

Identification of SNP markers associated with gut microbiome composition in chicken

A. Psifidi^{1,2}, M. Crotta², R.J. Pandit³, B. Fosso⁴, P.G. Koringa³, G. Limon², K. Boulton¹, C. Dadousis¹, G. Banos¹, J. Guitian², F.M. Tomley², D.N. Rank³, C.G. Joshi³, D. Hume¹ and D.P. Blake²

¹ The Roslin Institute & R(D)SVS, University of Edinburgh, UK, ² Royal Veterinary College, University of London, UK, apsifidi@rvc.ac.uk (Corresponding Author), ³ College of Veterinary Science and Animal Husbandry, Anand Agricultural University, India, ⁴ Institute of Biomembrans, Bioenergetics and Molecular Biotechnologies, Bari, Italy

Summary

Enteric microbiome composition and structure varies widely between individuals, and has been associated with multiple health conditions and differences in productivity in farm animals. Enteric microbial communities are clearly influenced by environmental factors and food, but host genotype may also have an impact. This study is part of an effort to comprehensively profile interactions between chicken genetic variation and the composition of the gut microbial ecosystem on a genome-wide scale. We collected blood and caecal lumen contents from 150 indigenous Indian Kadaknath chickens and 150 commercial-type Cobb400 broilers, raised in groups together under the same farm conditions. All chickens were genotyped with the 600K Affymetrix HD SNP array and the caecal lumen microbiome was defined using Illumina 16s rDNA amplicon sequencing. Single marker regression genome-wide association analysis was conducted. Using this dataset, we identified significant associations at genome-wide level ($P < 0.05$ Bonferroni corrected) between host genetic variation and gut microbiome composition.

Keywords: chickens, gut microbiome, GWAS, indigenous chickens, commercial broilers, genetic variation

Introduction

The importance of the caecal microbiome to chicken health and productivity has long been recognised, influencing food conversion, susceptibility to disease and colonisation by zoonotic pathogens (Stanley et al., 2014). The composition of the microbiome is influenced by multiple environmental factors (Stanley et al., 2014). For example, changes in host diet, intake of drugs and antibiotics, interactions with pathogens, social contact and interaction with the environment affect gut microbiome communities at taxonomic and functional levels. In addition, variation in immunity-related pathways has been shown to influence microbiome structure in humans and is also likely to affect chicken microbiomes. A breeding programme to promote good protective microbiota may be one approach to reduce dependence on antibiotic-mediated pathogen control. To test the feasibility of such a programme, we investigated the genomic architecture of gut microbiome composition using two chicken types: 1) an indigenous Indian breed, “Kadaknath”, that has been associated with improved

egg, meat and welfare traits (Haunshi et al., 2011), resistance to infectious disease, and immune parameters distinct from those of modern commercial chickens (Ramasami et al., 2010); and 2) a commercial broiler line, “Cobb400”, that has been extensively used in India and other Asian countries. We performed a genome-wide association analysis (GWAS) to identify single nucleotide polymorphisms (SNP), good candidate genes and biological pathways that associate with microbiome composition.

Material and methods

Populations

Kadaknath and Cobb400 chickens used in this study were raised together, under the same conditions, in 30 experimental farms. Half of the farms were located at the Central Poultry Research Station at AAU (Anand Agricultural University), Gujarat and the other half in commercial production units in the same region. Five individual chickens of each breed were hatched on the same day and reared in neighbouring pens in single poultry houses (independent replication). A deep litter system was employed using rice husk as a substrate in common with local practices.

Data

In total, blood samples from 300 birds (150 each Kadaknath and Cobb400) were spotted onto FTA cards. Both caeca were collected from each chicken and their contents were pooled and recovered into sterile cryovials containing Bacterial Protect RNA reagent. Genomic DNA extracted from the FTA cards was successfully genotyped using a high-density SNP array (600K, Affymetrix). Metagenomic DNA (mgDNA) was extracted from the caecal contents of each chicken. The V3 and V4 16S rRNA gene hypervariable regions were amplified from caecal mgDNA samples; amplicons were further processed for library preparation using Illumina’s Nextera XT library preparation kit (Illumina, USA). Sequencing was performed using an Illumina MiSeq desktop sequencer and generated files were analysed with the EBI Metagenomics (European Bioinformatics Institute Metagenomics) pipeline (Mitchell et al., 2016). Finally, microbiome data for each of the 300 birds were available as an Operational Taxonomic Units (OTU) table. Information on village, farm, diet, antibiotic administration, vaccination administration, use of water disinfectant, ventilation system, presence of fowl, presence of other animals, accessibility to visitors, sex and age were recorded.

Statistical Analyses

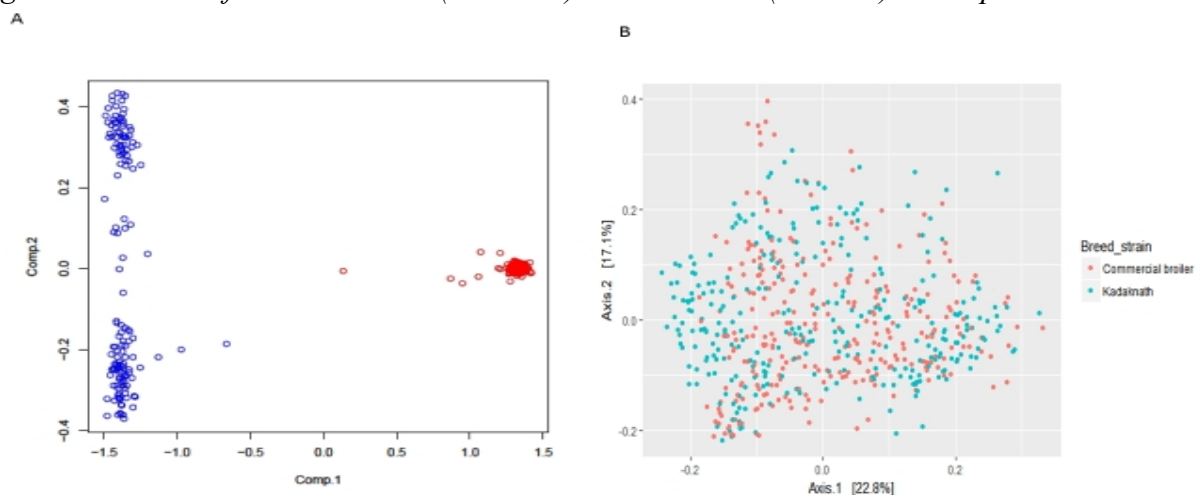
Initially, a multidimensional scaling analysis (MDS) was performed using an IBS distance matrix to identify the occurrence of any genetic differences between the two populations using the R/GenABEL package. Studying microbial communities is mainly based on the investigation of the diversity, a measure of the complexity of each community in terms of species (or OTUs) richness and evenness. Different measures of the within-sample microbial diversity (alpha-diversity) were calculated using the Vegan library from the R/Bioconductor package. To measure the between-sample differences in the microbial communities (beta-diversity), the distance metric UniFrac was used and the weighted Principal Coordinate Analysis (PCoA) performed on the obtained distance matrix using the R/Bioconductor package. The first five principal components of beta-diversity as well as alpha-diversity estimations have previously been used as phenotypes to study microbiome composition as a quantitative trait (Blekhman et al., 2015, Davenport et al., 2015). The Inverse Simpson index for alpha diversity and the first four beta diversity axes of the PCoA (that collectively

explained 58% of the microbiome variance) were used as individual microbiome phenotypes in the present study. All phenotypes were normally distributed. Individual chicken records were then adjusted for the fixed effects of sex, age, breed, diet, antibiotic administration, vaccination, water disinfectant, ventilation system, presence of fowl and other animals, accessibility to visitors and the random effect of farm. Data were pre-corrected for fixed effects and the residuals used as phenotypes in the ensuing GWAS. After quality control (the criteria for marker retention were: call rates 0.95, individuals with missing genotypes 0.05, Hardy-Weinberg equilibrium $P < 10^{-4}$, minor allele frequency 0.05), just over 400,000 SNPs remained for further analyses. The algorithm GEMMA was used to run the GWAS analyses based on a mixed model that included the genomic relationship matrix among-individuals as a random effect. After the Bonferroni correction, the $P < 0.05$ genome-wide significance threshold was set at 1.27×10^{-7} and the suggestive (i.e. one false positive per genome scan) significance threshold at 2.50×10^{-6} . Furthermore, using Ensembl, the significant SNPs were mapped on the reference genome (Galgal5) and a search for annotated genes within 0.5 MB windows around the significant SNPs was performed.

Results and Discussion

The initial MDS analysis revealed that the two populations were genetically distinct (Figure 1A). However, beta-diversity (PCoA) analysis revealed overlapping microbiome profiles for the two populations (Figure 1B). The first, second, third and fourth beta-diversity axes explained 23%, 17%, 10% and 8%, respectively, implying that the four axis are sufficient to study the microbiome composition.

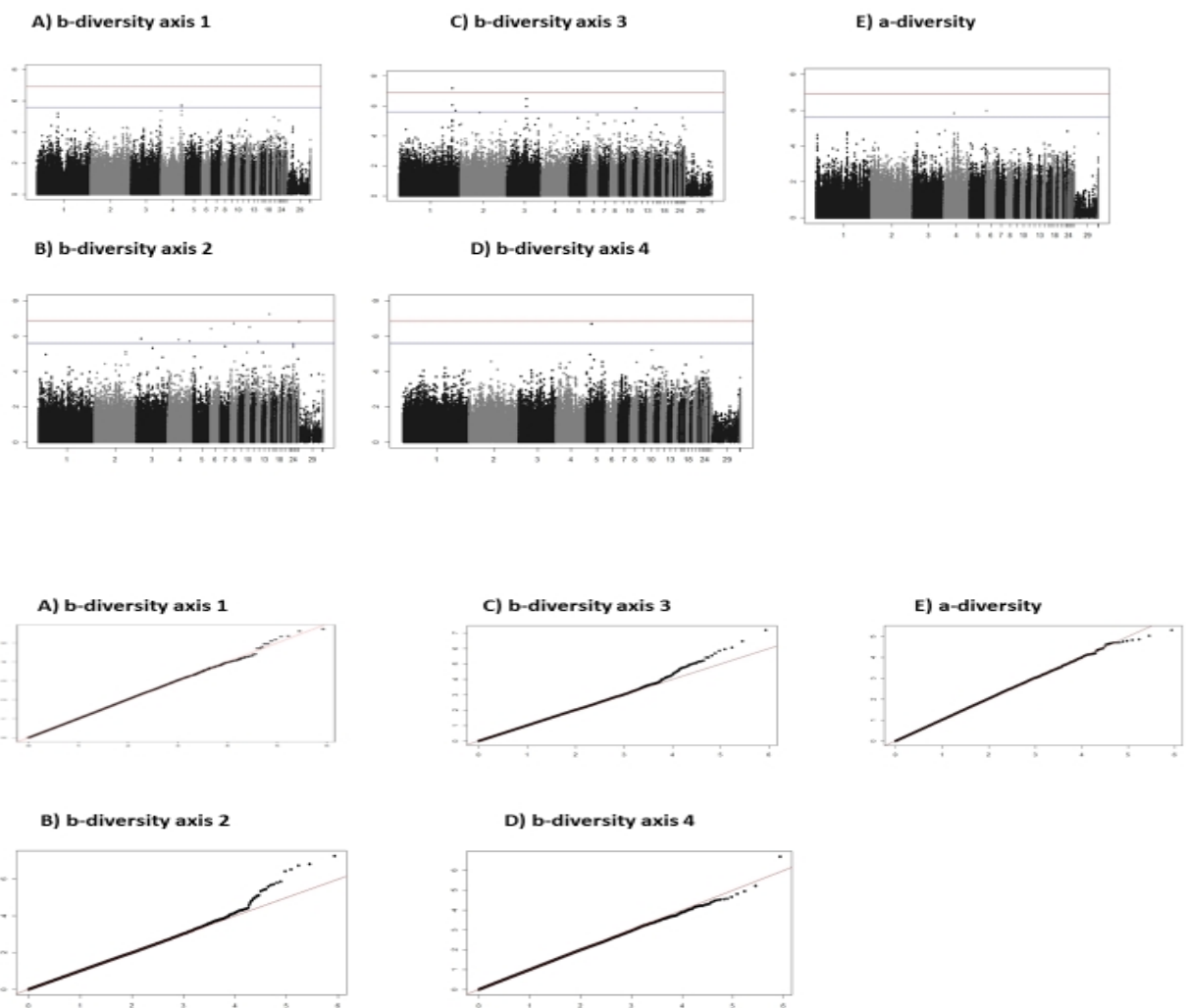
Figure 1. A) Multidimensional scaling analysis showing that Kadaknath (blue dots) and Cobb400 (red dots) are genetically distinct. B) Principal coordinate analysis showing that the gut microbiome of the Kadaknath (blue dots) and Cobb400 (red dots) overlap.



GWAS revealed: a) three SNPs located on the same region on chromosome 4 having suggestive genome-wide significant association with beta-diversity axis 1 (Figure 2A); b) two SNPs on chromosomes 14 and 28 with genome-wide significant association and nine SNPs on chromosomes 8, 10, 6, 3, 4, 12 and 24 having suggestive genome-wide significant association with beta-diversity axis 2 (Figure 2B). Interestingly, one of the SNPs on chromosome 4 was located close to those identified affecting beta-diversity axis 1; c) one SNP on chromosome 1 having genome-wide significant association with beta-diversity axis 3 and two suggestive

significant associations on chromosomes 3 and 11 (Figure 2C); d) one SNP on chromosome 5 with genome-wide significant association with beta-diversity axis 4 (Figure 2D; e) two SNPs on chromosomes 4 and 6 with a suggestive genome-wide significant association with alpha-diversity (Figure 2D). Many annotated genes involved in the immune response and metabolism were located close to the significant SNP markers.

Figure 2. Manhattan (MP) and Quantile-Quantile (Q-Q) plots for gut microbiome beta and alpha diversity indices in chickens; in the MP x-axis is chromosome number; y-axis is $-\log_{10}(P\text{-value})$; red and blue horizontal lines show the genome-wide and suggestive genome-wide significance thresholds, respectively; Q-Q plots (below) of observed P-values against the expected P-values for gut microbiome diversity indices.



Conclusion

To the best of our knowledge, this is the first chicken study investigating the genomic architecture of gut microbiome composition. Microbiome composition tested under the same environmental conditions was shown to be a heritable polygenic trait. These results support the possibility of breeding for protective microbiota in chickens. Future replication studies will focus also on the genotype-by-environment interaction by exploring different production

systems.

List of References

- Stanley D et al. 2014. Microbiota of the chicken gastrointestinal tract: influence on health, productivity and disease. *Applied microbiology and biotechnology*. 98(10):4301-10.
- Haunshi S et al. 2011. Characterization of two Indian native chicken breeds for production, egg and semen quality, and welfare traits. *Poult Sci.*;90(2):314-20.
- Ramasamy KT et al. 2010. Differential expression of Toll-like receptor mRNA in White Leghorn and indigenous chicken of India. *Vet Res Commun*.34(7):633-9.
- Mitchell A., et al. 2016. EBI metagenomics in 2016 --an expanding and evolving resource for the analysis and archiving of metagenomic data. *Nucleic acids research*,2016, 44(D1):595-603.
- Blekhman R. et al. (2015) Host genetic variation impacts microbiome composition across human body sites. *Genome Biology*201516:191. <https://doi.org/10.1186/s13059-015-0759-1>
- Davenport ER. et al. (2015) Genome-Wide Association Studies of the Human Gut Microbiota. *PLoS ONE* 10(11): e0140301. doi:10.1371/journal.pone.0140301