Analysis of Gut Microbial Communities and Resistance Genes in Pigs and Chickens in Central China

Yapei Rui \* and Gang Qiu \*†

College of Animal Science and Veterinary Medicine, Xinyang Agriculture and Forestry University, Xinyang 464000, China
\* Correspondence: 2007200008@xyafu.edu.cn
\† These authors contributed equally to this work.

Simple Summary: Objectives: The goal of this study was to reveal the gut microbiota of pigs and chickens in central China and the dynamics of antibiotic resistance genes carried by microorganisms. Methods: Free range and factory-farmed Gushi chickens and Huainan pigs were divided into eight groups. Faecal samples were collected from each group, and the metagenomic sequencing method was used to detect each group of samples. Results: The resistance genes showed the following trend, from high to low relative abundance: tetW was the highest, followed by tetW/N/W, then InuA; and others from high to low were mdtB, InuC, ANT6-IA, ErmB, mdtC, ErmQ, tetBP, vatE, evgS, acrB, cpxA, mefA, Escherichia coli-ampC, tetL, yojl, AcrF and mdtA. Conclusions: All groups administered enrofloxacin and oregano oil did not develop a drug-resistant phenotype during the 5-day treatment period, as grouped in this trial. In 2022, after Announcement No. 194 of the Ministry of Agriculture and Rural Affairs in China, the antimicrobial resistance (AMR) trend declined, but it did not fundamentally change, presumably due to the impact of environmental pollution caused by the long-term use of antimicrobials.

Abstract: Background: Basic data concerning the gut microbiota of the main animal husbandry breeds (pigs and chickens) are scarce in China. The dynamics of gut microbiota (pigs and chickens) in China and antibiotic resistance genes carried by microorganisms in the natural environment are unknown. Methods: Free range and factory-farmed Gushi chickens and Huainan pigs were divided into eight groups. Faecal samples were collected from each group, and the metagenomic sequencing method was used to detect each group of samples. Results: The resistance genes showed the following trend, from high to low relative abundance: tetW was the highest, followed by tetW/N/W, then InuA; and others from high to low were mdtB, InuC, ANT6-IA, ErmB, mdtC, ErmQ, tetBP, vatE, evgS, acrB, cpxA, mefA, Escherichia coli-ampC, tetL, yojl, AcrF and mdtA. All groups administered enrofloxacin and oregano oil did not develop a drug-resistant phenotype during the 5-day treatment period, as grouped in this trial. In 2022, after Announcement No. 194 of the Ministry of Agriculture and Rural Affairs in China, the antimicrobial resistance (AMR) trend declined, but it did not fundamentally change, presumably due to the impact of environmental pollution caused by the long-term use of antimicrobials.

Keywords: gut microbial communities; resistance genes; pig; chicken; China

1. Introduction

The discovery and widespread use of antimicrobials in the development of human civilization has an indelible meaning [1]. The extensive clinical application of antibiotics has greatly reduced the morbidity and mortality of bacterial infectious diseases in humans and animals [2], but due to the irregular use of antimicrobials in China (the largest producer and consumer of antimicrobials), severe bacterial resistance has emerged. Even the prevalence of cross-resistance [3] and multidrug resistance [4–11] has increased. In
recent years, to reduce the irrational use of antimicrobials and control drug resistance, the government of China has introduced a series of policy measures, including the National Action Plan for Containment of Bacterial Resistance. The Ministry of Agriculture and Rural Affairs issued the National Action Plan for Restraining Bacteria of Animal Origin and the National Five-Year Action Plan for Comprehensive Management of Veterinary Drugs (Antibacterial Drugs), aimed at curbing bacterial resistance. The antimicrobial resistance (AMR) of bacteria in China’s livestock and poultry breeding industry has begun to improve.

However, in the long term, AMR has become a major threat to human health in the 21st century. The animal breeding industry is generally believed to be one of the main sources of AMR, which causes resistant bacteria and genes to circulate through the entire chain of animals, food, environment, and humans, and threatens human health. At present, AMR is common in bacteria of animal origin. Multidrug resistant and even pandrug resistant strains are constantly appearing [12–16], which has a huge impact on food safety and public health [17].

Irrational use of antimicrobials promotes the emergence of AMR, resulting in changes in the composition and diversity of animal gut microbiota, which adversely affect host health [18]. Quantitative data on antibiotic resistance genes (ARGs) in pigs and chickens, the main food animals in China, are still lacking, and the relationship between ARGs and bacterial species in the gut microbiome of pigs and chickens has not been elucidated [19]. For most bacteria, metagenomic sequencing technology can predict AMR with high accuracy and has become an effective tool for analysing AMR [20].

The study of microorganisms has been based mainly on pure culture for the hundreds of years since Antoni van Leeuwenhoek invented the microscope. Among the trillions of microbial species, only 0.1–1% of the species are cultivable [21], which greatly limits the study of microbial diversity and molecular mechanisms of population resistance phenotypes.

Metagenomics is a method that was first proposed by Handelman [22] to study all the genomic information contained in the microbial population. Then, Kevin et al. defined metagenomics as a discipline that bypasses the isolation and culture of microbial individuals and applies genomics technology to study the microbial community in the natural environment. Metagenomics avoids the isolation and cultivation of microorganisms in the sample, provides a way to study microorganisms that cannot be isolated and cultivated, and more realistically reflects the composition and interaction of microorganisms in the sample to study their metabolic pathways and gene function at the molecular level [23].

In recent years, with the rapid development of sequencing technology and information technology, the use of next-generation sequencing technology to study metagenomics can quickly and accurately obtain a large amount of biological data and rich microbial research information [24–26]. Sequencing technology is an important means of characterisation that provides a powerful tool for better understanding the molecular mechanism of the microbial resistance phenotype.

According to the statistics of the World Food and Agriculture Organization, China’s number of live pigs, total pork production and poultry egg production ranks first in the world, and poultry egg production has ranked first in the world for many consecutive years. China has become one of the most important animal husbandry countries. The current understanding of the pig and chicken gut microbiome is limited, due mainly to the lack of comprehensive pig and chicken gut microbial reference gene sets and genome sets, which severely limits the mining and use of metagenomic sequencing data.

Fluoroquinolones are important antimicrobials used to treat Salmonella infections in poultry. Studies have shown that high-dose enrofloxacin (ENR, 100 mg/kg body weight (BW)) has a good killing effect on Salmonella enterica in chickens, but at the same time, the composition and structure of intestinal flora are greatly affected [27].
The observed intestinal ENR concentrations were shown to be both theoretically (based on pharmacokinetic and pharmacodynamic principles) and effectively (in vivo measurement) capable of significantly reducing the intestinal *E. coli* wild-type population [28].

In this study, the characteristics of drug resistance groups, drug-resistant bacteria and ARG contamination levels of pigs and chickens on different farms under different feeding modes were studied to provide a reference for the optimisation of antibiotic use in farms.

2. Methods

2.1. Group Test and Sample Collection

In this study, ten fresh faeces of Gushi chickens were collected from an intensive farm in Gushi County, Henan Province, P.R. China, on 15 July 2022, and ten fresh faeces of Gushi chickens were collected from a family farm. Ten fresh faeces of Huainan pigs were collected from an intensive Huainan pig farm, and 10 fresh faeces of Huainan pigs were collected from a mountain farm. The above collected subjects had not used antimicrobials in the past 2 months. On 19 July 2022, 10 faeces of the factory-bred Gushi chicken oregano oil group, 10 faeces of the enrofloxacin group (treatment amount), and 10 faeces of the factory-bred Huainan pig oregano oil group and enrofloxacin group (treatment amount) were collected, respectively. All animal-related work was carried out in strict accordance with the “Administrative Regulations on Laboratory Animals” formulated by the Ministry of Agriculture and Rural Affairs of China, and the project was approved by the Animal Ethics and Welfare Committee (AEWC) of Xinyang Agriculture and Forestry University, and then transported to the laboratory for extraction of total nucleic acid from the samples. Total DNA and total RNA were extracted from 80 samples.

Considering the consistency of animal species, age, weight, feeding place, feed and drinking water, the following test scheme was adopted in order to reduce the impact of the test on animal production. The sample size of this experiment was sufficient to reflect the basic distribution situation of the intestinal flora of the animals.

A. Free-range Gushi chicken without an additive group;
B. Factory-bred Gushi chicken oregano oil group (0.25 mL/kg BW);
C. Factory-bred Gushi chicken ENR group (5.0 mg/kg BW);
D. Free-range Huainan pigs without an additive group;
E. Factory-bred Huainan pigs without an additive group;
F. Factory-bred Huainan pigs (2021011803). All samples were stored in the sample preservation solution, mixed with pig oregano oil group (0.25 mL/kg BW);
G. Factory-bred Huainan pigs ENR Group (5.0 mg/kg BW)
H. Factory-bred Gushi chicken without an additive group.

Factory or intensive breeding was performed by industrial production in a factory farm in the following manner. First, the scale of factory breeding was large (>10,000). Second, the hybrid combination or mating line of animals with a consistent body shape, appearance, and balanced growth, were selected. Third, some mechanised and automatic equipment, such as automatic drinking fountains, automatic feeding troughs, and automatic dung cleaning, were used. Fourth, fewer staff, less land occupation, and high labor productivity, occurred. Fifth, scientific operation and management methods were used to organise production, so that production conditions and technological processes operated in accordance with standards and rules, including full price compound feed, modern epidemic prevention measures and all inflow and all outflow production process.

The family farm, or mountain farm, operated under the free-range breeding model. The family farm or mountain farm engaged in animal husbandry with an annual production of more than 500 pigs, more than 5000 chickens for broilers, more than 2000 laying hens, or more than 3000 mixed breeding poultry. The animals moved freely. The
animals were fed full price compound feed, and wild and green feed. Normal epidemic prevention and expulsion of parasites occurred.

The above animals were all adult animals. Pigs were in the fattening stage (5–8 months), and chickens were in the laying stage (19–28 weeks). The ENR group and oregano oil group were administered oregano orally twice a day for 5 days. The water intake of the animals was controlled in the experiment.

The raw samples or the extracted nucleic acid samples was transported cold (in dry ice, −78.5 °C) to the biotech company. The received samples were tested for eligibility. Qualified DNA samples were detected, and library construction and library detection were performed. Illumina NOVASEQ 6000 (Illumina, Inc., San Diego, CA, U.S.A.) was used to sequence the qualified library, and the raw data obtained by sequencing were used for later information analysis.

2.2. DNA Sample Detection

Library construction and sequencing yielded a sufficient amount of high-quality nucleic acids. The detection of DNA samples mainly included the analysis of DNA purity and integrity by agarose gel electrophoresis (AGE), and Qubit accurately quantified the DNA concentration.

2.3. Library Construction and Library Inspection

Qualified DNA samples were randomly broken into fragments of approximately 350 bp in length with a Covaris ultrasonic breaker, and the entire library was prepared through the steps of end repair, A-tail addition, sequencing adapter addition, purification, and PCR amplification.

After the library was constructed, Qubit 2.0 was used for preliminary quantification, the library was diluted to 2 ng/μL, and then an Agilent 2100 (Agilent, Santa Clara, CA, U.S.A.) was used to detect the insert size of the library. After the insert size met expectations, the Q-PCR method was used to determine the effective concentration of the library. Accurate quantification (library effective concentration >3 nM) was performed to ensure library quality.

2.4. On-Board Sequencing

After the library was qualified, the different libraries were pooled according to the requirements of effective concentration and target data volume, and then Illumina NOVASEQ 6000 sequencing was performed.

2.5. Information Analysis Process

2.5.1. Data Quality Control

There was a certain proportion of low-quality data in the raw data obtained by sequencing. To ensure the accuracy and reliability of subsequent information analysis results, quality control and host filtering of the raw data were first carried out to obtain valid data.

Reads that contained low-quality bases (quality value ≤38) that exceeded a certain percentage (set to 40 bp by default) were removed.
Reads with a certain proportion of N bases (default was 10 bp) were removed.
Reads whose overlap with the adapter exceeded a certain threshold (set to 15 bp by default) were removed.

2.5.2. Metagenome Assembly

Metagenome assembly started from clean data after the quality control of each sample was performed.

After preprocessing, clean data were obtained, and MEGAHIT assembly software was used for assembly analysis.
The assembly parameters were: presets meta-large.

The assembled scaffolds were interrupted from N junctions to obtain sequence fragments without N, called scaftigs (i.e., continuous sequences within scaffolds).

After quality control, the clean data of each sample were compared with the assembled scaftigs of each sample using Bowtie2 Version 2.4.5 software.

The alignment parameters were: --end-to-end, --sensitive, -I 200, -X 400.

For scaftigs generated by single-sample assembly, fragments below 500 bp were filtered out, and statistical analysis and subsequent gene prediction were performed.

2.5.3. Gene Prediction

Gene prediction started from scaftigs assembled from single samples. MetaGeneMark was used for gene prediction, and the genes predicted from each sample were combined to remove redundancy and construct a gene catalogue. Based on the clean data of each sample, the abundance information of the gene catalogue in each sample was obtained.

The basic steps of gene prediction and abundance analysis were as follows.

Starting from the scaftigs (≥500 bp) of each sample, MetaGeneMark was used to perform an open reading frame (ORF) prediction, and information of length less than 100 nt was filtered out.

The open reading frame (ORF) prediction results assembled by each sample were made de-redundant using CD-HIT software to obtain a nonredundant initial gene catalogue (here, in operation, the nucleic acid sequences encoded by nonredundant continuous genes are called genes). The default was identity 95%, coverage 90% for clustering, and the longest sequence was selected as the representative sequence.

The parameters were: -c 0.95, -G 0, -aS 0.9, -g 1, -d 0.

The clean data of each sample were compared with the initial gene catalogue by Bowtie2, and the number of reads of the gene in each sample was calculated.

The alignment parameters were: --end-to-end, --sensitive, -I 200, -X 400.

The genes that supported the number of reads ≤2 in each sample were filtered out, and the final gene catalogue (unigenes) for subsequent analysis was obtained.

Based on the number of reads and gene length in the alignment, the abundance information of each gene in each sample was calculated, and the calculation formula was as follows:

\[
G_k = \frac{r_k}{L_k} \times \frac{1}{\sum_{i=1}^{n} \frac{r_i}{L_i}}
\]

where \( r \) is the number of reads of the gene in the alignment, and \( L \) is the length of the gene.

Based on the abundance information of each gene in each sample in the gene catalogue, the basic information statistics, core–pan gene analysis, correlation analysis between samples, and Venn diagram analysis of the number of genes were performed.

2.5.4. Species Annotation

The species annotation information of each gene (Unigene) was obtained from the gene catalogue and compared with the MicroNR library, and the species abundance table of different taxonomic levels was obtained by combining the gene abundance table.

Based on the species abundance table and functional abundance table, the abundance clustering analysis, PCA and NMDS dimensionality reduction analysis, ANOSIM analysis and sample clustering analysis were able to be performed. When group information was available, multivariate statistical analysis of Metastat and LEfSe and comparative analysis of metabolic pathways were also able to be performed to explore the differences in species composition and functional composition among samples.
2.5.5. Resistance Gene Annotation

The gene catalogue and the Comprehensive Antibiotic Resistance Database (CARD) were used for annotation, the abundance distribution of resistance genes was obtained, and the species attribution and resistance mechanism of these resistance genes were found.

3. Results

3.1. Sequencing Data Preprocessing Results

The sequencing data preprocessing results are shown in Tables 1–3 and Figure 1.

Table 1. Data preprocessing statistics table.

<table>
<thead>
<tr>
<th>#Sample</th>
<th>InsertSize (bp)</th>
<th>RawData</th>
<th>CleanData</th>
<th>Clean_Q20</th>
<th>Clean_Q30</th>
<th>Clean_GC (%)</th>
<th>Effective (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>350</td>
<td>6633.18</td>
<td>6629.30</td>
<td>96.2</td>
<td>89.95</td>
<td>40.72</td>
<td>99.941</td>
</tr>
<tr>
<td>B</td>
<td>350</td>
<td>6982.84</td>
<td>6969.06</td>
<td>96.11</td>
<td>89.8</td>
<td>41.56</td>
<td>99.803</td>
</tr>
<tr>
<td>C</td>
<td>350</td>
<td>6714.52</td>
<td>6708.95</td>
<td>96.27</td>
<td>90.09</td>
<td>40.08</td>
<td>99.917</td>
</tr>
<tr>
<td>D</td>
<td>350</td>
<td>6155.96</td>
<td>6152.45</td>
<td>96.03</td>
<td>89.66</td>
<td>41.1</td>
<td>99.943</td>
</tr>
<tr>
<td>E</td>
<td>350</td>
<td>6822.24</td>
<td>6818.75</td>
<td>95.88</td>
<td>89.34</td>
<td>39.31</td>
<td>99.949</td>
</tr>
<tr>
<td>F</td>
<td>350</td>
<td>6528.38</td>
<td>6524.90</td>
<td>95.99</td>
<td>89.38</td>
<td>35.21</td>
<td>99.947</td>
</tr>
<tr>
<td>G</td>
<td>350</td>
<td>6884.02</td>
<td>6881.10</td>
<td>95.95</td>
<td>89.5</td>
<td>39.52</td>
<td>99.958</td>
</tr>
<tr>
<td>H</td>
<td>350</td>
<td>6645.17</td>
<td>6639.25</td>
<td>96.39</td>
<td>90.44</td>
<td>42.14</td>
<td>99.911</td>
</tr>
</tbody>
</table>

#Sample, sample name; InsertSize (bp), insert fragment length (default 350 bp library); RawData, the raw data of the computer; CleanData, the valid data obtained and filtered; Clean_Q20, percentage of bases with a sequencing error rate of less than 0.01 (quality value greater than 20) in CleanData; Clean_Q30, percentage of bases in CleanData with a sequencing error rate less than 0.001 (quality value greater than 30); Clean_GC (%), GC content of bases in CleanData; Effective (%), the percentage of effective data (CleanData) to raw data (RawData).

Table 2. The assembly results of each sample. Scaftig information (≥500 bp).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>77,097,982</td>
<td>50,858</td>
<td>1515.95</td>
<td>2087</td>
<td>625</td>
<td>304,611</td>
</tr>
<tr>
<td>B</td>
<td>103,806,670</td>
<td>79,760</td>
<td>1301.49</td>
<td>1565</td>
<td>598</td>
<td>94,301</td>
</tr>
<tr>
<td>C</td>
<td>53,150,756</td>
<td>34,873</td>
<td>1524.12</td>
<td>2127</td>
<td>628</td>
<td>182,450</td>
</tr>
<tr>
<td>D</td>
<td>114,767,070</td>
<td>76,524</td>
<td>1499.75</td>
<td>2000</td>
<td>634</td>
<td>269,282</td>
</tr>
<tr>
<td>E</td>
<td>91,187,139</td>
<td>64,129</td>
<td>1421.93</td>
<td>1804</td>
<td>619</td>
<td>395,129</td>
</tr>
<tr>
<td>F</td>
<td>98,440,700</td>
<td>58,234</td>
<td>1690.43</td>
<td>2689</td>
<td>644</td>
<td>185,631</td>
</tr>
<tr>
<td>G</td>
<td>85,667,957</td>
<td>51,678</td>
<td>1657.73</td>
<td>2510</td>
<td>637</td>
<td>297,861</td>
</tr>
<tr>
<td>H</td>
<td>77,519,231</td>
<td>54,238</td>
<td>1429.24</td>
<td>1830</td>
<td>617</td>
<td>157,857</td>
</tr>
</tbody>
</table>

Sample ID, indicates the sample name; Total Len. (bp), indicates the total length of the assembled scaftigs; No., indicates the total number of scaftigs assembled; Average Len. (bp), indicates the average length of scaftigs; N50 Len. (bp), indicates that the scaftigs were sorted by length and then summed from long to short. When the sum value reached 50% of the total length of scaftigs, the length value of scaftigs was found; N90 Len. (bp), indicates that the scaftigs are sorted by length and then summed from long to short. When the sum value reaches 90% of the total length of scaftigs, the length value of scaftigs was found; Max Len, indicates the length of the longest scaftigs assembled.
Table 3. Gene catalogue basic information. ORFs No. indicates the number of genes in the gene catalogue.

<table>
<thead>
<tr>
<th>ORFs No.</th>
<th>445,260</th>
</tr>
</thead>
<tbody>
<tr>
<td>integrity:end</td>
<td>82,047  (18.43%)</td>
</tr>
<tr>
<td>integrity:start</td>
<td>93,479  (20.99%)</td>
</tr>
<tr>
<td>integrity:all</td>
<td>237,785 (53.4%)</td>
</tr>
<tr>
<td>integrity:none</td>
<td>31,949  (7.18%)</td>
</tr>
<tr>
<td>Total Len. (Mb)</td>
<td>300.11</td>
</tr>
<tr>
<td>Average Len. (bp)</td>
<td>674.02</td>
</tr>
<tr>
<td>GC percent</td>
<td>39.44</td>
</tr>
</tbody>
</table>

integrity:end indicates the number and percentage of genes that only contain start codons; integrity:start indicates the number and percentage of genes that only contain stop codons; integrity:none indicates the number and percentage of genes with no start codon and no stop codon; integrity:all indicates the percentage of the number of complete genes (both start codons and stop codons); Total Len. (Mb) indicates the total length of genes in the gene catalogue, in millions; Average Len. (bp) indicates the average length of genes in the gene catalogue; GC percent indicates the predicted overall GC content of genes in the gene catalogue.

Figure 1. Scaf/FIG length distribution statistics of each sample (≥500 bp). The first vertical axis (Frequency (#)) represents the number of scaf/FIGs; the second vertical axis (Percentage (%)) represents the percentage of the number of scaf/FIGs; the horizontal axis represents the length of scaf/FIGs.
3.2. Metagenome Assembly (≥500 bp)

Based on the assembly results, the distribution of the length of scaffigs in each sample was counted and plotted, and the results are shown in Figure 1.

The gene catalogue length distribution is shown in Figure 2.

![UniqGene Length Distribution](image)

**Figure 2.** Gene catalogue length distribution. The first vertical axis (Frequency (†)) represents the number of scaffigs; the second vertical axis (Percentage (%)) represents the percentage of the number of scaffigs; the horizontal axis represents the length of scaffigs.

3.3. Core–Pan Gene Analysis

Based on the abundance table of genes in each sample, the number of genes in each sample can be obtained. By randomly extracting different numbers of samples, the number of genes between combinations of different numbers of samples can be obtained. Therefore, we constructed and mapped the dilution curve of the core and pan genes, and the results are shown in Figure 3.

![Dilution Curve](image)

**Figure 3.** Dilution curve of the core and pan genes. (a) Pan gene dilution curve. (b) Core gene dilution curve. The abscissa indicates the number of samples. The vertical axis represents the number of genes in the samples (1–8: A–G). 4E+5: 4 × 10^5, Same with the rest.
To examine the number of genes between the specified groups, the coexistence and unique information between different groups were analysed, and a Venn graph, or petals, were drawn. The results are shown in Figure 4.

Figure 4. Number of genes. In the figure, each circle represents a sample; the numbers in the overlapping parts of the circles represent the number of genes shared between the samples; the numbers without the overlapping part represent the unique number of genes of the sample.

3.4. Gene Number-Based Correlation Analysis between Samples

Biological repetition is necessary for biological experiments, and high-throughput sequencing technology is no exception. The correlation between the genetic abundance of the sample is an important indicator for testing the reliability of the experiment and whether the sample is selected. The closer to the correlation coefficient, the higher the similarities of the gene abatement mode between the samples. The results are shown in Figure 5.

Figure 5. Heatmap of the correlation coefficient between samples. In the figure, different colours represent the Spearman correlation coefficient; see the legend on the right for the relationship between the correlation coefficient and colour. The darker the colour is, the greater the absolute value of the correlation coefficient between samples. The right deviation of the ellipse indicates that the correlation coefficient is positive and the left deviation is negative. The flatter the ellipse, the larger the absolute value of the correlation coefficient.
3.5. **Microbial Species Annotation**

3.5.1. Basic Steps for Species Annotation

Genes were compared in various functional databases by DIAMOND software. Unigenes were compared with sequences for alignment (basic local alignment search tool (BLAST)p, e-value ≤ 1 × 10⁻⁵).

Alignment result filtering: for the alignment result of each sequence, the alignment result with e-value ≤ minimum e-value*10 was selected for subsequent analysis.

After filtering, since each sequence may have multiple alignment results and obtain multiple different species classification, to ensure its biological significance, the lowest common ancestor (LCA) algorithm (applied to the systematic classification of MEGAN software) was applied.

The taxonomic level before a branch represented the species annotation information of the sequence.

Based on the LCA annotation results and gene abundance table, the abundance information of each sample at each taxonomic level (Jiemenae, genus and species) was obtained. The abundance of a species in a sample was equal to the sum of gene abundances that the species annotated as the species.

Based on the LCA annotation results and the gene abundance table, the number of genes in each sample at each taxonomic level (Kingmenae, genus and species) was obtained. Among the genes of the species, the number of genes whose abundance was not 0 was presented.

Based on the abundance table at each taxonomic level (Jiemenae genus and species) and Krona analysis, the relative abundance overview was displayed, the abundance cluster heatmap was displayed, PCA and nonmetric multidimensional scaling (NMDS) dimensionality reduction were analysed, the ANOSIM difference between (within) groups was analysed, and Metastat and LEfSe multivariate statistical analyses of different species between groups were performed.

3.5.2. Overview of the Relative Abundance of Species

To comprehensively and intuitively display the relative abundance of different classification levels in each sample, we used Krona to display the species annotation results. The results are shown in Figure 6.
Figure 6. (a-h). Annotate results for species using Krona. In Figure 6, the circles represent different taxonomic levels (kingdom, phylum, class, order, family, genus, species) from inside to outside; the size of the sector represents the relative proportion of different species.

Studies have shown that the gut microbiota is individual-specific, and the composition of the microbiota of pigs under different feeding conditions and different intestinal parts is different.

Based on the relative abundance tables of different classification levels, the top 10 species with the largest relative abundance were selected in each group, the rest of the species were set as Others, and the species annotation results corresponding to each sample in the column chart of relative abundance at different taxonomic levels are shown in Figure 7.
Figure 7. The relative abundance of species was presented at the phylum and genus levels. (a). Histogram of relative abundance was shown at the gate level. (b) Histogram of relative abundance was presented at the genus level. The horizontal axis represents the sample name. The vertical axis represents the relative proportion of species annotated to a certain type. The legend on the right is the species category corresponding to each colour block.

3.5.3. Cluster Analysis of the Number and Relative Abundance of Annotated Genes

Based on the relative abundance tables of different classification levels, the top 35 genera in terms of abundance and their abundance information in each sample were selected to draw a heatmap, and were clustered at the species level to facilitate the display of results and information discovery to determine the species with more aggregation in the sample. The results are shown in Figure 8.

Figure 8. Cluster heatmap of the number and abundance of genes at the genus level. (a) Unannotated number statistical heatmap; the horizontal axis is the sample name; the vertical axis is the species information; and different colours represent the number of unigenes. (b) Cluster heatmap of relative abundance at the genus level; the horizontal direction is the sample information; the vertical direction is the species information. The cluster tree on the left in the figure is a species cluster tree. The value corresponding to the intermediate heatmap is the Z-value obtained after the relative abundance of species in each row is normalised; that is, the Z-value of a sample in a certain classification is the value obtained by dividing the difference between the relative abundance of the sample in that classification and the average relative abundance of all samples in that classification by the standard deviation of all samples in that classification.

3.5.4. Dimension Reduction Analysis Based on Species Abundance

At present, principal component analysis (PCA) and nonmetric multidimensional scaling (NMDS) are the main methods of dimension reduction analysis applied to ecological research. Among these methods, PCA is a dimension reduction analysis based
on a linear model that applies the method of variance decomposition to reduce the dimension of multidimensional data to extract the most important elements and structures in the data. PCA can extract two coordinate axes that reflect the differences between samples to the greatest extent to reflect the differences of multidimensional data on the two-dimensional coordinate map and then reveal the simple laws in the complex data background. NMDS is a nonlinear model whose purpose is to overcome the shortcomings of linear models and better reflect the nonlinear structure of ecological data. NMDS analysis is applied to reflect the species information contained in the sample in the form of points in the multidimensional space, while the difference between different points is reflected by the distance between points, which can reflect the differences between groups or within groups of samples. Based on the species abundance tables at different classification levels, we conducted PCA and NMDS analysis. If the species composition of the samples was more similar, then the distance between them was closer, as shown in the PCA and NMDS diagrams in Figure 9.

![PCA and NMDS results of species based on the phylum level.](image)

**Figure 9.** PCA and NMDS results of species based on the phylum level. (a) For phylum-level PCA, the abscissa represents the first principal component, and the percentage represents the contribution value of the first principal component to the sample difference. The ordinate represents the second principal component, and the percentage represents the contribution value of the second principal component to the sample difference. Each point in the figure represents a sample, and the samples in the same group are represented by the same colour. (b) For gate-level NMDS analysis, each point in the figure represents a sample, the distance between points represents the degree of difference, and the samples in the same group are represented by the same colour. When the stress is less than 0.2, NMDS analysis has a certain reliability.

### 3.5.5. Dimension Reduction Analysis of Bray–Curtis Distance Based on Species Abundance

Principal coordinate analysis (PCoA) extracts the most important elements and structures from multidimensional data through a series of eigenvalues and eigenvector ordering. We conducted PCoA based on the Bray–Curtis distance and selected the principal coordinate combination with the largest contribution rate for graphic display. If the sample distance was closer, the species composition structure was more similar. Therefore, the samples with high similarity in community structure tended to gather together, and the samples with large differences in community were far apart. The Bray–Curtis distance matrix was obtained based on species abundance tables at different taxonomic levels. The results of the PCoA analysis based on the gate level are shown in Figure 10.
Figure 10. Species PCoA results based on the phylum level. Note: the abscissa represents one principal component, the ordinate represents another principal component, and the percentage represents the contribution value of the principal component to the sample difference. Each point in the figure represents a sample, and the samples in the same group are represented by the same colour.

3.6. Cluster Analysis of Samples Based on Species Abundance

To study the similarity of different samples, a clustering tree of the samples was constructed by clustering the samples. The Bray–Curtis distance is the most common distance indicator used in systematic clustering. It is used mainly to describe the similarity between samples, and its size is the main basis for sample classification. Starting from the abundance table of genes in each sample, the Bray–Curtis distance matrix was used to perform cluster analysis among samples, and the clustering results were integrated with the relative abundance of species at the phylum level in each sample for display. The results are shown in Figure 11.

Figure 11. Clustering tree based on the Bray–Curtis distance. The left side is the Bray–Curtis distance clustering tree structure; the right side is the species relative abundance distribution map of each sample at the phylum level.
3.7. Linear Discriminant Analysis Effect Size (LEfSe) Analysis of Different Species between Groups

To screen the species biomarkers with significant differences between groups, first, the rank sum test method was used to detect the different species between different groups, and linear discriminant analysis (LDA) was used to achieve dimensionality reduction and evaluate the influence of the different species; that is, the LDA score was obtained. The LEfSe analysis results of species with differences between them included three parts, namely, the LDA value distribution histogram, the evolutionary branch diagram (phylogenetic distribution), and the abundance comparison diagram of biomarkers with significant differences between groups. The LDA value distribution map and evolutionary branch map of different species are shown in Figure 12.

![LDA value distribution and evolutionary clades](image)

**Figure 12.** Distribution of LDA values and evolutionary clades of different species. (a) is the LDA value distribution of different species. The LDA value distribution histogram shows the species whose LDA score is greater than the set value (the default setting is 4), that is, the biomarker with a significant difference between groups, and the length of the histogram represents the impact size of different species (i.e., LDA score). (b) shows the evolutionary clades of different species. The circles...
radiating from the inside to the outside represent the taxonomic level from phylum to genus (or species). Each small circle at a different taxonomic level represents a taxonomy at that level, and the diameter of the small circle is proportional to the relative abundance. Colouring principle: Species with no significant difference are uniformly coloured yellow, and the biomarkers of different species are coloured according to the group. The red nodes represent the microbial groups that play an important role in the red group, and the green nodes represent the microbial groups that play an important role in the green group. The microorganism species names represented by the English letters in the figure are shown in the legend on the right.

3.8. Annotation of Resistance Genes

Resistance genes are ubiquitous in both human intestinal microorganisms and other environmental microorganisms. The abuse of antimicrobials leads to irreversible changes in the microbial community in the human body and the environment, which poses risks to human health and the ecological environment. Therefore, research on resistance genes has received extensive attention from researchers. CARD is a resistance gene database that has emerged in recent years. It has the advantages of comprehensive information, user-friendliness, timely updates and maintenance. The core component of this database is antibiotic resistance ontology (ARO), which integrates information such as sequence, antibiotic resistance, mechanism of action and association between AROs and provides online interfaces between ARO and databases such as Protein Data Bank (PDB) and the National Center for Biotechnology Information (NCBI).

3.8.1. Basic Steps of Resistance Gene Annotation

Unigenes were compared with the CARD database using the resistance gene identifier (RGI) software provided by the CARD database (v2.0.1) (RGI has built-in BLASTp, and the results were scored by bitscore value comparison).

According to the comparison results of RGI and the abundance information of unigenes, the relative abundances of AROs were calculated.

Based on the abundance of ARO, the following were generated: the column diagram of abundance, the heatmap of abundance clustering, the circle diagram of abundance distribution, the analysis of ARO differences between groups, and the species attribution analysis of resistance genes (unigenes annotated to ARO) (for some AROs with longer names, we will display them in the form of the first three words and underlined abbreviations).

3.8.2. Overview of Resistance Gene Abundance

Based on the relative abundance table of resistance genes, we calculated the content and percentage of ARO in each sample and screened the top 20 AROs with the largest abundance. The results are shown in Figure 13.

![Figure 13](image-url)

**Figure 13.** Histogram of the abundance of different AROs in each sample. (a) The relative abundance of all genes of ARO in each sample, in ppm, which is the result of enlarging the original relative abundance data by 106 times. (b) Indicates the relative abundance of the top 20 AROs among all AROs, and the others are the sums of the relative abundances of the top 20 non-AROs.
To more intuitively observe the abundance ratio of ARO in each sample as a whole, and more intuitively display the overall distribution of ARO abundance, the ARO with the largest abundance of TOP10 was selected to be drawn as an overview circle, as shown in Figure 14.

Figure 14. Overview circle of resistance genes. The circle diagram is divided into two parts, with sample information on the right and ARO information on the left; Different colours in the inner circle represent different samples and AROs. The scale is the relative abundance, in ppm. The left side is the sum of the relative abundances of each ARO in a sample, and the right side is the sum of the relative abundances of each ARO in a sample. The left side of the outer ring is the relative percentage of each sample in an ARO, and the right side of the outer ring is the relative percentage of each ARO in a sample.

The resistance mechanism of resistance genes in the CARD database was classified, and the following resistance mechanism distribution map was drawn according to the relationship between the action mechanism of these resistance genes and species. The results are shown in Figure 15.
Figure 15. Overview circle of resistance mechanism and species. The circle diagram is divided into two parts: the right side is phylum-level species information, and the left side is resistance-mechanism information. Different colours in the inner circle indicate the resistance mechanisms of different species and resistances, and the scale is the number of genes. The left side is the sum of the number of resistance genes containing such resistance mechanisms in the species, and the right side is the sum of the number of resistance genes contained in the species in different resistance mechanisms. The left side of the outer circle is the relative proportion of resistance genes in each species to the resistance genes of their resistance mechanisms, and the right side of the outer circle is the relative proportion of resistance genes in each resistance mechanism to the resistance genes of their species.

The sequencing results clearly showed that the abundance and relative percentage of drug resistance genes in the eight typical samples showed a trend from high to low.


All groups administered enrofloxacin and oregano oil did not develop a drug-resistant phenotype during the 5-day treatment period, as grouped in this trial.

4. Discussion

One of the most important functions of the gut microbiota is to prevent bacterial overgrowth and infection by pathogenic microorganisms, mainly through the production
of antibacterial substances, site-occupying protection, and competition for nutrients with pathogenic bacteria.

Bacteriocins are antimicrobial peptides or complex proteins synthesised by the ribosomes of various Gram-positive and Gram-negative bacteria, and have bacteriostatic and bactericidal effects. For example, the bacteriocin-like substances secreted by lactic acid bacteria have broad-spectrum antibacterial effects and can inhibit the growth and reproduction of Gram-negative bacteria such as *Escherichia coli*. In addition, metabolites such as short-chain fatty acids (SCFAs) produced during bacterial metabolism can reduce intestinal pH, thereby inhibiting the growth and reproduction of exogenous bacteria. At the same time, the acidification of the intestinal environment promotes the acceleration of intestinal peristalsis, resulting in exogenous bacteria not interacting with the mucosa, and exogenous bacteria are excreted. The space-occupying protective effect of intestinal flora is to form a biofilm by closely combining with intestinal mucosal epithelial cells, thereby effectively preventing pathogenic bacteria from contacting the intestinal mucosa. Under the limited nutrient intake of the host, intestinal flora, especially symbiotic bacteria, have a natural competitive advantage for these nutrients, inhibiting the growth and reproduction of pathogenic bacteria.

The diversity and distribution of drug resistance groups in pig and chicken faeces under different feeding modes (free range and factory feeding) were systematically analysed.

Based on the basic situation of the intestinal flora of animals (pig and chicken) in central China, the drug-resistant phenotype presents dual traces of history and current practice. Tetracycline resistance remains the most abundant type of resistance, and this result is consistent with the findings of other Chinese scholars in other provinces [29,30].

Some factory-farmed pigs and chickens have higher levels of intestinal bacterial resistance than free-range pigs and chickens.

In this study, an integrated gene set of gut microbes was constructed using the metagenomic sequencing data of eight groups of swine and chicken gut microbes from different feeding patterns, combined with the gene sets of existing metagenomic sequencing data, which included 237,785 complete genes, of which 5% were unknown, expanding the comprehensiveness and completeness of the swine and chicken gut microbial gene sets. The genus-level dominant species represented by the 35 metagenomic assembled genomes were analysed by software. Using analytical software, we further revealed the diversity of swine and chicken gut microbiota composition and drug resistance phenotypes and identified differences in drug resistance phenotypes.

Comparative studies in animal samples provide important links to understanding the global distribution of ARGs and the spread of multidrug-resistant bacteria, resistance exchange networks, and how different habitats and phylogenetic relationships influence the evolutionary dynamics of global antimicrobial resistance.

Drug resistance in pathogenic bacteria is a growing threat to global health, and the development and spread of drug resistance in microorganisms is largely attributable to the misuse of antimicrobials.

AMR in animals is country-, or even, region-specific. A map of the global distribution of resistance shows resistance hotspots in northeastern China, northeastern India and northern Pakistan, while resistance is just beginning to emerge in central India and central and southern China, Kenya, Uruguay and southern Brazil [31–34]. In 2022, after the announcement of the Ministry of Agriculture and Rural Affairs No. 194, this trend declined, but it did not fundamentally change, presumably due to the impact of environmental pollution caused by the long-term use of antimicrobials.

There are differences in the intestinal microbiota under different feeding styles [35]. Animals receive different nutrients from free-range and factory-raised feed, which also has an impact on their gut flora. After group administration, the abundance of intestinal flora of animals in the two feeding modes showed same trend, and the abundance of AMR
genes had their own characteristics, which may be related to the long-term drug history, the AMR in the environment, and the feed of adult animals.

5. Conclusions
Metagenomic sequencing was conducted to study the relative changes of intestinal microflora and drug resistance genes in pigs and chickens in central China, after treatment amounts of enrofloxacin and oregano oil were added into the feed in factory and free range mode, respectively. The consistency of the trend of intestinal microflora changes and the existence of drug resistance genes were found after treatment, providing reference for rational drug use in the veterinary clinic. The results showed that tetW was the highest relative abundance resistance gene, followed by tetW/N/W, then lnuA; and others from high to low were mdtB, lnuC, ANT6-la, ErmB, mdtC, ErmQ, tetBP, vatE, evgS, acrB, cpxA, mefA, Escherichia coli-ampC, tetL, yofI, AcrF and mdtA.

Author Contributions: Conceptualization, G.Q. and Y.R.; methodology, Y.R.; software, G.Q.; validation, G.Q. and Y.R.; formal analysis, G.Q.; investigation, Y.R.; resources, G.Q.; data curation, Q.G.; writing—original draft preparation, Y.R.; writing—review and editing, G.Q.; visualization, G.Q.; supervision, G.Q.; project administration, G.Q.; funding acquisition, G.Q. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the following projects for their support: The Natural Science Foundation of Henan Province (212300410385, Qiu Gang), the Foundation of Central Laboratory of Xinyang Agriculture and Forestry University (FCL202013, Gang Qiu), the Key Scientific Research Project of Colleges and Universities of Henan Province (21A230016, Gang Qiu), the Innovative Research Team of Poultry Germplasm Resources Application and Healthy Breeding of Daihe Mountain Area in Xinyang Agriculture and Forestry University (XNKJTD-013) and the Key Scientific Research Project of Colleges and Universities of Henan Province (21B230009, Yapel Rui).

Institutional Review Board Statement: In this study, the research experiments conducted with animals were approved by the Ethical Committee and responsible authorities of our research organisation(s), following all guidelines, regulations, and legal and ethical standards required for animal experimentation.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analysed during this study were included in this published article. The raw sequence data reported in this paper were deposited in the Genome Sequence Archive (Genomics, Proteomics and Bioinformatics 2021) in the National Genomics Data Center (Nucleic Acids Res 2022), China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA008466, CRA008470), and are publicly accessible at https://ngdc.cncb.ac.cn/gsa.

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

References