity by Shannon–Weaver and Simpson indices. As shown in Table 3,  $\gamma$ -diversity represents the diversity across an entire landscape ( $\alpha + \beta$  diversity).  $\alpha$ -diversity corresponds to species diversity in sites/habitats at a local scale and  $\beta$ -diversity comprises species diversity among sites or habitats as shown in Table 3. We used 10 iterations for calculation of  $\alpha$  diversity and kept a 39,304 sequence depth for calculation of metrics. The  $\alpha$  diversity metrics ( $p$ -value) Excel file is provided in the Supplementary Material (Supplementary\_Excel\_Sheet\_1). We kept 97% sequence similarity to get an OTU, an operational definition of a species or group of species often used when only DNA sequence data is available. Chao1 returns the richness estimate for an OTU. The Shannon index takes into account the number and evenness of species, and the Simpson index represents the probability that two randomly selected individuals in the habitat will belong to the same species. Goods coverage is calculated as  $C = 1 - (s/n)$ , where  $s$  is the number of unique OTUs and  $n$  is the number of individuals in the sample. This index gives a relative measure of how well the sample represents the larger environment. As shown in Supplementary Figure S1a–c, few differences in chao1, Shannon, and Simpson were observed between the two stages (L vs. LD). However, the difference between the two groups—core microbiome and the group taxonomy results—can be seen. In particular, in the case of core microbiomes, the core OTU is 519 and the unique OTU is 102 among 621. *Ruminobacter* (0.40%), *Mitsuokella* (0.38%), *Lactobacillus\_CP000033.434247.435818* (0.12%), and *Veillonella* (0.08%) are at the genus level, whereas, principle coordinate plot, based on weighed unifracs distance matrices and  $\beta$  diversity analysis, was significantly clustered as L and LD, shown in Supplementary Figure S1d.

**Table 3.** Summary of community richness and diversity of all samples from both growth stages.

Sample Name	OTUs	Chao1	Shannon	Simpson	Goods Coverage
K4.5LD	349	399.833	5.999	0.961	0.995
K4.8LD	279	303.117	4.707	0.891	0.997
K4.4L	454	509.122	6.214	0.966	0.997
K4.10L	476	509.552	6.298	0.973	0.998
Y5.1LD	360	405.217	5.941	0.967	0.996
Y5.5LD	381	422.576	6.108	0.973	0.999
Y5.7LD	362	399.631	5.827	0.960	0.996
Y5.8LD	326	355.684	5.355	0.935	0.998
Y5.10LD	373	419.941	6.284	0.975	0.996
Y5.11LD	370	406.875	6.031	0.965	0.995
Y5.2L	418	456.478	5.756	0.958	0.997
Y5.3L	367	410.807	4.938	0.907	0.997
Y5.4L	382	429.275	6.051	0.973	0.997
Y5.6L	397	458.386	5.664	0.954	0.996
Y5.9L	310	351.437	4.998	0.923	0.998
Y5.12L	453	498.122	6.045	0.963	0.997

### 3.2. Taxonomic Assignment

We conducted stage wise taxonomy analysis of microbial composition by using RDP (Ribosomal Database Project) classifier, phylum, class, and family-level taxonomic annotations. Out of a total of 47 family, 158 genus, and 287 species sequences, unclassified reads were 7, 30, and 239, respectively; a complete list of taxonomic abundance percentages are provided in the Supplementary Material (Supplimentary\_Excel\_Sheet\_2). Overall, a major portion (>90%) of the bacterial composition belongs to *Bacteroidetes* and *Firmicutes*. Interestingly, the composition of *Bacteroidetes* increases from 37.28% to 45.98% with the age of pigs (L vs. LD), whereas the composition of *Firmicutes* decreases from 51.87% to 48.76% as shown in Figure 1. Furthermore, we investigated the changes among other microbial taxonomy distributions, such as the composition of Spirochaetes, which decreases from 9.13% to 1.19%. However, composition of Protobacteria increases from 0.71% to 3.78%. Genus-level comparison shows that *Anaerovibrio* is present as the highest population at LD stage as 7.53%, as compared to L as 0.94%. *AlloPrevotella* is present as 2.79% in L and 0.92% in LD. A Venn diagram of unique and shared OTUs between L and LD stages has been shown in Figure 1. A total of 519 OTUs were common in both growth stages. However, 60 and 43 OTUs are unique in L and LD, respectively. A complete Excel sheet containing the phylum-, class-, order-, family-, genus-, and species-wise population has been provided in the Supplementary Data (Supplimentary\_Excel\_Sheet\_3). The overall composition of microbiota has been shown in Figure 2.

**Figure 1.** Venn diagram of unique and shared OTUs between L and LD stages.

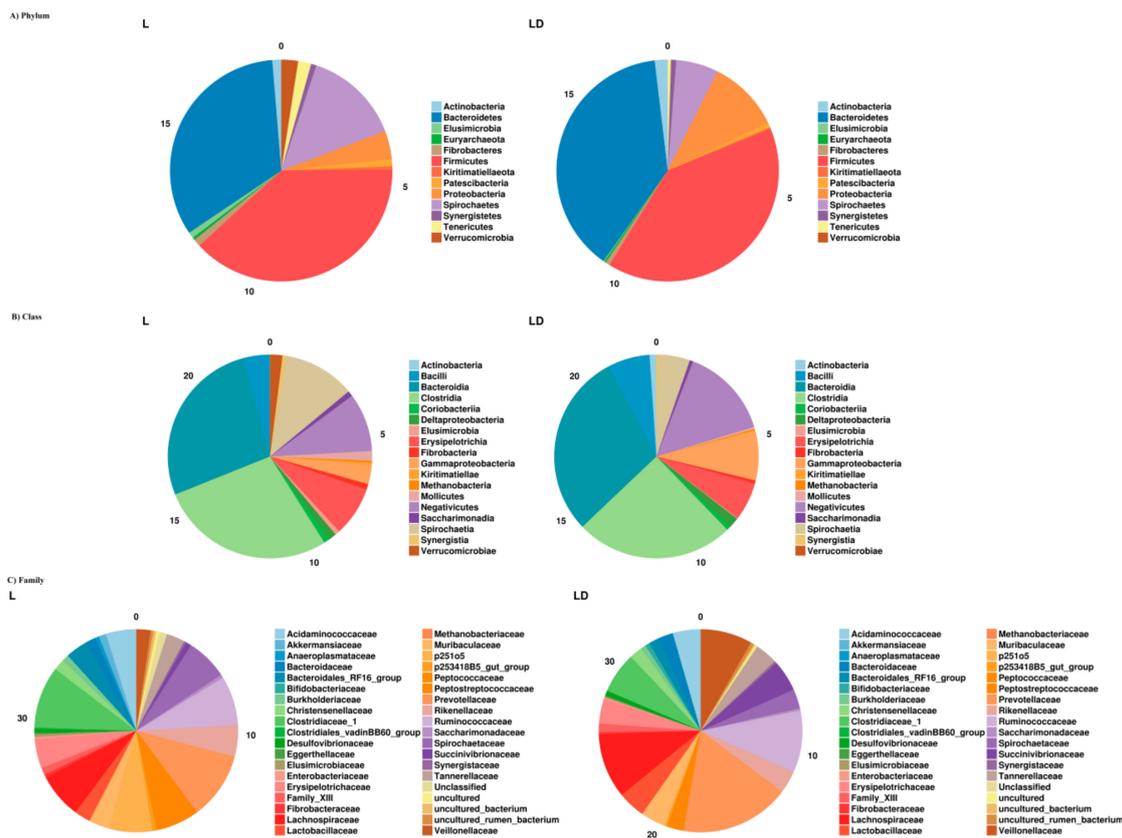


Figure 2. Pie chart presentation of taxonomic proportions (A) phylum- (B) class- (C) family-level classification.

### 3.3. Genus-Level Univariate Analysis

We conducted univariate analysis by ANOVA for identifying genus-level abundance. The significantly different genus is shown in a bar chart, and standard error is represented by error bars. The t-test was done for pairwise comparison as shown in Figure 3. The *p*-value threshold for significance was than 0.05 and found that some of the genera such as *Clostridium*, uncultured bacterium, *rikenellaceae*, *Prevotellaceae*, etc. are highly abundant in stage L, whereas in stage LD *Prevotella\_1*, *Succinivibrio*, *Roseburla*, *Anaerovirio*, etc. are differentially abundant. The ANOVA plots for phylum, class, family, order, and species levels are shown in Supplementary Figure S2a–d.

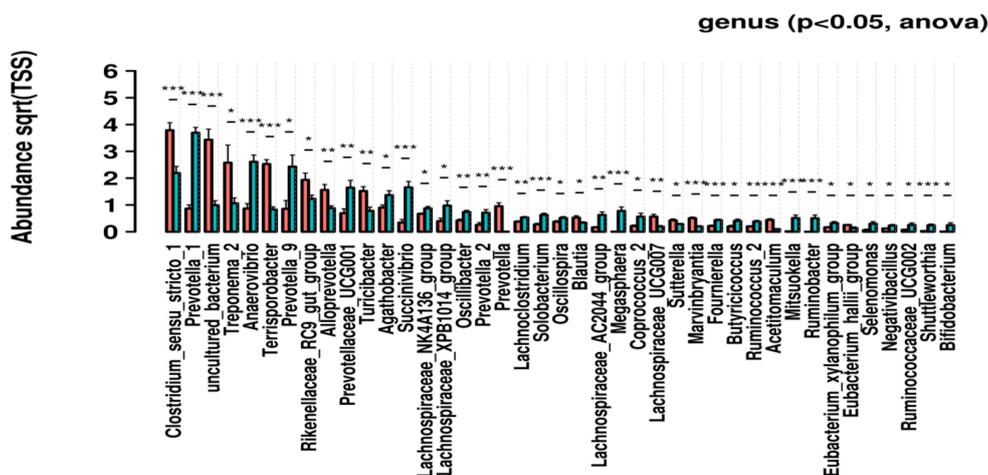
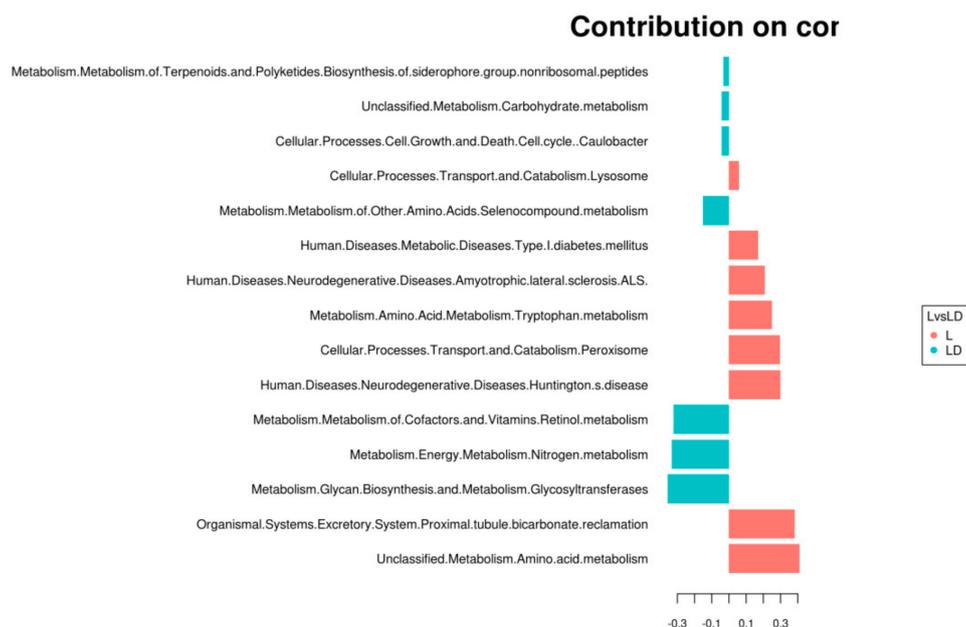


Figure 3. Genus-level univariate analysis through ANOVA test, Red color indicates the stage L, and blue is stage LD. X-axis representing the samples and Y-axis representing the relative OTU counts. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ , \*\*\*  $p \leq 0.001$ , \*\*\*\*  $p \leq 0.0001$ .

### 3.4. Statistical Analysis for Functional Annotation

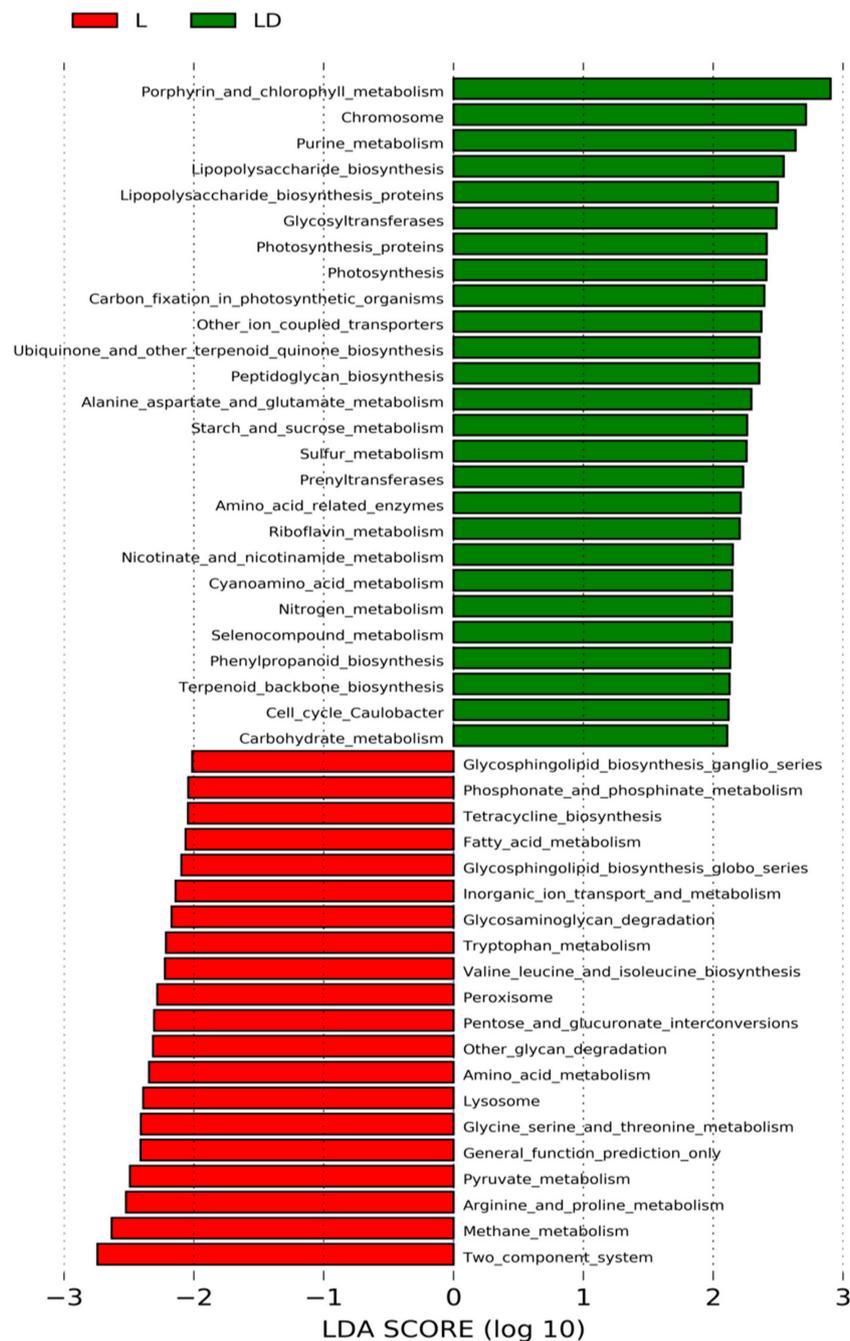
We have used mixMC for this study, as mixMC is being used extensively for microbial biomarker discovery [51]. It is a multivariate framework which can take into account sparsity and composition of the metagenomic data and establish associations between microbial populations and their habitat. The mixMC uses principal component analysis (PCA), and sparse partial least discriminant analysis (sPLS-DA). To predict the most prominent functional pathway that can describe best functional behavior of microbes present in the two different age groups of swine, we used sparse partial least discriminant analysis. We conducted analysis of the top 100 most abundant pathways present at  $\geq 2$  time points. As shown in Figure 4 (blue color indicates stage LD, and red color indicates stage L), the contribution plot shows the OTUs associated with both stages. The most consistently abundant pathways were compared with the metagenome sequence and shown as a violin plot. Amino acid metabolic pathways (glycine, serine, threonine, arginine, and proline) are highly abundant in stage L, whereas glycan biosynthesis, carbon fixation, sulfur metabolism, nitrogen metabolism, and genetic information are highly abundant in stage LD. Other associated plots such as PCA, CCA, NMDS, and RDA are shown in Supplementary Figure S3a–d. Functional profiling of a microbial community has been inferred on the basis of marker genes present in one or more samples. The OTU table represented the gene sequences of the marker gene with its relative abundance in each of the samples. Some of the stage-specific abundant metabolic pathways are shown in Figure 4.



**Figure 4.** Sparse partial least discriminant analysis, association contribution plot showing pathway association between two stages.

We have used linear discriminant analysis (LDA) effect size (LEfSe) algorithm and standard statistical tests to determine the functions of the OTUs. Based on LDA score, functional abundance has been inferred for both stages (Supplementary\_Excel\_Sheet\_4). LDA is similar to ANOVA and regression analysis, which attempts to produce one dependent variable as a linear combination. Differential abundances of selected metabolic activity and their associations with stages L and LD through ANOVA analysis is shown in Supplementary Figure S4. Comparative analysis of functional profiles through LEfSe shown in Figure 5, in stage LD, porphyrine and chlorophyll metabolism, lipopolysaccharide biosynthesis, photosynthesis etc. are highly enriched, while in stage L, lysosome, two-component systems, and amino acid metabolism are highly enriched. Furthermore, random forest (RF) method was used for functional profiling through feature selection. RF is an ensemble classification algorithm, which can combine the regression, by averaging the results of all trees and predicted the class as a

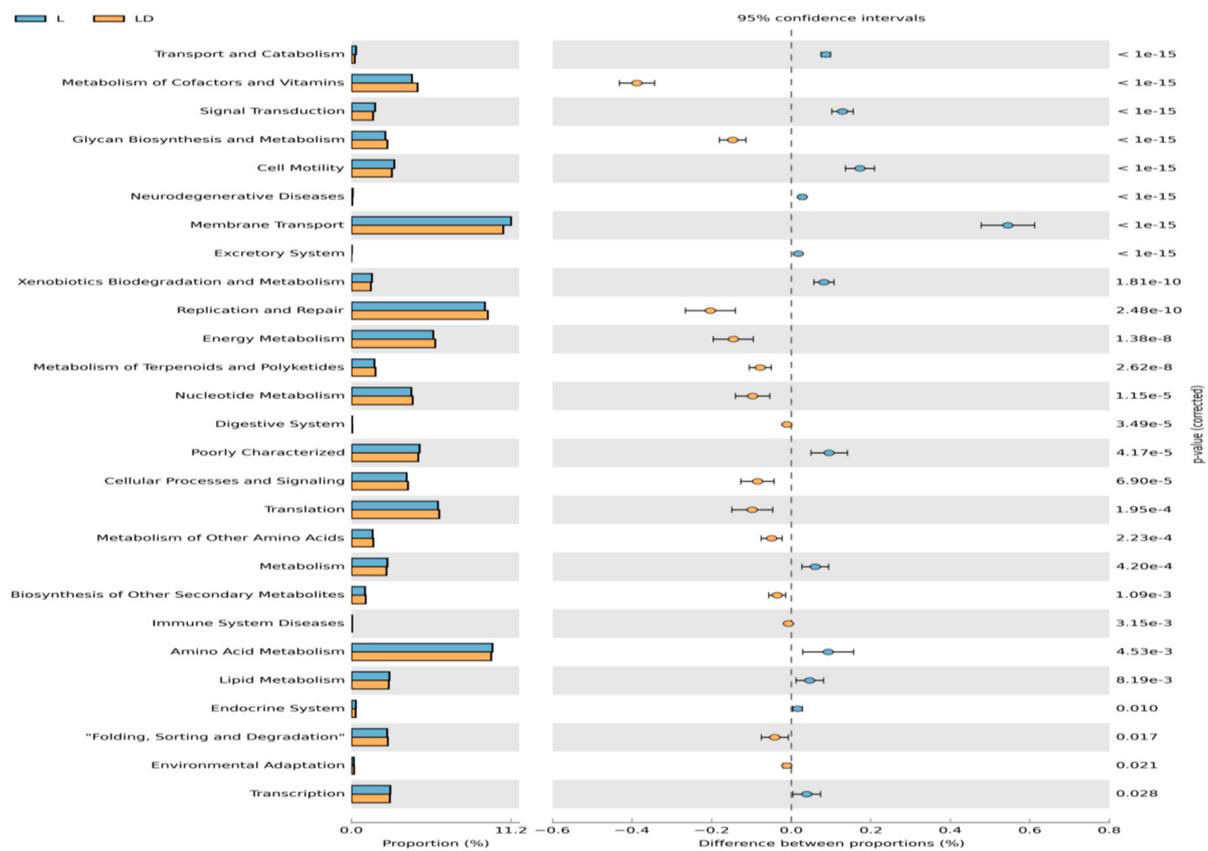
reflection of all class output. RF predicts that the pathways such as secretion system, DNA replication, two-component systems, bifolate metabolism, fructose and mannose metabolism, ABC (ATP-binding cassette) transporter systems, etc., are highly enriched (Supplementary Figure S5).



**Figure 5.** Comparative functional analysis through linear discriminant analysis (LDA) effect size (LEfSe) of microbiota of both stages. Histogram representation of LDA scores ( $(\log_{10}) > 2$ ) have been computed for differentially abundant pathways between the L and LD stage.

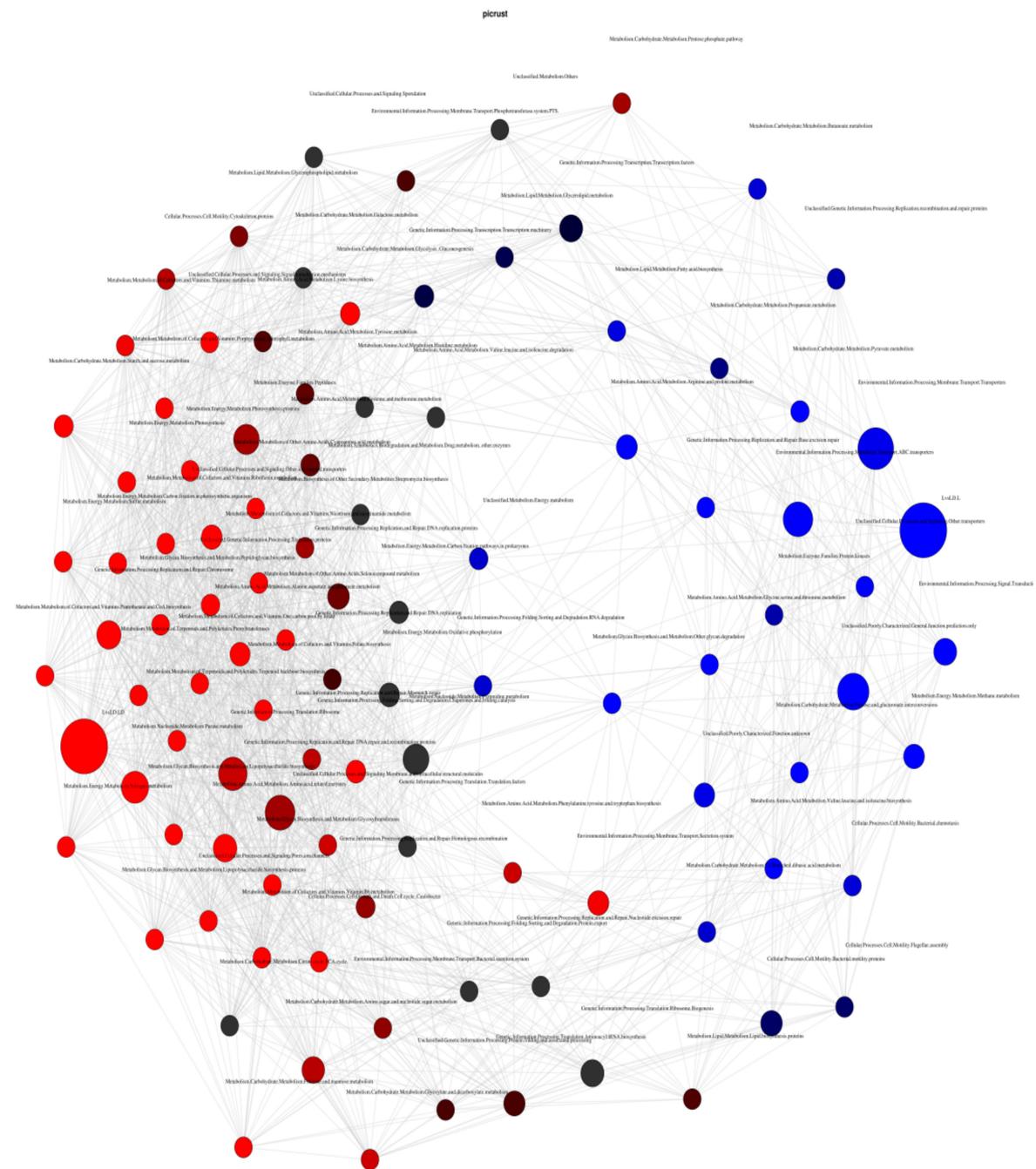
Comparative functional metagenome profiling through KEGG ontology between two stages of cecum content has also been established through STAMP software. As shown in Figure 6, purine metabolism is differentially enriched in stage L, whereas ABC-2-type transport systems, HPr kinase/phosphorylase, and glycolytic pathways are differentially enriched in stage LD at 95% confidence

intervals. Some other metabolic pathways such as HPr kinase, glucokinase,  $\beta$ -glucosidase, etc. were differentially enriched in LD stage, as shown in a yellow color bar in Figure 6.



**Figure 6.** Mean proportion and their differences in predicted functional metagenomes of the cecum microbiota. Blue color indicates stage L and yellow indicates stage LD.

Pearson’s correlation network is shown in Figure 7. Based on the top 100 functional pathways, we analyzed the entire predicted metabolic function between the two stages. Nodes have been colored based on their association with selected environmental variables (Pearson’s correlation). L correlations between nodes are presented in red, LD correlations in blue. Pathways involved in the hub node of L are environmental information, membrane transport, and ABC transport, whereas pathways involved in LD are genetic information, DNA replication and repair, and chaperones and folding.



**Figure 7.** Pearson’s correlation co-occurrence network of microbiota at both stages. Red indicates the L and blue indicates the LD.

**4. Discussion**

Multiple studies have been published about the microbial composition of the pig gut through high-throughput sequencing technologies as well as culture-based techniques [58–61]. It was suggested that microbial compositions are directly associated with genetics and aging of swine. It has been reported by multiple researchers that cecum bacterial population participates in regulation of harvesting of energy and is associated with body weight [59,62]. To the best of our knowledge, this is the first time there has been an attempt to uncover the microbiota of cecum contents at two different growth stages of Korean native swine. Exploration of bacterial taxonomy composition and their functional pattern at cecum of pig can facilitate an understanding of bacterial impact on host health as well as improving pig production. However, Pajarillo et al. in 2014 compared fecal microbial population

between pre-weaning and post-weaning piglets and also established the dominance of *Firmicutes* and *Bacteroidetes* in both stages [63]. Pigs from weaning stage to 10 weeks (Stage L) are termed as weaners, and from 10 weeks to around 26 weeks (Stage LD) are known as rearing pigs. In this study, 16 samples (8 from each) have been sequenced through 16s rRNA sequencing technologies, offering substantial information regarding their functional behavior. It was observed that at the age of more than 22 weeks, the microbial populations are relatively mature and consistent in cecum content [17]. The population of phyla, such as *Firmicutes* and *Bacteroidetes* are abundant in cecum content in all 16 samples of pig. Other phyla such as *Spirochaetes*, *Protobacteria*, *Verrucomicrobia*, etc. are also present in significant amounts. Whereas *Firmicutes* and *Spirochates* are most abundant in stage L, in contrast *Bacteroidetes* and *Proteobacteria* are most abundant in stage LD. Comparative assesment of OTUs between two growth stages showed that 60 and 43 OTUs are unique in L and LD, respectively. Such relative changes in microbial population may be due to stress or changes in the diet of pigs, as reported previously by Konstantinov et al., 2004, Mackie et al., 1999, and Poroyko et al., 2010 [64–66].

Our results are also supported by previous findings, as *Firmicutes* and *Bacteroidetes* were the most abundant phyla in the fecal content of pigs [1,5]. Genus-level comparison showed that *Clostridium sensu stricto 1* and *Treponema 2* are highly abundant in L as compared to LD. These findings also support the research reported by Lamendella et al., 2011 and Le Chatelier et al., 2013 that swine- and human-gut microbiomes are abundant with *Clostridium* and *Treponema* [1,67]. Genus-level univariate analysis also showed that *Prevotella*, *Clostridium*, and *Anaerovibrio* are abundant in swine metagenomes.

Functional annotations through PICRUST of the microbiota are essential to understand their involvement on the host metabolism and disease [68,69]. Purine metabolism (K01524) is differentially enriched in stage L microbiota, whereas ABC-2 type transport system AP-binding protein (K01990) is differentially enriched in stage LD. Kohl et al. in 2018 also reported that purine metabolism is differentially enriched in the cecum of small mammals [70]. Virkel et al. in 2018 [71] reviewed the role of ABC transporters in aging livestock animals through KEGG pathways. ATP-binding cassettes significantly influenced the bioavailability of multiple drugs. Similarly, through PICRUST and STAMP analysis, Yang Hui et al. in 2017 reported that glycine, serine, and threonine metabolism, and valine, leucine, and isoleucine degradation metabolism are significantly abundant [72].

Machine-learning approaches such as Orphelia, MGC, MetaGUN, etc. have been reported for gene prediction and taxonomy classification from metagenome data [73–76]. However, such approaches are still sparsely used for functional annotations. Here, we explored ANOVA, MixMC, LEfSe, RF, and Pearson's correlation to establish the functional significance between the two stages of pig cecum metagenome. Yang Hui, et al. in 2016 compared the functional pattern of microbiota of three different gut locations of swine through an ANOVA test [77]. Our ANOVA predictions also indicate that metagenomic functions are linked to the aging of swine. Kouchiwa, Takanori, et al. in 2012 showed that amino acid metabolism is linked to the aging of human, similarly to what we found in the case of the swine metagenome [78]. We used univariate as well as multivariate statistical tools to explore the functional pattern of the microbiome, and found both tools provided similar annotations. MixMC is a well-known multivariate microbial framework that allows annotation of the functional behavior of cecum microbes, and found that amino acid metabolism, cellular processes, and transport systems are abundant in stage L. Similarly in stage LD, glycan biosynthesis, energy metabolism, and nitrogen metabolism are significantly abundant [51].

Ten Zhen et al. in 2017 established the functional capacities with feed efficiency from cecum metagenomic data of swine through LDA effect size. Our results of LEfSe also supported their findings, as some of the metabolic pathways such as energy metabolism are differentially enriched in two different stages of cecum microbiome [79]. Multivariate analysis through PCA and sparse partial least discriminant analysis (sPLS) at both stages have also shown similar metabolic pathways, such as amino acid metabolism, human disease, and genetic information dominating in both stages. Our random-forest classifiers predicted that genetic information, chaperone, and folding pathway, amino acid metabolism, energy metabolism etc. are enriched. Similarly, Maltecca et al. in 2018 carried out

random-forest classification technique to decipher the growth prediction and carcass traits by using swine metagenome data [23]. Mach et al. in 2015 and Ramayo-Caldas et al. in 2016 reported that porcine gut microbes are associated with porcine growth, feed efficiency, and enterotype [80,81].

Metabolic functions identified through 16s rRNA analysis are only predictive in nature, as they are probable estimations, based on sequence similarity. Therefore, whole-metagenome analysis along with 16s rRNA analysis is required to know the exact pathways in which the identified microbes are involved. The network analysis may present a tendency of the predicted metabolic function between the two groups. We observed the clear cluster between two growth stages at a functional level through Pearson's correlation network. Similar analysis on soil microbial communities was reported by Barberan Albert et al., 2012 [82].

## 5. Conclusions

The composition of the cecum microbiome population of swine is strongly related with multiple factors such as developmental stages, immunity, diets, and environmental microbes. We established a critical resource of understanding for the metagenomic implications and bacterial abundance at different times of raising pigs, and emphasize how the microbiome contributes to age-related health. We found *Firmicutes* and *Bacteroidetes* phylum were abundant in both the stages. Interestingly, our findings show that composition of *Bacteroidetes* increases from L to LD, while *Firmicutes* compositions decrease from L to LD. Univariate genus analysis found *Prevotella*, *Clostridium*, and *Anaerovibrio* were significantly abundant. For the first time, we have used multiple machine-learning approaches such as multivariate mixMC, LEfSe, RF, and Pearson's correlation to explore the functional pattern of microbial populations, and found that amino acid metabolism and energy metabolism pathways were differentially significant at different stages of swine. Pearson's correlation network showed that genetic information and DNA repair pathways are actively involved while aging in swine. Our report about stage-specific cecum microbial profiling will also be a useful resource in pig health and pig production management systems.

### Nucleotide Sequence Availability

16s rRNA sequences of all 16 samples of this project have been submitted in NCBI SRA (Sequence Read Archive) submission portal and can be accessed through SRA accession PRJNA540190 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA540190>).

**Supplementary Materials:** The following are available online at <http://www.mdpi.com/2073-4425/10/5/382/s1>. Figure S1: Rarefaction plots of 16S rRNA amplicons for both stage samples; Figure 2. Univariate analysis through ANOVA; Figure 3. Plots of PCA, CCA, NMDS, RDA showing the multivariate statistical analysis; Figure 4. Differentially abundant selected metabolic activity and their association with stages L and LD through; Supplementary excel sheet 1–4.

**Author Contributions:** Conceptualization, H.K., K.S. and J.-E.P.; Data curation, H.K., W.P., K.K., J.P.; Formal analysis, H.K., K.S., W.P., J.-M.K. and J.-E.P.; Funding acquisition, J.-E.P.; Investigation, H.K., D.L. and J.-E.P.; Methodology, H.K. and J.-E.P.; Writing—review and editing, H.K., K.S., W.P., S.-H.L., B.-H.C., K.-T.L., E.-S.C., J.-M.K., D.L. and J.-E.P.

**Funding:** This work was carried out with the support of “Cooperative Research Program for Agriculture Science and Technology Development (Project title: Porcine epigenomic map construction and investigation of the imprinted genes, Project No. PJ01187601)” Rural Development Administration (RDA), Republic of Korea and also supported RDA Fellowship Program 2018 of National Institute of Animal Science, RDA, Republic of Korea.

**Acknowledgments:** The authors are thankful to Rural Development Administration (RDA), Republic of Korea, for supporting this research.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Lamendella, R.; Santo Domingo, J.W.; Ghosh, S.; Martinson, J.; Oerther, D.B. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. *BMC Microbiol.* **2011**, *11*, 103. [[CrossRef](#)]

2. Isaacson, R.; Kim, H.B. The intestinal microbiome of the pig. *Anim. Health Res. Rev.* **2012**, *13*, 100–109. [[CrossRef](#)] [[PubMed](#)]
3. Xu, J.; Gordon, J.I. Honor thy symbionts. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 10452–10459. [[CrossRef](#)]
4. Sonnenburg, J.L.; Angenent, L.T.; Gordon, J.I. Getting a grip on things: How do communities of bacterial symbionts become established in our intestine? *Nat. Immunol.* **2004**, *5*, 569–573. [[CrossRef](#)]
5. Looft, T.; Allen, H.K.; Cantarel, B.L.; Levine, U.Y.; Bayles, D.O.; Alt, D.P.; Henrissat, B.; Stanton, T.B. Bacteria, phages and pigs: The effects of in-feed antibiotics on the microbiome at different gut locations. *ISME J.* **2014**, *8*, 1566–1576. [[CrossRef](#)]
6. Muyzer, G.; Smalla, K. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) in microbial ecology. *Antonie van Leeuwenhoek* **1998**, *73*, 127–141. [[CrossRef](#)]
7. Strathdee, F.; Free, A. Denaturing Gradient Gel Electrophoresis (DGGE). In *DNA Electrophoresis: Methods and Protocols*; Makovets, S., Ed.; Humana Press: Totowa, NJ, USA, 2013; pp. 145–157.
8. Marsh, T.L. Terminal restriction fragment length polymorphism (T-RFLP): An emerging method for characterizing diversity among homologous populations of amplification products. *Curr. Opin. Microbiol.* **1999**, *2*, 323–327. [[CrossRef](#)]
9. Cornejo, P.; Azcon-Aguilar, C.; Miguel Barea, J.; Ferrol, N. Temporal temperature gradient gel electrophoresis (TTGE) as a tool for the characterization of arbuscular mycorrhizal fungi. *FEMS Microbiol. Lett.* **2004**, *241*, 265–270. [[CrossRef](#)] [[PubMed](#)]
10. Ronaghi, M.; Uhlén, M.; Nyrén, P. A sequencing method based on real-time pyrophosphate. *Science* **1998**, *281*, 363–365. [[CrossRef](#)] [[PubMed](#)]
11. Zoetendal, E.G.; Collier, C.T.; Koike, S.; Mackie, R.I.; Gaskins, H.R. Molecular ecological analysis of the gastrointestinal microbiota: A review. *J. Nutr.* **2004**, *134*, 465–472. [[CrossRef](#)]
12. Klindworth, A.; Pruesse, E.; Schweer, T.; Peplies, J.; Quast, C.; Horn, M.; Glöckner, F.O. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **2013**, *41*, e1. [[CrossRef](#)]
13. Quast, C.; Pruesse, E.; Yilmaz, P.; Gerken, J.; Schweer, T.; Yarza, P.; Peplies, J.; Glöckner, F.O. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **2012**, *41*, D590–D596. [[CrossRef](#)]
14. DeSantis, T.Z.; Hugenholtz, P.; Larsen, N.; Rojas, M.; Brodie, E.L.; Keller, K.; Huber, T.; Dalevi, D.; Hu, P.; Andersen, G.L. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl. Environ. Microbiol.* **2006**, *72*, 5069–5072. [[CrossRef](#)] [[PubMed](#)]
15. Turnbaugh, P.J.; Ley, R.E.; Mahowald, M.A.; Magrini, V.; Mardis, E.R.; Gordon, J.I. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **2006**, *444*, 1027–1031. [[CrossRef](#)] [[PubMed](#)]
16. Chakravorty, S.; Helb, D.; Burday, M.; Connell, N.; Alland, D. A detailed analysis of 16S ribosomal RNA gene segments for the diagnosis of pathogenic bacteria. *J. Microbiol. Methods* **2007**, *69*, 330–339. [[CrossRef](#)]
17. Kim, H.B.; Borewicz, K.; White, B.A.; Singer, R.S.; Sreevatsan, S.; Tu, Z.J.; Isaacson, R.E. Longitudinal investigation of the age-related bacterial diversity in the feces of commercial pigs. *Vet. Microbiol.* **2011**, *153*, 124–133. [[CrossRef](#)] [[PubMed](#)]
18. Miller, C.S.; Baker, B.J.; Thomas, B.C.; Singer, S.W.; Banfield, J.F. EMIRGE: Reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol.* **2011**, *12*, R44. [[CrossRef](#)]
19. Yuan, C.; Lei, J.; Cole, J.; Sun, Y. Reconstructing 16S rRNA genes in metagenomic data. *Bioinformatics* **2015**, *31*, i35–i43. [[CrossRef](#)] [[PubMed](#)]
20. Shah, N.; Tang, H.; Doak, T.G.; Ye, Y. Comparing bacterial communities inferred from 16S rRNA gene sequencing and shotgun metagenomics. In *Biocomputing 2011*; World Scientific: Kohala Coast, HI, USA, 2011; pp. 165–176.
21. Albanese, D.; Fontana, P.; De Filippo, C.; Cavalieri, D.; Donati, C. MICCA: A complete and accurate software for taxonomic profiling of metagenomic data. *Sci. Rep.* **2015**, *5*, 9743. [[CrossRef](#)]
22. Fiannaca, A.; La Paglia, L.; La Rosa, M.; Renda, G.; Rizzo, R.; Gaglio, S.; Urso, A. Deep learning models for bacteria taxonomic classification of metagenomic data. *BMC Bioinform.* **2018**, *19*, 198. [[CrossRef](#)]
23. Maltecca, C.; Lu, D.; Schillebeeckx, C.; McNulty, N.; Schawb, C.; Schull, C.; Tiezzi, F. Predicting Growth and Carcass Traits in Swine Using Metagenomic Data and Machine Learning Algorithms. *bioRxiv* **2018**, 363309.

24. Saraswati, S.; Sitaraman, R. Aging and the human gut microbiota—From correlation to causality. *Front. Microbiol.* **2015**, *5*, 764. [[CrossRef](#)]
25. Hu, J.; Nie, Y.; Chen, J.; Zhang, Y.; Wang, Z.; Fan, Q.; Yan, X. Gradual changes of gut microbiota in weaned miniature piglets. *Front. Microbiol.* **2016**, *7*, 1727. [[CrossRef](#)]
26. Langille, M.G.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkepile, D.E.; Thurber, R.L.V.; Knight, R.; et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013**, *31*, 814–821. [[CrossRef](#)] [[PubMed](#)]
27. Jolliffe, I. Principal component analysis. In *International Encyclopedia of Statistical Science*; Springer-Verlag: Heidelberg, Alemania, 2011; pp. 1094–1096.
28. Liu, Q. Variation partitioning by partial redundancy analysis (RDA). *Environmetrics* **1997**, *8*, 75–85. [[CrossRef](#)]
29. Ter Braak, C.J. Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. *Ecology* **1986**, *67*, 1167–1179. [[CrossRef](#)]
30. Hill, M.O.; Gauch, H.G. Detrended correspondence analysis: An improved ordination technique. *Vegetatio* **1980**, *42*, 47–58. [[CrossRef](#)]
31. Denman, S.E.; Martinez Fernandez, G.; Shinkai, T.; Mitsumori, M.; McSweeney, C.S. Metagenomic analysis of the rumen microbial community following inhibition of methane formation by a halogenated methane analog. *Front. Microbiol.* **2015**, *6*, 1087. [[CrossRef](#)]
32. Langille, M.G.; Meehan, C.J.; Koenig, J.E.; Dhanani, A.S.; Rose, R.A.; Howlett, S.E.; Beiko, R.G. Microbial shifts in the aging mouse gut. *Microbiome* **2014**, *2*, 50. [[CrossRef](#)]
33. Bokulich, N.A.; Subramanian, S.; Faith, J.J.; Gevers, D.; Gordon, J.I.; Knight, R.; Mills, D.A.; Caporaso, J.G. Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat. Methods* **2013**, *10*, 57–59. [[CrossRef](#)]
34. Magoč, T.; Salzberg, S.L. FLASH: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **2011**, *27*, 2957–2963. [[CrossRef](#)]
35. Edgar, R.C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **2010**, *26*, 2460–2461. [[CrossRef](#)]
36. Edgar, R.C. SEARCH\_16S: A new algorithm for identifying 16S ribosomal RNA genes in contigs and chromosomes. *bioRxiv* **2017**, 124131.
37. Li, W.; Fu, L.; Niu, B.; Wu, S.; Wooley, J. Ultrafast clustering algorithms for metagenomic sequence analysis. *Brief. Bioinform.* **2012**, *13*, 656–668. [[CrossRef](#)] [[PubMed](#)]
38. Lu, J.; Santo Domingo, J.; Shanks, O.C. Identification of chicken-specific fecal microbial sequences using a metagenomic approach. *Water Res.* **2007**, *41*, 3561–3574. [[CrossRef](#)]
39. Schloss, P.D.; Westcott, S.L.; Ryabin, T.; Hall, J.R.; Hartmann, M.; Hollister, E.B.; Lesniewski, R.A.; Oakley, B.B.; Parks, D.H.; Robinson, C.J.; et al. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* **2009**, *75*, 7537–7541. [[CrossRef](#)]
40. Hao, X.; Jiang, R.; Chen, T. Clustering 16S rRNA for OTU prediction: A method of unsupervised Bayesian clustering. *Bioinformatics* **2011**, *27*, 611–618. [[CrossRef](#)] [[PubMed](#)]
41. Price, M.N.; Dehal, P.S.; Arkin, A.P. FastTree: Computing large minimum evolution trees with profiles instead of a distance matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650. [[CrossRef](#)] [[PubMed](#)]
42. Plummer, E.; Twin, J.; Bulach, D.M.; Garland, S.M.; Tabrizi, S.N. A comparison of three bioinformatics pipelines for the analysis of preterm gut microbiota using 16S rRNA gene sequencing data. *J. Proteom. Bioinform.* **2015**, *8*, 283. [[CrossRef](#)]
43. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Pena, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [[CrossRef](#)]
44. Edgar, R. UCHIME2: Improved chimera prediction for amplicon sequencing. *BioRxiv* **2016**, 074252.
45. McDonald, D.; Clemente, J.C.; Kuczynski, J.; Rideout, J.R.; Stombaugh, J.; Wendel, D.; Wilke, A.; Huse, S.; Hufnagle, J.; Meyer, F.; et al. The Biological Observation Matrix (BIOM) format or: How I learned to stop worrying and love the ome-ome. *GigaScience* **2012**, *1*, 7. [[CrossRef](#)]
46. Caporaso, J.G.; Bittinger, K.; Bushman, F.D.; DeSantis, T.Z.; Andersen, G.L.; Knight, R. PyNAST: A flexible tool for aligning sequences to a template alignment. *Bioinformatics* **2009**, *26*, 266–267. [[CrossRef](#)]

47. McMurdie, P.J.; Holmes, S. phyloseq: An R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **2013**, *8*, e61217. [[CrossRef](#)]
48. Vázquez-Baeza, Y.; Pirrung, M.; Gonzalez, A.; Knight, R. EMPERor: A tool for visualizing high-throughput microbial community data. *Gigascience* **2013**, *2*, 16. [[CrossRef](#)]
49. Lozupone, C.; Knight, R. UniFrac: A new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* **2005**, *71*, 8228–8235. [[CrossRef](#)]
50. Holland, S.M. *Non-Metric Multidimensional Scaling (MDS)*; Department of Geology, University of Georgia: Athens, Greece, 2008.
51. Le Cao, K.-A.; Costello, M.-E.; Lakis, V.A.; Bartolo, F.; Chua, X.-Y.; Brazeilles, R.; Rondeau, P. MixMC: A multivariate statistical framework to gain insight into microbial communities. *PLoS ONE* **2016**, *11*, e0160169. [[CrossRef](#)]
52. Parks, D.H.; Tyson, G.W.; Hugenholtz, P.; Beiko, R.G. STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics* **2014**, *30*, 3123–3124. [[CrossRef](#)]
53. Parks, D.H.; Beiko, R.G. Identifying biologically relevant differences between metagenomic communities. *Bioinformatics* **2010**, *26*, 715–721. [[CrossRef](#)]
54. Cuevas, A.; Febrero, M.; Fraiman, R. An anova test for functional data. *Comput. Stat. Data Anal.* **2004**, *47*, 111–122. [[CrossRef](#)]
55. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
56. Faust, K.; Sathirapongsasuti, J.F.; Izard, J.; Segata, N.; Gevers, D.; Raes, J.; Huttenhower, C. Microbial co-occurrence relationships in the human microbiome. *PLoS Comput. Biol.* **2012**, *8*, e1002606. [[CrossRef](#)] [[PubMed](#)]
57. Freilich, S.; Kreimer, A.; Meilijson, I.; Gophna, U.; Sharan, R.; Ruppin, E. The large-scale organization of the bacterial network of ecological co-occurrence interactions. *Nucleic Acids Res.* **2010**, *38*, 3857–3868. [[CrossRef](#)]
58. Joyce, S.A.; Gahan, C.G. The gut microbiota and the metabolic health of the host. *Curr. Opin. Gastroenterol.* **2014**, *30*, 120–127. [[CrossRef](#)] [[PubMed](#)]
59. Looft, T.; Johnson, T.A.; Allen, H.K.; Bayles, D.O.; Alt, D.P.; Stedtfield, R.D.; Sul, W.J.; Stedtfield, T.M.; Chai, B.; Cole, J.R.; et al. In-feed antibiotic effects on the swine intestinal microbiome. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 1691–1696. [[CrossRef](#)]
60. Allen, H.K.; Looft, T.; Bayles, D.O.; Humphrey, S.; Levine, U.Y.; Alt, D.; Stanton, T.B. Antibiotics in feed induce prophages in swine fecal microbiomes. *MBio* **2011**, *2*, e00260-11. [[CrossRef](#)] [[PubMed](#)]
61. Allison, M.J.; Robinson, I.; Bucklin, J.; Booth, G. Comparison of bacterial populations of the pig cecum and colon based upon enumeration with specific energy sources. *Appl. Environ. Microbiol.* **1979**, *37*, 1142–1151. [[PubMed](#)]
62. Kim, H.B.; Isaacson, R.E. The pig gut microbial diversity: Understanding the pig gut microbial ecology through the next generation high throughput sequencing. *Vet. Microbiol.* **2015**, *177*, 242–251. [[CrossRef](#)]
63. Pajarillo, E.A.B.; Chae, J.-P.; Balolong, M.P.; Kim, H.B.; Kang, D.-K. Assessment of fecal bacterial diversity among healthy piglets during the weaning transition. *J. Gen. Appl. Microbiol.* **2014**, *60*, 140–146. [[CrossRef](#)]
64. Konstantinov, S.R.; Favier, C.F.; Zhu, W.Y.; Williams, B.A.; Klüß, J.; Souffrant, W.-B.; de Vos, W.M.; Akkermans, A.D.; Smidt, H. Microbial diversity studies of the porcine gastrointestinal ecosystem during weaning transition. *Anim. Res.* **2004**, *53*, 317–324. [[CrossRef](#)]
65. Mackie, R.I.; Sghir, A.; Gaskins, H.R. Developmental microbial ecology of the neonatal gastrointestinal tract. *Am. J. Clin. Nutr.* **1999**, *69*, 1035s–1045s. [[CrossRef](#)]
66. Poroyko, V.; White, J.R.; Wang, M.; Donovan, S.; Alverdy, J.; Liu, D.C.; Morowitz, M.J. Gut microbial gene expression in mother-fed and formula-fed piglets. *PLoS ONE* **2010**, *5*, e12459. [[CrossRef](#)] [[PubMed](#)]
67. Le Chatelier, E.; Nielsen, T.; Qin, J.; Prifti, E.; Hildebrand, F.; Falony, G.; Almeida, M.; Arumugam, M.; Batto, J.-M.; Kennedy, S.; et al. Richness of human gut microbiome correlates with metabolic markers. *Nature* **2013**, *500*, 541–546. [[CrossRef](#)] [[PubMed](#)]
68. Joice, R.; Yasuda, K.; Shafquat, A.; Morgan, X.C.; Huttenhower, C. Determining microbial products and identifying molecular targets in the human microbiome. *Cell Metab.* **2014**, *20*, 731–741. [[CrossRef](#)] [[PubMed](#)]
69. Cohen, L.J.; Kang, H.-S.; Chu, J.; Huang, Y.-H.; Gordon, E.A.; Reddy, B.V.B.; Ternei, M.A.; Craig, J.W.; Brady, S.F. Functional metagenomic discovery of bacterial effectors in the human microbiome and isolation of commendamide, a GPCR G2A/132 agonist. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E4825–E4834. [[CrossRef](#)] [[PubMed](#)]

70. Kohl, K.D.; Oakeson, K.F.; Orr, T.J.; Miller, A.W.; Forbey, J.S.; Phillips, C.D.; Dale, C.; Weiss, R.B.; Dearing, M.D. Metagenomic sequencing provides insights into the location of microbial detoxification in the gut of a small mammalian herbivore. *bioRxiv* **2018**, 299198.
71. Virkel, G.; Ballent, M.; Lanusse, C.; Lifschitz, A. Role of ABC transporters in veterinary medicine: Pharmaco-toxicological implications. *Curr. Med. Chem.* **2018**. [[CrossRef](#)]
72. Yang, H.; Huang, X.; Fang, S.; He, M.; Zhao, Y.; Wu, Z.; Yang, M.; Zhang, Z.; Chen, C.; Huang, L. Unraveling the fecal microbiota and metagenomic functional capacity associated with feed efficiency in pigs. *Front. Microbiol.* **2017**, *8*, 1555. [[CrossRef](#)] [[PubMed](#)]
73. Hoff, K.J.; Lingner, T.; Meinicke, P.; Tech, M. Orphelia: Predicting genes in metagenomic sequencing reads. *Nucleic Acids Res.* **2009**, *37*, W101–W105. [[CrossRef](#)]
74. Hoff, K.J.; Tech, M.; Lingner, T.; Daniel, R.; Morgenstern, B.; Meinicke, P. Gene prediction in metagenomic fragments: A large scale machine learning approach. *BMC Bioinform.* **2008**, *9*, 217. [[CrossRef](#)]
75. El Allali, A.; Rose, J.R. MGC: A metagenomic gene caller. *BMC Bioinform.* **2013**, *14*, S6. [[CrossRef](#)] [[PubMed](#)]
76. Liu, Y.; Guo, J.; Hu, G.; Zhu, H. Gene prediction in metagenomic fragments based on the SVM algorithm. *BMC Bioinform.* **2013**, *14*, S12. [[CrossRef](#)] [[PubMed](#)]
77. Yang, H.; Huang, X.; Fang, S.; Xin, W.; Huang, L.; Chen, C. Uncovering the composition of microbial community structure and metagenomics among three gut locations in pigs with distinct fatness. *Sci. Rep.* **2016**, *6*, 27427. [[CrossRef](#)]
78. Kouchiwa, T.; Wada, K.; Uchiyama, M.; Kasezawa, N.; Niisato, M.; Murakami, H.; Fukuyama, K.; Yokogoshi, H. Age-related changes in serum amino acids concentrations in healthy individuals. *Clin. Chem. Lab. Med.* **2012**, *50*, 861–870. [[CrossRef](#)] [[PubMed](#)]
79. Tan, Z.; Yang, T.; Wang, Y.; Xing, K.; Zhang, F.; Zhao, X.; Ao, H.; Chen, S.; Liu, J.; Wang, C. Metagenomic analysis of cecal microbiome identified microbiota and functional capacities associated with feed efficiency in landrace finishing pigs. *Front. Microbiol.* **2017**, *8*, 1546. [[CrossRef](#)]
80. Mach, N.; Berri, M.; Estellé, J.; Levenez, F.; Lemonnier, G.; Denis, C.; Leplat, J.J.; Chevaleyre, C.; Billon, Y.; Doré, J. Early-life establishment of the swine gut microbiome and impact on host phenotypes. *Environ. Microbiol. Rep.* **2015**, *7*, 554–569. [[CrossRef](#)] [[PubMed](#)]
81. Ramayo-Caldas, Y.; Mach, N.; Lepage, P.; Levenez, F.; Denis, C.; Lemonnier, G.; Leplat, J.-J.; Billon, Y.; Berri, M.; Doré, J.; et al. Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. *ISME J.* **2016**, *10*, 2973–2977. [[CrossRef](#)] [[PubMed](#)]
82. Barberán, A.; Bates, S.T.; Casamayor, E.O.; Fierer, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME J.* **2012**, *6*, 343–351. [[CrossRef](#)] [[PubMed](#)]

