Metagenomic Analysis of Chicken Gut Microbiota for Improving Metabolism and Health of Chickens — A Review

Ki Young Choi*, Tae Kwon Lee¹, and Woo Jun Sul*
Department of Systems Biotechnology, Chung-Ang University, Anseong 456-756, Korea

ABSTRACT: Chicken is a major food source for humans, hence it is important to understand the mechanisms involved in nutrient absorption in chicken. The microbial diversity is a central role in enhancing nutrient absorption and strengthening the immune system, thereby affecting both growth and health of chicken. There is little information on the diversity and functions of chicken GIT microbiota, its impact on the host, and the interactions between the microbiota and host. Here, we review the recent metagenomic strategies to analyze the chicken GIT microbiota composition and its functions related to improving metabolism and health. We summarize methodology of metagenomics in order to obtain bacterial taxonomy and functional inferences of the GIT microbiota and suggest a set of indicator genes for monitoring and manipulating the microbiota to promote host health in future. (Key Words: Chicken Gastrointestinal Tract, Microbiota, Metagenomics)

INTRODUCTION

Domestic chickens are the most common poultry and are valuable sources of protein for humans. Recently, the dual roles of microbiota in the gastrointestinal tract (GIT) have attracted attention: they support digestion of food and also provide nutrients. In addition, the GIT microbiota may play crucial roles in the health and immune system of animals. However, the diverse and complex structure of GIT microbiota hindered the understanding of the roles of GIT microbiota in chicken’s growth and health. As relatively few cultivated specimens of chicken GIT microbes are available mainly due to their requirement for anaerobic conditions similar to those in the GIT, researchers have turned to new culture-independent approaches, metagenomics, to predict nutritional and ecological roles of GIT microbes. Metagenomics is the study of genetic materials from environmental or host-associated microbiota to identify the microbial diversity and its functions. Previously, conventional, low-throughput Sanger sequencing method has been limited in metagenomic studies of GIT microbiota due to high sequencing costs and insufficient sequence data. Recently, newly developed sequencing platforms such as next-generation sequencing (NGS) have allowed the substantial researches into the diversity and functions of microbiota from the guts of various livestock animals. High-throughput NGS generates large volumes sequence data containing genetic information (Metzker, 2010), and this allows hypothesis-driven researches on chicken GIT microbiota, thereby highlighting the roles of previously unknown and rare microbial GIT species (Soin et al., 2006; Medinger et al., 2010). Furthermore, metagenomic data have raised new questions such as how microbiota stability and ecological shifts are influenced by nutrients and hosts (Andersson et al., 2008). Here, we review the current status on metagenomics-driven studies on chicken GIT’s microbiota and the involved methodology, and suggest possible approaches for better understanding of GIT microbiota to increase chicken productivity and overall health.

* Corresponding Author: Woo Jun Sul. Tel: +82-31-670-4707, Fax: +82-31-675-3108, E-mail: sulwj@cau.ac.kr
¹ Department of Environmental Engineering, Yonsei University, Wonju 220-710, Korea.
² These two authors contribute equally to this work.
Submitted Jan. 9, 2015; Revised Feb. 23, 2015; Accepted Mar. 31, 2015

Copyright © 2015 by Asian-Australasian Journal of Animal Sciences
This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
CHARACTERIZATION OF CHICKEN GIT MICROBIOTA

Bacterial community profiling based on 16S rRNA gene sequence and functional inference

When determining the composition of GIT microbiota using a gene-targeted-metagenomic approach, several marker genes such as those for ribosomal protein subunits, elongation factors, and RNA polymerase subunits represent distinct microbial populations. One of the most general choices for resolving the taxonomic composition of the microbial community is the hypervariable region in genes encoding 16S rRNA (Figure 1). Bacterial 16S rRNA gene possesses 9 hypervariable regions flanked by highly conserved regions, which are generally designated as polymerase chain reaction (PCR) primer sites. Sequence variations in hypervariable regions allow accurate bacterial taxonomic estimation by comparing against 16S rRNA gene sequences deposited in databases such as GreenGenes (DeSantis et al., 2006), the Ribosomal Database Project (Cole et al., 2014), and SILVA (Pruesse et al., 2007). Several bioinformatic pipelines of 16S rRNA gene NGS-based profiling (e.g., 16S amplicon sequencing: using 454 pyrosequencing, Illumina MiSeq, and Ion PGM Systems) have been proposed and designed for obtaining taxonomic information via processing raw sequences of 16S rRNA genes. These pipelines generally consist of the following steps: quality filtering of low-quality sequences and chimera checking (Edgar et al., 2011; Haas et al., 2011), removing pyrosequencing-oriented sequencing errors (Huse et al., 2007), and producing a cluster of nearly identical sequences referred to operational taxonomic units (OTUs) (Caporaso et al., 2010; Cole et al., 2014). OTUs are defined by two approaches i) unsupervised clustering on sequence similarity and ii) supervised clustering that directly assigns sequences to trained taxonomic clusters (Sul et al., 2011). For undefined microbiota, OTUs generated by unsupervised clustering are generally used to serve as individual phylotypes (Medinger et al., 2010). The representative sequences, which are either the most abundant sequences or sequence with the least distance summation within all sequences in the same OTUs, are then matched to those in the public database (e.g. the Ribosomal Database Project) to obtain taxonomic classification. This information has also been used to account for the degree of divergence between different communities or sample types. The latter approach is most useful for studying low-complexity microbiota and comparative analysis of the same type of samples under different experimental conditions.

In addition to providing taxonomic information, the OTUs provide information on population diversity, indicating richness and evenness of individual species in a sample as alpha diversity (Colwell, 2009). Both richness and evenness of the diversity of the microbiota are often important indicators of the health of livestock. For instance, the number of microbes in the gut decreases during

Figure 1. Metagenomic strategies for chicken GIT microbiota. GIT, gastrointestinal tract.
antibiotic treatment or nutrient imbalance (Lozupone et al., 2012). Several indices, including Chao, abundance-based coverage estimators, Shannon, and Simpson, calculate the abundance or distribution of OTUs within a particular population. For example, low richness suggests low number of species in the community and low evenness indicates that the sample consists of a few dominating taxa. These alpha diversity indices could be obtained using a bioinformatics pipeline such as MOTHUR (Hughes et al., 2011; Schloss et al., 2009) and Qiime (Caporaso et al., 2010). MOTHUR output consists of numeric data presented in a table or as rarefaction curves, which can then be compared with the values for the microbial community samples. Alpha diversity is generally considered to provide a poor reflection of microbial functional inference as an ecosystem. Although alpha diversity is not consistent with different samples, functionality can be maintained and replaced by the closely related and/or functionally similar species (Purschke et al., 2013). However, recent reports showed that alpha diversity is strongly correlated with diversity in a specific gene functions (Johnson et al., 2014). In addition, a few studies suggested that gut microbiota with high diversity could be more stable or healthier than those with low diversity (Clemente et al., 2012; Lozupone et al., 2012). These studies indicate that alpha diversity can contribute functional inference about the underlying microbiota mechanism.

Beta diversity is used when comparing multiple communities to determine the number of OTUs or taxa shared among them. Beta diversity integrates information about the degree of functional dissimilarity between multiple populations. Beta diversity calculated as a similarity index (e.g., Jaccard, Bray-Curtis) that shows shared populations between the microbial profiles of multiple communities, whereas alpha diversity shows a statistical summary of diversity in a single population (Martin et al., 2011). Two approaches to measure beta diversity are widely used: unweighted UniFrac (qualitative) and weighted UniFrac (quantitative) (Lurupone and Knight 2005; 2007). UniFrac is a method to calculate a distance between microbial communities using phylogenetic diversity. Many studies using UniFrac show that beta diversity is strongly influenced by environmental factors such as substrates, temperature, and pH.

Inflation of microbial diversity results could occur during a sequencing process due to sample preparation and preservation, DNA extraction, and choice of PCR primers (Lauber et al., 2010; Cruaud et al., 2014; Hang et al., 2014). The diversity results can be statistically varied even though the samples were stored at 4°C or lower. DNA extraction methods seem to affect the diversity results in different sample types. The choice of 16S rRNA gene’s hypervariable regions to be amplified by selected PCR primers introduces significant variation in the microbial diversities in the samples, because different regions show considerable differences in taxonomic coverage. These artifacts clearly indicate that the diversity results with different sample preparations should be carefully interpreted.

In diversity analysis, it is important to account for the contribution of sequencing errors on dramatically different outputs (Huse et al., 2007; Kunin et al., 2010). These errors can be caused by regions of 16S rRNA, read length and sequencing technology (Gilles et al., 2011). Recently, Illumina MiSeq and HiSeq are becoming widely popular sequencing platform for high-throughput sequencing since sequencing errors produced in these systems are easier to manage computationally than those from 454 pyrosequencing (Claesson et al., 2010; Luo et al., 2012). Another type of error that occurs during gene amplification is generation of chimeric sequences from two unique parental 16S rRNA gene sequences. Various computation programs have been developed to remove chimeric sequences by using reference sequences or self-query sequences (Edgar et al., 2011; Haas et al., 2011). Quality sequence filtering process is essential before analysis of microbial diversity.

**Functional gene-targeted metagenomics**

In contrast to 16S rRNA gene profiling, analysis of genes encoding proteins that function to regulate metabolic pathways can help to explore the diversity of particular ecological functions in the GIT sample (Figure 1). Further, protein-coding genes can be used as phylogenetic markers (most likely with 16S rRNA gene), although the high evolutionary rate of functional genes can mislead to linking with functional and taxonomic results (Case et al., 2007). Functional Gene Pipeline and Repository (FunGene: http://fungene.cme.msu.edu/) provides databases of biogeochemical, biodegradation, and antibiotic functional genes, as well as a pipeline that allows analysis of functional diversity based on OTUs (Fish et al., 2013). Using FunGene, the sequences are analyzed through two approaches: reference-based approaches or unsupervised approaches. The reference-based approaches use a tool, named FrameBot, to correct a sequencing error generated by matching the sequences against curated reference sequences (Wang et al., 2013). The unsupervised approaches cluster the aligned sequences generated by HHMER3, which is based on implementation of hidden Markov model for searching similar protein sequences and for making protein sequence alignment (Finn et al., 2011). All clustered sequences can be submitted into calculation of alpha and beta diversity, and classification of representative sequences extracted from each cluster using FunGene.
Approaches to further analysis of functional genes would be similar to analysis of 16S rRNA genes, but abundance of functional genes can be highly influenced by the function of interest in the ecosystem (Pereyra et al., 2010). Therefore, the accurate and fast quantification of functional genes in the ecosystem is a critical step to understand the potential of microbial function in the ecosystem. Multiplex PCR combined with q-PCR (Multiplex q-PCR) is an ideal tool to be used in this context, since not only quantitative results about the target functional gene, but also the distribution of multiple genes in the ecosystem is crucial for understanding microbial function in different ecosystems (Elmfro et al., 2000; Smith and Osborn, 2009). The multiplex q-PCR can amplify more than one targeted gene by addition of more than one pair of primers in the same reaction. This technique shows strong potential to save time and labor for multiple samples and genes targeted in different areas (Stedtfeld et al., 2008). For example, 18 antibiotic resistance types (total 141 primer sets) were successfully tested to evaluate the abundance and diversity of antibiotic resistance genes in swine gut microbiota using multiplex q-PCR (Looft et al., 2012). Recently, 3rd generation PCR technology, named digital PCR (dPCR), has developed for directly quantifying and clonally amplifying nucleic acids. Although dPCR is carried out a single reaction within a sample, the sample is distributed into a hundred thousands of partitions and PCR reactions are carried out in parallel. This feature provides for more accurately quantifying results than conventional q-PCR. Microarrays have been developed to use gene expression studies and monitor environmental processes. GeoChip is one of well-known comprehensive microarray chips in microbial ecology, especially useful to study ecosystem processes and functions (He et al., 2007). Current version of GeoChip contains more than 50,000 gene probes involved in biogeochemical cycles, metal reduction, resistance and organic compound degradation. It is a promising tool for high-throughput and cost-effective analysis in quantifying functional genes.

Although amplicon-based profiling is the most common approach to study environmental microbial diversity, it showed several limitations including amplicon size, primer (or probe) sensitivity, and errors during amplification and contamination. However, metagenome analysis without the PCR amplification step can be a powerful alternative to amplicon-based profiling for investigating all metabolic processes of microorganisms found in environmental samples. The process of metagenomic analysis compared with amplicon-based profiling is summarized in Figure 1. The simplest way to analyze the metagenome is by using short reads directly after the quality sequence filtering step (Thomas et al., 2012). These short reads can be used as taxonomic markers and matched against a marker gene database (e.g. Phylsift, EggNOG) and functional markers quantifying abundance of genetic pathways (Darling et al., 2014; Powell et al., 2014). These analyses provide taxonomic (e.g. MEGAN) and functional (e.g. KEGG pathway) contents of single data sets to rapidly search for interesting research targets and test hypothesis generated from experimental design. Alternatively, assembly algorithms can be used to reconstruct short reads into a sequence contig, which is a set of overlapping sequences representing the contiguous DNA fragment (Mende et al., 2012; Thomas et al., 2012). Although obtaining a complete individual genome from metagenomic sequences is still challenging unless high sequencing coverage, it is sufficient to characterize the major functions of the microbial communities as well as to identify their taxon by assigning to public genome reference database (Howe et al., 2014; Nielsen et al., 2014).

**THE DIFFERENT ROLES OF CHICKEN GIT MICROBIOTA IN NUTRIENT METABOLISM IN HOST**

The chicken GIT is like a machine for metabolizing nutrients. In most recent studies, researchers have not focused on culture-dependent studies but on culture-independent phylogenetic profiling against chicken GIT. Taking advantages of NGS, metagenomic approaches revealed the bacterial composition of the chicken GIT and also functional genes highly related to host digestion and nutrient absorption mechanisms. This technique brings us understanding of the taxonomic variety of the co-metabolizing micro-community within the host and the gene level of target nutrients metabolism. Therefore, we reviewed the recent reports about the candidates of target bacteria or genes to understand how they can help in host metabolism.

**Polysaccharide metabolism**

The caecum is a key organ for fermentation of various forms of polysaccharides and is highly associated with host health and performance (Stanley et al., 2014). Cecal microbiota play a central role in metabolizing polysaccharides because chickens do not possess a complete metabolic cycle for producing readily absorbable forms of polysaccharides. Recently, Sergeant et al. (2014) reported numerous polysaccharide degrading enzymes in the cecal metagenome, and Qu et al. (2008) showed that the genes for carbohydrate metabolism occupied about 20% genes in the GIT microbiota metagenome (Qu et al., 2008; Sergeant et al., 2014) including key enzymes like: carbohydrate esterase, polysaccharide lyase, and glycoside hydrolase which are lacking in chickens (Beckmann et al., 2006; Yeoman et al., 2012). During the digestion of
polysaccharides, GIT microbiota produced various short chain volatile fatty acids (SCFA), mostly acetate, while propionate, butyrate, valerate, isobutyrate, and isovalerate are other forms also found (Dunkley et al., 2007; Yeoman et al., 2012). The ratio of SCFA production was lower during the uptake and digestion of soluble polysaccharides than during the uptake and digestion of insoluble polysaccharides (Lei et al., 2012). In addition, carbohydrate metabolism-related proteins were most commonly found in the metaproteomic analysis of chicken GIT microbiota (Tang et al., 2014). Recently, Tang et al. (2014) showed the positive correlation between metagenomic and metaproteomic analyses for chicken microbiome composition and protein expression.

Nitrogen metabolism

Recently, Qu et al. (2008) reported that the environmental gene tags of the subsystems-based annotations (SEED) database showed that the GIT microbiota has genes involved in metabolism of protein and amino acids (9%), and nitrogen (1%) as nutrient sources (Qu et al., 2008). In case of host leaked or incompletely metabolized dietary protein, the GIT microbiota digested further amino acids as host nutrient because chickens unlike some other animals lack the enzymes required for complete nitrogen metabolism (Laslack and Zhao, 2011). As a result, effective nitrogen metabolism by GIT microbiota has produced healthier and more productive chickens (Stanley et al., 2014). When chicken GIT microbiota metabolized nitrogen sources, half of the nitrogen sources were converted into ammonia, which was excreted subsequently and this loss adversely affected the health of the chicken (McCubbin et al., 2002; Xin et al., 2011). Undissolved uric acid, which itself is not toxic, affects induction or degradation of bacterial proteins in chicken gut (Tang et al., 2014).

Fatty acids and lipids metabolism

Commercial chickens feed on grains which usually contain non-starch polysaccharides (NSPs) and the metabolism of NSPs produces SCFA. Despite the fact that Firmicutes, Bacteroides, and Proteobacteria were usually isolated or detected in the chicken ceca, Clostridiales contributed the most to SCFA metabolism (Oakley et al., 2014). SCFA can inhibit the growth of acid-sensitive pathogens by inducing low-pH environment and can improve both mineral absorption in host and growth of epithelial cells (Oakley et al., 2014; Sergeant et al., 2014). Generally, fatty acid and lipid metabolism related genes were detected to be only about 1% to 2% in the whole gut metagenome (Qu et al., 2008). Conjugated linoleic acid (CLA) is a by-product during metabolism of polyunsaturated linoleic acid and presence of CLA in the diet was reported to affect the death rate of embryo and quality of yolk (Aydin et al., 2001; Yeoman et al., 2012).

Host health

The presence of bacterial pathogens such as Salmonella, Clostridium, Campylobacter, Staphylococcus, and E. coli within chicken microbiota is a critical issue for host well-being (Oakley et al., 2014). Bacteria that are non-pathogenic to chickens and are found in chicken microbiota are sometimes pathogenic to humans (Newell et al., 2011). For example, recent reports suggest that most of Campylobacter infection cases in humans resulted from handling and eating raw or undercooked poultry meat (Kaakoush et al., 2014). Despite the presence of a pathogen (for humans) at a detectable level in the microbiota of healthy chickens, there were no significant changes in microbial community structure and diversity (Skraban et al., 2013). Similarly, even infection with a pathogen will not cause major changes in the remaining chicken resident microbiota (Videnska et al., 2013). Oakley et al. (2014) suggested that competitive exclusion is one of the most effective treatments for blocking pathogens, usually Salmonella, leading to enrichment in intestinal regulation (Oakley et al., 2014). The other approach is false signal binders, which mimic pathogen docking sites to gut epithelia that can prevent gut pathogens from recognizing and landing on these sites (Oakley et al., 2014).

The antibiotic effect on the GIT microbiome depends upon the dose of antibiotics and the age of host (Zhou et al., 2007; Stanley et al., 2014). In recent studies, no significant changes were detected in different flocks grown with or without antibiotics (Pedroso et al., 2013). Although, after discontinuing input of antibiotics, GIT microbiota was found to struggle to recover and to resist compositional change (Dethlefson et al., 2008; Stanley et al., 2014). The mannann-oligosaccharide (MOS) treatment raised the Firmicutes population, whereas, there was no effect on the Bacteroidetes population. An increased ratio of Firmicutes/Bacteroidetes is known to be related to obesity in mouse and human, owing to improved energy harvesting capacity for Firmicutes species (Turnbaugh et al., 2006; Stanley et al., 2014). MOS has been shown to reduce the abundance of E. coli (Baurhoo et al., 2007) and Salmonella in cecal contents of experimentally infected chickens (Fernandez et al., 2002). This reduction may be due to the ability of mannose in MOS to bind to mannose-binding lectins of Gram-negative bacteria expressing type 1 fimbriae, which consequently, reduces bacterial attachment to the intestinal epithelial cells (Ganner and Schatzmayer, 2012). The recent study suggested that MOS promotes a greater increase of the villi heights and goblet cell numbers of birds under mild environmental stress than when under antibiotic treatment. Finally, MOS can increase mucins, a
main component of mucus, production so that it can better maintain host intestinal health. The intestinal mucus thickness is a key for defending against invasion of enteric pathogens into epithelial cells and is reduced by antibiotics (Stanley et al., 2014). Unlike MOS, Fructooligosaccharide directly improves host performance, and while it does not inhibit pathogens activity, it promotes beneficial bacteria activity (Xu et al., 2003; Stanley et al., 2014). Several studies have reported that feed with probiotics containing Lactobacillus cultures can enhance chicken to gain weight more and create efficient feed absorption (Oakley et al., 2014). For example, two members of the Lactobacilli, Bifidobacterium thermophilum, and Enterococcus faecium, increased the intestinal muscle thickness and the jejuna villus perimeter and height (Chichlowski et al., 2007). Twelve Lactobacillus strains complex reduced serum cholesterol level in broilers (Jin et al., 1998). As shown in human and pig cases, Firmicutes also helps chickens to increase nutrient absorption, whereas Bacteriodetes decrease nutrient absorption (Jumppertz et al., 2011).

**FUTURE TRENDS**

Increase knowledge of the chicken GIT microbiome via culture-independent metagenomic analysis has helped to understand the dynamics of microbial communities and their role in metabolism and health state of chicken. Further GIT microbiota research will focus on finding the microbe’s role in mediating chicken growth under various environmental conditions, such as nutrients, stress and welfare to determine better quantity and quality of productivity. From metagenomic data of chicken GIT microbiota, identifying and tracking the key enzymatic genes (suggested in Table 1) related to metabolism and health states valuations may find an approach to improve the growth and productivity of chickens by combating deficiencies by administering in diet or doses, and introducing useful or less harmful bacteria possessing nutrient digesting genes. In addition, development of metagenomic approaches and bioinformatic tools will be crucial for progress in an analysis of chicken gut microbiota. By focusing on functional genes of chicken GIT microbiota, strategies for improving nutrient metabolism can be designed. For example, identifying the more effective polysaccharides degrading probiotic bacteria which then can be used to facilitate chicken metabolism. Strain-specific 16S rRNA gene sequencing with improved bioinformatic tools (Eren et al., 2014) can trace how each strain will be changed under various growth and environmental conditions and then monitor the abundance changes of target strains by food treatment. In addition, metabolism specific real-time PCR, performed with primer sets to detect each nutrient gene allows us to monitor the state of GIT microbiota. Metagenomic analyses in combination with other approaches will allow the identification of the microorganisms and real players among GIT microbiota that help in host metabolism and in maintaining its health.

**ACKNOWLEDGMENTS**

This research was supported by the Chung-Ang University Research Grants in 2013 and by a grant from the Strategic Initiative for Microbiomes in Agriculture and Food, Ministry of Agriculture, Food and Rural Affairs, Republic of Korea (as part of the [multi-ministerial] Genome Technology to Business Translation Program).

**REFERENCES**


Aydin, R., M. W. Pariza, and M. E. Cook. 2001. Olive oil prevents...
the adverse effects of dietary conjugated linoleic acid on chick hatchability and egg quality. J. Nutr. 131:800-806.


