

DISSERTATION | Valentin Peter Haas

GENOMIC AND MICROBIAL ANALYSES OF QUANTITATIVE TRAITS IN POULTRY



Genomic and Microbial Analyses of Quantitative Traits in Poultry

Dissertation

to obtain the doctoral degree of Agricultural Science (Dr. sc. agr.)

Faculty of Agricultural Science
University of Hohenheim
Institute of Animal Science

submitted by

Valentin Peter Haas

from

Heidelberg, Germany

Stuttgart - Hohenheim, 2023

Date of Oral Examination:

27.07.2023

Dean of the Faculty:

Prof. Dr. Ralf Vögele

Examination Committee:

Chairperson of the Oral Examination:

Prof. Dr. Thilo Streck

Supervisor and Reviewer:

Prof. Dr. Jörn Bennewitz

Second Reviewer:

Prof. Dr. Klaus Wimmers

Additional Examiner:

Jun.-Prof. Dr. Amélia Camarinha-Silva

TABLE OF CONTENTS

SUMMARY	9
SUMMARY (GERMAN)	11
GENERAL INTRODUCTION	13
CHAPTER ONE	18
Mapping genes for phosphorus utilization and correlated traits using a 4k SNP linkage map in Japanese quail (<i>Coturnix japonica</i>)	
CHAPTER TWO.....	42
Composition of the ileum microbiota is a mediator between the host genome and phosphorus utilization and other efficiency traits in Japanese quail (<i>Coturnix japonica</i>)	
CHAPTER THREE	74
Inferring causal structures of gut microbiota diversity and feed efficiency traits in poultry using Bayesian learning and genomic structural equation models	
CHAPTER FOUR.....	98
Invited review: Toward the use of host microbiota interplay in poultry and pig breeding	
GENERAL DISCUSSION	116
Microbial Analysis of Bone Ash Traits	117
Use of Microbial Data in Quantitative Analyses.....	118
Subordinate Microbial Features	120
Hologenomic Modeling	122
Gut Microbiota Development.....	122
Concluding Remarks.....	123
ACKNOWLEDGEMENT	128
CURRICULUM VITAE	129
EIDESSTÄTTLICHE VERSICHERUNG	130

SUMMARY

Feed and nutrient efficiency will become increasingly important in poultry production in the coming years. In addition to feed efficiency, particular attention is paid to phosphorus (P) in nonruminants. Especially growing animals have a high demand of P but through the low usability of plant-based P sources for nonruminants, mineral P is added to their feeds. Due to worldwide limited mineral P sources, the high environmental impact of P in excretions and high supplementation costs, a better utilization of P from feed components is required. Animals' P utilization (PU) is known to be influenced by the host genetics and by gastrointestinal microbiota. The overall aim of this thesis was to investigate the relationships between host genetics, gastrointestinal microbiota composition and quantitative traits with the focus on PU and related traits in F2 cross Japanese quail (*Coturnix japonica*). Japanese quail represent a model species for agriculturally important poultry species.

In **Chapter one**, a genetic linkage map for 4k genome-wide distributed SNPs in the study design was constructed and quantitative trait loci (QTL) linkage mapping for performance as well as bone ash traits using a multi-marker regression approach was conducted. Several genome-wide significant QTL were mapped, and subsequent single marker association analyses were performed to find trait associated marker within the significant QTL regions. The analyses revealed a polygenic nature of the traits with few significant QTL and many undetectable QTL. Some overlapping QTL regions for different traits were found, which agreed with the genetic correlations between the traits. Potential candidate genes within the discovered QTL regions were identified and discussed.

Chapter two provided a new perspective on utilization and efficiency traits by incorporating gastrointestinal microbiota and investigated the links between host genetics, gastrointestinal microbiota and quantitative traits. We demonstrated the host genetic influences on parts of the microbial colonization localized in the ileum by estimating heritabilities and mapping QTL regions. From 59 bacterial genera, 24 showed a significant heritability and six genome-wide significant QTL were found. Structural equation models (SEM) were applied to determine causal relationships between the heritable part of the microbiota and efficiency traits. Furthermore, accuracies of different microbial and genomic trait predictions were compared and a hologenomic selection approach was investigated based on the host genome and the heritable part of the ileum microbiota composition. This chapter confirmed the indirect influence of host genetics via the microbiota composition on the quantitative traits.

Chapter three further extended the approaches to identify causalities from **chapter two**. Bayesian learning algorithms were used to discover causal networks. In this approach,

SUMMARY

microbial diversity was considered as an additional quantitative trait and analyzed jointly with the efficiency traits in order to model and identify their directional relationships. The detected directional relationships were confirmed using SEM and extended to SEM association analyses to separate total SNP effects on a trait into direct or indirect SNP effects mediated by upstream traits. This chapter showed that up to one half of the total SNP effects on a trait are composed of indirect SNP effects via mediating traits. A method for detecting causal relationships between microbial and efficiency traits was established, allowing separation of direct and indirect SNP effects.

Chapter four includes an invited review on the major genetic-statistical studies involving the gut microbiota information of nonruminants. The review discussed the analyses conducted in **chapter one to three** and places the analyses published in these chapters in the context of other statistical approaches. **Chapter four** completed the microbial genetic approaches published to date and discussed the potential use of microbial information in poultry and pig breeding.

The **general discussion** includes further results not presented in any of the chapters and discusses the general findings across the chapters.

SUMMARY (GERMAN)

In der Geflügelproduktion wird die Futter- und Nährstoffeffizienz in den kommenden Jahren immer wichtiger werden. Neben der klassischen Futtereffizienz wird vor allem Phosphor (P) bei Nichtwiederkäuern eine besondere Rolle spielen. Nichtwiederkäuer haben hauptsächlich während der Wachstumsphase einen hohen P-Bedarf, der aufgrund der geringen Verwertbarkeit pflanzlicher P-Quellen, mittels mineralischen P in den Futtermitteln ergänzt wird. Im Hinblick auf die weltweit begrenzten mineralischen P-Vorkommen, der hohen Umweltbelastung durch tierische P-Ausscheidungen und den hohen Kosten der Supplementierung, ist eine bessere Verwertung von P aus Futterkomponenten erforderlich. Es ist bekannt, dass die Phosphorverwertung (PU) durch die Genetik der Tiere und durch die gastrointestinale Mikrobiota beeinflusst wird. Ziel dieser Arbeit war es, die Beziehungen zwischen der Wirtsgenetik, der Zusammensetzung der gastrointestinalen Mikrobiota und quantitativen Merkmalen zu untersuchen. Hierbei lag der Schwerpunkt auf der PU der japanischen Wachtel (*Coturnix japonica*) als Modellspezies für landwirtschaftlich wichtige Geflügelarten. In den Analysen wurden japanische Wachteln einer F2-Kreuzung verwendet.

In **Kapitel eins** wurde eine genetische Kopplungskarte für 4k genomweit verteilte SNPs im Studiendesign erstellt und quantitative Merkmalsgenorte (engl.: quantitative trait loci, QTL) Kopplungsanalysen für Leistungs- und Knochenaschemerkmale unter Verwendung eines Multimarker-Regressionsansatzes durchgeführt. Mehrere genomweit signifikante QTL-Regionen wurden kartiert, und mittels Einzelmarker-Assoziationsanalysen konnten merkmalsassoziierte Marker innerhalb der signifikanten QTL-Regionen gefunden werden. Die Analysen ergaben einen polygenen Charakter der Merkmale mit wenigen signifikanten QTL und vielen nicht nachweisbaren QTL. Einige sich für verschiedene Merkmale überschneidende QTL-Regionen konnten gefunden werden, was mit den genetischen Korrelationen zwischen den verwendeten Merkmalen übereinstimmte. Potenzielle Kandidatengene innerhalb der entdeckten QTL-Regionen wurden zudem identifiziert und diskutiert.

Kapitel zwei bot eine neue Perspektive auf Verwertungs- und Effizienzmerkmale durch die Einbeziehung der gastrointestinalen Mikrobiota als Informationsquelle. In diesem Kapitel wurde die Verbindungen zwischen Wirtsgenetik, Magen-Darm-Mikrobiota und quantitativen Merkmalen näher untersucht. Wir wiesen mittels geschätzten Heritabilitäten und kartierten QTL-Regionen die wirtsgenetischen Einflüsse auf Teile der im Ileum lokalisierten mikrobiellen Besiedlung nach. Von insgesamt 59 untersuchten Bakteriengattungen zeigten 24 eine signifikante Heritabilität und es wurden sechs genomweit signifikante QTL-Regionen gefunden. Unter der Verwendung von Strukturgleichungsmodellen (engl.: structural equation models, SEM) konnten kausale Beziehungen zwischen dem vererbten Teil der Mikrobiota

und Effizienzmerkmalen bestätigt werden. Darüber hinaus wurden die Genauigkeiten verschiedener mikrobieller und genomischer Merkmalsvorhersagen verglichen und ein hologenomischer Selektionsansatz auf Grundlage des Wirtsgenoms und des vererbaren Anteils der Zusammensetzung der Ileum-Mikrobiota untersucht. Dieses Kapitel bestätigte den indirekten Einfluss der Wirtsgenetik über die Zusammensetzung der Magen-Darm-Mikrobiota auf die untersuchten quantitativen Merkmale.

In **Kapitel drei** wurden die kausalen Ansätze aus **Kapitel zwei** weiter ausgebaut. Es wurden Bayessche Lernalgorithmen verwendet, um kausale Netzwerke zwischen verschiedenen Merkmalen zu entdecken. Bei diesem Ansatz wurde die mikrobielle Diversität als zusätzliches quantitatives Merkmal betrachtet und gemeinsam mit den Effizienzmerkmalen analysiert, um deren direktionale Beziehungen zu modellieren und zu ermitteln. Die zuvor ermittelten direktionalen Beziehungen wurden danach mittels SEM bestätigt und auf SEM-Assoziationsanalysen ausgedehnt. Die SEM-Assoziationsanalysen ermöglichten die Unterteilung totaler SNP-Effekte auf ein Merkmal in direkte oder indirekte SNP-Effekte, die durch vorgeschaltete Merkmale vermittelt werden. Dieses Kapitel zeigte, dass sich die totalen SNP-Effekte eines Merkmals bis zur Hälfte aus indirekten SNP-Effekten über vermittelnde Merkmale zusammensetzen können. Eine Methode zum Nachweis kausaler Beziehungen zwischen mikrobiellen und Effizienzmerkmalen wurde entwickelt, was eine Trennung von direkten und indirekten SNP-Effekten ermöglicht.

Kapitel vier enthält eine Übersicht über die wichtigsten genetisch-statistischen Studien die Informationen zur Darm-Mikrobiota von Nichtwiederkäuern mit einbeziehen. Die Übersicht diskutiert die in den **Kapiteln eins** bis **drei** durchgeführten Analysen und stellt diese in den Kontext weiterer statistischer Ansätze. Kapitel vier vervollständigt die bisher veröffentlichten mikrobiologischen genetischen Ansätze und erörtert die mögliche Nutzung mikrobiologischer Informationen in der Geflügel- und Schweinezucht.

Die **allgemeine Diskussion** enthält weitere Ergebnisse, die in keinem der Kapitel vorgestellt wurden und erörtert die allgemeinen Erkenntnisse in den einzelnen Kapiteln.

GENERAL INTRODUCTION

In view of the growing world population and the simultaneously increasing demand for food, environmentally friendly, efficient and sustainable food production is needed (FAO, 2014; Ulian et al., 2020). A more efficient use of forage grown in the field is a necessary aspect for future livestock farming (Michalk et al., 2019). In livestock-dense regions in particular, feeding-related oversupply is accompanied by negative environmental effects (Herrero et al., 2015). In addition to general feed efficiency, the impacts of nitrogen and phosphorus discharges associated with nonruminant livestock production are also environmentally relevant (Steinfeld et al., 2006).

Phosphorus (P) is an essential mineral for all living organisms and is primarily stored in the form of phytin acids and salt phytate in plant-based feed components (Eeckhout and Paepe, 1994; Rodehutschord, 2016). Growing nonruminants have a high P demand (Williams et al., 2000) but have a comparably low endogenous phytase activity in the gastrointestinal tract (GIT) (Maenz and Classen, 1998). For this reason, mineral P (phosphate rock) or phytase enzymes are added to nonruminant feeds. Considering the globally limited mineral P sources located in only five countries (Cordell et al., 2009; Neset and Cordell, 2012), the environmental impact of additional supplementation (Campbell et al., 2017), and the high costs of mineral P or phytase enzyme supplementation, the P utilization (PU) of animals is the focus of this thesis.

The GIT of poultry harbors a large community of microbial settlers (Apajalahti et al., 2004). With the estimation of microbiabilities, the proportion of the phenotypic variance of a trait explained by the microbiota composition can be assessed (Difford et al., 2018). In addition to host genetics, parts of the gastrointestinal microbiota influence efficiency and performance traits (e.g., Wen et al., 2019, Zhou et al., 2022). For some performance traits studied, the proportion of phenotypic variance explained by the microbiota is equal to or even higher than the additive genetic proportion (e.g., Vollmar et al., 2020, Khanal et al., 2021). The genetic and microbial influences on PU have been demonstrated, with a significant heritability of 0.14 and a significant microbiability of 0.15 for quail (Beck et al., 2016; Vollmar et al., 2020). Previous studies also found significant heritabilities for P bioavailability in broilers and chickens (Ankra-Badu et al., 2010; Zhang et al., 2003). In these studies, PU or P bioavailability were genetically and phenotypically correlated with performance traits.

Besides the investigated PU, feed efficiency of the quail was evaluated. General feed efficiency can be divided into digestive and metabolic efficiency (Martin et al., 2021; Puillet et al., 2016). Digestive efficiency reflects the ability to absorb ingested feed nutrients from the GIT, whereas metabolic efficiency reflects the redistribution of absorbed nutrients into animal products. Microbial settlers in the GIT may be important drivers of digestive efficiency depending on the considered trait. The microbial influence on PU can be explained by the fact that the intestinal

phytase activity is low but microorganisms in the GIT can produce phytase (Maenz and Classen, 1998; Rodehutscord, 2017). Therefore, selective breeding for improved PU and performance traits based on the heritable part of the ileum microbiota, as proposed in Weishaar et al. (2020), may be beneficial.

In the present thesis, quantitative traits of the model species Japanese quail (*Coturnix japonica*) were studied. The Japanese quail was used because of its short generation interval, small body size, low space requirements, and good comparability to other poultry (Kayang et al., 2004; Mills et al., 1997; Rodehutscord and Dieckmann, 2005). The overall aim of this thesis was to better understand the relationships between host genetics, gastrointestinal microbiota composition and quantitative traits in the dataset of an F2 design of Japanese quail.

In **chapter one**, a 4k SNP linkage map of Japanese quail genotypes is constructed. Based on this linkage map, quantitative trait loci (QTL) linkage mapping is carried out for performance and bone ash traits to learn more about the genetic architecture of these traits.

Chapter two studies the potential influences of host genetics on microbial features of the GIT by calculating heritabilities and performing QTL linkage analyses for microbial genera. By using structural equation models (SEM), unidirectional relationships between heritable microbial genera and quantitative traits are investigated. A combination of microbial best linear unbiased prediction (MBLUP) and genomic best linear unbiased prediction (GBLUP) is used to investigate the utility of hologenomic selection for the traits under consideration.

Chapter three focuses on the directional, i.e., the causal, relationship between different quantitative traits and microbial alpha diversity of the ileum. Directional relationships in a pool of traits are investigated with the help of Bayesian networks, and by using SEM, the recursive relationships are analyzed in detail. Finally, SEM association analyses are used to split the total SNP effects of a trait into direct and indirect SNP effects, with the latter mediated by upstream traits in the network.

Chapter four is an invited review article that explains the current microbial genetic-statistical approaches in breeding research, addresses open questions and puts the published analyses of this thesis in the context of other studies. This chapter describes the current and prospective usability of microbiota information in pig and poultry breeding.

The thesis ends with a **general discussion**. This section shows additional results of the analyses of bone ash traits and genomic analyses of microbial features at the phylum level and debates the usability of microbial features in Bayesian network analyses.

References

- Ankra-Badu, G. A., G. M. Pesti, and S. E. Aggrey. 2010. Genetic interrelationships among phosphorus, nitrogen, calcium, and energy bioavailability in a growing chicken population. *Poult. Sci.* 89(11):2351–2355. doi:10.3382/ps.2010-00870.
- Apajalahti, J., A. Kettunen, and H. Graham. 2004. Characteristics of the gastrointestinal microbial communities, with special reference to the chicken. *Worlds Poult. Sci. J.* 60(2):223–232. doi:10.1079/wps20040017.
- Beck, P., H.-P. Piepho, M. Rodehutschord, and J. Bennewitz. 2016. Inferring relationships between Phosphorus utilization, feed per gain, and bodyweight gain in an F2 cross of Japanese quail using recursive models. *Poult. Sci.* 95(4):764–773. doi:10.3382/ps/pev376.
- Campbell, B. M., D. J. Beare, E. M. Bennett, J. M. Hall-Spencer, J. S. I. Ingram, F. Jaramillo, R. Ortiz, N. Ramankutty, J. A. Sayer, and D. Shindell. 2017. Agriculture production as a major driver of the Earth system exceeding planetary boundaries. *Ecol. Soc.* 22(4). doi:10.5751/ES-09595-220408.
- Cordell, D., J.-O. Drangert, and S. White. 2009. The story of phosphorus: Global food security and food for thought. *Glob. Environ. Chang.* 19(2):292–305. doi:10.1016/j.gloenvcha.2008.10.009.
- Difford, G., D. R. Plichta, P. Løvendahl, J. Lassen, S. J. Noel, O. Højberg, A.-D. G. Wright, Z. Zhu, L. Kristensen, H. B. Nielsen, B. Guldbandsen, and G. Sahana. 2018. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genet.* 14(10):e1007580. doi:10.1371/journal.pgen.1007580.
- Eeckhout, W., and M. de Paepe. 1994. Total phosphorus, phytate-phosphorus and phytase activity in plant feedstuffs. *Anim. Feed Sci. Technol.* 47(1-2):19–29. doi:10.1016/0377-8401(94)90156-2.
- FAO. 2014. Building a common vision for sustainable food and agriculture: Principles and approaches. Food and Agriculture Organization of the United Nations, Rome.
- Herrero, M., S. Wiersenius, B. Henderson, C. Rigolot, P. Thornton, P. Havlík, I. de Boer, and P. J. Gerber. 2015. Livestock and the environment: What have we learned in the past decade? *Annu. Rev. Environ. Resour.* 40(1):177–202. doi:10.1146/annurev-environ-031113-093503.
- Kayang, B. B., A. Vignal, M. Inoue-Murayama, M. Miwa, J. L. Monvoisin, S. Ito, and F. Minvielle. 2004. A first-generation microsatellite linkage map of the Japanese quail. *Anim. Genet.* 35(3):195–200. doi:10.1111/j.1365-2052.2004.01135.x.

- Khanal, P., C. Maltecca, C. Schwab, J. Fix, and F. Tiezzi. 2021. Microbiability of meat quality and carcass composition traits in swine. *J. Anim. Breed. Genet.* 138(2):223–236. doi:10.1111/jbg.12504.
- Maenz, D. D., and H. L. Classen. 1998. Phytase activity in the small intestinal brush border membrane of the chicken. *Poult. Sci.* 77(4):557–563. doi:10.1093/ps/77.4.557.
- Martin, P., V. Ducrocq, P. Faverdin, and N. C. Friggens. 2021. Invited review: Disentangling residual feed intake - Insights and approaches to make it more fit for purpose in the modern context. *J. Dairy Sci.* 104(6):6329–6342. doi:10.3168/jds.2020-19844.
- Michalk, D. L., D. R. Kemp, W. B. Badgery, J. Wu, Y. Zhang, and P. J. Thomassin. 2019. Sustainability and future food security - A global perspective for livestock production. *Land Degrad. Dev.* 30(5):561–573. doi:10.1002/ldr.3217.
- Mills, A. D., L. L. Crawford, M. Domjan, and J. M. Faure. 1997. The behavior of the Japanese or domestic quail *Coturnix japonica*. *Neurosci. Biobehav. Rev.* 21(3):261–281. doi:10.1016/S0149-7634(96)00028-0.
- Neset, T.-S. S., and D. Cordell. 2012. Global phosphorus scarcity: Identifying synergies for a sustainable future. *J. Sci. Food Agric.* 92(1):2–6. doi:10.1002/jsfa.4650.
- Puillet, L., D. Réale, and N. C. Friggens. 2016. Disentangling the relative roles of resource acquisition and allocation on animal feed efficiency: Insights from a dairy cow model. *Genet. Sel. Evol.* 48(1):72. doi:10.1186/s12711-016-0251-8.
- Rodehutschord, M. 2016. Chapter 10 Interactions between minerals and phytate degradation in poultry – challenges for phosphorus digestibility assays. In: M. Rodehutschord, C. L. Walk, I. Kühn, H. H. Stein, and M. T. Kidd, editors, *Phytate destruction: Consequences for precision animal nutrition*. Wageningen Academic Publishers, Wageningen. p. 167–178.
- Rodehutschord, M. 2017. Advances in understanding the role of phytate in phosphorus and calcium nutrition of poultry. In: T. J. Applegate, editor, *Achieving sustainable production of poultry meat, volume 2: breeding and nutrition*. p. 165–80.
- Rodehutschord, M., and A. Dieckmann. 2005. Comparative studies with three-week-old chickens, turkeys, ducks, and quails on the response in phosphorus utilization to a supplementation of monobasic calcium phosphate. *Poult. Sci.* 84(8):1252–1260. doi:10.1093/ps/84.8.1252.
- Steinfeld, H., P. Gerber, T. D. Wassenaar, Nations, Food and Agriculture Organization of the United, V. Castel, M. Rosales, M. R. M., and C. de Haan. 2006. *Livestock's long*

- shadow: Environmental issues and options. Food and Agriculture Organization of the United Nations, Rom.
- Ulian, T., M. Diazgranados, S. Pironon, S. Padulosi, U. Liu, L. Davies, M.-J. R. Howes, J. S. Borrell, I. Ondo, O. A. Pérez-Escobar, S. Sharrock, P. Ryan, D. Hunter, M. A. Lee, C. Barstow, Ł. Łuczaj, A. Pieroni, R. Cámara-Leret, A. Noorani, C. Mba, R. Nono Womdim, H. Muminjanov, A. Antonelli, H. W. Pritchard, and E. Mattana. 2020. Unlocking plant resources to support food security and promote sustainable agriculture. *Plants People Planet* 2(5):421–445. doi:10.1002/ppp3.10145.
- Vollmar, S., R. Wellmann, D. Borda-Molina, M. Rodehutschord, A. Camarinha-Silva, and J. Bennewitz. 2020. The gut microbial architecture of efficiency traits in the domestic poultry model species Japanese quail (*Coturnix japonica*) assessed by mixed linear models. *G3 (Bethesda)* 10(7):2553–2562. doi:10.1534/g3.120.401424.
- Weishaar, R., R. Wellmann, A. Camarinha-Silva, M. Rodehutschord, and J. Bennewitz. 2020. Selecting the hologenome to breed for an improved feed efficiency in pigs - A novel selection index. *J. Anim. Breed. Genet.* 137(1):14–22. doi:10.1111/jbg.12447.
- Wen, C., W. Yan, C. Sun, C. Ji, Q. Zhou, D. Zhang, J. Zheng, and N. Yang. 2019. The gut microbiota is largely independent of host genetics in regulating fat deposition in chickens. *ISME J.* 13(6):1422–1436. doi:10.1038/s41396-019-0367-2.
- Williams, B., S. Solomon, D. Waddington, B. Thorp, and C. Farquharson. 2000. Skeletal development in the meat-type chicken. *Br. Poult. Sci.* 41(2):141–149. doi:10.1080/713654918.
- Zhang, W., S. E. Aggrey, G. M. Pesti, H. M. Edwards, and R. I. Bakalli. 2003. Genetics of phytate phosphorus bioavailability: Heritability and genetic correlations with growth and feed utilization traits in a randombred chicken population. *Poult. Sci.* 82(7):1075–1079. doi:10.1093/ps/82.7.1075.
- Zhou, Q., F. Lan, S. Gu, G. Li, G. Wu, Y. Yan, X. Li, J. Jin, C. Wen, C. Sun, and N. Yang. 2022. Genetic and microbiome analysis of feed efficiency in laying hens. *Poult. Sci.*:102393. doi:10.1016/j.psj.2022.102393.

CHAPTER ONE

Mapping genes for phosphorus utilization and correlated traits using a 4k SNP linkage map in Japanese quail (*Coturnix japonica*)

S. Vollmar*, V. Haas*, M. Schmid*, S. Preuß*, R. Joshi†, M. Rodehutschord* and J. Bennewitz*

*Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

†Department of Animal and Aquacultural Sciences, Faculty of Biosciences, Norwegian University of Life Sciences, As N-1432, Norway

Address for correspondence: S. Vollmar, Institute of Animal Science, University of Hohenheim, Stuttgart D-70599, Germany. E-mail: solveig.vollmar@uni-hohenheim.de

Published in *Animal Genetics* (2021), Volume 52 (1), pp. 90-98

<https://doi.org/10.1111/age.13018>

Summary

A large F2 cross with 920 Japanese quail was used to map QTL for phosphorus utilization, calcium utilization, feed per gain and body weight gain. In addition, four bone ash traits were included, because it is known that they are genetically correlated with the focal trait of phosphorus utilization. Trait recording was done at the juvenile stage of the birds. The individuals were genotyped genome-wide for about 4k SNPs and a linkage map constructed, which agreed well with the reference genome. QTL linkage mapping was performed using multimarker regression analysis in a line cross model. Single marker association mapping was done within the mapped QTL regions. The results revealed several genome-wide significant QTL. For the focal trait phosphorus utilization, a QTL on chromosome CJA3 could be detected by linkage mapping, which was substantiated by the results of the SNP association mapping. Four candidate genes were identified for this QTL, which should be investigated in future functional studies. Some overlap of QTL regions for different traits was detected, which is in agreement with the corresponding genetic correlations. It seems that all traits investigated are polygenic in nature with some significant QTL and probably many other small-effect QTL that were not detectable in this study.

Keywords: feed utilization, Japanese quail, linkage map, quantitative trait loci

Introduction

Phosphorus is an essential mineral for all living organisms. It is important for energy metabolism, nucleic acid synthesis, enzyme activity and bone mineralization. Most of the phosphorus in plant seeds and feedstuffs produced thereof is present as phytic acid and its salts, called phytates (Eeckhout & Paepe 1994). Owing to low endogenous phytase activity in the digestive tract of poultry, phytate-P sources can only partially be utilized. Therefore, poultry diets are usually supplemented with mineral phosphorus, often in combination with exogenous phytase, which results in additional costs. Additionally, global mineral phosphorus resources are limited, and the phosphorus in excreta has an environmental impact. Therefore, it is desirable to minimize mineral phosphorus supplementation without compromising animal health and performance. Thus, high phosphorus utilization (PU) by animals is desirable.

Japanese quail (*Coturnix japonica*) has long been an important model species in poultry studies because of its short generation intervals, small body size, which results in a smaller space requirement (Kayang *et al.* 2004; Cheng *et al.* 2010), and similarity to other poultry species (Stock & Bunch 1982; Shibusawa *et al.* 2001). A recent study implemented an F2 experimental design with approximately 1000 Japanese quail and phenotyped the F2 individuals for PU and related traits (Beck *et al.* 2016a). The coefficient of variation for PU was 0.11, which indicated substantial variation, with a heritability of 0.14 (± 0.06). By applying structural equation models some complex relationships of PU were detected with body weight gain and feed per gain ratio (Beck *et al.* 2016a). A subsequent study of the ileum microbiota composition of those birds estimated a significant microbiability for PU (Borda-Molina *et al.* 2020; Vollmar *et al.* 2020). In addition, ileal transcriptome profiles, miRNA–mRNA and gut microbiome interactions of subsets of quails with divergent PU have been studied (Oster *et al.* 2020; Ponsuksili *et al.* 2020).

Because calculation of PU involves quantitative measurement of feed intake and excretion over several days, PU is a very-hard-to-measure trait in a routine breeding enterprise. Therefore, proxy traits and genetically correlated traits are desirable and convenient to measure. Bone ash traits are features that have been used to determine the bioavailability of phosphorus in quail (Vali & Jalali 2011) and chicken (Li *et al.* 2017). Several bone ash traits were analyzed using samples from the experiment of Beck *et al.* (2016a) and the genetic correlations with PU were estimated, which were between 0.5 and 0.6 (Künzel *et al.* 2019). Thus, it might be possible to consider bone ash traits as proxy traits to breed for the improvement of PU.

Until now, it has been largely unknown whether the genetic variance of PU is caused by many QTL with small effects or if there are some large QTL that might be of special interest for breeding purposes. Therefore, the aim of this study was to map the QTL associated with the

focal trait PU as well as other performance traits and bone ash traits in Japanese quail using an F2 cross. The individuals were genotyped genome-wide with 4k SNPs, and we used these data to establish a linkage map and subsequently to conduct QTL linkage and association mapping.

Material and Methods

Experimental design

The experiment was conducted in accordance with the German Animal Welfare Legislation approved by the Animal Welfare Commissioner of the University Hohenheim (approval number S371/13TE). An F2 cross of Japanese quail (*C. japonica*) was established. The details of the F2 design can be found in Beck *et al.* (2016a), and only the essential steps are described in the following. The founder lines were divergently selected for social reinstatement behavior in an earlier experiment conducted at the INRA, France (Mills & Faure 1991). The selection of these founder individuals is thus not related to the focal trait PU. Twelve males from founder line A (B) were mated to 12 females from founder line B (A) to produce the F1 generation. From this generation, 17 males and 34 females were selected, and one male was mated with two females, resulting in 920 F2 individuals. These individuals belonged to 34 full-sib families and 17 paternal half-sib families, with approximately the same family size. A low-P-content diet was provided to allow the quails to exhibit their full PU potential. The diet did not contain mineral P supplement or phytase.

Trait records

Body weight gain (BWG) was calculated as the difference in body weight at days 10 and 15. Feed per gain ratio (F:G) was calculated as feed intake (FI) within this 5-day period divided by BWG. PU and calcium utilization (CaU) were calculated for this period based on quantitative intake and excretion of the elements as described in Beck *et al.* (2016a). The quails were slaughtered on day 15, and the right tibia and the right foot were preserved. The total amount of ash in the tibia and foot (TA and FA) as well as ash concentrations in the dry matter of the bones (TA% and FA%) were recorded as described in detail in Künzel *et al.* (2019). Descriptive statistical parameters, heritabilities and trait abbreviations are provided in Table 1. The heritabilities of the traits were estimated by Beck *et al.* (2016a) and Künzel *et al.* (2019) using mixed linear animal models.

DNA collection and SNP genotyping

One milliliter of blood was collected from each animal using EDTA-K tubes and stored at -20°C until DNA extraction was performed using the Maxwell 16 Blood DNA Purification Kit (Promega). The DNA concentration was adjusted to 50 ng/ μl to ensure consistent measurements. Using a customer's Illumina iSelect chip, we genotyped 5388 SNPs. The SNP markers were mapped through the chicken genome using the method described in Recoquillay

et al. (2015), as no quail genome was available at the time of genotyping. The following criteria were applied to filter the genotypes: one or more conflicting genotypes between parent and offspring, a MAF ≤ 0.03 , an SNP call frequency ≤ 0.9 and cluster separation ≤ 0.4 . This led to the exclusion of 842 SNPs. Furthermore, we rejected SNPs on the sex chromosomes Z or W and in the linkage group (LG) LGE22C19W28_E50C23 or E64 (information obtained from the *C. japonica* reference genome assembly (NCBI GCA_001577835.1)). This filtering resulted in a total of 3986 SNP markers for further analysis.

Table 1 Descriptive statistics and heritabilities of the traits.

Trait ^{1,2}	Abbreviation	Unit	Minimum	Maximum	Mean	h^2 (SE)
Phosphorus utilization ⁺	PU	%	21.49	87.43	71.41	0.14 (0.06)
Calcium utilization [*]	CaU	%	19.42	84.31	60.56	0.17 (≤ 0.10)
Feed per gain ⁺	F:G	g/g	1.21	3.92	1.78	0.12 (0.06)
Feed intake [*]	FI	g	16.11	62.35	42.65	0.11 (≤ 0.10)
Body weight gain ⁺	BWG	g	5.80	37.85	24.50	0.09 (0.14)
Tibia ash (mg) [*]	TA	mg	19.20	83.50	45.82	0.23 (≤ 0.10)
Tibia ash (%) [*]	TA%	%	35.53	55.71	45.26	0.23 (≤ 0.10)
Foot ash (mg) [*]	FA	mg	19.60	83.60	44.76	0.34 (≤ 0.10)
Foot ash (%) [*]	FA%	%	12.10	21.91	17.30	0.31 (≤ 0.10)

¹ From days 10–15 of life.

² Measurements and heritabilities from Beck *et al.* (2016a)[†] and Künzel *et al.* (2019)^{*} and SEs are in parentheses.

Linkage map construction

The linkage mapping software LEP-MAP2 (Rastas *et al.* 2015) was used to build a sex-averaged Japanese quail map. The software uses pedigree and marker information to assign SNP markers to LGs and computes the likelihood of the marker order within each LG using standard hidden Markov models (Rastas *et al.* 2013; Rastas *et al.* 2015). In the first step, the module *SeparateChromosomes* was used to assign markers to the LG. We used the option LOD = 1–20 to test lodLimits with a sizeLimit = 5 so that LGs with fewer than five markers were removed. A lodLimit of 5 resulted in 27 LGs with 3975 markers assigned to them. The remaining markers were assigned to LGs by using the module *JoinSingles* with lodLimit = 1–15 and lodDifference = 2. A lodLimit = 1 was selected because there was no difference compared with other lodLimits in terms of results, and an additional nine SNPs could be assigned. The module *OrderMarkers* orders the markers within each LG. This step was replicated five times to select the best order with the highest likelihood. The module was run with the options polishWindow = 30, filterWindow = 10 (both parameters are used for speeding up the computations), numThreads = 10 (maximum number of threads to use), useKosambi = 1 (using Kosambi mapping function), minError = 0.15 (because genotyping errors can lead to large map distances) and sexAveraged = 1 (to compute the sex-averaged map distances).

To compare the calculated genetic map with the reference genome *C. japonica* (NCBI GCA_001577835.1), the flanking sequences for each SNP were aligned by performing BLAST searches of the reference genome. This led to the assignment of our LGs to the chromosomes. These assignments were used throughout the rest of the study.

QTL linkage and association analysis

A line cross model was applied in this study. For this purpose, we used the package RQTL2 (Broman *et al.* 2019). This program was developed for inbred line crosses. We estimated the F_{ST} value for each SNP in the two founder populations using eq (8) in Weir & Cockerham (1984). Subsequently, we selected only those SNPs with an $F_{ST} > 0.23$, which comprised approximately half of the SNPs, and the selected SNPs were used for QTL linkage mapping. This filtering ensured that the assumptions regarding the inbred founder lines made by the software were approximately fulfilled. In addition, we selected only those chromosomes with > 40 SNPs because we applied multimarker linkage mapping. These two filter steps resulted in 1968 SNPs that were used for QTL linkage mapping on 19 chromosomes. Subsequently, we applied the RQTL2 software package and estimated QTL genotype probabilities for each F2 individual and each marker position. These probabilities were used in a regression analysis to map the QTL. We included the hatches as fixed effects in the regression model. The LOD score was used as a test statistic, and correction for multiple testing was done using the permutation test (10 000 permutations). We considered two significance criteria for each trait, i.e. 1 and 5% genome-wide significance (LOD scores 4.9–5.9 and 4.2–4.7 respectively). The QTL support intervals (SI) were approximated using the LOD drop off method with a drop of 1.5 LOD (Manichaikul *et al.* 2006). The upper and lower bounds of the SI were extended by 5 cM to be conservative. Because the assumptions of the linkage QTL mapping approach regarding the inbred founder lines were only approximated fulfilled (i.e. not every SNP with $F_{ST} > 0.23$ was divergently fixated in the two founder lines), we conducted an SNP association analyses. For this purpose, we tested all markers within the SI (i.e. also those with an $F_{ST} < 0.23$) for trait associations to support the presence of a QTL. We repeated this process for each SNP within the intervals separately by applying a mixed linear model using the software GCTA (Yang *et al.* 2011). The hatches were considered as fixed effects, and correction for putative population stratification effects was performed by including a random animal effect based on a genomic relationship matrix that was calculated using all markers except those on the chromosome under consideration (i.e. the leave-one-chromosome-out option in gcta). As only those markers in the SI were tested for associations, no correction for multiple testing was performed.

To identify positional candidate genes in the 0.5 Mbp regions up- and downstream of significant SNPs, we used Genome Data viewer from NCBI and the reference genome assembly (NCBI GCA_001577835.1).

Results

Construction of the linkage map

The summary of the linkage map is shown in Table 2, and a list of SNPs with their chromosomal position was made public available (see Data availability statement). The linkage map is plotted in Fig. S1. A total number of 3975 SNPs were assigned to 27 LGs. The map covers 1735 cM with individual LG lengths that range from approximately 3 cM [*C. japonica* chromosome (CJA) 28] to 253 cM (CJA1) (Table 2). The number of markers per chromosome varied from 5 (CJA25) to 769 SNPs (CJA1), and the average density was 0.81 markers per cM across all chromosomes. We estimated a high correlation between the genetic (cM) position of the calculated linkage map and the physical (bp) position of the reference genome assembly (NCBI GCA_001577835.1), ranging from 0.88 to 0.99 (Table 2). Overall, the order of the markers of the genetic map agreed well with the order of the physical positions of the reference genome. No LGs could be assigned to chromosome 16, because this chromosome is poorly characterized so far and no SNP could be assigned to it. Figure 1 shows the comparison between the physical (bp) and genetic map (cM) for chromosome 2, and some outliers are visible. These outlier markers were either identified at other positions within an LG compared with the reference genome or had positions that were not yet known. The comparisons of the remaining chromosomes are shown in Fig. S1.

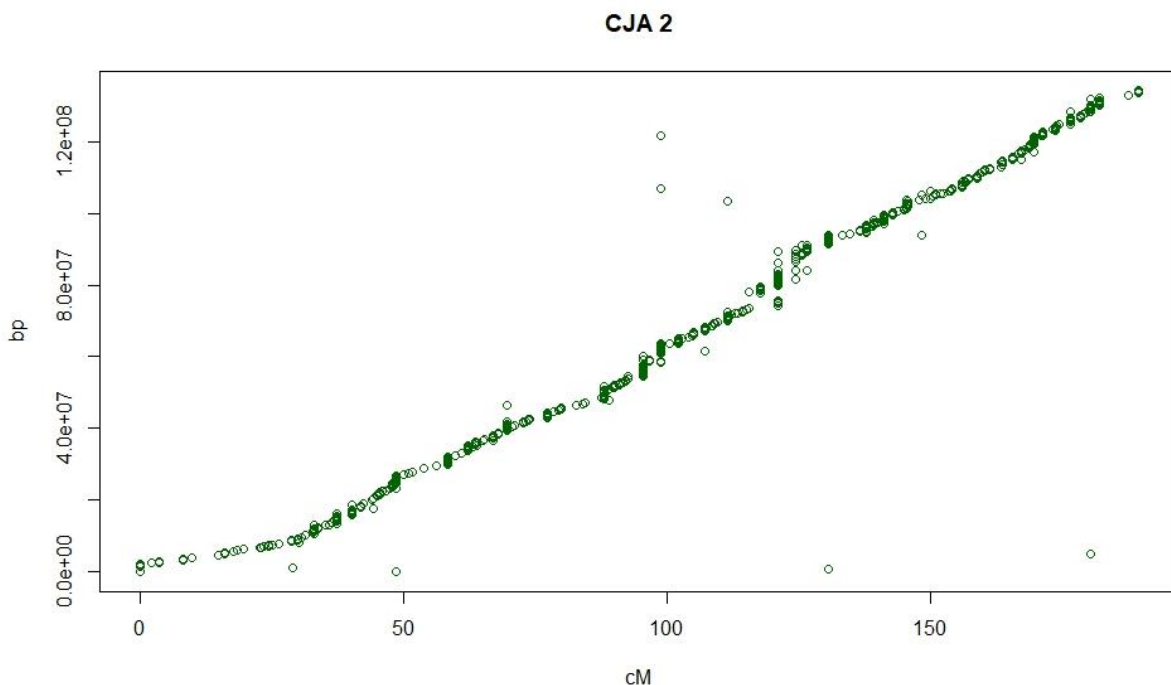


Figure 1 Plot of SNPs that were assigned to chromosome 2. The y-axis shows the physical position (bp), which is based on the reference genome assembly (NCBI GCA_001577835.1), and the x-axis shows the genetic position (cM). Note that the SNP positions at 0 bp refer to an as yet unknown position.

Table 2 Numbers of markers (n SNPs) on each chromosome (*Coturnix japonica*, CJA), length in cM and in Mb, average number of markers per cM and per Mb, and correlation between the cM and the bp positions.

CJA	n SNPs	Length		Markers		Correlation
		cM	Mb ¹	per cM	per Mb	
1	769	253.09	177	0.33	0.23	0.96
2	650	189.57	136	0.29	0.20	0.98
3	457	153.11	101	0.34	0.22	0.94
4	436	116.80	83	0.27	0.19	0.98
5	278	97.64	54	0.35	0.19	0.89
6	145	64.45	32	0.44	0.19	0.97
7	152	68.45	34	0.45	0.20	0.93
8	138	54.70	27	0.40	0.15	0.92
9	121	53.67	21	0.44	0.15	0.99
10	88	44.22	19	0.50	0.17	0.96
11	91	43.40	18	0.48	0.18	0.95
12	90	47.72	17	0.53	0.15	0.99
13	63	38.74	16	0.61	0.19	0.96
14	75	47.11	13	0.63	0.16	0.92
15	55	46.04	12	0.84	0.19	0.98
16	-	-	0.3	-	-	-
17	52	45.19	9	0.87	0.14	0.98
18	40	38.77	10	0.97	0.13	0.95
19	56	45.78	9	0.82	0.12	0.97
20	62	51.92	13	0.84	0.20	0.96
21	26	31.08	6	1.20	0.15	0.96
22	18	39.11	4	2.17	0.13	0.96
23	27	39.93	5	1.48	0.13	0.96
24	36	46.08	6	1.32	0.15	0.98
25	5	3.33	3	0.67	0.10	0.88
26	19	37.30	5	2.33	0.14	0.95
27	17	35.41	5	2.08	0.19	0.89
28	9	2.77	4	0.31	0.13	0.95
total	3975	1735.36	839.30	-	-	-
average	147	64.27	29.98	0.81	0.17	0.95

¹ Size in Mb based on the reference genome *Coturnix japonica* 2.0 (NCBI GCA_001577835.1).

Identification of QTL

The test statistic plots of the analyzed chromosomes are shown in Figs. 2 & 3 for the performance and the bone ash traits respectively. A total of 21 QTL (eight QTL for 1%, 13 QTL for 5%) were mapped for all traits at a 1% (5%) genome-wide significance level. For all traits, QTL could be mapped, except for F:G. A detailed description of the QTL is given in Table 3. For PU, we identified one QTL on CJA3, and for BWG, we found one QTL on CJA3, whereas all other traits were associated with two or more QTL (Table 3, Figs. 2 & 3). Some SI overlapped for several traits. For example, the SI on CJA3 for PU and FA% and the SI on CJA4 for CaU, TA and FA overlapped (Table 3).

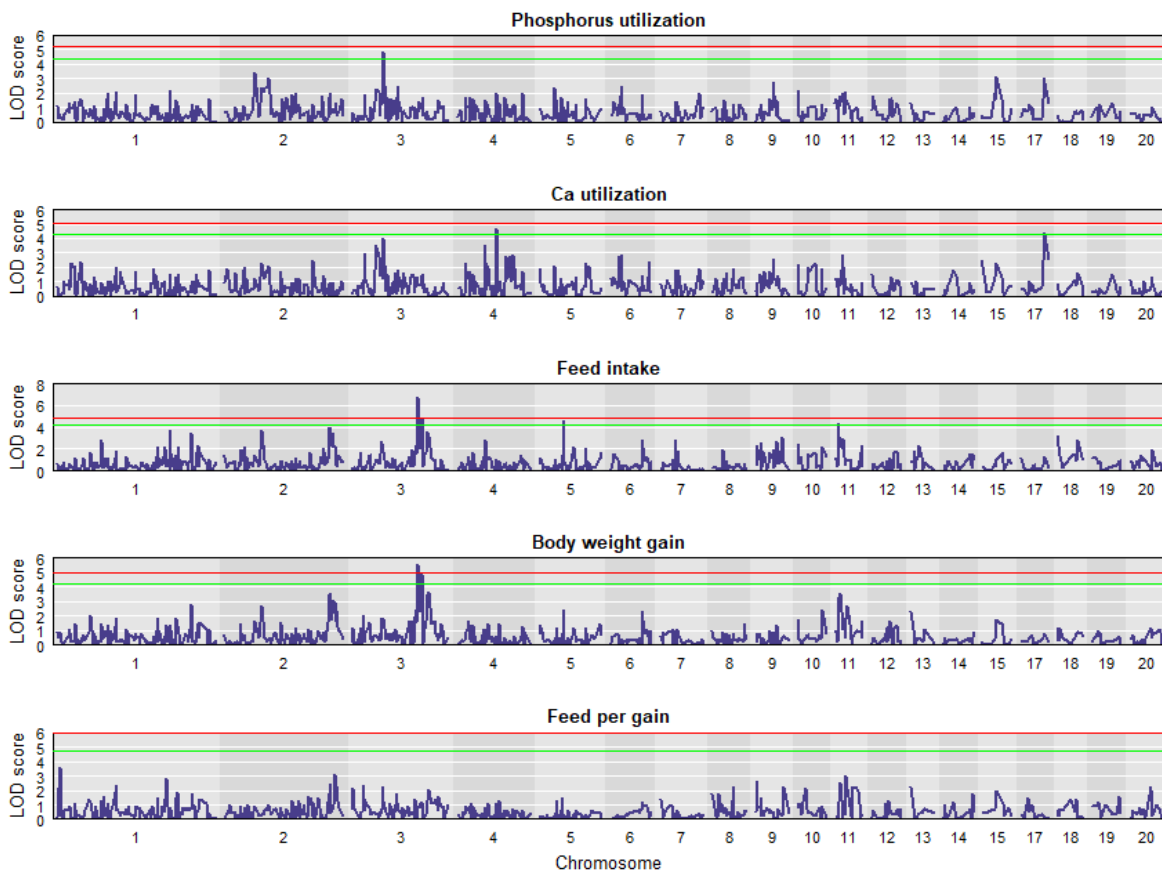


Figure 2 Plot of the QTL linkage mapping scan of growth and efficiency traits with LOD score test statistics. The green and red lines correspond to genome-wide significance levels of 5 and 1% respectively.

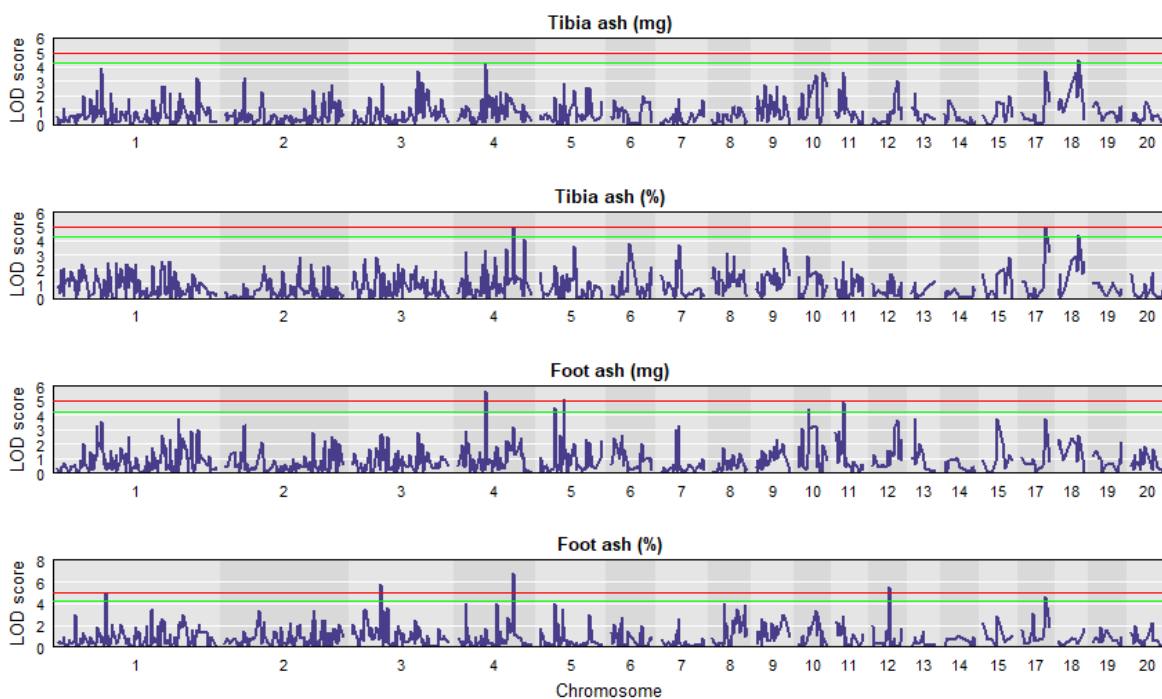


Figure 3 Plot of the QTL linkage mapping scan of bone ash traits with LOD score test statistics. The green and red lines correspond to genome-wide significance levels of 5 and 1% respectively.

The results of the SNP association analyses are shown as the numbers of significant SNPs in the QTL regions, and the significant SNPs are listed in Table S1. A total of 127 SNPs were shown to be significant in QTL regions for all traits. Significant SNPs were found in all QTL regions for the traits PU, CaU, FI, BWG and TA (Table 3, Table S1). Although the SI on CJA3 overlapped for PU and FA%, no significant identical SNPs could be found in this region (Table S1). PU was associated with five significant SNPs, and FA% was associated with three SNPs on CJA3. Several SNPs were significantly associated with several traits. The QTL on CJA3 for FI shared five significant SNPs with the QTL for BWG (id12506 at 91 cM, id10670 at 95 cM, id10683 at 97 cM, id06748 at 101 cM and id14876 at 102 cM). Nine SNPs on CJA11 (id06872 and id32446 at 7 cM, id15452 at 11 cM, id32451, id07827, id09706, id05659 and id08551 at 13 cM, and id05029 at 16 cM) were significant within the QTL region for FI and FA (Table S1). One SNP (id08651 at 39 cM) on CJA18 was significant within the QTL region for TA and TA%. No other common significant SNP similarities could be found despite the presence of overlapping SI.

Table 3 Trait specific positions of significant QTL (Pos) on the chromosomes (CJA), LOD score test statistics (LOD) at the 1% (**) and 5% (*) genome-wide significance level, and the corresponding support intervals (SI). SI_low and SI_high = beginning and end of the support interval respectively, with the number (*n*) of significant SNPs identified by the association analysis.

Trait ¹	CJA	Pos	LOD	SI_low	SI_high	<i>n</i> of SNPs
PU	3	48.9	4.82*	35.89	65.71	5
CaU	4	62.6	4.64*	31.71	75.64	14
	17	36.8	4.35*	20.21	57.69	6
FI	3	104.6	6.73**	90.27	119.49	7
	5	38.6	4.63*	25.99	51.10	4
BWG	11	5.5	4.34*	0.00	27.44	10
	3	104.6	5.59**	90.27	124.58	12
TA	4	44.7	4.23*	31.71	57.82	28
	18	30.6	4.43*	6.11	49.06	1
TA%	4	88.6	4.98*	72.78	120.10	0
	17	36.8	4.95*	20.21	57.69	1
	18	30.6	4.38*	6.11	49.06	1
FA	4	45.2	5.63**	32.68	57.82	0
	5	38.6	5.05**	11.00	51.10	0
	10	15.5	4.45*	0.00	42.25	6
	11	13.2	4.92*	0.00	27.44	10
	1	77.1	4.99**	63.95	90.56	4
FA%	3	44.2	5.76**	31.22	58.10	3
	4	88.6	6.74**	72.78	103.79	12
	12	27.5	5.45**	13.05	41.16	0
	17	36.8	4.62*	4.58	57.69	3

¹ For trait abbreviations, see Table 1.

Candidate genes associated with PU, performance and bone ash traits

We identified numerous genes in a 0.5 Mbp region up- and downstream of the significant SNPs in all QTL regions. For the PU QTL on CJA3 we identified 73 positional genes (see Table S2). Of these genes, 51 have known functions. No functional annotation analyses were conducted. No SNP within exon regions could be identified. Therefore, we looked for SNPs that were either intronic or obvious and were related to metabolic processes in which phosphorus might play a role. This filtering led to four genes (from the initial 73; Table 4), which were discussed in detail.

Table 4 Genes and their functions¹ and positions in the reference genome within the PU QTL region of CJA3.

Official gene symbol	Gene name	Function ¹
<i>BMP2</i>	<i>Bone morphogenetic protein 2</i>	Ligand of TGF-beta superfamily, induces cartilage and bone formation
<i>PLCB1</i>	<i>Phospholipase Cβ1</i>	Hydrolyze phospholipids into fatty acids and other lipophilic molecules, catalyzes the formation of inositol-1,4,5-trisphosphate and diacylglycerol from phosphatidylinositol-4,5-bisphosphate, uses calcium as a cofactor, involved in intracellular transduction of many extracellular signals
<i>PLCB4</i>	<i>Phospholipase Cβ4</i>	Hydrolyze phospholipids into fatty acids and other lipophilic molecules, catalyzes the formation of inositol-1,4,5-trisphosphate and diacylglycerol from phosphatidylinositol-4,5-bisphosphate, uses calcium as a cofactor, involved in intracellular transduction of many extracellular signals
<i>TGFB2</i>	<i>Transforming Growth Factor β2</i>	Involved in TGF- β -2 chains, Involved in many processes, e.g. cell differentiation, growth, or morphogenesis processes

¹ According to GeneCards and UniProt.

Discussion

In previous studies, we analyzed the impact of the quail genome (Beck *et al.* 2016a; Künzel *et al.* 2019), ileum microbiota composition (Borda-Molina *et al.* 2020; Vollmar *et al.* 2020) and transcriptomic profiles (Oster *et al.* 2020) as well as miRNA–mRNA and gut microbiota (Ponsuksili *et al.* 2020) on the focal trait PU and other related traits. Preliminary QTL mapping was done on few chromosomes and markers, without reporting any clear signals (Beck *et al.* 2016b). Hence, a thorough QTL mapping has not been done previously. This study filled this gap by conducting QTL linkage and association mapping for these traits in the same experimental design. The results clearly showed that all of the investigated traits are polygenic

in nature and are associated with several significant QTL as well as many other small-effect QTL that were not detectable.

Linkage map

Until the publication of the reference genome in 2016, only a few low-density genetic maps were available. The map calculations were based on ALFP markers (Roussot *et al.* 2003) or microsatellites (Kayang *et al.* 2004), or both types of markers (Kikuchi *et al.* 2005). Recoquillay *et al.* (2015) were the first to calculate a genome-wide linkage map based on SNP markers. Our genetic map agreed well with the map from Recoquillay *et al.* (2015). Next, based on the reference genome *C. japonica* 2.0 (NCBI GCA_001577835.1), genome assemblies for other quail species (Wu *et al.* 2018) and Japanese quail (Morris *et al.* 2020) were developed. As our experiment with several full- and half-sib families and approximately 1000 animals across three generations can be seen as a powerful linkage mapping design, we developed a further linkage map.

The coverage and density of SNP markers were low for some LGs (Table 2, Fig. S1). This is especially noticeable for the smaller LGs (assigned to CJA25 and 28) with fewer than 10 markers. This is a result of the chosen `sizeLimit = 5` in the module `SeparateChromosome` of the software `lepmap2`, as a larger `sizeLimit` resulted in a larger number of markers that could not be assigned to any LG. In addition, this `sizeLimit` was chosen to obtain the best fit based on the karyotype of Japanese quail. The genome of Japanese quail is closely related to that of the domestic chicken (*Gallus gallus domesticus*) (Wu *et al.* 2018) and shows a typical avian species karyotype that includes 10 pairs of macrochromosomes and numerous small microchromosomes (Schmid *et al.* 1989; Zlotina *et al.* 2019).

After comparison with the reference genome (e.g. Fig. 1), only a few markers could not be assigned to physical positions. However, most marker positions in the LGs were consistent with the chromosomes of the reference genome. The good fit of the map is also demonstrated by the high correlation of the linkage and physical marker positions (Table 2). Overall, the present linkage map seems to be of good quality and consistent with the reference genome. This justified the use of this map for the QTL linkage analysis.

QTL results and candidate genes

Our study adds new information for QTL in Japanese quail and provides novel QTL affecting PU, i.e. the PU QTL on CJA3 (Table 3). Owing to the use of different methods, experimental designs and trait definitions and recordings, a sophisticated comparison of QTL linkage mapping results across studies is difficult and thus was not performed in this study. QTL associated with other traits in Japanese quail have been reported by Minvielle *et al.* (2005), Esmailizadeh *et al.* (2012), Ori *et al.* (2014), Sohrabi *et al.* (2012), Recoquillay *et al.* (2015) and Knaga *et al.* (2018).

Some trait interrelationships could be identified by studying the genetic and phenotypic correlations (Beck *et al.* 2016a; Künzel *et al.* 2019) as well as the overlapping of the QTL SI (Table 3). For example, on CJA3, we detected QTL associated with PU and FA% in the same chromosomal region (Table 3). These traits are genetically correlated (0.46) (Künzel *et al.* 2019). On CJA4, we mapped QTL associated with CaU, TA and FA (Table 3), and these traits also showed substantial genetic correlations. The strong genetic correlation between BWG and FI (approximately 0.87, Künzel *et al.* 2019) can be partly explained by the QTL on CJA3, which mapped to both traits (Table 3).

Two of the four most interesting candidate genes in the PU QTL on CJA3 (Table 4) are *transforming growth factor- β 2 (TGFB2)* and *bone morphogenetic protein 2 (BMP2)*. Both genes are members of the TGFB superfamily (Iqbal *et al.* 2018; Loozen *et al.* 2019), which is known to encode multifunctional growth factors involved in cell differentiation, growth and morphogenesis processes (Li *et al.* 2003; Darzi Niarami *et al.* 2014). *TGFB2* is also involved in the mitogen-activated protein kinase (MAPK) signaling pathway (Kyoto Encyclopedia of Genes and Genomes, KEGG). This pathway is associated with many tissue-building and -rebuilding processes in organisms. The other two candidate genes are *phospholipase C β 1* and *4 (PLCB1 and PLCB4)* (Table 4). According to KEGG analysis, they are involved in a broad spectrum of biological processes, including inositol phosphate metabolism, the calcium signaling pathway, the phosphatidylinositol signaling system, the GnRH signaling pathway and the Wnt signaling pathway. Involvement in inositol phosphate-related pathways is of specific interest, because phytate provided the main source of P in the diet. Variation in PU likewise was caused by differences in digestive phytate breakdown, thus providing a different amounts of inositol and inositol phosphates for the quail's metabolism. Also far-reaching and as an example, the Wnt signaling pathway is known to be involved in bone metabolism, which supports the connection of PU and bone ash traits (Robling 2013; Maeda *et al.* 2019; Ponsuksili *et al.* 2020). This partially explains the genetic correlation of the traits.

Conclusion

The experimental design used in this study proved to be powerful for the calculation of an SNP linkage map. Several genome-wide significant QTL could be mapped by linkage and subsequent association analyses. It seems that the focal trait PU and the other performance and bone ash traits are polygenic in nature and are associated with some significant QTL and probably many other small-effect QTL that were not detectable in this study. Some overlap of QTL regions for different traits was detected, which is in agreement with the corresponding genetic correlations. For PU, a QTL on CJA3 could be detected by linkage mapping, which was substantiated by the results of the SNP association mapping. Four candidate genes were identified for this QTL, which should be investigated in further functional studies.

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (BE3703/12-1 and CA1708/2-1) and was part of the research unit P-FOWL (FOR 2601). The reference quail genome (GCA_001577835.2) was established by the International Quail Genome Consortium and McDonnell Genome Institute, Washington University School of Medicine. Open access funding was enabled and organized by Projekt DEAL.

Data availability statement

Genotype and phenotype data, pedigree information and the genetic map are available through OSF and can be accessed at <https://osf.io/57nty/>.

References

- Beck P., Piepho H.-P., Rodehutsord M. & Bennewitz J. (2016a) Inferring relationships between phosphorus utilization, feed per gain, and bodyweight gain in an F2 cross of Japanese quail using recursive models. *Poultry Science* 95, 764–73.
- Beck P., Stratz P., Preuß S., Pitel F., Recoquillay J., Duval E., Rodehutsord M. & Bennewitz J. (2016b) Linkage mapping of quantitative trait loci for phosphorus utilization and growth related traits in an F2-cross of Japanese quail (*Coturnix japonica*). *European Poultry Science* 80. <https://www.european-poultry-science.com/Linkage-mapping-of-quantitative-trait-loci-for-phosphorus-utilization-and-growth-related-traits-in-an-Fsub2sub-cross-of-Japanese-quail-span-classws-name-Coturnix-japonica,QUIEPTUwNTg4NjkmTUIEPTE2MTAxNA.html>.
- Borda-Molina D., Roth C., Hernández-Arriaga A., Rissi D., Vollmar S., Rodehutsord M., Bennewitz J. & Camarinha-Silva A. (2020) Effects on the ileal microbiota of phosphorus and calcium utilization, bird performance, and gender in Japanese Quail. *Animals* 10, 885.
- Broman K.W., Gatti D.M., Simecek P., Furlotte N.A., Prins P., Sen Ś., Yandell B.S. & Churchill G.A. (2019) R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics* 211, 495–502.
- Cheng K.M., Bennett D.C. & Mills A.D. (2010) The Japanese Quail. In: *The UFAW Handbook on the Care and Management of Laboratory and Other Research Animals*, 8th edn (Ed. by J.K. Kirkwood & R. Hubrecht), pp. 655–673. Wiley-Blackwell, Chichester, West Sussex, Ames, IA.
- Darzi Niarami M., Masoudi A.A. & Torshizi R.V. (2014) Association of single nucleotide polymorphism of GHSR and TGFB2 genes with growth and body composition traits in sire and dam lines of a broiler chicken. *Animal Biotechnology* 25, 13–22.
- Eeckhout W. & de Paepe M. (1994) Total phosphorus, phytate-phosphorus and phytase activity in plant feedstuffs. *Animal Feed Science and Technology* 47, 19–29.
- Esmailizadeh A.K., Baghizadeh A. & Ahmadizadeh M. (2012) Genetic mapping of quantitative trait loci affecting bodyweight on chromosome 1 in a commercial strain of Japanese quail. *Animal Production Science* 52, 64–8.
- Iqbal M., Zhang H., Mehmood K. *et al.* (2018) Icariin: a potential compound for the recovery of tibial dyschondroplasia affected chicken via up-regulating BMP-2 expression. *Biological Procedures Online* 20, 15.

- Kayang B.B., Vignal A., Inoue-Murayama M., Miwa M., Monvoisin J.L., Ito S. & Minvielle F. (2004) A first-generation microsatellite linkage map of the Japanese quail. *Animal Genetics* 35, 195–200.
- Kikuchi S., Fujima D., Sasazaki S., Tsuji S., Mizutani M., Fujiwara A. & Mannen H. (2005) Construction of a genetic linkage map of Japanese quail (*Coturnix japonica*) based on AFLP and microsatellite markers. *Animal Genetics* 36, 227–31.
- Knaga S., Siwek M., Tavaniello S., Maiorano G., Witkowski A., Jezewska-Witkowska G., Bednarczyk M. & Zieba G. (2018) Identification of quantitative trait loci affecting production and biochemical traits in a unique Japanese quail resource population. *Poultry Science* 97, 2267–77.
- Künzel S., Bennewitz J. & Rodehutschord M. (2019) Genetic parameters for bone ash and phosphorus utilization in an F2 cross of Japanese quail. *Poultry Science* 98, 4369–72.
- Li H., Deeb N., Zhou H., Mitchell A.D., Ashwell C.M. & Lamont S.J. (2003) Chicken quantitative trait loci for growth and body composition associated with transforming growth factor-beta genes. *Poultry Science* 82, 347–56.
- Li X., Zhang D. & Bryden W.L. (2017) Calcium and phosphorus metabolism and nutrition of poultry: are current diets formulated in excess? *Animal Production Science* 57, 2304–10.
- Loozen L.D., Kruijt M.C., Kragten A.H.M., Schoenfeldt T., Croes M., Oner C.F., Dhert W.J.A. & Alblas J. (2019) BMP-2 gene delivery in cell-loaded and cell-free constructs for bone regeneration. *PLOS ONE* 14, e0220028.
- Maeda K., Kobayashi Y., Koide M., Uehara S., Okamoto M., Ishihara A., Kayama T., Saito M. & Marumo K. (2019) The regulation of bone metabolism and disorders by Wnt signaling. *International Journal of Molecular Sciences* 20, 5525.
- Manichaikul A., Dupuis J., Sen S. & Broman K.W. (2006) Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 174, 481–9.
- Mills A.D. & Faure J.-M. (1991) Divergent selection for duration of tonic immobility and social reinstatement behavior in Japanese quail (*Coturnix coturnix japonica*) chicks. *Journal of Comparative Psychology* 105, 25–38.
- Minvielle F., Kayang B.B., Inoue-Murayama M., Miwa M., Vignal A., Gourichon D., Neau A., Monvoisin J.-L. & Ito S. (2005) Microsatellite mapping of QTL affecting growth, feed consumption, egg production, tonic immobility and body temperature of Japanese quail. *BMC Genomics* 6, 87.

- Morris K.M., Hindle M.M., Boitard S. *et al.* (2020) The quail genome: insights into social behaviour, seasonal biology and infectious disease response. *BMC Biology* 18, 14.
- Ori R.J., Esmailizadeh A.K., Charati H., Mohammadabadi M.R. & Sohrabi S.S. (2014) Identification of QTL for live weight and growth rate using DNA markers on chromosome 3 in an F2 population of Japanese quail. *Molecular Biology Reports* 41, 1049–57.
- Oster M., Reyer H., Trakooljul N., Weber F.M., Xi L., Muráni E., Ponsuksili S., Rodehutschord M., Bennewitz J. & Wimmers K (2020) Ileal transcriptome profiles of Japanese quail divergent in phosphorus utilization. *International Journal of Molecular Sciences* 21, 2762.
- Ponsuksili S., Reyer H., Hadlich F. *et al.* (2020) Identification of the key molecular drivers of phosphorus utilization based on host miRNA-mRNA and gut microbiome interactions. *International Journal of Molecular Sciences* 21, 2818.
- Rastas P., Calboli F.C.F., Guo B., Shikano T. & Merilä J. (2015) Construction of ultra-dense linkage maps with Lep-MAP2: stickleback F2 recombinant crosses as an example // Data from: Construction of ultra-dense linkage maps with Lep-MAP2: stickleback F2 recombinant crosses as an example. *Genome Biology and Evolution* 8, 78–93.
- Rastas P., Paulin L., Hanski I., Lehtonen R. & Auvinen P. (2013) Lep-MAP: fast and accurate linkage map construction for large SNP datasets. *Bioinformatics (Oxford, England)* 29, 3128–34.
- Recoquillay J., Pitel F., Arnould C. *et al.* (2015) A medium density genetic map and QTL for behavioral and production traits in Japanese quail. *BMC Genomics* 16, 10.
- Robling A.G. (2013) The expanding role of Wnt signaling in bone metabolism. *Bone* 55, 256–7.
- Roussot O., Fève K., Plisson-Petit F., Pitel F., Faure J.-M., Beaumont C. & Vignal A. (2003) AFLP linkage map of the Japanese quail *Coturnix japonica*. *Genetics Selection Evolution* 35, 559–72.
- Schmid M., Enderle E., Schindler D. & Schempp W. (1989) Chromosome banding and DNA replication patterns in bird karyotypes. *Cytogenetics and Cell Genetics* 52, 139–46.
- Shibusawa M., Minai S., Nishida-Umehara C., Suzuki T., Mano T., Yamada K., Namikawa T. & Matsuda Y. (2001) A comparative cytogenetic study of chromosome homology between chicken and Japanese quail. *Cytogenetics and Cell Genetics* 95, 103–9.

- Sohrabi S.S., Esmailizadeh A.K., Baghizadeh A., Moradian H., Mohammadabadi M.R., Askari N. & Nasirifar E. (2012) Quantitative trait loci underlying hatching weight and growth traits in an F2 intercross between two strains of Japanese quail. *Animal Production Science* 52, 1012.
- Stock A.D. & Bunch T.D. (1982) The evolutionary implications of chromosome banding pattern homologies in the bird order Galliformes. *Cytogenetics and Cell Genetics* 34, 136–48.
- Vali N. & Jalali M.A. (2011) Influence of different levels of phytase enzyme on Japanese quail (*Coturnix japonica*) performance. *Agriculturae Conspectus Scientificus* 76, 387–390.
- Vollmar S., Wellmann R., Borda-Molina D., Rodehutschord M., Camarinha-Silva A. & Bennewitz J. (2020) The gut microbial architecture of efficiency traits in the domestic poultry model species Japanese quail (*Coturnix japonica*) assessed by mixed linear models. *Genes Genomes Genetics* 10, 2553–62.
- Weir B.S. & Cockerham C.C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358.
- Wu Y., Zhang Y., Hou Z. *et al.* (2018) Population genomic data reveal genes related to important traits of quail. *GigaScience* 7, 1–16.
- Yang J., Lee S.H., Goddard M.E. & Visscher P.M. (2011) GCTA: a tool for genome-wide complex trait analysis. *American Journal of Human Genetics* 88, 76–82.
- Zlotina A., Maslova A., Kosyakova N., Al-Rikabi A.B.H., Liehr T. & Krasikova A. (2019) Heterochromatic regions in Japanese quail chromosomes: comprehensive molecular-cytogenetic characterization and 3D mapping in interphase nucleus. *Chromosome Research* 27, 253–70.

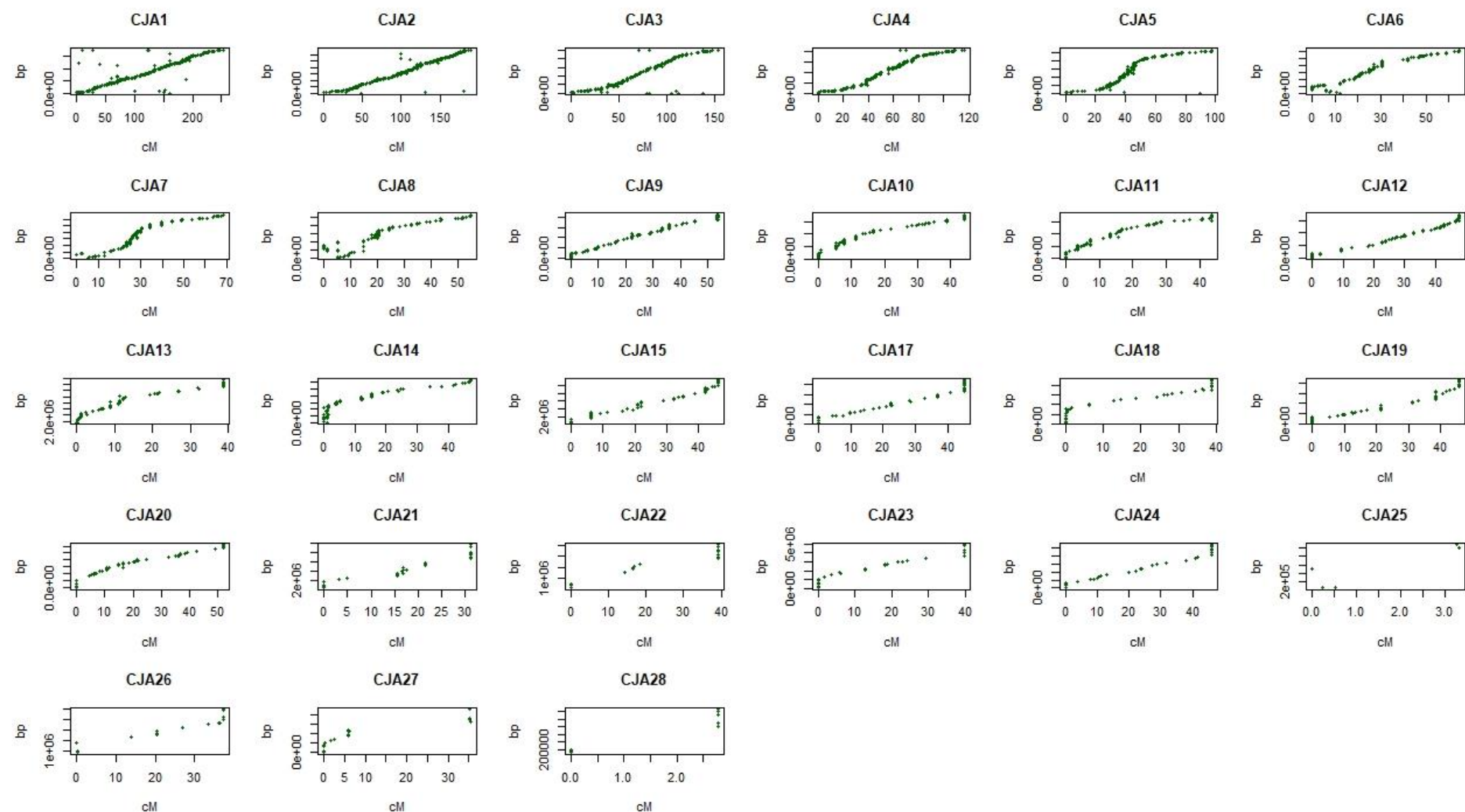
Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Comparison of genetic and physical maps.

Table S1. Summary of trait-associated markers ($P \leq 0.05$) within the significant QTL regions.

Table S2. Summary of 73 identified genes for PU in a 0.5 Mbp region up- and downstream of significant SNPs.



Supplemental Figure S1 Comparison of genetic and physical maps. The y-axis shows the physical position (bp), which is based on the reference genome assembly (NCBI GCA_001577835.1), and the x-axis shows the genetic position (cM). Note that the SNP positions at 0 bp refer to an as yet unknown position.

Supplemental Table S1 Summary of trait-associated markers ($P \leq 0.05$) within the significant QTL regions.

Trait ^a	CJA	Marker ID	p-value	cM	Matched markers with other traits
PU	3	id10557	0.041	37.70	
	3	id01656	0.042	37.70	
	3	id14384	0.036	37.70	
	3	id14623	0.012	41.86	
	3	id10574	0.034	43.22	
CaU	4	id06563	0.027	46.69	
	4	id09235	0.048	65.59	
	4	id04477	0.048	65.59	
	4	id14317	0.025	65.59	
	4	id12663	0.012	65.59	
	4	id16422	0.042	65.59	
	4	id12664	0.042	65.59	
	4	id05313	0.043	68.33	
	4	id10860	0.034	68.33	
	4	id03398	0.042	68.33	
	4	id05311	0.042	68.33	
	4	id03657	0.040	71.08	
	4	id12682	0.017	74.72	
	4	id10870	0.048	75.27	
	17	id11675	0.002	28.28	
	17	id23952	0.019	36.76	
	17	id03867	0.044	45.19	
	17	id05727	0.013	45.19	
	17	id03250	0.013	45.19	
	17	id13349	0.036	45.19	
FI	3	id12506	0.003	90.82	BWG
	3	id10670	0.001	95.39	BWG
	3	id07565	0.038	97.35	
	3	id10683	0.019	97.35	BWG
	3	id06748	0.022	100.81	BWG
	3	id14876	0.036	101.52	BWG
	3	id29726	0.049	112.08	
	5	id21765	0.016	30.87	
	5	id10931	0.025	35.87	
	5	id32246	0.035	36.37	
	5	id32730	0.032	36.93	
	11	id06872	0.024	7.32	FA
	11	id32446	0.027	7.32	FA
	11	id15452	0.003	11.30	FA
	11	id32451	0.005	13.21	FA
	11	id07827	0.011	13.21	FA
	11	id09706	0.011	13.21	FA
	11	id05659	0.009	13.21	FA
	11	id08551	0.009	13.21	FA
	11	id05029	0.023	15.58	FA
11	id32793	0.039	16.83		
BWG	3	id12506	0.009	90.82	FI

CHAPTER ONE

	3	id10670	0.002	95.39	FI
	3	id19121	0.009	97.34	
	3	id14875	0.030	97.35	
	3	id19122	0.030	97.35	
	3	id10683	0.008	97.35	FI
	3	id03367	0.028	97.35	
	3	id14232	0.046	100.81	
	3	id06748	0.039	100.81	FI
	3	id14876	0.042	101.52	FI
	3	id10530	0.028	109.31	
	3	id12548	0.013	120.70	
TA	4	id02692	0.013	38.31	
	4	id05290	0.007	38.31	
	4	id11068	0.007	38.31	
	4	id10783	0.008	38.31	
	4	id03383	0.021	38.69	
	4	id07589	0.028	38.69	
	4	id04753	0.028	38.69	
	4	id02319	0.028	38.69	
	4	id06108	0.006	38.69	
	4	id09188	0.006	38.69	
	4	id03382	0.031	38.69	
	4	id25452	0.049	38.69	
	4	id31883	0.031	38.86	
	4	id14009	0.019	39.51	
	4	id20374	0.010	39.51	
	4	id09186	0.028	39.51	
	4	id23126	0.031	39.51	
	4	id10788	0.028	39.51	
	4	id06761	0.044	39.51	
	4	id16557	0.044	39.51	
	4	id10790	0.044	39.51	
	4	id16986	0.044	39.51	
	4	id04323	0.019	39.76	
	4	id09192	0.019	40.13	
	4	id08248	0.038	44.21	
	4	id25646	0.028	46.87	
	4	id09217	0.039	47.02	
	4	id07604	0.049	49.98	
	18	id08651	0.030	38.77	TA%
TA%	17	id06643	0.001	40.83	
	18	id08651	0.029	38.77	TA
FA	10	id07281	0.049	0.88	
	10	id13413	0.026	5.22	
	10	id05642	0.011	16.49	
	10	id09667	0.043	16.49	
	10	id07288	0.027	32.10	
	10	id32850	0.004	35.04	
	11	id06872	0.023	7.32	FI
	11	id32446	0.017	7.32	FI
	11	id15452	0.012	11.30	FI

11	id32451	0.011	13.21	FI
11	id07827	0.021	13.21	FI
11	id09706	0.023	13.21	FI
11	id05659	0.022	13.21	FI
11	id08551	0.022	13.21	FI
11	id05029	0.041	15.58	FI
11	id32374	0.042	16.82	
FA%	1	id33073	0.008	71.94
	1	id03827	0.024	76.45
	1	id09976	0.013	80.66
	1	id11811	0.027	80.85
	3	id10552	0.020	32.13
	3	id02658	<0.001	32.13
	3	id14857	0.046	48.40
	4	id17823	0.008	74.72
	4	id09240	0.008	74.72
	4	id32711	0.032	74.72
	4	id10866	0.008	74.72
	4	id25649	0.026	77.60
	4	id04769	0.024	79.63
	4	id09248	0.045	83.66
	4	id00486	0.030	88.64
	4	id26718	0.042	90.26
	4	id13708	0.019	91.30
	4	id09265	0.040	103.54
	4	id27230	0.034	103.64
	17	id33074	0.047	9.73
	17	id00247	0.002	17.16
	17	id19277	0.036	36.74

^a For trait abbreviations, see Table 1.

Supplemental Table S2 Summary of 73 identified genes for PU in a 0.5 Mbp region up- and downstream of significant SNPs.

Trait ^a	Gene	Gene location in bp		SNP	Location (distance to genes in bp)	SNP location	
		Start	End			cM	bp
PU	PAK5	11950702	12094008				
	LAMP5	12095782	12106823				
	PLCB4	12108448	12289519				
	PLCB1	12329777	12674893	id14384	intronic	37.703	12540995
	LOC107310957	12675334	12679908				
	TMX4	12718163	12743981				
	HAO1	12748177	12774489				
	LOC107310956	12762766	12764953				
	LOC107310980	13000416	13110455				
	BMP2	13149896	13155078				
	LOC116652411	13174497	13239523				
	LOC107310917	13273456	13326531				
	FERMT1	13326689	13344740				
	LRRN4	13349146	13357235				
	CCT4	13362618	13368311				
	FAM161A	13368497	13376521				
	DTD1	13376027	13383387				
	LOC116653205	13383435	13385218				
	SEC23B	13384133	13398061				
	POLR3F	13399683	13405682				
	DZANK1	13405772	13421996				
	BIRC5	13422245	13423482				
	KAT14	13424420	13438934				
	LOC107310984	13438522	13439387				
	OVOL2	13447675	13454039				
	MGME1	13456378	13461301				
	SNX5	13461501	13477385	id10557	upstream (13110)	37.701	13490495
	RRBP1	13509416	13529344	id10557	downstream (18920)	37.701	13490495
	DSTN	13532567	13541879				
	LOC116653173	13541962	13562948				
	LOC107310933	13564800	13567019				
	LOC107310890	13584169	13587185				
	SYNDIG1	13662551	13678715	id01656	upstream (13979)	37.703	13692694
	LOC107310995	13727528	13728659	id01656	downstream (34834)	37.703	13692694
	CST7	13770487	13778397				
	APMAP	13782846	13793306				
	ACSS1	13793652	13823150				
	TTBK1	13841207	13915358				
	SLC22A7	13918080	13930431				

TTL	13931216	13945307				
POLR1B	13945877	13958409				
PTCRA	13959083	13960082				
CNPY3	13960406	13962555				
GNMT	13962792	13964620				
PEX6	13965569	13977199				
VSX1	13980174	13982882				
ENTPD6	13996794	14007427				
MAL	14010292	14014218				
MRPS5	14017074	14048607				
LOC116653272	14150276	14150648				
SLC8A1	14175324	14272642				
LOC116653182	14184512	14190798				
LOC107311002	16816696	16819420				
TGFB2	16895806	16958385				
RRP15	16964138	16991977				
SPATA17	17175446	17258347				
GPATCH2	17258251	17354587	id10574	intronic	43.223	17293862
LOC107311059	17363606	17392868				
LOC116653249	17447035	17456715				
ESRRG	17456684	17830994				
LOC107311023	17471465	17474542				
LOC116653250	17678879	17681797				
USH2A	17848612	18204761				
KCTD3	18205056	18228494				
LOC107311086	18240156	18254071				
KCNK2	18303701	18471703	id14623	intronic	41.856	18451011
LOC107311089	18421856	18471849	id14623	intronic	41.856	18451011
CENPF	18473857	18508763				
PTPN14	18517965	18615002				
SMYD2	18616279	18643814				
LOC107311062	18650875	18657208				
PROX1	18724302	18774251				
LOC116653231	18899477	19022498				

^a For trait abbreviations, see Table 1.

CHAPTER TWO

Composition of the ileum microbiota is a mediator between the host genome and phosphorus utilization and other efficiency traits in Japanese quail (*Coturnix japonica*)

Valentin Haas*, Solveig Vollmar, Siegfried Preuß, Markus Rodehutschord, Amélia Camarinha-Silva and Jörn Bennewitz

*Correspondence: valentin.haas@uni-hohenheim.de

Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

Published in *Genetics Selection Evolution* (2022), Volume 54:20

<https://doi.org/10.1186/s12711-022-00697-8>

Abstract

Background: Phosphorus is an essential nutrient in all living organisms and, currently, it is the focus of much attention due to its global scarcity, the environmental impact of phosphorus from excreta, and its low digestibility due to its storage in the form of phytates in plants. In poultry, phosphorus utilization is influenced by composition of the ileum microbiota and host genetics. In our study, we analyzed the impact of host genetics on composition of the ileum microbiota and the relationship of the relative abundance of ileal bacterial genera with phosphorus utilization and related quantitative traits in Japanese quail. An F2 cross of 758 quails was genotyped with 4k genome-wide single nucleotide polymorphisms (SNPs) and composition of the ileum microbiota was characterized using target amplicon sequencing. Heritabilities of the relative abundance of bacterial genera were estimated and quantitative trait locus (QTL) linkage mapping for the host was conducted for the heritable genera. Phenotypic and genetic correlations and recursive relationships between bacterial genera and quantitative traits were estimated using structural equation models. A genomic best linear unbiased prediction (GBLUP) and microbial (M)BLUP hologenomic selection approach was applied to assess the feasibility of breeding for improved phosphorus utilization based on the host genome and the heritable part of composition of the ileum microbiota.

Results: Among the 59 bacterial genera examined, 24 showed a significant heritability (nominal $p \leq 0.05$), ranging from 0.04 to 0.17. For these genera, six genome-wide significant

QTL were mapped. Significant recursive effects were found, which support the indirect host genetic effects on the host's quantitative traits via microbiota composition in the ileum of quail. Cross-validated microbial and genomic prediction accuracies confirmed the strong impact of microbial composition and host genetics on the host's quantitative traits, as the GBLUP accuracies based on the heritable microbiota-mediated components of the traits were similar to the accuracies of conventional GBLUP based on genome-wide SNPs.

Conclusion: Our results revealed a significant effect of host genetics on composition of the ileal microbiota and confirmed that host genetics and composition of the ileum microbiota have an impact on the host's quantitative traits. This offers the possibility to breed for improved phosphorus utilization based on the host genome and the heritable part of composition of the ileum microbiota.

Background

The poultry industry is a fast-growing sector of the global food supply. For economic and environmental impact reasons, feed efficiency and nutrient efficiency have received considerable attention in poultry research. Phosphorus (P) is an essential nutrient with finite global mineral resources and an enormous environmental impact due to its excretion in animal faeces [1–3]. In plant seeds, P is primarily stored in the form of phytic acid (myo-inositol hexaphosphate, InsP6) [4]. Better utilization of P from feed components is a desirable goal, which requires the action of phytase and other phosphatase enzymes that catalyze the stepwise cleavage of P from InsP6 in the digestive tract [5, 6]. Poultry are known to have comparably low endogenous phytase activity. Nevertheless, native InsP degradation in the gastrointestinal tract (GIT) can occur by the action of phytases and phosphatases that originate from the endogenous mucosa of the GIT from some vegetable feed components, or from the gut microbiota (reviewed in [7]).

It is well known that microbiota composition in the GIT of livestock is influenced by environmental factors, such as diet or housing conditions. However, numerous literature results indicate that host genetics has also an effect on GIT microbial colonization; and significant heritabilities have been reported for the relative abundance of bacterial genera in the cow rumen [8, 9], for bacterial genera and operational taxonomic units (OTU) level in the pig colon [10, 11], and for cecal and fecal microbial species in chickens [12–14].

A previous study on Japanese quails under standardized feeding and housing conditions showed that P utilization (PU) varied substantially between birds and had a heritability of 0.14 [15]. Borda-Molina et al. [16] detected differences in the relative abundance of different microbial genera between quails from the same population with high- versus low-PU. In a subsequent study, we confirmed that PU and related traits are strongly influenced by the hosts'

composition of the ileal microbiota and we estimated a significant microbiability with an order of magnitude similar to heritability [17]. We assessed the microbial architecture of PU and related host traits by applying microbiome-wide association analysis (MWAS) and found that they were polymicrobial, with many trait-associated bacterial genera, but none of the genera had an exceptionally large effect [17]. Subsequently, we used the same dataset to map quantitative trait loci (QTL) for PU in the quail genome. The quails were genotyped with 4k single nucleotide polymorphisms (SNPs) and several significant QTL were identified [18].

In MWAS, it is assumed that the microbiota composition is the cause of the variation in host's quantitative traits [17], but this might not be true because microbiota composition and quantitative traits could be influenced by the same host QTL and thus may be correlated via a common set of host QTL. Structural equation models (SEM) [19] are multivariate mixed model equations that account for recursive relationships between traits and allow separation of direct and indirect genetic effects that are responsible for the genetic relationship between traits. If a recursive relationship exists, QTL that directly affect one trait may indirectly affect the second trait via the recursive relationship between the traits. Saborío-Montero et al. [8] used SEM to identify significant polymicrobial recursive interactions between rumen microbiota and methane emissions in cattle. Tiezzi et al. [20] also confirmed recursive effects of fecal microbiota composition on fat deposition in pigs, using SEM.

In Weishaar et al. [21], a genome-based selection index to improve a quantitative trait in the host was developed that considered the hologenome, i.e., both the host genome and microbial metagenome of microbiota composition in the GIT. The selection index included estimated breeding values for the direct effect of the host genome on the trait and for the indirect effect mediated by microbiota composition. The core aspect of the method is a reference population with trait-recorded animals that have been genotyped for a SNP-chip and characterized for microbiota composition. A microbial mixed model was used to estimate the effect of the animal's microbiota on the trait. Subsequently a genomic mixed linear model was applied to predict the SNP effects for the estimated animal microbiota effect. Weishaar et al. [21] successfully applied this model to a small pig dataset on feed efficiency.

Using the same quail dataset as in our previous studies [17, 18], the aim of the current study was to analyze the impact of host genetics on composition of the ileum microbiota and the relationship of ileal bacterial genera with PU and related traits, i.e. body weight gain (BWG), feed intake (FI) and feed per gain (F:G). For this purpose, we estimated heritabilities of the relative abundance of bacterial genera, and the correlations and recursive relationships between the relative abundance of the bacterial genera and these four traits, and performed QTL linkage mapping for the relative abundance of bacterial genera. Subsequently, the

hologenomic selection approach developed by Weishaar et al. [21] was applied to assess the feasibility of breeding for improved PU based on the host genome and the heritable part of the composition of the ileum microbiota.

Methods

Experimental design

Details of the experimental design are in Beck et al. [15] and, thus, only the most relevant aspects are presented in the following. This animal experiment was performed according to the requirements of the German Animal Welfare Legislation and was approved by the Animal Welfare Commissioner of the University of Hohenheim (approval number S371/13TE). An F2 population of Japanese quail (*Coturnix japonica*) was established based on two divergent lines selected for social reinstatement behavior [22]. Twelve males and 12 females from each founder line were mated to generate the F1 generation. Seventeen roosters and 34 hens from the F1 generation were randomly selected and mated (one male with two females), resulting in 920 F2 individuals. These F2 birds were phenotyped between 10 and 15 days of age, while the birds were provided with a corn-soybean meal-based diet without mineral P or phytase supplements. A diet with an overall low P content was chosen to evaluate the PU potential of the quails.

Sample collection, SNP genotyping, and characterization of the ileum microbiota

The focal trait of this experiment was PU, which was calculated based on total P intake and P excretion, as well as based on FI during the experimental period. Quail BWG was quantified as the difference in body weight between days 10 and 15. The F:G ratio was computed as the FI during this 5-day period divided by the BWG. The quails were slaughtered at 15 days of age to collect ileum samples for further analysis. The birds were incubated and slaughtered on 12 different days, which were treated as test days in the statistical analysis. Estimates of the phenotypic and genetic correlations between the four recorded traits are in Beck et al. [15] and Künzel et al. [23].

DNA preparation, 4k SNP genotyping, and construction of a genome-wide linkage map are described in detail by Vollmar et al. [18]. In brief, all birds were genotyped for 5388 SNPs and the following criteria were applied to filter the genotypes: SNPs with one or more conflicting genotypes between parent and offspring, a minor allele frequency (MAF) ≤ 0.03 , a SNP call frequency ≤ 0.9 , and a cluster separation ≤ 0.4 were removed. We also excluded SNPs on the sex chromosomes Z and W. Finally, 3986 SNPs remained for further analysis.

Analyses of the composition of the ileum microbiota were performed by targeted amplicon sequencing, as described in Borda-Molina et al. [16]. Sequences were clustered into OTU at $> 97\%$ similarity. In total, 1188 OTU with an average relative abundance higher than 0.0001%

and a sequence length greater than 250 bp were used in further analyses. Representative sequences were manually identified with the seqmatch function of the RDP database [24]. The output taxonomy table followed the confidence threshold cut-off value for each taxonomic level as defined by Yarza et al. [25]: genus (94.5%), family (86.5%), order (82.0%), class (78.5%) and phylum (75.0%) [25]. Due to the use of a strict quality filter on the sequences, several samples were excluded. The final dataset included data on 758 quails with SNP genotypes, microbiota composition characteristics, and trait records (PU, FI, BWG, and F:G).

Statistical analyses

Transformation of microbial data

We used two microbial classifications for the statistical analyses, i.e., microbial genus and OTU. The latter was used to build the microbial relationship matrix \mathbf{M} (see below). Genera data were filtered for a minimum of 0.01% of the average relative abundance of each genus. This filtering step reduced the number of genera from 200 to 59. Because the distribution of the relative abundance of each microbial genus deviated remarkably from a Gaussian distribution, we applied a Box–Cox transformation with a specific lambda for each genus. The lambda was determined by a grid search to maximize the likelihood function of a normal distribution, following Box and Cox [26]:

$$f(\mathbf{y}) = \begin{cases} \frac{y^{\lambda-1}}{\lambda} (\lambda \neq 0) \\ \log y (\lambda = 0) \end{cases},$$

where \mathbf{y} is a vector of the relative abundances of each microbial genus to be transformed, and λ is the transformation parameter determined for each genus, which ranged from -2 to 0.505.

Mixed linear models for microbial composition

The following statistical analyses using a mixed linear model were performed in R Studio (Version 3.5.3) [27], ASReml R (Version 3.0) [28], and ASReml 4.1 [29]:

$$\mathbf{y} = \mu \mathbf{1} + \mathbf{Z}_{\mathbf{td}} \mathbf{td} + \mathbf{Z}_{\mathbf{a}} \mathbf{a} + \mathbf{e}, \quad (1)$$

where \mathbf{y} is a vector of the transformed relative abundances of each genus, μ is the trait mean, and $\mathbf{1}$ is a vector of 1s; \mathbf{td} is a vector of the random test day effects, assumed to follow a normal distribution $\mathbf{td} \sim N(0, \mathbf{I}\sigma_{\mathbf{td}}^2)$, where $\sigma_{\mathbf{td}}^2$ is the variance, \mathbf{I} is the identity matrix, and $\mathbf{Z}_{\mathbf{td}}$ is the design matrix; \mathbf{a} is a vector of the random animal effects, assumed to follow a normal distribution $\mathbf{a} \sim N(0, \mathbf{A}\sigma_{\mathbf{a}}^2)$, where \mathbf{A} is the pedigree-based relationship matrix and $\sigma_{\mathbf{a}}^2$ the additive genetic variance, and $\mathbf{Z}_{\mathbf{a}}$ is the design matrix. We chose to use pedigree instead of SNP genotypes here, because of the limited number of SNPs in the study. Finally, \mathbf{e} is a vector of random residuals, assumed to follow a normal distribution $\mathbf{e} \sim N(0, \mathbf{I}\sigma_{\mathbf{e}}^2)$, where $\sigma_{\mathbf{e}}^2$ is the variance.

Using this mixed linear model, heritability (h_y^2) of each microbial genus was estimated as $h_y^2 = \frac{\sigma_a^2}{\sigma_p^2}$, with $\sigma_p^2 = \sigma_a^2 + \sigma_{td}^2 + \sigma_e^2$. Significance of the heritabilities was tested by conducting a likelihood ratio test on the random animal effects. The test statistic was computed as $D = 2[\log(L_2) - \log(L_1)]$, where L_2 is the likelihood of the full Model (1) and L_1 that of Model (1) without random animal effects, and is distributed as a chi-square with one degree of freedom under the null hypothesis of zero heritability. All microbial genus heritabilities with a nominal p value ≤ 0.05 were used for further analyses. To estimate the number of false positives among the significant heritabilities, we calculated the false discovery rate (FDR) q value [30], assuming 59 comparisons. The FDR q value of the significant heritability with the largest p value provided an estimate of the proportion of false positives among the significant heritabilities of the microbial genera.

Bivariate analyses of microbial composition with host traits using structural equation models

We estimated the phenotypic correlations as Pearson correlations between each significant heritable genus ($p \leq 0.05$) and each of the four host traits (PU, FI, BWG, and F:G) using the function `cor.test()` in R [27]. Subsequently, bivariate SEM were applied to trait-genus combinations with a significant phenotypic correlation ($p \leq 0.05$) in order to estimate phenotypic (r_p) and genetic correlations (r_g) and to reveal the biological link between heritable microbiota and host traits. Of the four host traits, only FI is assumed to affect microbiota compositions, e.g. by the rate of food passage in the gastrointestinal tract. However, for the sake of simplicity, a recursive relationship between the heritable microbial genera and each of the four traits was assumed. The following bivariate recursive mixed linear model was applied, using the notation of Rosa et al. [31]:

$$\mathbf{y} = (\mathbf{\Lambda} \otimes \mathbf{I}_n)\mathbf{y} + \mathbf{Z}_{td}\mathbf{td} + \mathbf{Z}_a\mathbf{a} + \mathbf{e}, \quad (2)$$

where \mathbf{y} is the vector of the phenotypic records of the two analyzed traits for the n individuals in the identity matrix \mathbf{I}_n . The off-diagonal 2x2 matrix $\mathbf{\Lambda}$ contains the structural coefficients $\lambda_{i,j}$, which express the rate of change of trait i (PU, FI, BWG or F:G) as a result of the recursive influence of trait j , i.e., the relative abundance of the microbial genera in the ileum:

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 0 \\ \lambda_{i,j} & 0 \end{bmatrix}.$$

The remaining terms are as defined in Model (1). The joint distribution of \mathbf{td} , \mathbf{a} , and \mathbf{e} was as follows:

$$\begin{bmatrix} \mathbf{td} \\ \mathbf{a} \\ \mathbf{e} \end{bmatrix} \sim N \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \mathbf{T} \otimes \mathbf{I}_m & 0 & 0 \\ 0 & \mathbf{G} \otimes \mathbf{A} & 0 \\ 0 & 0 & \mathbf{R} \otimes \mathbf{I}_n \end{bmatrix} \right\}, \quad (3)$$

where \mathbf{T} , \mathbf{G} , and \mathbf{R} are the test day, additive-genetic, and residual variance-covariance matrices of the system of equations [19]. Identity matrices \mathbf{I}_m and \mathbf{I}_n have dimensions equal to the number of test days (m) and the number of individuals (n), respectively. Matrix \mathbf{A} is the numerator relationship matrix, and \otimes is the Kronecker product. In Model (3), we assumed a diagonal residual covariance matrix to ensure the identifiability [31]. The following transformations were applied to obtain genetic parameters that correspond to those of mixed linear models without recursive effects [19, 31, 32]:

$$\begin{aligned} \mathbf{T}^* &= (\mathbf{I} - \mathbf{\Lambda})^{-1} \mathbf{T} (\mathbf{I} - \mathbf{\Lambda})'^{-1} \\ \mathbf{G}^* &= (\mathbf{I} - \mathbf{\Lambda})^{-1} \mathbf{G} (\mathbf{I} - \mathbf{\Lambda})'^{-1} \\ \mathbf{R}^* &= (\mathbf{I} - \mathbf{\Lambda})^{-1} \mathbf{R} (\mathbf{I} - \mathbf{\Lambda})'^{-1} \\ \mathbf{P}^* &= \mathbf{T}^* + \mathbf{G}^* + \mathbf{R}^* \end{aligned} \quad (4)$$

Phenotypic and genetic correlations were estimated based on these matrices (\mathbf{T}^* , \mathbf{G}^* , \mathbf{R}^*) using standard notations. Standard errors were estimated using the method described by Beck et al. [15]. Note that it would be possible to estimate genetic parameters from the \mathbf{T} , \mathbf{G} and \mathbf{R} matrices, which can be interpreted as ‘system parameters’ [19] that control the ‘system’ of the traits. Comparison of estimates of genetic parameters based on \mathbf{G} versus \mathbf{G}^* would also shed light on the nature of the genetic correlation, i.e. the extent to which it is driven by pleiotropy or by indirect effects. However, since this would result in many additional parameters with large standard errors due to the limited sample size, estimates based on the \mathbf{T} , \mathbf{G} and \mathbf{R} matrices are not shown.

QTL linkage analyses of microbial genera

For QTL mapping, we used the R package R/qtI2 [33], which was originally set up for inbred crosses. Because the founders in our study were not inbred, this assumption was not fulfilled. Therefore, we calculated the QTL genotype probabilities for each F2 individual and each chromosomal position using the R package MAPfastR [34], which was developed for outbred line crosses, which were then transferred to R/qtI2. Genome scans were performed using regression of the phenotypes on two QTL genotype probability-derived regression variables, representing the QTL additive and dominant effects [33]. The software did not allow the inclusion of random nuisance effects, other than a residual, or classification effects. Therefore, the effects of test days were included as dummy covariates in the model. The resulting logarithm of the odds (LOD) scores per cM were used as test statistics. To address the problem of multiple testing, a permutation test (10,000 permutations) was applied to derive 5 and 10% genome-wide significance thresholds for each microbial genus. Support intervals (SI) for QTL

position were determined by using the 1.5 LOD drop-off method, which corresponds approximately to a 95% confidence interval [35].

Within the SI for each identified QTL, all markers were evaluated for trait association using the single-marker association mapping approach implemented in the software package GCTA [36]. The model regressed the phenotypes on the number of copies of the 1-allele at the SNP (i.e. 0, 1, or 2 copies) and included test days as dummy covariates and the random animal genetic effect with a SNP-derived covariance matrix, as implemented in the software using the LOCO option. No correction for multiple testing was performed during the association analysis within the SI because the number of SNPs within a SI was usually small, reducing the problem of multiple testing in genome-wide association analysis.

Genomic and microbial trait predictions

To evaluate the hologenomic selection index proposed by Weishaar et al. [21], we first applied a microbial linear mixed model to each quantitative trait as follows [17]:

$$\mathbf{y} = \mu\mathbf{1} + \mathbf{Z}_{td}\mathbf{td} + \mathbf{I}\mathbf{k} + \mathbf{e}, \quad (5)$$

where \mathbf{y} is the vector of observations of one of the four performance traits (PU, FI, BWG, or F:G) for n animals and \mathbf{k} is the vector of the random microbiota effects of all animals, assumed to be distributed as $\mathbf{k} \sim N(0, \mathbf{M}\sigma_k^2)$, where σ_k^2 is the microbial variance and \mathbf{M} is the microbial relationship matrix calculated as described in Camarinha-Silva et al. [10]. The remaining terms are as defined in Model (1).

Estimates of the microbial animal effects $\hat{\mathbf{k}}$ for each trait from Model (5) were then used as observations in the following genomic prediction model, as proposed by Weishaar et al. [21]:

$$\hat{\mathbf{k}} = \mu\mathbf{1} + \mathbf{I}\mathbf{m} + \mathbf{e}, \quad (6)$$

where μ is the overall mean, \mathbf{m} is the vector of random animal genetic effects, assumed distributed $\mathbf{m} \sim N(0, \mathbf{G}\sigma_m^2)$, where \mathbf{G} is the genomic covariance matrix, estimated using the 4k SNP genotypes following method 1 of VanRaden [37] and σ_m^2 is the genomic variance of the estimated microbiota effects; and \mathbf{e} is the vector of residuals with variance σ_e^2 . Heritability of the microbiota-mediated trait $\hat{\mathbf{k}}$ was calculated as $h_k^2 = \frac{\sigma_m^2}{\sigma_m^2 + \sigma_e^2}$. Significance tests for estimates of heritability were performed by likelihood ratio tests.

Three types of predictions were performed and evaluated using cross-validation, two genomic predictions and on microbial prediction. Model (1) was used to obtain genomic best linear unbiased predictions (GBLUP), but with the \mathbf{A} matrix replaced by the \mathbf{G} matrix. Model (5) was used to obtain microbial (M)BLUP [10]. For GBLUP of the microbiota-mediated part of the trait,

Model (5) was used to obtain estimates of the random microbiota effects of the animals for each of the four traits, which were subsequently used as observations in Model (6).

Microbial and genomic predictions were evaluated using cross-validation with 500 repetitions, with variance components fixed at their estimated values. For each repetition, a reference population of 80% of the animals was randomly selected to estimate the effects of OTU and/or SNPs. The remaining 20% of animals were used as the validation population for prediction of animal effects. The average Pearson correlation between the predicted animal effects and the observed animal phenotypes across replications were used as the accuracy of prediction, with confidence intervals calculated from the 2.5 and 97.5% quantiles of the 500 correlations.

Results

Heritabilities, correlations, and structural coefficients

Among the 59 bacterial genera examined, 24 had a significant estimate of heritability ($p \leq 0.05$) (Table 1), with estimates ranging from 0.04 to 0.17. The highest estimates were for *Clostridium sensu stricto*, *Lactobacillus* and *Bifidobacterium*, at 0.17, 0.12 and 0.10, respectively. All but one of the heritable genera belonged to the Firmicutes and Actinobacteria phyla. The average relative abundance of the heritable genera ranged from 0.01 to 24.33%.

Table 2 shows estimates of genetic (r_g) and phenotypic (r_p) correlations and of the structural coefficients $\lambda_{PU,Genus}$, $\lambda_{FI,Genus}$, $\lambda_{BWG,Genus}$ and $\lambda_{F:G,Genus}$ between the microbial genera and each of the four traits, for all genera with a significant Pearson correlation ($p \leq 0.05$) and a significant structural coefficient ($p \leq 0.05$) between the genera and each of the four traits. Genetic correlations estimates between the considered genera and the four phenotypes had large standard errors due to the limited number of animals and ranged from -0.19 (*Enterococcus*) to 0.52 (*Lactococcus*) for PU, from -0.03 (*Microbacterium*) to 0.47 (*Leuconostoc*) for FI, from -0.69 (*Enterococcus*) to 0.48 (*Lactococcus*) for BWG, and from -0.38 (*Lactococcus*) to 0.54 (*Enterococcus*) for F:G. Ranges for estimates of phenotypic correlations were narrower, from -0.09 (*Streptococcus*) to 0.14 (*Bacillus*) for PU, from 0.17 (*Microbacterium*) to 0.31 (*Leuconostoc*) for FI, from -0.16 (*Macroccoccus*) to 0.25 (*Leuconostoc*) for BWG, and from -0.38 (*Lactococcus*) to 0.17 (*Enterococcus*) for F:G. Estimates of the standardized recursive effects, $\lambda_{i,Genus}$, ranged from -0.039 ($\lambda_{PU,Enterococcus}$) to 0.081 ($\lambda_{PU,Bacillus}$) for PU, from 0.028 ($\lambda_{FI,Microbacterium}$) to 0.095 ($\lambda_{FI,Bacillus}$) for FI, from -0.056 ($\lambda_{BWG,Streptococcus}$) to 0.102 ($\lambda_{BWG,Bacillus}$) for BWG, and from -0.064 ($\lambda_{F:G,Lactococcus}$) to 0.058 ($\lambda_{F:G,Enterococcus}/\lambda_{F:G,Staphylococcus}$) for F:G. The complete information for all heritable genera is in Additional file 1: Table S1, Additional file 2: Table S2, Additional file 3: Table S3, and Additional file 4: Table S4.

Table 1 Bacterial genera with significant heritability estimates ($p \leq 0.05$).

Phylum	Genus	Average relative abundance (%)	Heritability	SE	p value	False discovery rate
Actinobacteria	<i>Bifidobacterium</i>	0.48	0.10	0.05	< 0.001	< 0.001
Firmicutes	<i>Clostridium sensu stricto</i>	14.11	0.17	0.07	< 0.001	< 0.001
Firmicutes	<i>Lactobacillus</i>	24.33	0.12	0.05	< 0.001	< 0.001
Firmicutes	<i>Macrococcus</i>	0.23	0.06	0.03	< 0.001	0.002
Proteobacteria	<i>Escherichia/Shigella</i>	14.17	0.09	0.05	0.001	0.008
Actinobacteria	<i>Cutibacterium</i>	0.06	0.08	0.04	0.002	0.019
Firmicutes	<i>Aerococcus</i>	0.47	0.08	0.04	0.003	0.025
Firmicutes	<i>Bacillus</i>	0.08	0.06	0.03	0.006	0.042
Firmicutes	<i>Staphylococcus</i>	0.31	0.05	0.03	0.006	0.042
Firmicutes	<i>Tyzzereella</i>	0.08	0.07	0.04	0.007	0.042
Firmicutes	Unc. <i>Lachnospiraceae</i>	0.38	0.06	0.03	0.010	0.051
Firmicutes	<i>Enterococcus</i>	3.75	0.06	0.04	0.011	0.055
Actinobacteria	<i>Corynebacterium</i>	0.47	0.06	0.04	0.012	0.055
Actinobacteria	<i>Curtobacterium</i>	0.01	0.06	0.04	0.014	0.058
Firmicutes	<i>Streptococcus</i>	8.25	0.08	0.05	0.022	0.088
Actinobacteria	<i>Microbacterium</i>	0.02	0.05	0.03	0.026	0.094
Firmicutes	<i>Sellimonas</i>	0.03	0.05	0.03	0.028	0.094
Firmicutes	<i>Leuconostoc</i>	0.12	0.04	0.03	0.029	0.094
Firmicutes	<i>Ruminococcus 2</i>	0.28	0.05	0.03	0.030	0.094
Firmicutes	<i>Anaerofilum</i>	0.02	0.04	0.03	0.040	0.108
Firmicutes	<i>Anaerostipes</i>	0.06	0.04	0.03	0.040	0.108
Firmicutes	<i>Lactococcus</i>	0.14	0.04	0.03	0.040	0.108
Actinobacteria	<i>Corynebacterium</i>	0.15	0.05	0.04	0.043	0.109
Firmicutes	<i>Subdoligranulum</i>	0.07	0.05	0.03	0.049	0.121

SE standard errors

Table 2 Estimates of genetic (r_g) and phenotypic (r_p) correlations and of regression coefficients $\lambda_{i,Genus}$ between the considered genera and each of the four traits.

Trait	Genus ^a	r_g	SE	r_p	SE	$\lambda_{i,Genus}$ ^b	SE
P utilization	<i>Bacillus</i>	0.36	0.35	0.14	0.04	0.081	0.026
	<i>Enterococcus</i>	-0.19	0.38	-0.09	0.04	-0.039	0.016
	<i>Escherichia/Shigella</i>	-0.07	0.35	-0.07	0.04	-0.026	0.014
	<i>Lactococcus</i>	0.52	0.36	0.13	0.04	0.055	0.020
	<i>Leuconostoc</i>	0.48	0.35	0.13	0.04	0.070	0.024
	<i>Streptococcus</i>	-0.14	0.37	-0.09	0.04	-0.030	0.013
Feed intake	<i>Bacillus</i>	0.35	0.34	0.29	0.06	0.095	0.023
	<i>Curtobacterium</i>	0.17	0.36	0.20	0.05	0.029	0.009
	<i>Lactococcus</i>	0.37	0.37	0.27	0.06	0.057	0.018
	<i>Leuconostoc</i>	0.47	0.34	0.31	0.07	0.082	0.021
	<i>Microbacterium</i>	-0.03	0.40	0.17	0.05	0.028	0.011
Body weight gain	<i>Bacillus</i>	0.35	0.39	0.23	0.05	0.102	0.025
	<i>Curtobacterium</i>	0.41	0.37	0.19	0.04	0.035	0.010
	<i>Enterococcus</i>	-0.69	0.33	-0.15	0.04	-0.054	0.015
	<i>Lactococcus</i>	0.48	0.40	0.24	0.05	0.076	0.019
	<i>Leuconostoc</i>	0.47	0.39	0.25	0.05	0.097	0.023
	<i>Macrococcus</i>	0.15	0.37	-0.16	0.07	-0.024	0.012
	<i>Streptococcus</i>	0.16	0.42	-0.11	0.04	-0.056	0.013
Feed per gain	<i>Aerococcus</i>	0.25	0.41	0.12	0.04	0.038	0.011
	<i>Clostridium sensu stricto</i>	0.15	0.37	-0.11	0.04	-0.050	0.013
	<i>Cutibacterium</i>	-0.25	0.41	0.09	0.04	0.027	0.011
	<i>Enterococcus</i>	0.54	0.40	0.17	0.04	0.058	0.016
	<i>Escherichia/Shigella</i>	0.01	0.41	-0.09	0.04	-0.035	0.014
	<i>Lactococcus</i>	-0.38	0.43	-0.07	0.05	-0.064	0.020
	<i>Ruminococcus 2</i>	-0.33	0.47	0.11	0.04	0.039	0.015
	<i>Sellimonas</i>	-0.27	0.48	0.10	0.04	0.014	0.006
	<i>Staphylococcus</i>	-0.13	0.43	0.14	0.04	0.058	0.020
<i>Streptococcus</i>	0.32	0.43	0.12	0.04	0.038	0.013	

SE standard errors

^a All genera with a significant heritability ($p \leq 0.05$), significant Pearson correlation ($p \leq 0.05$) and significant structural coefficient ($p \leq 0.05$) between the genus and the considered trait

^b in units σ_p

QTL mapping results

The QTL linkage mapping results are presented as genome scan plots for each heritable microbial genus with significant QTL in Fig. 1. For clarity, only the first 23 *Coturnix japonica* chromosomes (CJA) are shown within the plots, since no significant peaks were observed for the other chromosomes. As described in Vollmar et al. [18], none of the genotyped SNPs were located on CJA16. Six QTL with genome-wide significance thresholds of 5 and 10% were found across all genera (Table 3). Significant peaks were detected for the microbial genera

Aerococcus on CJA3, for *Bacillus* on CJA2, for *Cutibacterium* on CJA2, for *Escherichia/Shigella* on CJA24, for *Ruminococcus 2* on CJA3, and for *Streptococcus* on CJA5.

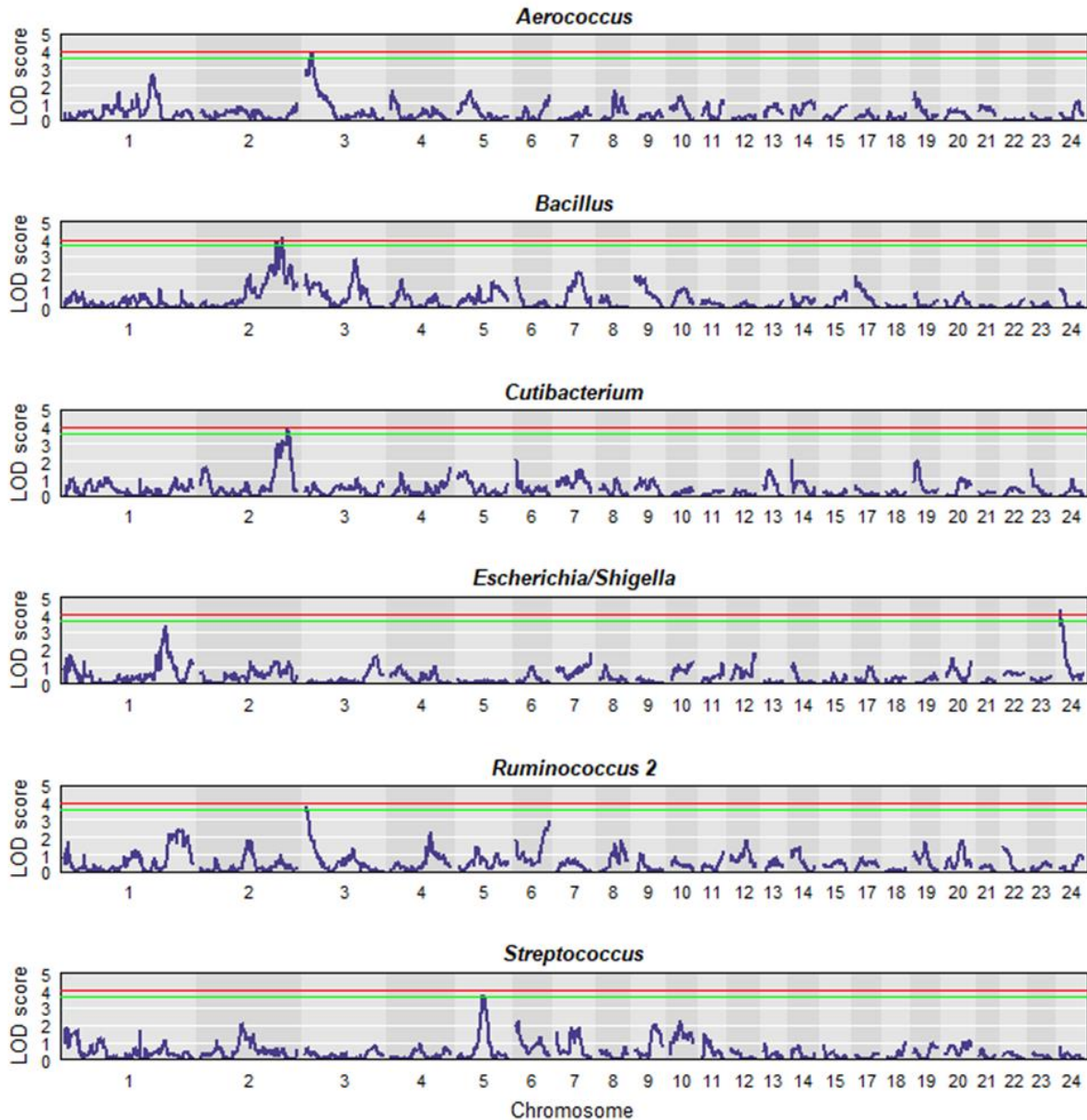


Figure 1 Plots of the QTL linkage mapping scan of heritable genera with significant QTL. QTL linkage mapping scan plots of heritable genera (p value ≤ 0.05) with significant QTL. The LOD score is the test statistic, and the red and green lines correspond to genome-wide significance levels of 5 and 10%, respectively.

Results of the SNP-based association analyses for SNPs within the SI regions of the six identified QTL are in Table 3. Significant SNPs were found for each SI region, for a total of 103 significant SNPs, as listed in Additional file 5: Table S5. Due to overlapping SI, significant SNPs were shared between *Bacillus* and *Cutibacterium* on CJA2 and between *Aerococcus* and *Ruminococcus 2* on CJA3 (see Additional file 5: Table S5).

Table 3 Results of the QTL linkage mapping.

Trait	CJA	Pos (cM)	LOD	Support interval borders (cM)		Number of significant SNPs
				Low	High	
<i>Aerococcus</i>	3	12	4.00**	0	19	16
<i>Bacillus</i>	2	160	4.14**	147	164	24
<i>Cutibacterium</i>	2	171	3.83 *	147	178	34
<i>Escherichia/Shigella</i>	24	0	4.31**	0	8	2
<i>Ruminococcus 2</i>	3	2	3.74 *	0	10	10
<i>Streptococcus</i>	5	50	3.78 *	44	57	17

Pos: positions in cM of 5% (**) and 10% (*) genome-wide significant QTL on *Coturnix japonica* chromosomes (CJA), with LOD score test statistics (LOD) and the corresponding QTL support intervals (SI), in cM. SI_low and SI_high represent the beginning and the end of the SI, respectively, and significant SNPs ($p \leq 0.05$) are obtained from the SNP-trait association analysis. The corresponding genetic linkage map can be found in Vollmar et al. [18]

Genomic and microbial trait predictions

Estimates of heritability for the animal microbiota effects for the four traits based on Model (6) were 0.07 (SE = 0.04, p value = 0.020) for PU, 0.14 (SE = 0.05, p value ≤ 0.001) for FI, 0.06 (SE = 0.04, p value = 0.020) for BWG, and 0.03 (SE = 0.03, p value = 0.267) for F:G. For all traits, except F:G, the estimate of heritability of animal microbiota effects was significant ($p \leq 0.05$). The results of the cross-validation of microbial and genomic predictions are in Table 4. Genomic predictions, \hat{g} , and genomic predictions of the microbiota-mediated part of the traits, \hat{m} , had similar correlations with the trait phenotypes. Average correlations between the microbial predictions, \hat{k} , and the trait phenotypes were slightly higher than GBLUP accuracies for PU and FI, and markedly higher for BWG and F:G.

Table 4 Mean accuracy and confidence interval (CI) of the genomic and microbial trait predictions.

Trait	MBLUP		GBLUP		Microbiota-mediated GBLUP	
	Accuracy	95% CI	Accuracy	95% CI	Accuracy	95% CI
PU	0.22	0.09:0.35	0.18	0.05:0.32	0.16	0.01:0.31
FI	0.31	0.17:0.43	0.24	0.10:0.35	0.22	0.07:0.38
BWG	0.34	0.20:0.46	0.13	-0.01:0.25	0.14	-0.03:0.29
F:G	0.31	0.10:0.47	0.10	-0.05:0.23	0.07	-0.07:0.23

Estimated accuracy of the MBLUP and GBLUP of the trait observations and GBLUP of the microbiota-mediated part of the trait observations

PU P utilization, FI feed intake, BWG body weight gain, F:G feed per gain

Discussion

In previous studies, we investigated the impact of host genetics [15, 18, 23] and of the composition of the ileum microbiota [16, 17] on PU and related traits in Japanese quail. To complement these studies, here we modeled the microbiota composition as a host trait and

investigated how the microbiota composition and the host genome can predict the four host phenotypic traits considered in this study.

It is well known that gut microbial colonization is determined by the environmental and genetic background of animals. External factors, such as diet, husbandry, photoperiod, and litter can overlay or mask the effects of host genetics [38–41]. To reduce external influences on gut microbiota and to ensure comparability of animals, standardized housing and management conditions were used for all animals in this study. The microbiota composition DNA samples used in this study originated from an experiment that took place several years ago [15], and at that time, the importance of having control samples of feed, water, litter, DNA extraction, etc. was underestimated.

A large fraction of the bacterial genera showed a significant estimate of heritability of their relative abundance (Table 1). The three genera with the highest heritability estimates were *Clostridium sensu stricto* (0.17), *Lactobacillus* (0.12), and *Bifidobacterium* (0.10). These heritability estimates are lower than those calculated by Camarinha-Silva et al. [10] and Estellé et al. [42] for ileal bacterial genera in pigs and by Org et al. [43] in mice, but a solid comparison across species is questionable. Mignon-Grasteau et al. [13] estimated moderate heritabilities for relative abundance of members of the genera *Lactobacillus* and *Clostridium* in the ceca of chickens.

To examine the relationship between the quantitative traits and composition of the ileal microbiota, we applied SEM, which separate the direct and indirect genetic relationships between the considered traits [44]. Direct genetic effects are due to pleiotropic QTL or linkage disequilibrium between QTL [44]. In the present study, indirect genetic effects were due to the recursive effect of bacterial genera on the host phenotype traits. The selected equation models were modeled in a simplified way and no competing models were compared, in part because the size of the dataset was too limited to do this in a thorough manner. It has to be acknowledged that the assumption that the microbiota affects host phenotype traits is questionable, especially for FI, and, therefore, these results have to be interpreted with some caution. Interactions between microbiota and quantitative traits by using structural equation modeling have rarely been studied, except by Saborío-Montero et al. [8] for methane emission in cattle and by Tiezzi et al. [20] for fat deposition in pigs. Nonzero estimates for $\lambda_{PU,Genus}$, $\lambda_{FI,Genus}$, $\lambda_{BWG,Genus}$, and $\lambda_{F:G,Genus}$ (Table 2) indicate that there is a functional link between the four traits and some microbial genera, and the small standard errors confirm the correct directional assumption of the recursive relationship between the bacterial genera and the performance traits.

The main genera that were found to be involved in the recursive relationships with the host traits were *Bacillus*, *Lactococcus* and *Leuconostoc*, albeit with weak signals. Negative recursive interactions with PU were shown for the three genera *Enterococcus*, *Escherichia/Shigella* and *Streptococcus*, and positive recursive interactions were shown between these three same genera and F:G (Table 2). A negative interaction between *Enterococcus* and BWG was also found and a joint positive recursive effect was identified for *Curtobacterium* with both FI and BWG. For F:G, the negative interaction with *Escherichia/Shigella* and *Lactococcus* is worth mentioning, as well as the positive interaction with *Enterococcus*, and *Streptococcus* (Table 2). Our results are in agreement with those of Vollmar et al. [17], who reported an association of *Bacillus* and *Leuconostoc* with PU using a MWAS, and also with those of Borda-Molina et al. [16], who confirmed a positive phenotypic association of the relative abundance of *Bacillus* and *Leuconostoc* in the ileum with PU. Lactic acid bacteria are known to be phytase degraders and some species of the genus *Bacillus* showed extracellular phytase activity that might improve PU efficiency [45, 46].

The phenotypic correlations between the bacterial genera and each of the four performance traits were within a low to medium range (Table 2). Because of the limited number of animals in our study, estimates of genetic correlations (Table 2) had large standard errors and they should be viewed as a supplement to the main results of our study.

To date, only a few QTL studies have been conducted in quail [18, 47–49], with QTL mapped for different behavioral and performance traits. However, several authors have investigated host QTL for microbial colonization of the GIT in other species. For instance, Mignon-Grasteau et al. [13] performed QTL analyses of microbial genera in the ceca of chickens. In our study, six significant host QTL for microbial composition in quail were detected (Fig. 1). One of these significant QTL was for *Bacillus*, on CJA2. Relative abundance of *Bacillus* was most highly correlated and showed the highest recursive relationships with PU, FI, and BWG. Interestingly, the SI of a previously identified QTL for growth rate on this chromosome [48] overlaps with the SI of the QTL on CJA2 for *Bacillus*. Similarly, Essa et al. [50] identified a QTL for BWG on chicken chromosome 2. The SI for the QTL on CJA2 for *Bacillus* also overlapped with that of the QTL for *Cutibacterium* (Table 3) and several common significant SNPs were detected (see Additional file 5: Table S5). The SI of the QTL identified for *Aerococcus* on CJA3 overlapped with that of *Ruminococcus 2*, and several common SNPs were also detected for this overlapping region. On CJA5, the SI of a QTL for *Streptococcus* overlapped with the SI for QTL for FI and foot ash reported in Vollmar et al. [18].

All microbial genera for which significant QTL were mapped were also significantly correlated or significantly linked via recursive interaction with at least one of the host traits, as described

above. Thus, it can be assumed that the QTL detected on CJA2 for *Bacillus* directly influences the microbial colonization of the ileum with *Bacillus* and indirectly influences PU, FI, and BWG. The use of SEM for QTL mapping may also be of interest for future studies [51].

Several studies have revealed possible mechanisms for the positive or negative effects of bacterial genera on quantitative traits. Among the bacterial genera with the highest effects on the four traits analyzed here, the genus *Bacillus* was most important. It is used as a probiotic in chickens and improves performance traits [52, 53], positively affects the immune system [54–56], increases digestive enzyme activity [55–57], and synthesizes phytases [58]. A previous study has reported that some strains of *Leuconostoc* have weak enzymatic activities, including the formation of acid phosphatase [59]. In addition, an immunomodulatory activity due to induced cytokine production has been reported in chicken [60]. Some subspecies of *Lactococcus* supplied in the diet of broilers have resulted in a lower F:G, increased body weight, and reduced mortality, with positive effects on the immune system [61, 62] and carcass quality [63]. As noted above, relative abundance of the genus *Enterococcus* has a negative impact on PU, FI and BWG. In humans, some members of this genus are considered to be opportunistic pathogens due to their antibiotic resistance [64, 65]. In chickens, these bacteria can lead to increased one-day mortality [66], and the formation of toxic metabolites by bacterial metabolization of protein has also been reported [67]. Relative abundance of the genera *Staphylococcus* and *Streptococcus* also negatively influenced the traits analyzed here, and have been shown to cause different diseases and affect poultry health with, depending on the bacterial species, several clinical observations that range from drowsiness and poor feed intake to increased mortality (reviewed in [68]).

Results from the microbial and genomic predictions (Table 4) confirmed the strong impact of the host genome and microbiota composition on the analyzed host traits. The two-step procedure proposed by Weishaar et al. [21] to estimate breeding values for the microbiota-mediated part of a trait was also successful in our study, in particular for PU and FI. The GBLUP accuracy for the microbiota-mediated part of the host phenotype was only slightly lower than the prediction accuracy of the conventional GBLUP. One explanation for this results might be that the genetic effect of the host on the trait mediated by the microbiota is much stronger than the direct genetic effect of the host. To substantiate this hypothesis, we fitted Model (1) with an additional random animal effect with the microbiota-based covariance matrix \mathbf{M} for PU. Compared to a model with only the microbiota effect, the estimate of microbiability for PU remained at almost the same level (0.15) but the estimate of heritability dropped from 0.12 to 0.07 (results not shown). A similar pattern was observed by Difford et al. [9] in a study on dairy cattle rumen microbiota composition and methane production. This clearly shows that fitting both random effects simultaneous is beneficial but that assuming a zero covariance between

the two random effects is too simplistic. How to model both effects simultaneously and how to interpret the results from such models biologically is an ongoing research topic [69–71] but this is outside the scope of our study.

Conclusions

We detected a significant genetic effect of the host on composition of the ileum microbiota in quail. Among the 59 bacterial genera, 24 showed a significant heritability of their relative abundance. The estimated correlations of the bacterial genera with the four host traits analyzed (PU, FI, BWG, and F:G) and the calculated recursive effects from the SEM confirmed the recursive relationship between the relative abundance of individual bacterial genera and these traits. Several significant QTL were identified for microbiota composition, which were supported by trait-associated SNPs (associated with PU, FI, BWG, or F:G). The application of microbial and genomic mixed linear models allowed accurate prediction of PU and the related traits. In particular, applying these models made it possible to predict the microbiota-mediated part of the traits, demonstrating the feasibility of hologenomic selection.

Declarations

Ethics approval and consent to participate

This animal experiment was accomplished according to the requirements of the German Animal Welfare Legislation and was approved by the Animal Welfare Commissioner of the University of Hohenheim (approval number S371/13TE).

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors thank two anonymous reviewers for critical and helpful comments on an earlier version of the manuscript.

Authors' contributions

VH and SV performed the statistical analyses. ACS generated the microbiota dataset and SP the SNP genotypes. VH and JB wrote the paper. JB, MR, and ACS initiated and supervised the study. All authors read and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, BE3703/12-1 and CA1708/2-1) and was part of the research unit P-FOWL (FOR 2601).

Availability of data and materials

Not applicable.

References

1. Pavlov E, Aschar-Sobbi R, Campanella M, Turner RJ, Gómez-García MR, Abramov AY. Inorganic polyphosphate and energy metabolism in mammalian cells. *J Biol Chem.* 2010;285:9420–8.
2. Campbell BM, Beare DJ, Bennett EM, Hall-Spencer JM, Ingram JSI, Jaramillo F, et al. Agriculture production as a major driver of the earth system exceeding planetary boundaries. *Ecol Soc.* 2017;22:8.
3. Cordell D, Drangert J-O, White S. The story of phosphorus: global food security and food for thought. *Glob Environ Change.* 2009;19:292–305.
4. Rodehutschord M. Interactions between minerals and phytate degradation in poultry—challenges for phosphorus digestibility assays. In: Walk CL, Kühn I, Stein HH, Kidd MT, Rodehutschord M, editors. *Phytate destruction: consequences for precision animal nutrition.* Wageningen: Wageningen Academic Publisher; 2016. p. 167–78.
5. Zeller E, Schollenberger M, Witzig M, Shastak Y, Kühn I, Hoelzle LE, et al. Interactions between supplemented mineral phosphorus and phytase on phytate hydrolysis and inositol phosphates in the small intestine of broilers. *Poult Sci.* 2015;94:1018–29.
6. Sommerfeld V, Schollenberger M, Kühn I, Rodehutschord M. Interactive effects of phosphorus, calcium, and phytase supplements on products of phytate degradation in the digestive tract of broiler chickens. *Poult Sci.* 2018;97:1177–88.
7. Rodehutschord M. Advances in understanding the role of phytate in phosphorus and calcium nutrition of poultry. In: Applegate T, editor. *Achieving sustainable production of poultry meat, Volume 2: Breeding and nutrition.* Cambridge: Burleigh Dodds Science Publishing; 2017. p. 165–80.
8. Saborío-Montero A, Gutiérrez-Rivas M, García-Rodríguez A, Atxaerandio R, Goiri I, López de Maturana E, et al. Structural equation models to disentangle the biological relationship between microbiota and complex traits: methane production in dairy cattle as a case of study. *J Anim Breed Genet.* 2020;137:36–48.
9. Difford GF, Plichta DR, Løvendahl P, Lassen J, Noel SJ, Højberg O, et al. Host genetics and the rumen microbiome jointly associate with methane emissions in dairy cows. *PLoS Genet.* 2018;14:e1007580.
10. Camarinha-Silva A, Maushammer M, Wellmann R, Vital M, Preuss S, Bennewitz J. Host genome influence on gut microbial composition and microbial prediction of complex traits in pigs. *Genetics.* 2017;206:1637–44.

11. Bergamaschi M, Maltecca C, Schillebeeckx C, McNulty NP, Schwab C, Shull C, et al. Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. *Sci Rep.* 2020;10:10134.
12. Meng H, Zhang Y, Zhao L, Zhao W, He C, Honaker CF, et al. Body weight selection affects quantitative genetic correlated responses in gut microbiota. *PLoS One.* 2014;9:e89862.
13. Mignon-Grasteau S, Narcy A, Rideau N, Chantry-Darmon C, Boscher M-Y, Sellier N, et al. Impact of selection for digestive efficiency on microbiota composition in the chicken. *PLoS One.* 2015;10:e0135488.
14. Zhao L, Wang G, Siegel P, He C, Wang H, Zhao W, et al. Quantitative genetic background of the host influences gut microbiomes in chickens. *Sci Rep.* 2013;3:1163.
15. Beck P, Piepho H-P, Rodehutschord M, Bennewitz J. Inferring relationships between phosphorus utilization, feed per gain, and bodyweight gain in an F2 cross of Japanese quail using recursive models. *Poult Sci.* 2016;95:764–73.
16. Borda-Molina D, Roth C, Hernández-Arriaga A, Rissi D, Vollmar S, Rodehutschord M, et al. Effects on the ileal microbiota of phosphorus and calcium utilization, bird performance, and gender in Japanese quail. *Animals (Basel).* 2020;10:885.
17. Vollmar S, Wellmann R, Borda-Molina D, Rodehutschord M, Camarinha-Silva A, Bennewitz J. The gut microbial architecture of efficiency traits in the domestic poultry model species Japanese quail (*Coturnix japonica*) assessed by mixed linear models. *G3 (Bethesda).* 2020;10:2553–62.
18. Vollmar S, Haas V, Schmid M, Preuß S, Joshi R, Rodehutschord M, et al. Mapping genes for phosphorus utilization and correlated traits using a 4k SNP linkage map in Japanese quail (*Coturnix japonica*). *Anim Genet.* 2021;52:90-8.
19. Gianola D, Sorensen D. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics.* 2004;167:1407–24.
20. Tiezzi F, Fix J, Schwab C, Shull C, Maltecca C. Gut microbiome mediates host genomic effects on phenotypes: a case study with fat deposition in pigs. *Comput Struct Biotechnol J.* 2020;19:530–44.
21. Weishaar R, Wellmann R, Camarinha-Silva A, Rodehutschord M, Bennewitz J. Selecting the hologenome to breed for an improved feed efficiency in pigs - a novel selection index. *J Anim Breed Genet.* 2020;137:14–22.

22. Mills AD, Faure JM. Divergent selection for duration of tonic immobility and social reinstatement behavior in Japanese quail (*Coturnix coturnix japonica*) chicks. *J Comp Psychol.* 1991;105:25–38.
23. Künzel S, Bennewitz J, Rodehutschord M. Genetic parameters for bone ash and phosphorus utilization in an F2 cross of Japanese quail. *Poult Sci.* 2019;98:4369–72.
24. Cole JR, Wang Q, Fish JA, Chai B, McGarrell DM, Sun Y, et al. Ribosomal database project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* 2014;42:D633-42.
25. Yarza P, Yilmaz P, Pruesse E, Glöckner FO, Ludwig W, Schleifer K-H, et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat Rev Microbiol.* 2014;12:635–45.
26. Box GEP, Cox DR. An analysis of transformations. *J R Stat Soc B Stat Met.* 1964;26:211–52.
27. R Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2019.
28. Butler DG, Cullis BR, Gilmour AR, Gogel BJ. ASReml-R reference manual: mixed models for S language environments. Brisbane: Queensland Government, Department of Primary Industries and Fisheries; 2009.
29. Gilmour AR, Gogel BJ, Cullis BR, Welham SJ, Thompson R. ASReml user guide release 4.1 functional specification. Hemel Hempstead: VSN International Ltd; 2015.
30. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc B Stat Met.* 1995;57:289–300.
31. Rosa GJM, Valente BD, Los Campos G de, Wu X-L, Gianola D, Silva MA. Inferring causal phenotype networks using structural equation models. *Genet Sel Evol.* 2011;43:6.
32. Valente BD, Rosa GJM, Los Campos G de, Gianola D, Silva MA. Searching for recursive causal structures in multivariate quantitative genetics mixed models. *Genetics.* 2010;185:633–44.
33. Broman KW, Gatti DM, Simecek P, Furlotte NA, Prins P, Sen Ś, et al. R/qtl2: software for mapping quantitative trait loci with high-dimensional data and multiparent populations. *Genetics.* 2019;211:495–502.

34. Nelson RM, Nettelblad C, Pettersson ME, Shen X, Crooks L, Besnier F, et al. MAPfastR: quantitative trait loci mapping in outbred line crosses. *G3 (Bethesda)*. 2013;3:2147–9.
35. Manichaikul A, Dupuis J, Sen S, Broman KW. Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics*. 2006;174:481–9.
36. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.
37. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci*. 2008;91:4414–23.
38. Hieke A-SC, Hubert SM, Athrey G. Circadian disruption and divergent microbiota acquisition under extended photoperiod regimens in chicken. *PeerJ*. 2019;7:e6592.
39. Hubert SM, Al-Ajeeli M, Bailey CA, Athrey G. The role of housing environment and dietary protein source on the gut microbiota of chicken. *Animals (Basel)*. 2019;9:1085.
40. Shang QH, Liu SJ, He TF, Liu HS, Mahfuz S, Ma XK, et al. Effects of wheat bran in comparison to antibiotics on growth performance, intestinal immunity, barrier function, and microbial composition in broiler chickens. *Poult Sci*. 2020;99:4929–38.
41. Wang L, Lilburn M, Yu Z. Intestinal microbiota of broiler chickens as affected by litter management regimens. *Front Microbiol*. 2016;7:593.
42. Estellé J, Mach N, Ramayo-Caldas Y, Levenez F, Lemonnier G, Denis C, et al. The influence of host's genetics on the gut microbiota composition in pigs and its links with immunity traits. In *Proceedings of the 10th World Congress of Genetics Applied to Livestock Production: 17-22 August 2014; Vancouver; 2014*.
43. Org E, Parks BW, Joo JWJ, Emert B, Schwartzman W, Kang EY, et al. Genetic and environmental control of host-gut microbiota interactions. *Genome Res*. 2015;25:1558–69.
44. Valente BD, Rosa GJM, Gianola D, Wu X-L, Weigel K. Is structural equation modeling advantageous for the genetic improvement of multiple traits? *Genetics*. 2013;194:561–72.
45. Künzel S, Borda-Molina D, Zuber T, Hartung J, Siegert W, Feuerstein D, et al. Relative phytase efficacy values as affected by response traits, including ileal microbiota composition. *Poult Sci*. 2021;100:101133.

46. Fu S, Sun J, Qian L, Li Z. Bacillus phytases: present scenario and future perspectives. *Appl Biochem Biotech.* 2008;151:1–8.
47. Recoquillay J, Pitel F, Arnould C, Leroux S, Dehais P, Moréno C, et al. A medium density genetic map and QTL for behavioral and production traits in Japanese quail. *BMC Genomics.* 2015;16:10.
48. Knaga S, Siwek M, Tavaniello S, Maiorano G, Witkowski A, Jezewska-Witkowska G, et al. Identification of quantitative trait loci affecting production and biochemical traits in a unique Japanese quail resource population. *Poult Sci.* 2018;97:2267–77.
49. Ori RJ, Esmailzadeh AK, Charati H, Mohammadabadi MR, Sohrabi SS. Identification of QTL for live weight and growth rate using DNA markers on chromosome 3 in an F2 population of Japanese quail. *Mol Biol Rep.* 2014;41:1049–57.
50. Essa BH, Suzuki S, Nagano AJ, Elkholya SZ, Ishikawa A. QTL analysis for early growth in an intercross between native Japanese Nagoya and White Plymouth Rock chicken breeds using RAD sequencing-based SNP markers. *Anim Genet.* 2021;52:232–6.
51. Pegolo S, Momen M, Morota G, Rosa GJM, Gianola D, Bittante G, et al. Structural equation modeling for investigating multi-trait genetic architecture of udder health in dairy cattle. *Sci Rep.* 2020;10:7751.
52. Li C-L, Wang J, Zhang H-J, Wu S-G, Hui Q-R, Yang C-B, et al. Intestinal morphologic and microbiota responses to dietary Bacillus spp. in a broiler chicken model. *Front Physiol.* 2019;9:1968.
53. Abdel-Moneim A-ME, Selim DA, Basuony HA, Sabic EM, Saleh AA, Ebeid TA. Effect of dietary supplementation of Bacillus subtilis spores on growth performance, oxidative status, and digestive enzyme activities in Japanese quail birds. *Trop Anim Health Prod.* 2020;52:671–80.
54. Bai K, Huang Q, Zhang J, He J, Zhang L, Wang T. Supplemental effects of probiotic Bacillus subtilis fmbJ on growth performance, antioxidant capacity, and meat quality of broiler chickens. *Poult Sci.* 2017;96:74–82.
55. Li Gong, Wang B, Mei X, Xu H, Qin Y, Li W, et al. Effects of three probiotic Bacillus on growth performance, digestive enzyme activities, antioxidative capacity, serum immunity, and biochemical parameters in broilers. *Anim Sci J.* 2018;89:1561–71.
56. Fazelnia K, Fakhraei J, Yarahmadi HM, Amini K. Dietary supplementation of potential probiotics *Bacillus subtilis*, *Bacillus licheniformis*, and *Saccharomyces cerevisiae* and synbiotic improves growth performance and immune responses by modulation in

- intestinal system in broiler chicks challenged with *Salmonella Typhimurium*. *Probio Antimicrob Proteins*. 2021;13:1081-92.
57. Wang Y, Heng C, Zhou X, Cao G, Jiang L, Wang J, et al. Supplemental *Bacillus subtilis* DSM 29784 and enzymes, alone or in combination, as alternatives for antibiotics to improve growth performance, digestive enzyme activity, anti-oxidative status, immune response and the intestinal barrier of broiler chickens. *Brit J Nutr*. 2021;125:494-507.
 58. Latorre JD, Hernandez-Velasco X, Wolfenden RE, Vicente JL, Wolfenden AD, Menconi A, et al. Evaluation and selection of *Bacillus* species based on enzyme production, antimicrobial activity, and biofilm synthesis as direct-fed microbial candidates for poultry. *Front Vet Sci*. 2016;3:95.
 59. Paula AT de, Jeronymo-Geneviva AB, Silva LF, Todorov SD, Franco BDGM, Penna ALB. *Leuconostoc mesenteroides* SJRP55: a potential probiotic strain isolated from Brazilian water buffalo mozzarella cheese. *Ann Microbiol*. 2015;65:899–910.
 60. Seo BJ, Rather IA, Kumar VJR, Choi UH, Moon MR, Lim JH, et al. Evaluation of *Leuconostoc mesenteroides* YML003 as a probiotic against low-pathogenic avian influenza (H9N2) virus in chickens. *J Appl Microbiol*. 2012;113:163–71.
 61. Zhang T, Xie J, Zhang M, Fu N, Zhang Y. Effect of a potential probiotics *Lactococcus garvieae* B301 on the growth performance, immune parameters and caecum microflora of broiler chickens. *J Anim Physiol Anim Nutr*. 2016;100:413–21.
 62. Fajardo P, Pastrana L, Méndez J, Rodríguez I, Fuciños C, Guerra NP. Effects of feeding of two potentially probiotic preparations from lactic acid bacteria on the performance and faecal microflora of broiler chickens. *Sci World J*. 2012;2012:562635.
 63. Mujnisa A, Gustina L, Natsir A, Hasan S. Dosage effects of *Lactococcus lactis* ssp. *lactis* 2 as a probiotic on the percentage of carcass, abdominal fat content and cholesterol level in broilers. *Int J Poult Sci*. 2018;17:100–5.
 64. Hollenbeck BL, Rice LB. Intrinsic and acquired resistance mechanisms in enterococcus. *Virulence*. 2012;3:421–33.
 65. Maasjost J, Mühldorfer K, Cortez de Jäckel S, Hafez HM. Antimicrobial susceptibility patterns of *Enterococcus faecalis* and *Enterococcus faecium* isolated from poultry flocks in Germany. *Avian Dis*. 2015;59:143–8.
 66. Gregersen RH, Petersen A, Christensen H, Bisgaard M. Multilocus sequence typing of *Enterococcus faecalis* isolates demonstrating different lesion types in broiler breeders. *Avian Pathol*. 2010;39:435–40.

CHAPTER TWO

67. King MD, Guentzel MN, Arulanandam BP, Lupiani B, Chambers JP. Proteolytic bacteria in the lower digestive tract of poultry may affect avian influenza virus pathogenicity. *Poult Sci.* 2009;88:1388–93.
68. Logue CM, Andreasen CB, Borst LB, Eriksson H, Hampson DJ, Sanchez S, et al. Other bacterial diseases. In: Swayne DE, Boulianne M, Logue CM, McDougald LR, Nair V, Suarez DL, de Wit S, Grimes T, Johnson D, Kromm M, Prajitno TY, Rubinoff I, Zavala G, editors. *Diseases of poultry*. 14th edition. Hoboken: Wiley Blackwell; 2020. pp. 995–1085.
69. Christensen OF, Börner V, Varona L, Legarra A. Genetic evaluation including intermediate omics features. *Genetics.* 2021;219:iyab130
70. Saborío-Montero A, Gutiérrez-Rivas M, López-García A, García-Rodríguez A, Atxaerandio R, Goiri I, et al. Holobiont effect accounts for more methane emission variance than the additive and microbiome effects on dairy cattle. *Livest Sci.* 2021;250:104538.
71. Pérez-Enciso M, Zingaretti LM, Ramayo-Caldas Y, Los Campos G de. Opportunities and limits of combining microbiome and genome data for complex trait prediction. *Genet Sel Evol.* 2021;53:65.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12711-022-00697-8>.

Additional file 1: Table S1. Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{PU,Genus}$. Correlations and regression coefficients between PU and Genus with significant heritability ($p \leq 0.05$) The standard errors (SE) presented in parentheses and $\lambda_{PU,Genus}$ in units σ_p . ¹P utilization – Genus with significant heritability ($p \leq 0.05$).

Additional file 2: Table S2. Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{FI,Genus}$. Correlations and regression coefficients between FI and Genus with significant heritability ($p \leq 0.05$) The standard errors (SE) presented in parentheses and $\lambda_{FI,Genus}$ in units σ_p . ¹Feed intake – Genus with significant heritability ($p \leq 0.05$).

Additional file 3: Table S3. Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{BWG,Genus}$. Correlations and regression coefficients between BWG and Genus with significant heritability ($p \leq 0.05$) The standard errors (SE) presented in parentheses and $\lambda_{BWG,Genus}$ in units σ_p . ¹Body weight gain – Genus with significant heritability ($p \leq 0.05$).

Additional file 4: Table S4. Correlations and regression coefficients between F:G and Genus with significant heritability ($p \leq 0.05$) The standard errors (SE) presented in parentheses and $\lambda_{F:G,Genus}$ in units σ_p . ¹Feed per gain – Genus with significant heritability ($p \leq 0.05$).

Additional file 5: Table S5. Trait-associated markers from GCTA within the significant QTL regions. Summary of the trait-associated markers from GCTA ($p \leq 0.05$) within the significant QTL regions. In addition, all markers that are significantly associated with another characteristic are listed in the last column. ¹in cM.

Additional file 1: Table S1 Correlations and regression coefficients between PU and Genera with significant heritability.

Traits ¹	r_g	(SE)	r_p	(SE)	$\lambda_{PU,Genus}$	(SE)
PU - <i>Aerococcus</i>	-0.628	(0.307)	-0.021	(0.039)	0.019	(0.011)
PU - <i>Anaerostipes</i>	0.312	(0.397)	0.075	(0.040)	0.008	(0.005)
PU - <i>Bacillus</i>	0.357	(0.349)	0.141	(0.040)	0.081	(0.026)
PU - <i>Bifidobacterium</i>	0.191	(0.320)	-0.002	(0.042)	-0.003	(0.017)
PU - <i>Clostridium sensu stricto</i>	-0.283	(0.315)	-0.025	(0.040)	0.009	(0.013)
PU - <i>Corynebacterium</i>	-0.660	(0.389)	0.003	(0.039)	0.021	(0.010)
PU - <i>Corynebacterium</i>	-0.490	(0.357)	-0.012	(0.041)	0.023	(0.014)
PU - <i>Curtobacterium</i>	0.188	(0.382)	0.074	(0.038)	0.017	(0.010)
PU - <i>Cutibacterium</i>	0.222	(0.356)	-0.001	(0.040)	-0.006	(0.011)
PU - <i>Enterococcus</i>	-0.192	(0.380)	-0.085	(0.038)	-0.039	(0.016)
PU - <i>Escherichia/Shigella</i>	-0.065	(0.347)	-0.065	(0.040)	-0.026	(0.014)
PU - <i>Lactobacillus</i>	0.112	(0.329)	-0.038	(0.047)	-0.009	(0.010)
PU - <i>Lactococcus</i>	0.523	(0.355)	0.129	(0.041)	0.055	(0.020)
PU - <i>Leuconostoc</i>	0.480	(0.349)	0.134	(0.042)	0.070	(0.024)
PU - <i>Macrococcus</i>	-0.130	(0.336)	-0.006	(0.053)	0.021	(0.012)
PU - <i>Microbacterium</i>	-0.022	(0.401)	0.042	(0.039)	0.018	(0.012)
PU - <i>Ruminococcus 2</i>	0.462	(0.359)	0.027	(0.040)	0.007	(0.015)
PU - <i>Sellimonas</i>	0.416	(0.368)	0.056	(0.039)	0.005	(0.006)
PU - <i>Staphylococcus</i>	0.839	(0.264)	0.019	(0.045)	-0.029	(0.020)
PU - <i>Streptococcus</i>	-0.143	(0.366)	-0.086	(0.038)	-0.030	(0.013)
PU - <i>Subdoligranulum</i>	0.534	(0.354)	0.052	(0.039)	0.005	(0.006)
PU - <i>Tyzzera</i>	-0.384	(0.378)	0.018	(0.040)	< 0.001	(< 0.001)
PU - Unc. <i>Lachnospiraceae</i>	0.292	(0.364)	0.053	(0.040)	0.022	(0.014)

Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{PU,Genus}$ between PU and Genera with significant heritability ($p \leq 0.05$). The standard errors (SE) presented in parentheses and $\lambda_{PU,Genus}$ in units σ_p . ¹ P utilization – Genus with significant heritability ($p \leq 0.05$).

Additional file 2: Table S2 Correlations and regression coefficients between FI and Genera with significant heritability.

Traits ¹	r_g (SE)	r_p (SE)	$\lambda_{FI,Genus}$ (SE)
FI - <i>Aerococcus</i>	-0.210 (0.358)	-0.073 (0.048)	0.007 (0.010)
FI - <i>Anaerostipes</i>	0.414 (0.394)	0.125 (0.059)	0.001 (0.004)
FI - <i>Bacillus</i>	0.347 (0.340)	0.290 (0.059)	0.095 (0.023)
FI - <i>Bifidobacterium</i>	0.093 (0.323)	0.138 (0.061)	0.021 (0.015)
FI - <i>Clostridium sensu stricto</i>	-0.205 (0.318)	-0.017 (0.047)	0.028 (0.012)
FI - <i>Corynebacterium</i>	-0.430 (0.387)	-0.052 (0.054)	0.022 (0.009)
FI - <i>Corynebacterium</i>	-0.258 (0.367)	-0.100 (0.058)	0.013 (0.012)
FI - <i>Curtobacterium</i>	0.168 (0.360)	0.195 (0.047)	0.029 (0.009)
FI - <i>Cutibacterium</i>	0.053 (0.350)	0.122 (0.056)	0.015 (0.010)
FI - <i>Enterococcus</i>	-0.576 (0.310)	-0.077 (0.041)	-0.022 (0.014)
FI - <i>Escherichia/Shigella</i>	-0.380 (0.334)	-0.061 (0.053)	0.006 (0.012)
FI - <i>Lactobacillus</i>	-0.097 (0.314)	-0.202 (0.084)	0.001 (0.010)
FI - <i>Lactococcus</i>	0.366 (0.365)	0.270 (0.064)	0.057 (0.018)
FI - <i>Leuconostoc</i>	0.466 (0.338)	0.307 (0.065)	0.082 (0.021)
FI - <i>Macrococcus</i>	0.291 (0.312)	-0.226 (0.103)	0.003 (0.012)
FI - <i>Microbacterium</i>	-0.026 (0.395)	0.165 (0.051)	0.028 (0.011)
FI - <i>Ruminococcus 2</i>	0.425 (0.396)	0.084 (0.054)	-0.002 (0.013)
FI - <i>Sellimonas</i>	0.340 (0.395)	0.119 (0.052)	0.002 (0.005)
FI - <i>Staphylococcus</i>	0.563 (0.327)	0.168 (0.080)	-0.020 (0.019)
FI - <i>Streptococcus</i>	0.116 (0.372)	-0.054 (0.043)	-0.038 (0.012)
FI - <i>Subdoligranulum</i>	0.369 (0.399)	0.074 (0.046)	0.004 (0.005)
FI - <i>Tyzzarella</i>	-0.264 (0.376)	0.083 (0.054)	< 0.001 (< 0.001)
FI - Unc. <i>Lachnospiraceae</i>	0.334 (0.368)	0.112 (0.054)	0.008 (0.013)

Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{FI,Genus}$ between FI and Genera with significant heritability ($p \leq 0.05$). The standard errors (SE) presented in parentheses and $\lambda_{FI,Genus}$ in units σ_p . ¹ Feed intake – Genus with significant heritability ($p \leq 0.05$).

Additional file 3: Table S3 Correlations and regression coefficients between BWG and Genera with significant heritability.

Traits ¹	r_g	(SE)	r_p	(SE)	$\lambda_{BWG,Genus}$	(SE)
BWG - <i>Anaerofilum</i>	0.202	(0.464)	-0.044	(0.040)	-0.009	(0.001)
BWG - <i>Anaerostipes</i>	0.319	(0.472)	0.057	(0.047)	< 0.001	(0.005)
BWG - <i>Bacillus</i>	0.351	(0.392)	0.232	(0.048)	0.102	(0.025)
BWG - <i>Bifidobacterium</i>	0.116	(0.364)	0.056	(0.050)	-0.001	(0.017)
BWG - <i>Clostridium sensu stricto</i>	-0.385	(0.355)	0.044	(0.042)	0.051	(0.013)
BWG - <i>Corynebacterium</i>	-0.444	(0.435)	-0.048	(0.045)	0.010	(0.010)
BWG - <i>Corynebacterium</i>	-0.160	(0.415)	-0.110	(0.047)	-0.013	(0.013)
BWG - <i>Curtobacterium</i>	0.413	(0.374)	0.185	(0.041)	0.035	(0.010)
BWG - <i>Cutibacterium</i>	0.143	(0.394)	0.043	(0.047)	-0.003	(0.011)
BWG - <i>Enterococcus</i>	-0.685	(0.332)	-0.149	(0.038)	-0.054	(0.015)
BWG - <i>Escherichia/Shigella</i>	-0.583	(0.379)	0.007	(0.045)	0.033	(0.013)
BWG - <i>Lactobacillus</i>	0.213	(0.355)	-0.092	(0.064)	-0.002	(0.010)
BWG - <i>Lactococcus</i>	0.479	(0.399)	0.236	(0.050)	0.076	(0.019)
BWG - <i>Leuconostoc</i>	0.471	(0.392)	0.248	(0.051)	0.097	(0.023)
BWG - <i>Macrococcus</i>	0.150	(0.371)	-0.161	(0.072)	-0.024	(0.012)
BWG - <i>Microbacterium</i>	0.292	(0.427)	0.115	(0.044)	0.015	(0.012)
BWG - <i>Ruminococcus 2</i>	0.455	(0.460)	-0.001	(0.045)	-0.021	(0.014)
BWG - <i>Sellimonas</i>	0.342	(0.454)	0.033	(0.045)	-0.005	(0.006)
BWG - <i>Staphylococcus</i>	0.458	(0.395)	0.045	(0.062)	-0.042	(0.020)
BWG - <i>Streptococcus</i>	0.157	(0.416)	-0.109	(0.039)	-0.056	(0.013)
BWG - <i>Subdoligranulum</i>	0.449	(0.471)	0.003	(0.041)	-0.005	(0.005)
BWG - <i>Tyzzarella</i>	-0.312	(0.414)	0.071	(0.045)	< 0.001	(< 0.001)
BWG - <i>Unc. Lachnospiraceae</i>	0.358	(0.426)	0.020	(0.045)	-0.014	(0.014)

Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{BWG,Genus}$ between BWG and Genera with significant heritability ($p \leq 0.05$). The standard errors (SE) presented in parentheses and $\lambda_{BWG,Genus}$ in units σ_p . ¹ Body weight gain – Genus with significant heritability ($p \leq 0.05$).

Additional file 4: Table S4 Correlations and regression coefficients between F:G and Genera with significant heritability.

Traits ¹	r_g	SE	r_p	SE	$\lambda_{F:G,Genus}$	SE
F:G - <i>Aerococcus</i>	0.247	(0.408)	0.120	(0.038)	0.038	(0.011)
F:G - <i>Anaerofilum</i>	-0.410	(0.453)	0.043	(0.038)	0.002	(0.001)
F:G - <i>Bacillus</i>	-0.224	(0.424)	-0.052	(0.043)	-0.068	(0.026)
F:G - <i>Bifidobacterium</i>	-0.084	(0.379)	0.081	(0.041)	0.030	(0.017)
F:G - <i>Clostridium sensu stricto</i>	0.151	(0.368)	-0.107	(0.038)	-0.050	(0.013)
F:G - <i>Corynebacterium</i>	0.300	(0.475)	0.024	(0.039)	0.008	(0.010)
F:G - <i>Corynebacterium</i>	-0.018	(0.445)	0.053	(0.041)	0.033	(0.014)
F:G - <i>Curtobacterium</i>	-0.578	(0.370)	-0.065	(0.039)	-0.021	(0.010)
F:G - <i>Cutibacterium</i>	-0.254	(0.407)	0.089	(0.040)	0.027	(0.011)
F:G - <i>Enterococcus</i>	0.539	(0.397)	0.165	(0.036)	0.058	(0.016)
F:G - <i>Escherichia/Shigella</i>	0.012	(0.405)	-0.086	(0.039)	-0.035	(0.014)
F:G - <i>Lactobacillus</i>	-0.521	(0.326)	-0.087	(0.047)	0.013	(0.010)
F:G - <i>Lactococcus</i>	-0.383	(0.431)	-0.067	(0.045)	-0.064	(0.020)
F:G - <i>Leuconostoc</i>	-0.201	(0.448)	-0.054	(0.046)	-0.080	(0.023)
F:G - <i>Macrococcus</i>	0.246	(0.381)	-0.004	(0.060)	0.047	(0.011)
F:G - <i>Microbacterium</i>	-0.595	(0.395)	0.001	(0.039)	0.006	(0.013)
F:G - <i>Ruminococcus 2</i>	-0.334	(0.466)	0.106	(0.039)	0.039	(0.015)
F:G - <i>Sellimonas</i>	-0.273	(0.476)	0.102	(0.039)	0.014	(0.006)
F:G - <i>Staphylococcus</i>	-0.125	(0.432)	0.137	(0.043)	0.058	(0.020)
F:G - <i>Streptococcus</i>	0.321	(0.425)	0.118	(0.037)	0.038	(0.013)
F:G - <i>Tyzzarella</i>	0.149	(0.435)	-0.011	(0.040)	< 0.001	(< 0.001)
F:G - Unc. <i>Lachnospiraceae</i>	-0.247	(0.448)	0.099	(0.039)	-0.002	(0.014)

Genetic correlations r_g , phenotypic correlations r_p and regression coefficients $\lambda_{F:G,Genus}$ between F:G and Genera with significant heritability ($p \leq 0.05$). The standard errors (SE) presented in parentheses and $\lambda_{F:G,Genus}$ in units σ_p . ¹ Feed per gain – Genus with significant heritability ($p \leq 0.05$).

Additional file 5: Table S5 Trait-associated markers from GCTA within the significant QTL regions.

Trait	CJA	Marker ID	p value	Pos ¹	Markers matched with other traits
<i>Aerococcus</i>	3	id00986	< 0.001	0.000	<i>Ruminococcus 2</i>
	3	id29156	0.027	0.000	
	3	id00575	< 0.001	0.000	
	3	id13672	0.007	0.360	
	3	id15191	0.028	1.830	
	3	id07523	0.016	3.082	
	3	id10549	0.033	3.082	
	3	id02839	0.030	3.082	
	3	id14154	0.028	8.013	
	3	id12388	0.019	8.024	
	3	id24815	0.003	9.433	
	3	id18519	0.009	11.689	
	3	id03758	0.021	16.649	
	3	id13768	0.014	17.290	
	3	id05996	0.008	17.290	
3	id12382	0.008	17.290		
<i>Bacillus</i>	2	id09016	0.003	147.885	<i>Cutibacterium</i>
	2	id05956	0.023	148.328	
	2	id32858	0.006	148.948	
	2	id12291	0.013	150.053	
	2	id33001	0.014	150.574	
	2	id25447	0.008	150.911	
	2	id30476	0.021	151.894	
	2	id03102	0.002	154.051	
	2	id00923	0.003	154.051	
	2	id04053	0.002	157.198	
	2	id33819	0.008	157.198	
	2	id10445	0.002	157.198	
	2	id13471	0.001	158.582	
	2	id06720	0.001	158.587	
	2	id02198	< 0.001	158.750	
	2	id04692	0.041	159.867	
	2	id01497	0.002	160.198	
	2	id10454	< 0.001	160.198	
	2	id02833	0.003	160.198	
	2	id02131	0.002	160.317	
2	id12311	0.028	161.185		
2	id08123	0.028	161.185		
2	id09029	0.011	163.378		
2	id09032	0.032	163.648		
<i>Cutibacterium</i>	2	id33001	0.032	150.574	<i>Bacillus</i>
	2	id04053	0.015	157.198	<i>Bacillus</i>
	2	id10445	0.049	157.198	<i>Bacillus</i>
	2	id18905	0.023	157.199	
	2	id06720	0.043	158.587	<i>Bacillus</i>
	2	id02198	0.039	158.750	<i>Bacillus</i>
	2	id01497	0.045	160.198	<i>Bacillus</i>
	2	id10454	0.009	160.198	<i>Bacillus</i>

2	id02833	0.011	160.198	<i>Bacillus</i>	
2	id02131	0.039	160.317	<i>Bacillus</i>	
2	id09032	0.035	163.648	<i>Bacillus</i>	
2	id01918	0.037	165.518		
2	id13472	0.037	165.518		
2	id32230	0.045	168.860		
2	id15179	0.032	168.860		
2	id32700	0.041	168.871		
2	id17401	0.014	169.669		
2	id23477	0.013	169.669		
2	id28643	0.013	169.669		
2	id08129	0.011	169.669		
2	id10474	0.007	170.978		
2	id31089	0.001	171.280		
2	id12342	0.003	171.280		
2	id08131	0.001	171.280		
2	id15180	0.020	171.280		
2	id03809	0.002	171.280		
2	id01498	0.002	171.280		
2	id09043	0.017	173.562		
2	id32704	0.044	173.563		
2	id10482	0.021	173.563		
2	id17403	0.044	173.567		
2	id13854	0.003	174.374		
2	id00622	0.001	176.564		
2	id17758	0.014	176.564		
<hr/>					
<i>Escherichia/Shigella</i>	24	id10195	0.042	0.000	
	24	id13533	0.041	0.000	
<hr/>					
<i>Ruminococcus 2</i>	3	id00986	< 0.001	0.000	<i>Aerococcus</i>
	3	id18286	0.039	0.000	
	3	id00575	0.001	0.000	<i>Aerococcus</i>
	3	id13672	0.009	0.360	<i>Aerococcus</i>
	3	id06525	0.001	0.360	
	3	id07523	0.009	3.082	<i>Aerococcus</i>
	3	id02839	0.014	3.082	<i>Aerococcus</i>
	3	id34021	0.002	3.082	
	3	id14154	0.008	8.013	<i>Aerococcus</i>
	3	id12388	0.009	8.024	<i>Aerococcus</i>
<hr/>					
<i>Streptococcus</i>	5	id14925	< 0.001	44.417	
	5	id08312	0.035	44.614	
	5	id15246	0.026	45.341	
	5	id03766	0.005	45.342	
	5	id17506	0.015	45.342	
	5	id01172	0.038	45.342	
	5	id04785	0.007	45.342	

Summary of the trait-associated markers from GCTA ($p \leq 0.05$) within the significant QTL regions. In addition, all markers that are significantly associated with another characteristic are listed in the last column. ¹ in cM.

CHAPTER THREE

Inferring causal structures of gut microbiota diversity and feed efficiency traits in poultry using Bayesian learning and genomic structural equation models

Valentin Haas^{*,1}, Markus Rodehutschord^{*}, Amélia Camarinha-Silva^{*} and Jörn Bennewitz^{*}

^{*} Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

¹ Corresponding author: valentin.haas@uni-hohenheim.de

Published in *Journal of Animal Science* (2023), Volume 101, skad044

<https://doi.org/10.1093/jas/skad044>

Lay Summary

Feed efficiency and phosphorus efficiency are of increasing importance in poultry breeding. It was frequently shown that next to the birds' genomes also the gut microbiota composition is important for these efficiency traits. The gut microbiota composition is a mediator between the genomes of the birds and their efficiency traits. In the present study, an approach was taken to consider the animal's gut microbiota diversity, efficiency traits, and the genomes of the animals together in a causal network to decipher the mediator role between the traits. Growing Japanese quail were used as model species. A stable network could be established that placed the diversity of the gut microbiota composition at the forefront, with direct and indirect links to other traits like phosphorus utilization and retention, feed per gain ratio, and growth. Together with genome scans, the results confirmed the mediator role of the gut microbiota composition because several traits associated variants affected the efficiency traits directly and indirectly via the gut microbiota composition.

Teaser Text

This study investigated the complex interrelationship between gut microbiota composition and feed and phosphorus efficiency traits in Japanese quail. The results offered new insight into the traits' hierarchical structure and shared genetic architecture.

Abstract

Feed and phosphorus efficiency are of increasing importance in poultry breeding. It has been shown recently that these efficiency traits are influenced by the gut microbiota composition of the birds. The efficiency traits and the gut microbiota composition are partly under control of the host genome. Thus, the gut microbiota composition can be seen as a mediator trait between the host genome and the efficiency traits. The present study used data from 749 individuals of a Japanese quail F2 cross. The birds were genotyped for 4k SNP and trait recorded for phosphorus utilization (PU) and phosphorus retention (PR), body weight gain (BWG) and feed per gain ratio (F:G). The gut microbiota composition was characterized by targeted amplicon sequencing. The alpha diversity was calculated as the Pielou's evenness index (J'). A stable Bayesian network was established using a Hill-Climbing learning algorithm. Pielou's evenness index was placed as the most upstream trait and BWG as the most downstream trait, with direct and indirect links via PR, PU, and F:G. The direct and indirect effects between J' , PU, and PR were quantified with structural equation models, which revealed a causal link from J' to PU and from PU to PR. Quantitative trait loci (QTL) linkage mapping revealed three genome-wide significant QTL regions for these traits with in total 49 trait-associated SNP within the QTL regions. Structural equation model association mapping separated the total SNP effect for a trait into a direct effect and indirect effects mediated by upstream traits. Although the indirect effects were in general small, they contributed to the total SNP effect in some cases. This enabled us to detect some shared genetic effects. The method applied allows for the detection of shared genetic architecture of quantitative traits and microbiota compositions.

Keywords: Bayesian network, causal network, gut microbiota composition, phosphorus efficiency traits, poultry breeding, structural equations

List of Abbreviations

BWG, body weight gain; CJA, *Coturnix japonica* chromosome; F:G, feed per gain ratio; GWAS, genome-wide association study; HC, Hill-Climbing; High J' , 25 birds with the greatest Pielou's evenness index; HPD, highest posterior density; J' , Pielou's alpha diversity index; LOD, logarithm of the odds; Low J' , 25 birds with the lowest Pielou's alpha diversity index; MCMC, Markov chain Monte Carlo; OTU, operational taxonomic units; P, phosphorus PR, phosphorus retention; PSD, posterior SD; PU, phosphorus utilization; QTL, quantitative trait loci; SEM, structural equation models; SI, support interval; SNP, single-nucleotide polymorphism

Introduction

It was repeatedly shown that the gut microbiota composition influences feed and nutrient efficiency traits in pigs (reviewed by Maltecca et al. (2020)) and poultry (Vollmar et al., 2020).

Because the gut microbiota composition is partly under host genetic control, it might be used to improve genomic breeding value estimation (Christensen et al., 2021) or selection indices (Weishaar et al., 2020). Hence, the gut microbiota composition can be seen as a mediator between the host genome and efficiency traits, and it might be that the genetic variation observed for efficiency traits is the sum of direct and indirect genetic variation mediated by the heritable part of the microbiota composition. One option to analyze such a relationship is to apply structural equation models (SEM), as introduced in livestock quantitative genetics by Gianola and Sorensen (2004). Structural equation models provide a functional link between traits and allow for the prediction of one trait from other upstream traits (Valente et al., 2010). These classes of models have been frequently used in livestock quantitative genetics (Dhakal et al., 2016; Inoue et al., 2016; Okamura et al., 2020; Pegolo et al., 2020; Pegolo et al., 2021) and recently in research on genetic relationships between traits and gut microbiota (Saborío-Montero et al., 2020; Tiezzi et al., 2021; Haas et al., 2022). In the latter studies the microbiota composition was considered as multiple traits and relative abundancies of single features (e.g., bacteria genera or operational taxonomic units (OTU)) were included one at a time in an SEM. Currently there is no index available that includes all these features, which might also be hampered by the fact that the relative abundancies are compositional data. Alternatively, alpha diversity measures of the gut microbiota are usually calculated using the relative abundance of all microbiota features. Hence, these diversity measures can be seen as one way of reducing the dimensionality towards one single trait. Significant heritabilities of alpha diversity measures of pig fecal microbiota composition and genetic correlation with production traits were recently reported by Lu et al. (2018), Aliakbari et al. (2021) and Déru et al. (2022). These studies also discussed the usefulness of alpha diversity measures as a potential selection target in a breeding scheme with the aim to promote correlated traits.

In the SEM studies considering the microbiota composition mentioned above, the causal link between the traits considered were chosen based on assumptions on biological knowledge, i.e., the microbiota affects the trait. However, a more detailed investigation of this relationship including multiple traits would be beneficial. For this purpose, inductive causation (IC) algorithms (Verma and Pearl, 2022) or Bayesian networks were proposed (Chickering, 2013), which suggest causal structures and establish a network based on the observed data. Bayesian networks are probabilistic models that can be used to graphically visualize dependencies between different variables, i.e., quantitative traits, (Heckerman, 2008) with directed relationships represented by arrows between traits. A network can be viewed as factorization of a joint probability distribution, where the conditional probability distributions at each node represent the factors and the graph structure represents their combination method

(Daly and Qiang, 2007). In resulting graphs, the nodes represent the variables, and the edges represent the probabilistic dependencies between them.

Momen et al. (2019) extended SEM towards genome-wide association studies (GWAS). These SEM-GWAS enable to separate a total SNP effect for a trait into a direct effect and indirect effects mediated by upstream traits in the inferred network. The application of SEM-GWAS adds more information to the shared genetic architecture of multiple quantitative traits within a network. Pegolo et al. (2020) used Bayesian networks to establish a network between dairy cattle udder health traits and applied SEM and subsequently SEM-GWAS to these traits. This kind of extended SEM studies would be of interest to shed more light on the complex relationship between gut microbiota composition and host quantitative traits, but has to the best of our knowledge not systematically applied in this field.

In previous studies, we detected a substantial impact of the gut microbiota composition on feed and phosphorus (P) efficiency in Japanese quail (Vollmar et al., 2020) and found that both the microbiota composition and the host efficiency traits were significantly heritable (Beck et al., 2016; Haas et al., 2022). Based on this, we established some causal structures between selected features of the microbiota composition and some host traits using prior biological information using bivariate SEM (Haas et al., 2022). This approach was rather simplified in a sense that only one microbiota feature was considered at time and no relationship between quantitative traits were modelled, i.e., the effect of one microbiota feature on one quantitative trait was investigated.

The aim of the present study was (i) to determine the causal links between quail gut microbiota alpha diversity and several efficiency traits of the quail with a Bayesian learning algorithm and graphical models, (ii) to use the learned structure to compute structural coefficients with SEM for three quantitative traits and the alpha diversity, and (iii) to perform SEM association analysis to split total SNP effects for a trait into direct, and indirect effects.

Material and Methods

This animal experiment was performed according to the requirements of the German Animal Welfare Legislation and approved by the Animal Welfare Commissioner of the University of Hohenheim (approval number S371/13TE).

Experimental design and data collection

Details regarding the design of the underlying study can be found in Beck et al. (2016), and only the essentials are described here. We used a dataset of an F2 cross consisting of 749 Japanese quail (*Coturnix japonica*) that were fed a diet with marginal P concentration. Phenotyping took place between 10th and 15th day of life. The traits P utilization (PU), feed per gain ratio (F:G), P retention (PR), and body weight gain (BWG) were recorded, for details see

Beck et al. (2016). After the test period, the quails were slaughtered to collect ileum digesta samples and characterized the individual microbiota composition by targeted amplicon sequencing. Sequence reads were clustered into OTU with a 97% similarity, as described in detail by Borda-Molina et al. (2020). In total, 1,188 OTU with an average relative abundance > 0.0001% and a sequence length > 250 bp were used. As a measure of the OTU diversity observed within individuals, we calculated Pielou's evenness index (J') (Pielou, 1966), using Primer 7 software (Clarke and Warwick, 2001):

$$J' = \frac{H'}{H'_{\max}} = \frac{H'}{\log S},$$

with $H' = -\sum_i p_i \log(p_i)$, where H' is the measurement of the Shannon diversity index, and p_i is the proportion of the total count arising from the i th OTU, and S is the total number of OTU.

The birds were genotyped genome-wide with 4k SNP, for which a genetic linkage map was available (Vollmar et al., 2021). Slaughtering and breeding took place on twelve different days, resulting in twelve test days in the statistical analysis. Scatterplots of raw phenotypes for the traits PU, F:G, PR, and BWG with the microbiota trait J' are produced using R Studio 4.1.1 (R Core Team, 2021).

Statistical analyses

Genetic parameter estimation

The statistical analysis started with a multivariate estimation of genetic parameters of the host traits (heritability, genetic and phenotypic correlations) using the following Bayesian animal model in R Studio 4.1.1 (R Core Team, 2021):

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e} \text{ (Model 1)},$$

where \mathbf{y} is the vector of the phenotypes of the $t = 5$ traits (J' , PU, F:G, PR, and BWG), \mathbf{b} is the vector of test day effects (12 test days) with flat priors, \mathbf{a} is the vector with random additive genetic effects, \mathbf{e} is the vector of residuals, and \mathbf{X} and \mathbf{Z} are the known design matrices. The vectors \mathbf{a} and \mathbf{e} were assumed to have independent Gaussian distributions $\mathbf{a} \sim N(0, \Sigma_a \otimes \mathbf{A})$ and $\mathbf{e} \sim N(0, \Sigma_e \otimes \mathbf{I})$, where \mathbf{A} is the pedigree-based relationship matrix, \mathbf{I} is the identity matrix for residuals, and Σ_a and Σ_e are the $t \times t$ variance-covariance matrices of genetic and residual effects. The operator \otimes denotes the Kronecker product. For the additive genetic and residual variances, flat and uninformative priors were used. The MCMCglmm package (Hadfield, 2010) was used to obtain the posterior distributions by applying Markov chain Monte Carlo (MCMC) sampling. The chain length was 750,000 iterations, the burn-in 150,000 and the thinning interval 300. The genetic parameters were estimated using standard notations from the posterior distributions of the variance and covariance components, and the highest posterior

density (HPD) intervals (95%) were determined. In addition, the residuals from this model were stored, which were needed for the Bayesian network analysis (see below).

QTL linkage analyses

Quantitative trait loci (QTL) mapping was done using a linkage mapping approach by Vollmar et al. (2021) for the traits PU, F:G, and BWG. Linkage mapping was applied because the number of SNPs was too low for a comprehensive GWAS. To complete the QTL mapping, we applied the same linkage mapping procedure as Vollmar et al. (2021) for the remaining two traits J' and PR, using the R package R/qtl2 (Broman et al., 2019), which was developed for inbred crosses. Because the Japanese quail are not an inbred cross, we selected the SNP from the marker list according to $F_{st} > 0.23$ in the two founder lines. This ensured a strong divergent allele frequency across the two founder lines and helped to fulfill the requirement of divergent allele fixation approximately. The package R/qtl2 estimated the QTL genotype probabilities for each marker position and each F2 individual and performed genome scans using regression of the phenotypes on these probabilities. Only additive QTL effects were considered. The test day effects were included in the regression analyses as fixed effects. The logarithm of the odds (LOD) score was used as test statistic. Permutation test with 10,000 permutations was applied to obtain genome-wide significance threshold values. We considered two significance limits, i.e., 1% and 5% genome-wide significance, and QTL support intervals (SI) were obtained using 1.5 LOD drop-off, which corresponds approximately to a 95% confidence interval (Manichaikul et al., 2006). The bounds of the SI were extended slightly in both directions (Vollmar et al., 2021) to be more conservative. Within the SI for the identified QTL, all SNP were tested for trait associations using single-marker association mapping implemented in GCTA using the LOCO option to control for population stratification (Yang et al., 2011). The existence of the QTL was inferred from the linkage mapping controlling for genome-wide multiple testing using permutations. Therefore, correcting for multiple testing for these single marker association analyses was restricted to the number of tested SNPs within the SI and this was done using a Bonferroni approach.

Bayesian network

To obtain the Bayesian network of the five traits, we used the score-based Hill-Climbing (HC) learning algorithm (Daly and Qiang, 2007) as implemented in the R package bnlearn 4.7 (Scutari, 2010) and visualized it graphically with the R package qgraph 1.9 (Epskamp et al., 2012). We used the residuals obtained from Model 1 to avoid that the traits are dependent based on systematic genomic or test day effects, which would bias the network (Rosa et al., 2011). We used the bootstrapping procedure with 50,000 bootstrap samples, in which the strength of the arcs (proportion that there is an arc between), and the proportion of direction were estimated.

SEM and SEM association analysis

Based on the Bayesian network analysis results, we selected $t = 3$ traits, i.e., J', PU, and PR for SEM analysis. This reduction of the number of traits was necessary, because the dataset was too small for a complete SEM analysis for all five traits in the network, which resulted in convergence problems (not shown). The following SEM implemented in ASReml 4.1 (Gilmour et al., 2015) was applied (using the notation of Rosa et al. (2011)):

$$\mathbf{y} = (\mathbf{\Lambda} \otimes \mathbf{I}_n)\mathbf{y} + \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \text{ (Model 2),}$$

where \mathbf{y} is the vector of scaled (mean = 0, SD = 1) trait records (J', PU, and PR) for the n individuals in the identity matrix \mathbf{I}_n and $\mathbf{\Lambda}$ is a off-diagonal matrix ($t \times t$) of structural coefficients $\lambda_{i,j}$, which expresses the change of trait i through the recursive influence of trait j . The following structure was assumed based on the Bayesian network:

$$\mathbf{\Lambda} = \begin{bmatrix} 0 & 0 & 0 \\ \lambda_{PU,J'} & 0 & 0 \\ \lambda_{PR,J'} & \lambda_{PR,PU} & 0 \end{bmatrix}.$$

The remaining terms are defined above. The residual variances in \mathbf{R} were assumed to be independent to ensure identifiability (Rosa et al., 2011).

Model 2 was extended towards accounting for single SNP that were within the QTL SI detected in the QTL linkage mapping. The model for this SEM association analysis was adapted from Momen et al. (2019) and Pegolo et al. (2020) as:

$$\mathbf{y} = (\mathbf{\Lambda} \otimes \mathbf{I}_n)\mathbf{y} + \mathbf{Ws} + \mathbf{Xb} + \mathbf{Za} + \mathbf{e} \text{ (Model 3),}$$

where \mathbf{W} is a $n \times t$ matrix of the genotype codes (coded as 0, 1 or 2) of the marker under consideration and \mathbf{s} is the $t \times 1$ vector with SNP effects. The other terms are as defined above. This model allows for the split of the total SNP effect on a focal trait into a direct effect and in indirect effects. The latter ones are defined as those effects on a focal trait that are mediated by other traits that are upstream in the network. They were calculated by multiplying path coefficients with direct SNP effects of upstream traits, see Supplementary Material S1 for a full presentation of each trait. The total effect of a SNP for the focal trait was the sum of the direct effect and all indirect effects.

Results**Heritabilities and correlations**

The scatterplots (Fig. 1) indicate a relationship between the traits BWG, PU, PR and F:G with J'. This can also be seen from the phenotypic correlations, which are presented along with genetic correlations and heritabilities in Table 1. The traits PU, PR, and BWG were negatively

correlated with J' . The correlation between J' and F:G is close to zero. Some genetic correlations are substantial, especially between PU and PR, BWG and PR, and J' and BWG. The heritability estimates range from 0.11 (F:G) to 0.26 (PR). For some of the traits the heritability was estimated previously univariately with REML (Beck et al., 2016) and estimates agree with the values reported here.

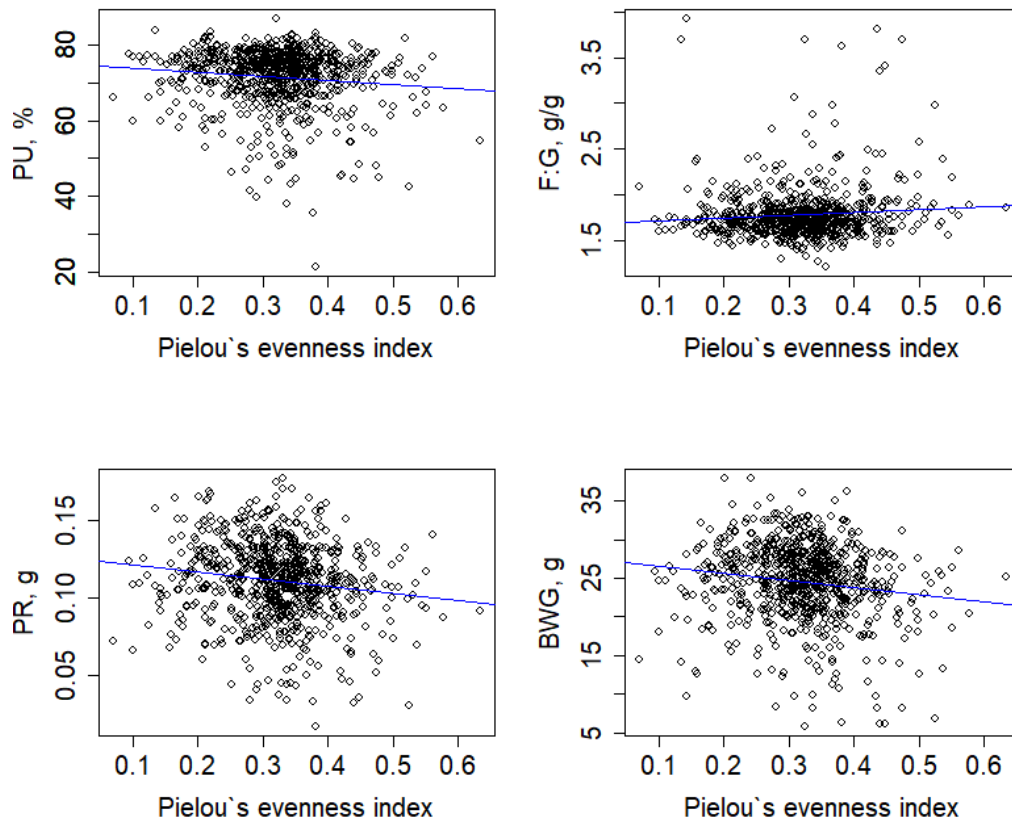


Figure 1 Scatterplot comparisons of the traits phosphorus utilization (PU), feed per gain ratio (F:G), phosphorus retention (PR), and body weight gain (BWG) and the host microbiota trait Pielou's evenness index (J') and regression line between traits.

Table 1 Heritability (diagonal), genetic (below the diagonal), and phenotypic (above the diagonal) correlations between the phenotypic traits, with the posterior SD in parenthesis from the multivariate model (Model 1).

Traits ^a	J'	PU	F:G	PR	BWG
J'	0.19 (0.08)*	-0.10 (0.04)*	0.06 (0.04)	-0.15 (0.05)*	-0.16 (0.05)*
PU	-0.17 (0.32)	0.14 (0.07)*	-0.38 (0.04)*	0.79 (0.02)*	0.57 (0.03)*
F:G	-0.06 (0.34)	0.03 (0.40)	0.11 (0.08)*	-0.26 (0.05)*	-0.62 (0.05)*
PR	-0.46 (0.23)*	0.58 (0.22)*	0.41 (0.29)	0.26 (0.12)*	0.82 (0.02)*
BWG	-0.62 (0.22)*	0.37 (0.33)	0.23 (0.38)	0.81 (0.11)*	0.17 (0.11)*

^a J' : Pielou's evenness index, PU: Phosphorus utilization, F:G: Feed per gain ratio, PR: Phosphorus retention, BWG: Body weight gain;

* 95% highest posterior density interval not including 0.

Bayesian network and structural equation coefficients

The Bayesian network, established with the residuals from Model 1 for the five traits, is shown in Fig. 2. Bootstrap sample proportions are given for the edges and the direction of the edges, respectively. The recursive effect from J' to PU was added manually after network discovery as the relationship was too weak from the Bayesian network. This was justified based on previous biological justification, as we found strong associations in our previous studies between the microbiota composition and PU (Borda-Molina et al., 2020; Vollmar et al., 2020). In general, the network is supported by bootstrap analysis. Pielou's alpha diversity index J' affects PR directly, and BWG indirectly, mediated via PR. All other traits affect BWG directly. Phosphorus utilization affects F:G and PR directly and BWG indirectly via these two traits.

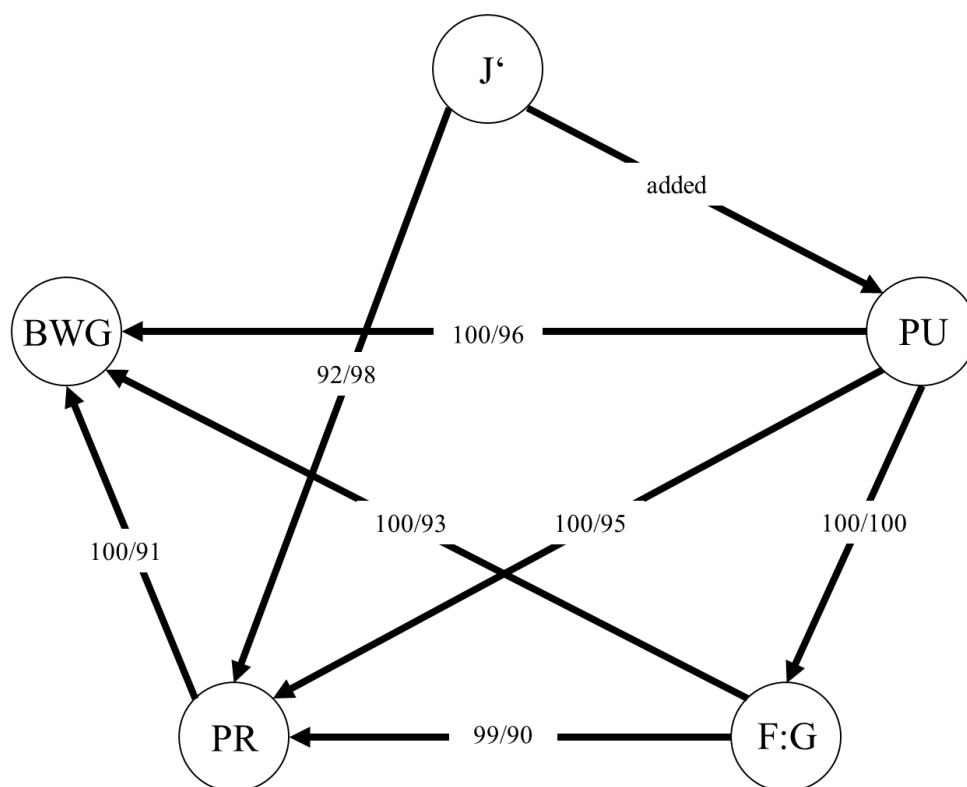


Figure 2 Bayesian network structure, resulting from the Hill-Climbing learning algorithm (J': Pielou's evenness index, PU: Phosphorus utilization, F:G: Feed per gain ratio, PR: Phosphorus retention, BWG: Body weight gain). The path J' → PU was added based on prior biological knowledge. The numbers are bootstrap sample proportions (50,000 samples), indicating an existing arc between the variables (left figures) and the proportion of this direction of the arc (right figures).

The structural coefficients $\lambda_{i,j}$ estimated with Model 2 are shown in Table 2. Pielou's evenness index J' had a slight negative impact on PU and no impact on PR. The impact of PU on PR was positive and substantial.

Table 2 Structural coefficients $\lambda_{i,j}$ from the trivariate SEM (Model 2), with standard error (SE) and P -value.

Structural coefficient ^a	Path	Lambda	SE	$P_{nominal}$
$\lambda_{PU,J'}$	$J' \rightarrow PU$	-0.08	0.03	0.03
$\lambda_{PR,J'}$	$J' \rightarrow PR$	-0.01	0.02	0.74
$\lambda_{PR,PU}$	$PU \rightarrow PR$	0.79	0.02	< 0.01

^a λ = structural coefficient between the traits (J': Pielou's evenness index; PU: Phosphorus utilization; PR: Phosphorus retention)

QTL mapping results

The plot of the test statistic obtained from the QTL linkage mapping for the traits included in the SEM analyses are shown in Fig. 3. For J' there is a significant QTL located on *Coturnix japonica* chromosome (CJA) 13, for PU on CJA3, and for PR also on CJA3.

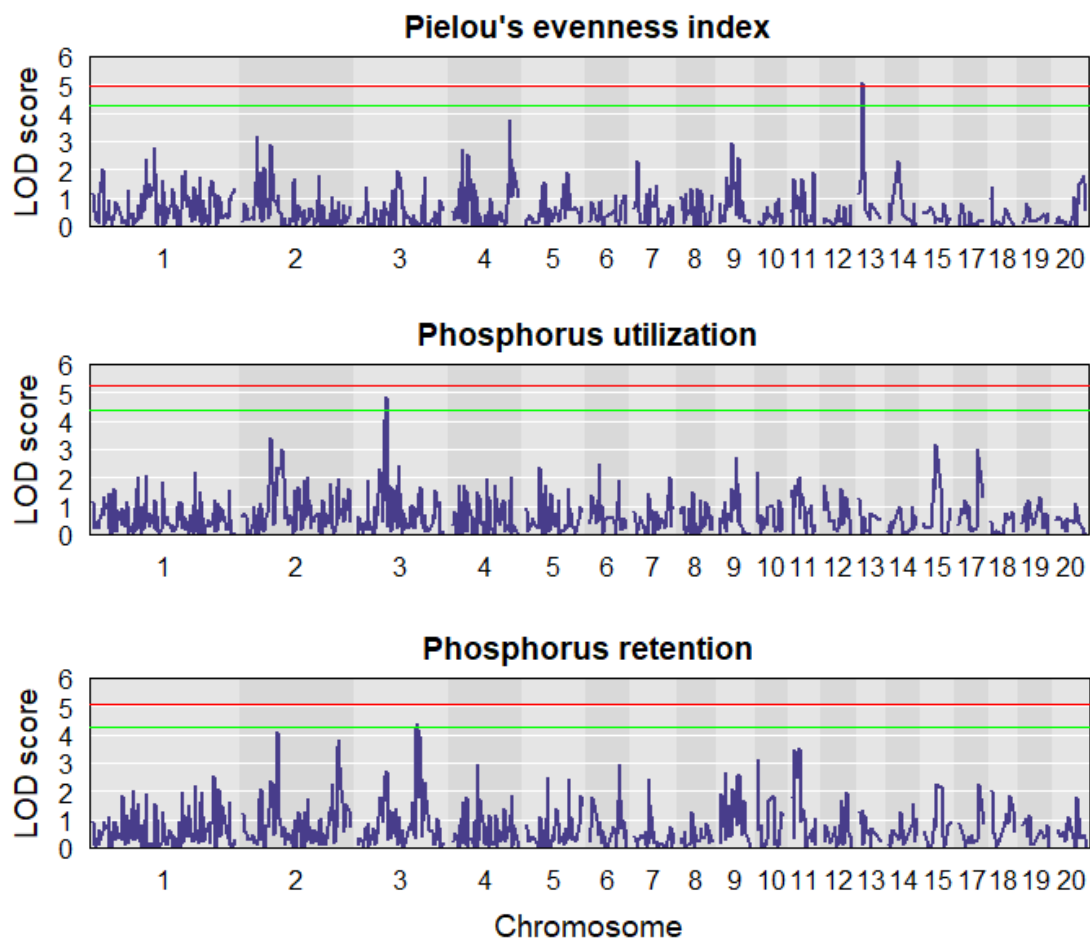


Figure 3 Plot of the QTL linkage mapping result of Pielou's evenness index (J'), phosphorus utilization (PU), and phosphorus retention (PR) with LOD score as test statistic. The red and green lines correspond to genome-wide significance levels of 1% and 5%, respectively. The PU results are taken from Vollmar et al. (2021).

The results for the three significant QTL with chromosomal position, test statistic, SI, and the number of significant trait-associated SNP from GCTA are shown in Table 3. The SI ranges from 12.6 cM (J') to 29.8 cM (PU). The QTL regions on CJA3 for trait PU and PR do not overlap. In total, 49 SNP within the SI were trait-associated ($P_{nominal} \leq 0.05$), see Supplementary Table S2.

Table 3 QTL linkage mapping results position (Pos) in cM of the 1% (**) and 5% (*) genome-wide significant QTL on Coturnix japonica chromosome (CJA), with the LOD score, the Support interval borders, the number of significant trait-associated SNPs from GCTA ($P_{nominal} \leq 0.05$) and the number of Bonferroni adjusted significant SNPs ($P_{Bonferroni} \leq 0.05$).

Trait ^a	CJA	Pos (cM)	LOD	Support interval borders (cM)		Number of nominal significant SNPs	Number of adjusted significant SNPs
				low	high		
J'	13	7.1	5.07**	1.3	13.9	8	1
PU	3	48.9	4.82*	35.9	65.7	17	1
PR	3	104.4	4.36*	97.8	117.1	24	10

^a J': Pielou's evenness index, PU: Phosphorus utilization, PR: Phosphorus

SNP effects

The general structure of the SEM association analysis for the 49 significant SNP detected in the QTL mapping and the reduced network with estimated structural coefficients are shown in Fig. 4. Each SNP has a direct effect on the traits ($s_{J'}$, s_{PU} , s_{PR}) and for PU and PR also indirect effects, see Supplementary Material S1.

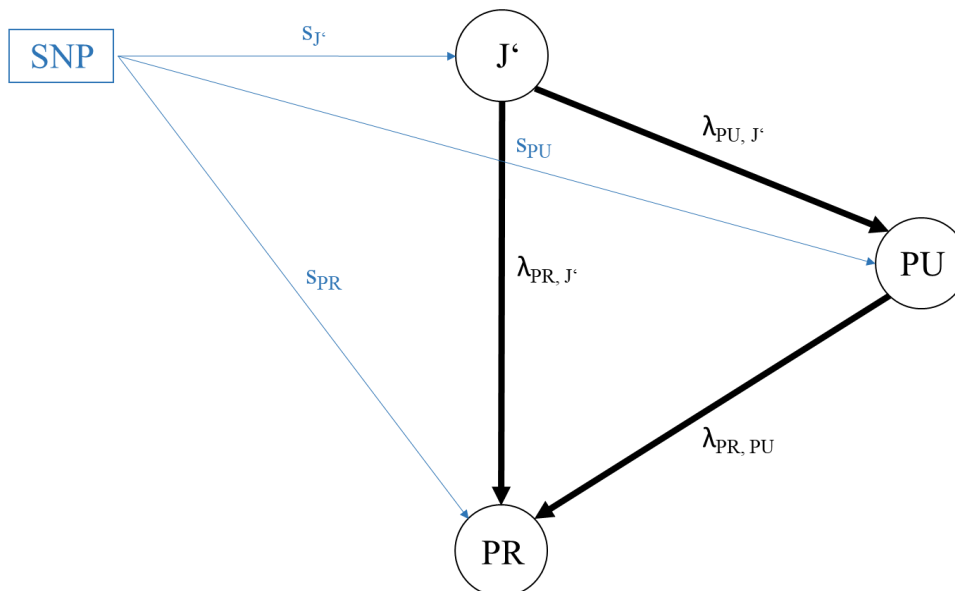


Figure 4 General structure of SEM association analysis with structural coefficients ($\lambda_{PU, J'}$; $\lambda_{PR, PU}$; $\lambda_{PR, J'}$) (J': Pielou's evenness index; PU: Phosphorus utilization; PR: Phosphorus retention) and direct SNP effects ($s_{J'}$, s_{PU} , s_{PR}).

Figure 5 shows a section of the SNP with their effects on PU and PR in their respective QTL regions, for all those with significant direct SNP effect ($P_{nominal} < 0.05$) (calculated via Model 3). A complete list can be found in Supplementary Table S3.

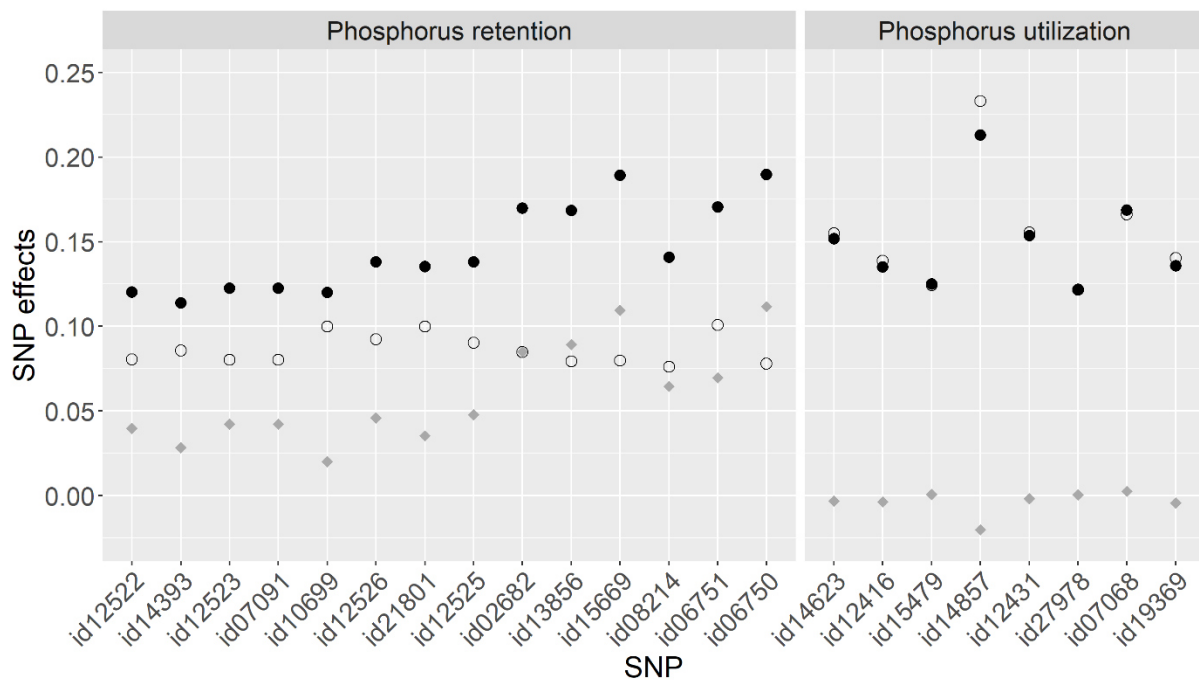


Figure 5 Plots of direct (circle), summed indirect (rhombus), and total effects (black dot) of the SNP within the support interval (SI) of the CJA3 QTL for phosphorus retention (PR) (right), and CJA3 for phosphorus utilization (PU) (left), obtained by using SEM association analysis. The plot shows all SNP with a significant direct SNP effect ($P_{nominal} < 0.05$).

Discussion

This study analyzed the interrelationship between ileal microbiota composition, host genetics, and quantitative efficiency traits in Japanese quail using Bayesian networks and SEM. It extended previous work using simplified assumptions, i.e., the causal structure among selected features of microbiota composition and quantitative traits was given a priori (Haas et al., 2022). In addition, the microbiota composition was modelled using Pielou's alpha diversity index instead of the average relative abundances of bacteria genera or OTU. This allows for modelling of the microbiota composition based on community species evenness as a biological indicator. The heritability estimates of the diversity index J' indicates a significant host genetic effect. Similar values were reported in pigs using different alpha diversity indexes in the ileum (Tang et al., 2020) and fecal samples (Lu et al., 2018; Aliakbari et al., 2021; Déru et al., 2022). These studies reported also negative genetic correlations between alpha diversity measures and body weight gain, which agrees with the negative genetic correlation between J' and BWG found in this study.

The Bayesian HC learning algorithm (Scutari, 2010) identified a causal structure among the residuals obtained from Model 1 of the traits. The residuals were chosen as input variables in the network analysis to exclude putative confounding common genetic factors, as suggested by Rosa et al. (2011) and Pegolo et al. (2020). The connection J' on PU was chosen based on the strong effect of the microbiota on PU, as detected in our previous study (Borda-Molina et al., 2020; Vollmar et al., 2020). This external information given to the algorithm ensured a stable network of the traits, as indicated by the bootstrap sampling results. The hierarchy within this network is that the microbiota diversity J' is the most upstream and BWG the most downstream trait and the other three are mediator traits. More precisely, J' directly influenced PU and PR, and indirectly F:G (mediated by PU), BWG (mediated by PU and PR). This is biologically meaningful because P supply of the quail was marginal and an increase in PU at such supply level should increase retention of P and overall growth of the birds. The Bayesian network results agrees with the correlations between J' and the other traits. This points to the presence of some microbiota features that promote these traits. Increased frequencies of these features might reduce the diversity index J' but improve the other traits in a desired direction, from a breeder's perspective (note that a reduced F:G is also desired because this improves feed efficiency). To identify these microbiota features, we used the 25 animals with the greatest (High J') and the 25 with the lowest (Low J') Pielou's alpha diversity index and compared the microbiota composition between these groups at the phylum level and at the bacterial genus level. At the phylum level (Fig. 6) Firmicutes had significant ($P < 0.0001$) greater abundances in the Low J' (98.3%) than in the High J' (66.6%) group. The phylum Proteobacteria (1.4% and 16.4% in the Low J' and High J' group, respectively, $P < 0.0001$), Actinobacteria (0.3% in the Low J' and 9.8% in the High J' group, $P < 0.0001$), and Bacteroidetes (0.1% in the Low J' line versus 7.1% in the High J' line, $P < 0.002$) had significant lower levels of relative abundances in the Low J' as in the High J' group. This agrees with Vollmar et al. (2020), who found, that a greater abundance of Firmicutes increased and greater abundance of Proteobacteria decreased PU. At the bacteria genus level, we found a similar pattern (Fig. 7). Particularly worth mentioning are *Lactobacillus* (46.6% in the Low J' and 16.0% in the High J' group, $P < 0.001$), and *Candidatus Arthromitus* (35.5% in the Low J' line versus 11.9% in the High J' line, $P < 0.003$), *Escherichia/Shigella* (1.4% and 15.4% in the Low and High J' group, respectively, $P < 0.003$), and *Enterococcus* (0.9% and 6.6% in the Low and High J' line, respectively, $P < 0.01$). These findings at genus level agrees with Borda-Molina et al. (2020) who found distinct differences in relative abundance of some bacteria between low PU and high PU phenotypes (e.g., *Candidatus Arthromitus*). This indicates that differences in the quantities of P accreted by the birds are mainly driven by individual capabilities to utilize P from the digestive tract. Most of the P contained in the diet was present in the form of phytate (Beck et al., 2016), and

endogenous enzymes (by microbes or intestinal epithelia) are needed to make the P available for absorption (Rodehutsord et al., 2022).

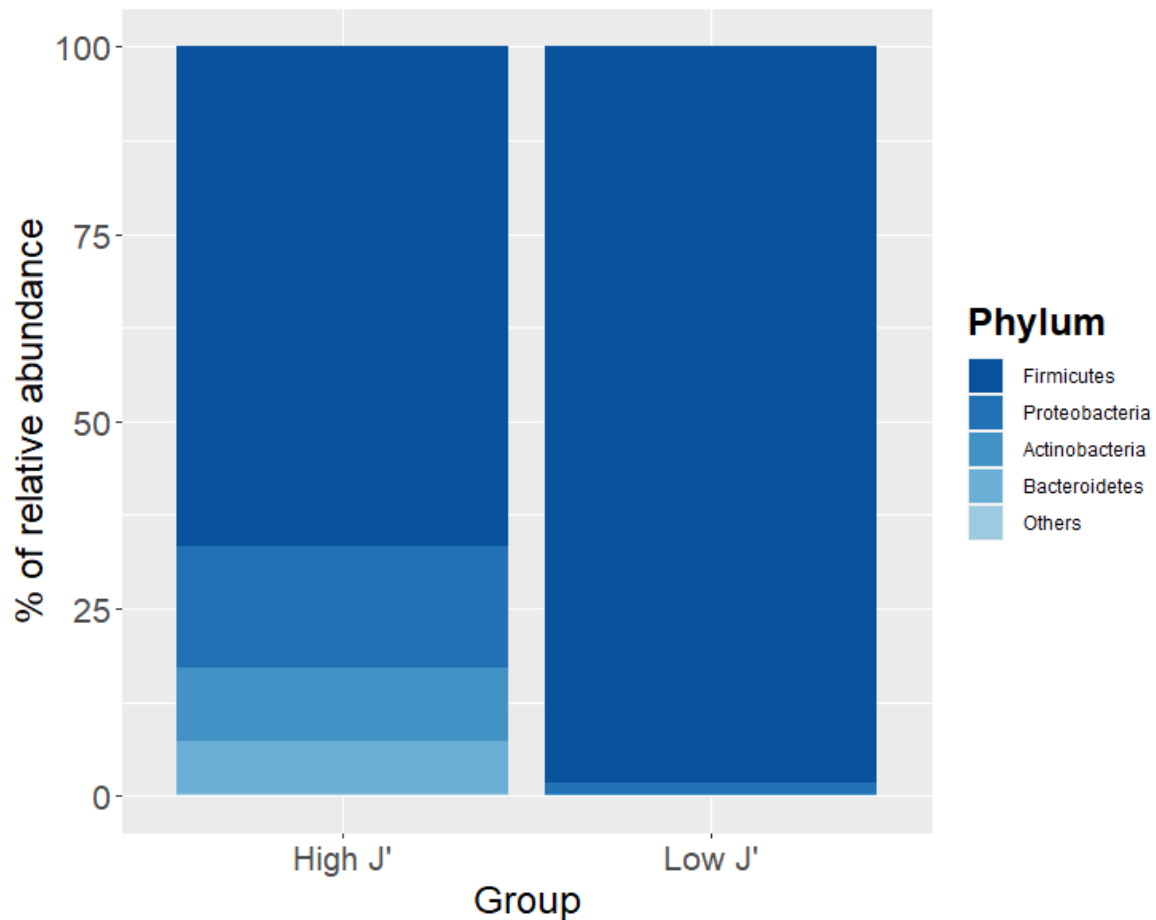


Figure 6 Relative abundance percentage of phyla of the 25 animals with the greatest Pielou's alpha diversity index (High J') and the 25 animals with the lowest Pielou's alpha diversity index (Low J'). The term 'Others' are the phyla with a relative abundance < 1% in both groups.

The dataset was underpowered to apply SEM to all five traits included in the Bayesian network analysis. Therefore, only a subset of three traits was selected from the network and analyzed with SEM. The results from this reduced network SEM revealed a directional effect of J' on PU and PU on PR, i.e., significant estimates for $\lambda_{PU,J'}$ and for $\lambda_{PR,PU}$, respectively. The directional effect of J' on PR, as suggested by the network, could not be confirmed by the SEM, because $\lambda_{PR,J'}$ was small and not significant. Thus, it seems biologically plausible that J' was found to directly affect PU but not PR.

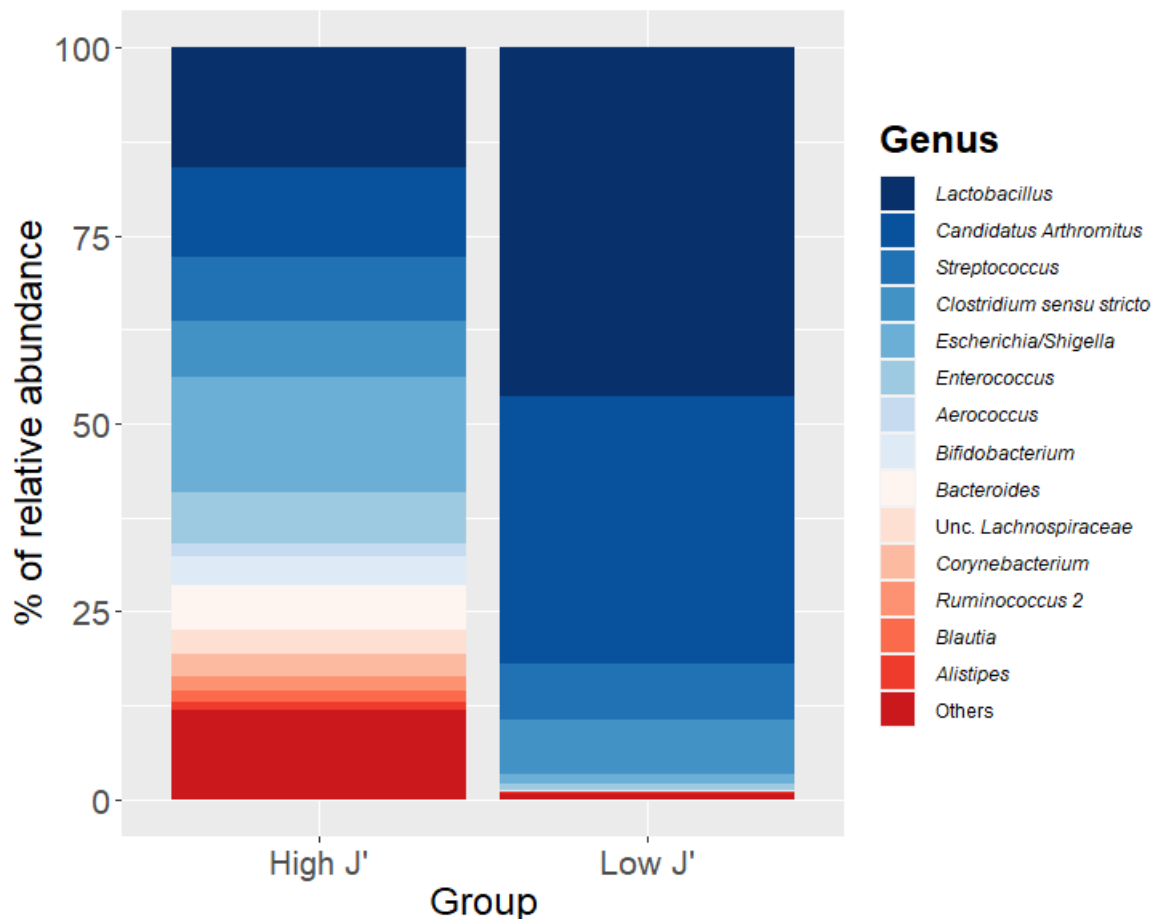


Figure 7 Relative abundance percentage of Genera of the 25 animals with the greatest Pielou's alpha diversity index (High J') and the 25 animals with the lowest Pielou's alpha diversity index (Low J'). The term 'Others' are the genera with a relative abundance < 1% in both Groups.

The SEM (Model 2) extension towards SNP association analysis (Model 3) revealed some shared genetic effects of the three traits that became not obvious from the single trait QTL or SNP association analysis. This can be seen from the QTL regions and the underlying trait-associated SNP, which differ for the traits. However, the SEM association analysis detected some indirect SNP effects for the network downstream trait PR that were mediated by the network upstream traits J' and PU. The total SNP effects on PR of the SNP within QTL region can be explained in about one-third by the indirect SNP effects. In some cases, the indirect effect was larger than the direct effect. The indirect SNP effect with the largest share is explained by the upstream trait PU, according to the amount of recursive influence (Supplementary Table S3). The SNP effects of PU are explained exclusively by the direct SNP effects.

In conclusion, our approach considers jointly the gastrointestinal microbiota diversity, host genetics, and quantitative traits. It has been shown that Bayesian networks are well suited to

deciphering complicated biological systems, also with the consideration of host microbiota data. With the additional use of structural equation models and their combination with association studies, total SNP effects could be divided into a direct effect and indirect effects. Including the host microbiota in such causal networks makes it possible to understand better the relationship between the gastrointestinal microbiota, host genetics, and efficiency traits.

Disclosures

The authors declare no conflict of interest.

Acknowledgements

This research was funded by the Deutsche Forschungsgemeinschaft (DFG, BE3703/12-2) and was part of the research unit P-FOWL (FOR 2601).

Literature Cited

- Aliakbari, A., O. Zemb, Y. Billon, C. Barilly, I. Ahn, J. Riquet, and H. Gilbert. 2021. Genetic relationships between feed efficiency and gut microbiome in pig lines selected for residual feed intake. *J. Anim. Breed. Genet.* 138(4):491–507. doi:10.1111/jbg.12539.
- Beck, P., H.-P. Piepho, M. Rodehutsord, and J. Bennewitz. 2016. Inferring relationships between Phosphorus utilization, feed per gain, and bodyweight gain in an F2 cross of Japanese quail using recursive models. *Poult. Sci.* 95(4):764–773. doi:10.3382/ps/pev376.
- Borda-Molina, D., C. Roth, A. Hernández-Arriaga, D. Rissi, S. Vollmar, M. Rodehutsord, J. Bennewitz, and A. Camarinha-Silva. 2020. Effects on the Ileal Microbiota of Phosphorus and Calcium Utilization, Bird Performance, and Gender in Japanese Quail. *Animals (Basel)* 10(5):885. doi:10.3390/ANI10050885.
- Broman, K. W., D. M. Gatti, P. Simecek, N. A. Furlotte, P. Prins, Ś. Sen, B. S. Yandell, and G. A. Churchill. 2019. R/qtl2: Software for Mapping Quantitative Trait Loci with High-Dimensional Data and Multiparent Populations. *Genetics* 211(2):495–502. doi:10.1534/genetics.118.301595.
- Chickering, D. M. 2013. A Transformational Characterization of Equivalent Bayesian Network Structures. arXiv preprint arXiv:1302.4938.
- Christensen, O. F., V. Börner, L. Varona, and A. Legarra. 2021. Genetic evaluation including intermediate omics features. *Genetics* 219(2). doi:10.1093/genetics/iyab130.
- Clarke, K. R., and R. M. Warwick. 2001. *Change in Marine Communities: An Approach to Statistical Analysis and Interpretation*. PRIMER-E: Plymouth, UK.
- Daly, R., and S. Qiang. 2007. Methods to accelerate the learning of bayesian network structures.: *Proceedings of the 2007 UK Workshop on Computational Intelligence*. Imperial College., London.
- Déru, V., A. Bouquet, O. Zemb, B. Blanchet, M. L. de Almeida, L. Cauquil, C. Carillier-Jacquin, and H. Gilbert. 2022. Genetic relationships between efficiency traits and gut microbiota traits in growing pigs being fed with a conventional or a high-fiber diet. *J. Anim. Sci.* 100(6). doi:10.1093/jas/skac183.
- Dhakal, K., F. Tiezzi, J. S. Clay, and C. Maltecca. 2016. Causal relationships between clinical mastitis events, milk yields and lactation persistency in US Holsteins. *Livest. Sci.* 189:8–16. doi:10.1016/j.livsci.2016.04.015.

- Epskamp, S., A. O. J. Cramer, L. J. Waldorp, V. D. Schmittmann, and D. Borsboom. 2012. qgraph: Network Visualizations of Relationships in Psychometric Data. *J. Stat. Soft.* 48(4):1–18. doi:10.18637/jss.v048.i04.
- Gianola, D., and D. Sorensen. 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics* 167(3):1407–1424. doi:10.1534/genetics.103.025734.
- Gilmour, A. R., B. J. Gogel, B. R. Cullis, S. J. Welham, and R. Thompson. 2015. ASReml User Guide Release 4.1 Functional Specification. VSN International Ltd, Hemel Hempstead, HP1 1ES, UK.
- Haas, V., S. Vollmar, S. Preuß, M. Rodehutschord, A. Camarinha-Silva, and J. Bennewitz. 2022. Composition of the ileum microbiota is a mediator between the host genome and phosphorus utilization and other efficiency traits in Japanese quail (*Coturnix japonica*). *Genet. Sel. Evol.* 54(1). doi:10.1186/s12711-022-00697-8.
- Hadfield, J. D. 2010. MCMC Methods for Multi-Response Generalized Linear Mixed Models: The MCMCglmmR Package. *J. Stat. Soft.* 33(2):1–22. doi:10.18637/jss.v033.i02.
- Heckerman, D. 2008. A Tutorial on Learning with Bayesian Networks. In: D. E. Holmes, editor, *Innovations in Bayesian networks: Theory and applications*. Springer, Berlin, Heidelberg. p. 33–82.
- Inoue, K., B. D. Valente, N. Shoji, T. Honda, K. Oyama, and G. J. M. Rosa. 2016. Inferring phenotypic causal structures among meat quality traits and the application of a structural equation model in Japanese Black cattle. *J. Anim. Sci.* 94(10):4133–4142. doi:10.2527/jas.2016-0554.
- Lu, D., F. Tiezzi, C. Schillebeeckx, N. P. McNulty, C. Schwab, C. Shull, and C. Maltecca. 2018. Host contributes to longitudinal diversity of fecal microbiota in swine selected for lean growth. *Microbiome* 6(1):4. doi:10.1186/s40168-017-0384-1.
- Maltecca, C., M. Bergamaschi, and F. Tiezzi. 2020. The interaction between microbiome and pig efficiency: A review. *J. Anim. Breed. Genet.* 137(1):4–13. doi:10.1111/jbg.12443.
- Manichaikul, A., J. Dupuis, S. Sen, and K. W. Broman. 2006. Poor performance of bootstrap confidence intervals for the location of a quantitative trait locus. *Genetics* 174(1):481–489. doi:10.1534/genetics.106.061549.
- Momen, M., M. T. Campbell, H. Walia, and G. Morota. 2019. Utilizing trait networks and structural equation models as tools to interpret multi-trait genome-wide association studies. *Plant Methods* 15(1):107. doi:10.1186/s13007-019-0493-x.

- Okamura, T., K. Ishii, M. Nishio, G. J. M. Rosa, M. Satoh, and O. Sasaki. 2020. Inferring phenotypic causal structure among farrowing and weaning traits in pigs. *Anim. Sci. J.* 91(1):e13369. doi:10.1111/asj.13369.
- Pegolo, S., M. Momen, G. Morota, G. J. M. Rosa, D. Gianola, G. Bittante, and A. Cecchinato. 2020. Structural equation modeling for investigating multi-trait genetic architecture of udder health in dairy cattle. *Sci. Rep.* 10(1):7751. doi:10.1038/s41598-020-64575-3.
- Pegolo, S., H. Yu, G. Morota, V. Bisutti, G. J. M. Rosa, G. Bittante, and A. Cecchinato. 2021. Structural equation modeling for unraveling the multivariate genomic architecture of milk proteins in dairy cattle. *J. Dairy Sci.* 104(5):5705–5718. doi:10.3168/jds.2020-18321.
- Pielou, E. C. 1966. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* 13:131–144. doi:10.1016/0022-5193(66)90013-0.
- R Core Team. 2021. R: A Language and Environment for Statistical Computing: R Foundation for Statistical Computing, Vienna, Austria.
- Rodehutschord, M., V. Sommerfeld, I. Kühn, and M. R. Bedford. 2022. Phytases: Potential and Limits of Phytate Destruction in the Digestive Tract of Pigs and Poultry. In: M. R. Bedford, G. Partridge, C. L. Walk, M. Hruby, C. Evans, H. Irving, J. Vehmanperä, K. Juntunen, J. Patience, Q. Li, A. Petry, J. Lee, K. Brown, A. Cowieson, D. Menezes-Blackburn, R. Greiner, and U. Konietzny, editors, *Enzymes in Farm Animal Nutrition*. CABI, GB. p. 124–152.
- Rosa, G. J. M., B. D. Valente, G. de Los Campos, X.-L. Wu, D. Gianola, and M. A. Silva. 2011. Inferring causal phenotype networks using structural equation models. *Genet. Sel. Evol.* 43:6. doi:10.1186/1297-9686-43-6.
- Saborío-Montero, A., M. Gutiérrez-Rivas, A. García-Rodríguez, R. Atxaerandio, I. Goiri, E. López de Maturana, J. A. Jiménez-Montero, R. Alenda, and O. González-Recio. 2020. Structural equation models to disentangle the biological relationship between microbiota and complex traits: Methane production in dairy cattle as a case of study. *J. Anim. Breed. Genet.* 137(1):36–48. doi:10.1111/jbg.12444.
- Scutari, M. 2010. Learning Bayesian Networks with the bnlearn R Package. *J. Stat. Soft.* 35(3):1–22. URL <http://www.jstatsoft.org/v35/i03/>.
- Tang, S., Y. Xin, Y. Ma, X. Xu, S. Zhao, and J. Cao. 2020. Screening of Microbes Associated With Swine Growth and Fat Deposition Traits Across the Intestinal Tract. *Front. Microbiol.* 11:586776. doi:10.3389/fmicb.2020.586776.

- Tiezzi, F., J. Fix, C. Schwab, C. Shull, and C. Maltecca. 2021. Gut microbiome mediates host genomic effects on phenotypes: a case study with fat deposition in pigs. *Comput. Struct. Biotechnol. J.* 19(2-3):530–544. doi:10.1016/j.csbj.2020.12.038.
- Valente, B. D., G. J. M. Rosa, G. de Los Campos, D. Gianola, and M. A. Silva. 2010. Searching for recursive causal structures in multivariate quantitative genetics mixed models. *Genetics* 185(2):633–644. doi:10.1534/genetics.109.112979.
- Verma, T. S., and J. Pearl. 2022. Equivalence and Synthesis of Causal Models. In: *Probabilistic and Causal Inference*. ACM, New York, NY, USA. p. 221–236.
- Vollmar, S., V. Haas, M. Schmid, S. Preuß, R. Joshi, M. Rodehutsord, and J. Bennewitz. 2021. Mapping genes for phosphorus utilization and correlated traits using a 4k SNP linkage map in Japanese quail (*Coturnix japonica*). *Anim. Genet.* doi:10.1111/age.13018.
- Vollmar, S., R. Wellmann, D. Borda-Molina, M. Rodehutsord, A. Camarinha-Silva, and J. Bennewitz. 2020. The Gut Microbial Architecture of Efficiency Traits in the Domestic Poultry Model Species Japanese Quail (*Coturnix japonica*) Assessed by Mixed Linear Models. *G3 (Bethesda)* 10(7):2553–2562. doi:10.1534/g3.120.401424.
- Weishaar, R., R. Wellmann, A. Camarinha-Silva, M. Rodehutsord, and J. Bennewitz. 2020. Selecting the hologenome to breed for an improved feed efficiency in pigs-A novel selection index. *J. Anim. Breed. Genet.* 137(1):14–22. doi:10.1111/jbg.12447.
- Yang, J., S. H. Lee, M. E. Goddard, and P. M. Visscher. 2011. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88(1):76–82. doi:10.1016/j.ajhg.2010.11.011.

Supplementary Materials

Supplementary Material S1: Structure to estimate the indirect and the total SNP effects.

Pielou's evenness index (J'). Only the effects exist for J', as the Bayesian network algorithm did not find any upstream traits. The total effect of the SNP under consideration within the SI of the QTL for J' on this trait correspond to its direct effect:

$$\text{Direct}_{s \rightarrow J'} = s_{J'}$$

$$\text{Total}_{s \rightarrow J'} = \text{Direct}_{s \rightarrow J'} = s_{J'}$$

Phosphorus utilization (PU). The total SNP effect of the trait PU can be decomposed into the direct and one indirect effect, mediated by J' (J' → PU) with the structural coefficient $\lambda_{PU, J'}$.

$$\text{Direct}_{s \rightarrow PU} = s_{PU}$$

$$\text{Indirect}(1)_{s \rightarrow PU} = \lambda_{PU, J'} * s_{J'}$$

$$\text{Total}_{s \rightarrow PR} = \text{Direct}_{s \rightarrow PR} + \text{Indirect}(1)_{s \rightarrow PU} = s_{PU} + \lambda_{PU, J'} * s_{J'}$$

Phosphorus retention (PR). The effects of each SNP on PR can be broken down as follows based on the SEM association analysis results, considering the structure coefficients $\lambda_{PR, J'}$, $\lambda_{PU, J'}$, and $\lambda_{PR, PU}$ from the network:

$$\text{Direct}_{s \rightarrow PR} = s_{PR}$$

$$\text{Indirect}(1)_{s \rightarrow PR} = \lambda_{PR, J'} * s_{J'}$$

$$\text{Indirect}(2)_{s \rightarrow PR} = \lambda_{PR, PU} * s_{PU}$$

$$\text{Indirect}(3)_{s \rightarrow PR} = \lambda_{PU, J'} * \lambda_{PR, PU} * s_{J'}$$

$$\begin{aligned} \text{Total}_{s \rightarrow PR} &= \text{Direct}_{s \rightarrow PR} + \text{Indirect}(1)_{s \rightarrow PR} + \text{Indirect}(2)_{s \rightarrow PR} + \text{Indirect}(3)_{s \rightarrow PR} \\ &= s_{PR} + \lambda_{PR, J'} * s_{J'} + \lambda_{PR, PU} * s_{PU} + \lambda_{PU, J'} * \lambda_{PR, PU} * s_{J'} \end{aligned}$$

Supplementary Table S2: Trait-associated markers from GCTA ($P_{nominal} \leq 0.05$) within the significant QTL regions on *Coturnix japonica* chromosome (CJA) from the QTL linkage analyses.

Trait	CJA	Marker ID	$P_{nominal}$	Pos ^a	$P_{Bonferroni}$ ^b
Pielou's evenness index	13	id13986	0.04	1.30	0.33
	13	id13985	0.01	2.33	0.09
	13	id34020	0.01	2.33	0.08
	13	id11613	< 0.01	6.26	0.02
	13	id17634	0.05	11.07	0.38
	13	id07862	0.02	11.07	0.14
	13	id07864	0.03	12.23	0.25
	13	id28086	0.02	12.23	0.18
Phosphorus utilization	3	id14623	0.03	41.86	0.43
	3	id12416	0.01	44.17	0.09
	3	id02399	0.04	47.18	0.73
	3	id09078	0.03	48.39	0.56
	3	id15479	0.02	48.39	0.32
	3	id23567	0.05	48.40	0.77
	3	id03947	0.04	48.40	0.68
	3	id14857	< 0.01	48.40	0.03
	3	id12431	0.02	48.91	0.32
	3	id27978	0.02	48.91	0.32
	3	id21747	0.02	53.21	0.31
	3	id21189	0.02	53.21	0.37
	3	id07068	< 0.01	54.48	0.07
	3	id19369	0.02	55.21	0.36
	3	id21668	0.02	55.23	0.39
	3	id10612	0.01	62.85	0.17
3	id06736	0.02	64.48	0.26	
Phosphorus retention	3	id12522	< 0.01	102.77	0.10
	3	id14393	0.01	102.77	0.31
	3	id12523	< 0.01	103.23	0.10
	3	id07091	< 0.01	103.23	0.10
	3	id10699	< 0.01	104.44	0.07
	3	id10698	0.01	104.44	0.34
	3	id09145	< 0.01	104.52	0.07
	3	id12526	< 0.01	104.58	0.02
	3	id21801	< 0.01	104.58	0.02
	3	id12525	< 0.01	104.58	0.02
	3	id20838	< 0.01	105.53	0.07
	3	id32871	< 0.01	107.50	0.10
	3	id06546	0.04	107.97	0.96
	3	id03760	0.01	107.97	0.12
	3	id10706	0.01	107.97	0.12
	3	id02682	< 0.01	109.45	< 0.01
	3	id13856	< 0.01	109.45	0.02
	3	id23496	< 0.01	109.45	0.05
	3	id00547	< 0.01	109.73	0.02
	3	id15669	< 0.01	112.08	0.02
3	id08214	0.02	112.08	0.46	
3	id06751	0.01	112.72	0.14	

CHAPTER THREE

3	id10714	< 0.01	112.72	0.05
3	id06750	< 0.01	112.72	0.02

^a in cM; ^b Bonferroni adjusted *P*-value for the number of SNP within the respective support interval.

Supplementary Table S3: Direct, indirect, and total SNP effects of the SNP within the QTL region from the QTL mapping results calculated with SEM association analysis (Model 3) on trait phosphorus utilization (PU) and phosphorus retention (PR).

Marker ID	QTL	Direct	Indirect(1)	Indirect(2)	Indirect(3)	Total	P_{nominal}^a
id14623	PU	0.155	-0.003			0.152	0.01
id12416		0.139	-0.004			0.135	0.02
id02399		0.110	-0.002			0.109	0.15
id09078		0.094	0.005			0.100	0.13
id15479		0.124	0.001			0.125	0.03
id23567		0.109	-0.002			0.106	0.16
id03947		0.110	-0.001			0.109	0.16
id14857		0.233	-0.020			0.213	0.03
id12431		0.155	-0.002			0.154	0.03
id27978		0.122	0.000			0.122	0.03
id21747		0.137	0.005			0.142	0.15
id21189		0.101	-0.006			0.095	0.09
id07068		0.166	0.002			0.169	0.01
id19369		0.140	-0.004			0.136	0.05
id21668		0.136	0.004			0.141	0.15
id10612		0.107	0.004			0.111	0.19
id06736		0.096	0.004			0.100	0.21
id12522	PR	0.080	0.000	0.038	0.001	0.120	0.02
id14393		0.086	0.000	0.030	-0.001	0.114	0.02
id12523		0.080	0.000	0.041	0.001	0.123	0.02
id07091		0.080	0.000	0.041	0.001	0.123	0.02
id10699		0.100	0.000	0.018	0.002	0.120	< 0.01
id10698		0.064	0.001	0.055	0.004	0.124	0.08
id09145		0.060	0.000	0.086	0.003	0.149	0.12
id12526		0.092	0.000	0.046	0.000	0.138	0.01
id21801		0.100	0.000	0.033	0.002	0.135	< 0.01
id12525		0.090	0.000	0.047	0.000	0.138	0.01
id20838		0.060	0.000	0.085	0.002	0.147	0.12
id32871		0.070	0.000	0.059	0.004	0.133	0.05
id06546		0.068	0.000	0.030	0.001	0.099	0.06
id03760		0.050	0.000	0.035	-0.002	0.083	0.17
id10706		0.044	0.000	0.041	0.002	0.087	0.23
id02682		0.085	0.000	0.083	0.002	0.170	0.01
id13856		0.079	0.000	0.088	0.001	0.168	0.02
id23496		0.065	0.000	0.087	0.000	0.152	0.07
id00547		0.069	0.000	0.088	0.002	0.159	0.05
id15669		0.080	0.000	0.107	0.002	0.189	0.02
id08214		0.076	0.000	0.063	0.001	0.141	0.02
id06751		0.101	0.000	0.070	0.000	0.170	0.03
id10714		0.048	0.000	0.095	0.002	0.145	0.18
id06750		0.078	0.000	0.109	0.002	0.190	0.02

^a calculated P -value from Model 3 for the direct SNP effects

CHAPTER FOUR

Invited review: Toward the use of host microbiota interplay in poultry and pig breeding

Valentin Haas* and Jörn Bennewitz

Institute of Animal Science, University of Hohenheim, Stuttgart, Germany

* Corresponding author: valentin.haas@uni-hohenheim.de

Published in *Lohmann Information* (2023)

<https://lohmann-breeders.com/lohmanninfo/toward-the-use-of-host-microbiota/>

Abstract

The influence of microorganisms in the gastrointestinal tracts of nonruminants has long been considered nonexistent or almost nonexistent. However, past studies have demonstrated the colonization of gut microbiota and its potential influence on efficiency traits in a wide variety of livestock species. Since efficiency traits are currently becoming more popular, the present study addresses the question of how such microbial information from nonruminants can be used in scientific approaches to animal genetics and breeding in the subject area of animal efficiency and performance. The main current statistical methods and models that consider gut microbial colonization will be explained. Ways that quantify microbial influences on quantitative traits, that explain hologenomic (host genome plus microbiota genome) approaches and that consider causal relationships between microbial features along with efficiency traits and host genetics, are presented here.

Introduction

The efficiency of livestock production is of growing interest, especially in terms of resource conservation, environmental protection, animal welfare and food-feed competition. Feed efficiency can be divided into digestive and metabolic efficiency (Martin et al., 2021; PUILLET et al., 2016). Digestive efficiency reflects the ability to absorb nutrients from ingested feed in the gastrointestinal tract (GIT) into the bloodstream, whereas metabolic efficiency is affected by allocation and reallocation processes, i.e., the conversion of absorbed nutrients into animal products.

The GITs of poultry harbor a variety of microorganisms (Apajalahti et al., 2004) influenced by external factors such as litter or diet (Borda-Molina et al., 2018; Kers et al., 2018) and by host genetics (Haas et al., 2022; Meng et al., 2014; Mignon-Grasteau et al., 2015; Wen et al., 2021; Zhao et al., 2013). These members of the bacterial microbiota are not silent roommates, as they live in symbiosis with the host and are therefore involved in a number of processes in the digestive efficiency pathway, e.g., efficiency, utilization of nutrients, immune system and animal health (Maki et al., 2019; Rodehutsord et al., 2022; Stanley et al., 2014; Yadav and Jha, 2019). Figure 1 shows a summary of factors influencing the colonization of the GITs by microorganisms.

The consideration of microorganisms in the GITs of nonruminants can help improve the digestibility of various nutrients, as well as performance traits, and it seems to be beneficial to consider gastrointestinal microbiota in animal breeding (e.g., Haas et al., 2022; Khanal et al., 2020; Lu et al., 2018; Maltecca et al., 2019; Weishaar et al., 2020). A deeper look at noteworthy statistical approaches in animal breeding that consider the bacterial colonization in the GITs of nonruminants (pigs and poultry) is compiled in this review. This review covers major microbial and genetic approaches but does not claim to be exhaustive.

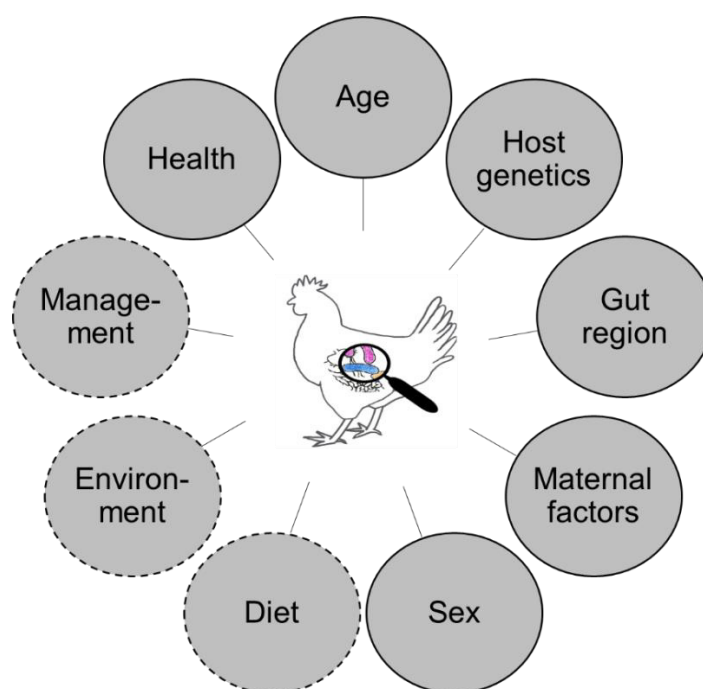


Figure 1 Factors influencing the animal gut microbiota. Full circles represent the host effects, and dashed circles represent the external influencing factors.

Microbial Data

The interplay between the host, its gut microbiota and efficiency is complex and not yet fully understood. Nevertheless, the use of microbiota data has arrived in animal science and

breeding. To use microbiota in genetic-statistical analyses, a powerful source of information is needed. In studies, mainly targeted amplicon sequencing of the 16S ribosomal RNA (rRNA) gene is applied (Borda-Molina et al., 2018; Maltecca et al., 2020). For this, the DNA is extracted, and a specific target region (amplicons) of the small ribosomal subunit RNA gene is amplified. The 16S rRNA genes are conserved in microorganisms and contain a hypervariable region that allows for delineation of the microorganisms. The resulting sequences are then clustered into operational taxonomic units (OTUs) based on bioinformatic processing using a similarity threshold, or the sequences are differentiated into amplicon sequence variants (ASVs) based on single nucleotide changes. Subsequently, the sequences for subordinate ranks (e.g., phylum, genus) can be taxonomically assigned. Therefore, deep characterization of microbiota communities and their quantification via relative abundances can be achieved.

Another common technique is metagenomic shotgun sequencing (whole-genome sequencing) (Borda-Molina et al., 2018; Maltecca et al., 2020; Pérez-Cobas et al., 2020). This approach allows the parallel sequencing of DNA from all microorganisms in the ecosystem (GIT sample) with a high degree of coverage for species differentiation. Metagenomics enables the collection of genomes and their corresponding genes and allows for the characterization of potential bacterial functions.

Since inexpensive and efficient sequencing methods to quantify the gut microbiota exist, the number of host microbiota studies has increased (Guevarra et al., 2019). The taxonomic classification of microbial data from the phylum to strain level is shown in Figure 2.

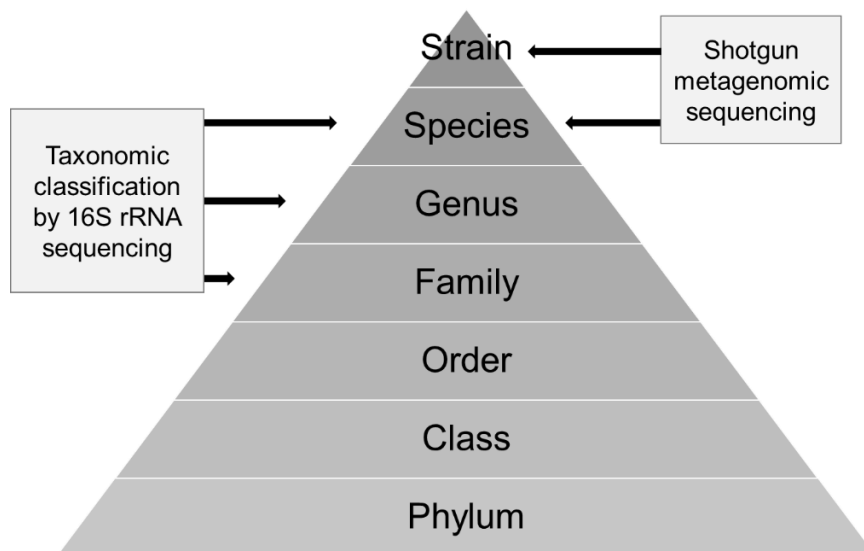


Figure 2 Taxonomic classification of microbial data from the phylum to strain level with the characterization potential.

The GIT of poultry has various sections, with different digestion functions, and thus a differentiated settlement of microorganisms. The concentration of microbes increases in number from section to section and reaches its maximum in the paired ceca and the rectum (Yadav and Jha, 2019), and the distribution of the species in the different GIT sections is also different due to different habitat conditions (Yegani and Korver, 2008). For example, the crop, jejunum, and ileum are more likely to be home to bacteria of the family *Lactobacillaceae* and the caeca to the family *Ruminococcaceae* (Witzig et al., 2015).

Microbial data are given as compositional data (in relative abundances in percent), which means that the data per animal are multivariate and have a unit sum, i.e., an animal has 100 percent over all detected microbial features. Therefore, many animals in the microbiota dataset have a relative abundance of zero for some microbial characteristics, which complicates the use of microbial data for some analyses.

Heritabilities and host genetic architecture of microbial features

As stated above, the microbial colonization of the GITs of nonruminants is influenced by host genetics. Some studies have estimated significant narrow sense heritabilities for different GIT sections and different microbial features. Without any differentiation of the study environments and designs, the heritabilities differ at the GIT section and the microbial database used. For example, heritabilities for bacterial genera in pig colon samples are in the range of 0.32 to 0.57 (Camarinha-Silva et al., 2017) and bacterial genera in fecal samples are in the range of 0.07 to 0.33 (Chen et al., 2018); in pig colons at the OTU level, the range is between 0.03 to 0.55 (Bergamaschi et al., 2020). Similar significant heritabilities could be found for chicken microbiota in fecal samples at the genus level, ranging from 0.21 to 0.79 (Meng et al., 2014), and Wen et al. (2021) reported different heritabilities in different sections of the GIT at the genus level: in the duodenum from 0.44 to 0.62, in the jejunum from 0.31 to 0.44, in the ileum from 0.38 to 0.79, in the cecum from 0.36 to 0.87, and in the feces from 0.41 to 0.71. Lower heritabilities were reported in quail ileum at the genus level, from 0.04 to 0.17 (Haas et al., 2022). Additionally, heritabilities of the diversity (i.e., alpha diversity index) of the microbiota composition in pigs and poultry are shown in the literature in a range of 0.15 to 0.26, depending on the animal species, the respective GIT section, and the study design (Aliakbari et al., 2021; Déru et al., 2022a; Haas et al., 2023; Lu et al., 2018).

In addition to narrow-sense heritability, genetic correlation is also an important parameter in animal breeding. Several studies have estimated low to high genetic correlations between different microbial features and host efficiency traits. Aliakbari et al. (2021) and Déru et al. (2022a) found, for example, significant genetic correlations in pigs, Mignon-Grasteau et al. (2015) in chickens, and Haas et al. (2022) in Japanese quail. Some studies have reported

significant genomic regions for different microbial features (e.g., as genera, OTU, or alpha diversity index) and different animal species by using genome-wide association studies (GWAS) and quantitative trait loci (QTL) analyses (e.g., Haas et al., 2022; Haas et al., 2023; Wen et al., 2021).

Microbial relationship matrix \mathbf{M} and microbiability

In animal breeding, heritability of the narrow sense is used to describe the relationship between the additive genotypic values of a trait and the phenotypic recording in a population. To quantify the relationship between the GIT microbiota and the recorded phenotype, the parameter microbiability was introduced by Difford et al. (2016). Microbiability describes the part of the phenotypic variation of a trait that is explained by the microbial composition in the GIT. The approach to calculate microbiability, i.e., the proportion of microbial variance in the phenotypic variance, is equivalent to calculating heritability using the random animal effect in a linear mixed model. However, the random animal effect is modeled by a covariance structure from a microbial relationship matrix \mathbf{M} rather than the pedigree-based/genomic relationship matrix used in the calculation of heritability. Different methods exist to build a microbial relationship matrix (He et al., 2022). A commonly used form is:

$$\mathbf{M} = 1/N\mathbf{X}\mathbf{X}^T,$$

with matrix \mathbf{X} ($n \times N$ matrix, with n the number of animals and N number of microbial features), which contains the standardized and log-transformed abundances of the microbial features (Camarinha-Silva et al., 2017).

The animal microbiota correlation between two quantitative traits can be estimated with bivariate or multivariate microbial linear mixed models. Here, one does not use the pedigree-based/genomic relationship matrix as in genetic correlations but the microbial relationship matrix \mathbf{M} . Medium to high animal microbiota correlations were found between different efficiency traits in quail (Vollmar et al., 2020) and pigs (Aliakbari et al., 2022; Déru et al., 2022b).

Many studies have confirmed that the microbiome explains a substantial part of the phenotypic variation in efficiency traits (e.g., Verschuren et al., 2020 in pigs, Vollmar et al., 2020 in quail, and Wen et al., 2019 in chickens), and the animal microbial correlations are mostly significant and at high levels. This enables the use of gut microbiota as a potential predictor of complex traits in animals.

Microbiome-wide association analysis (MWAS)

Classical GWAS are used to detect trait-associated SNPs via mixed linear models. It is a common tool for the detection of genomic regions involved in the expression of quantitative

traits (Schmid and Bennewitz, 2017). A method to detect trait-associated microbial features is microbiome-wide association studies (MWAS) (Tiezzi et al., 2021; Vollmar et al., 2020). The relative abundance of a single microbial feature can be implemented as a fixed effect (fixed covariate) in a mixed linear model. Vollmar et al. (2020) modeled the random genetic animal effect in MWAS to model the population structure as in a GWAS.

The MWAS approach was designed to find parts of the respective microbiota class that are associated with the trait under consideration. The study of Vollmar et al. (2020) clearly showed that some substantial peaks could be found at the genus level. It should be noted that some microbial parts affected more than one efficiency trait, and some bacteria contributed more than others to the overall phenotypic variance of one trait. However, a polymicrobial influence is emerging in all traits. This method can be used to identify trait-associated bacterial features. Aliakbari et al. (2022) confirmed the polymicrobial MWAS results for feed efficiency and performance traits in pigs, and Wang et al. (2022) found higher effects for some trait-associated genera on the fat composition of pigs. It seems as if the identified microbial features are causal. However, the use of compositional data limits the identification of causality using MWAS results (Vollmar et al., 2020).

Microbial trait predictions

When the microbiability of a quantitative trait is greater than zero, microbiome information can be used to predict the trait phenotypes by microbial best linear unbiased prediction (MBLUP) (Camarinha-Silva et al., 2017) or by machine learning approaches (e.g., Maltecca et al., 2019). This requires a large reference population comparable to genomic selection.

In most studies, MBLUP has similar or even higher prediction accuracy than comparable genomic BLUP for the same feed efficiency characteristics (e.g., Verschuren et al., 2020, Weishaar et al., 2020, Haas et al., 2022). The main difference between microbial and genomic trait predictions is that the microbial predictions are not stable because the microbiota composition varies with the section of the GIT, animal age, and environmental factors (see Figure 1) (Maltecca et al., 2020; Weishaar et al., 2020). Part of the microbial composition is determined by host genetics. Hence, there exists an overlap when selecting either microbial or genomic prediction values (Ross and Hayes, 2022).

Hologenomic Approach

Host genetics and the gastrointestinal microbiota can influence quantitative traits. Simultaneously, the microbiota composition in the GIT is influenced by host genetics, which means that there exists a relationship between these three features (Figure 3). In most of the studies, these pathways were considered separately, either as genomic or microbial phenotypic predictions. However, we know that genetic and microbial variance are not

independent, and therefore, to consider both direct genetic effects on a phenotype and indirect genetic effects via the microbiota, it may be beneficial to consider them together in a hologenomic approach. Therefore, the hologenome represents the entirety of the DNA of the host (host genome) and the DNA of the gut microbiota (metagenome) (Bordenstein and Theis, 2015; Estellé, 2019).

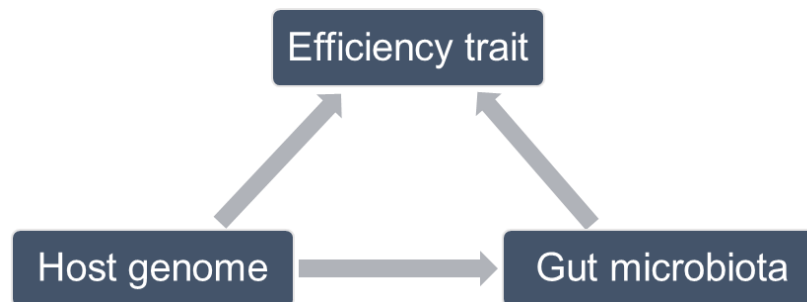


Figure 3 Relationships between efficiency traits, host genetics, and gut microbiota.

The simplest conceivable method would be to put the genomic and microbial relationship matrices together in a mixed linear model, as was done, for example, in Aliakbari et al. (2022) and Déru et al. (2022b). However, the problem is that interactions with independent covariates between markers and microbial features are not computable. Several studies have developed different methods to address this problem, and a brief overview is given in the following.

Weishaar et al. (2020) developed a two-step method that first used a microbial linear mixed model to predict the animal microbiota effect of the respective trait from microbial abundances and then a genomic linear mixed model to predict SNP effects for the previously predicted animal microbiota effect of the trait. This method does not consider the genetic effects that directly affect the corresponding trait (not via the microbiome). For this reason, the authors used a selection index to separate the proportion of genes that influence the microbiota effects of a trait and the proportion of genes that directly influence the trait but not the microbes. This addresses the genes that explain the trait not only directly through a change in metabolic pathways but also indirectly through a change in the composition of the gut microbiota. This method was implemented to place more weight on the breeding values explained by the digestive pathway rather than the metabolic pathway, and selective breeding becomes possible. A related method was developed by Christensen et al. (2021). In their study, the authors decomposed the breeding value of a quantitative trait into two genetic effects by directly estimating the microbiota-mediated breeding value and the residual breeding value of a trait, i.e., the genetic effects of the trait, without the genetic effects via the microbiome. The difference between the two methods is that Christensen et al. (2021) estimated the residual breeding value in a linear mixed model, whereas Weishaar et al. (2020) did so in a final step via the selection index. Khanal et al. (2020), Pérez-Enciso et al. (2021) and Qadri et al. (2022)

considered both microbial features and host genetics together in one model and implemented different forms of interactions between host genomes and microbiomes.

The hologenomic selection approach mentioned above might enable the differentiated selection of the acquisition and reallocation process in animal efficiency. The goal is to adapt the trait-specific microbial colonization in the GIT to put more weight on the acquisition pathway for better utilization or usability from feed components that can then be used in the bloodstream. This should somewhat reduce selection for metabolic efficiency, which creates a discrepancy between animal performance and health (Huber, 2018). However, metabolites produced by microorganisms can also alter the metabolic pathways of animals by producing short chain fatty acids. Thus, microbiota can also affect the metabolic effect, not just the digestive effect of efficiency (Weishaar et al., 2020). Hologenomic approaches considering only single traits and the interactions with other phenotypic expressions are of course not considered, and this still requires much research, but a first foundation to implement this in breeding practice has been developed. In general, across the different studies, an improvement in the accuracy of trait prediction was observed when microbial data were included (e.g., Déru et al., 2022b; Khanal et al., 2020). However, there is still much room for improvement in the statistical representation of the covariance between the gastrointestinal microbiota and host genetics in terms of phenotypic traits.

The composition of the gut microbiota is partially under the genetic control of the host and can be used to improve genomic predictions of efficiency traits. Thus, it appears that the gut microbiota acts as a mediator between host genetics and efficiency traits, and it seems that the host genetic effect of efficiency traits is composed of a direct genetic effect of the trait and an indirect genetic effect via the modeling of the gut microbiota composition.

Are Gut Microbes the Causal Drivers of Efficiency Traits?

A tool to differentiate direct and indirect effects between traits is structural equation models (SEM), as introduced in animal breeding by Gianola and Sorensen (2004). These authors used structural coefficients between trait combinations in multivariate mixed model equations to estimate the rate of change of trait i through the recursive influence of trait j , i.e., the rate of change of trait i by the change of one unit of trait j (Figure 4).

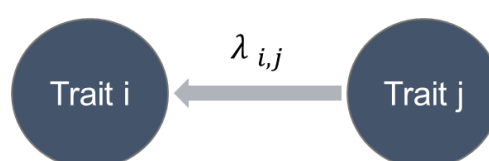


Figure 4 Recursive (directional) relationship between traits.

The combination of phenotypes, host genetics, and microbial data in structural equation concepts was done in most cases by using a linear mixed model approach, where the microbial features were considered as phenotypic trait records. First, it was published by Saborío-Montero et al. (2020) in methane emissions in dairy cows and by Haas et al. (2022) in phosphorus utilization and related traits in Japanese quail as a model species. Low to high unidirectional relationships were found depending on the host species and traits under consideration. Tiezzi et al. (2021) also confirmed direct and microbiome-mediated host genomic effects on backfat traits in swine by using GWAS and MWAS in a causal framework.

Looking at a pool of different correlated phenotypes, dependencies or causal relationships between the traits are not easy to identify, e.g., does the animal consume more food to perform as genetically predisposed or does the animal perform because of increased feed intake. Bayesian networks can be used to detect such complex directional relationships between characteristics with a large dataset. Model residuals are used to remove any confounding variables such as gender, herd effect, test day effect, additive genetic effects, etc., that might distort the network (Rosa et al., 2011). By using a Bayesian learning algorithm, we found a stable causal network for three efficiency traits (body weight gain, phosphorus, and calcium utilization), one bone ash trait (tibia ash), and two microbial phyla (*Firmicutes* and *Actinobacteria*) in an F2-cross of 750 Japanese quail (Figure 5). Numbers on the arrows are bootstrap sample proportions (50,000 samples) indicating an existing arc (figures on the left) and the proportion of the direction (figures on the right). The causal structure between the efficiency and ash traits is logical from a biological point of view, but the connection with microbial features is novel.

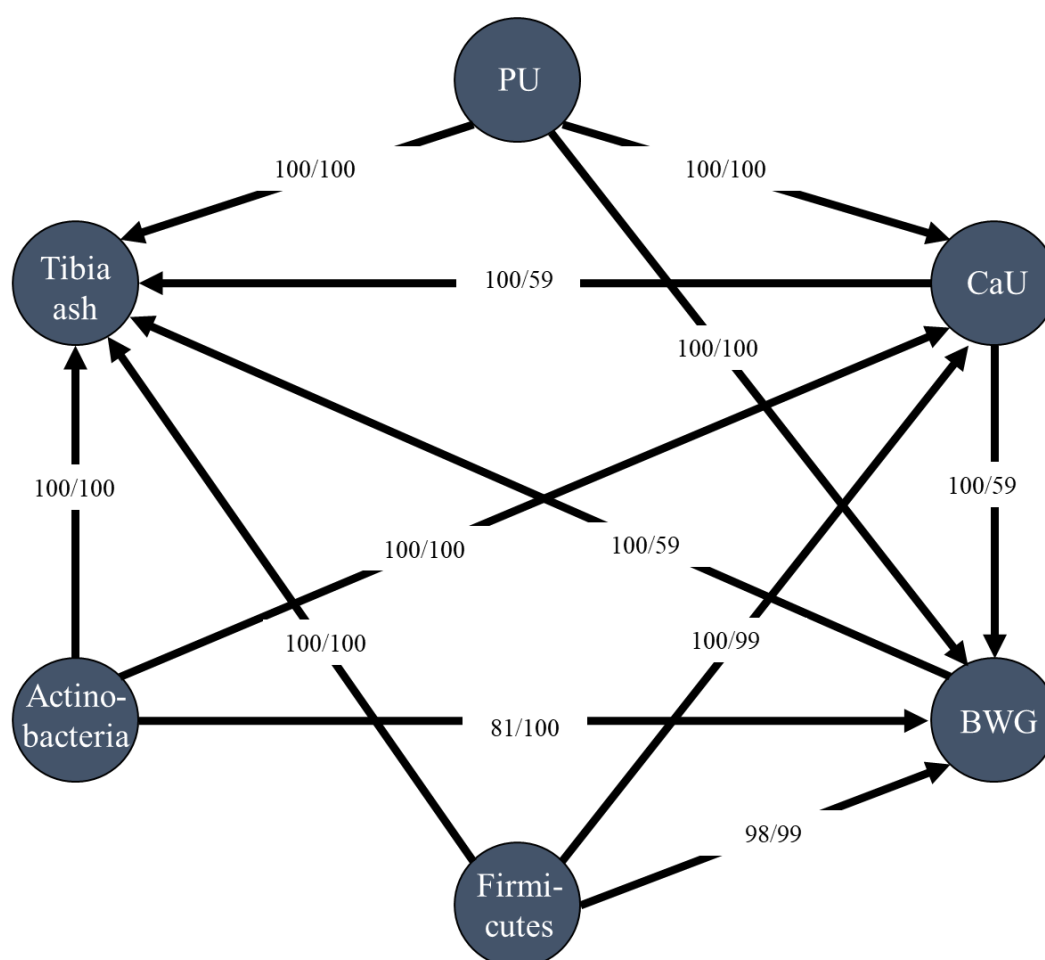


Figure 5 Causal network obtained from a Bayesian network learning algorithm (PU: phosphorus utilization, CaU: calcium utilization, BWG: body weight gain).

Such causal networks can then be used in conjunction with SEM to quantify unidirectional relationships, which shows how host genomes can be used to map and understand causal structures among quantitative and microbiota traits. These results of the SEM can then be extended to SEM-GWAS on the basis of the causal network in connection to a GWAS, where the differentiation of SNP effects into direct and indirect effects can be performed (Momen et al., 2019; Pegolo et al., 2020). Many indirect SNP effects, which directly influence an upstream trait and exert an indirect influence via the recursive relationship, are probably not detected by a classical GWAS (Tiezzi et al., 2021).

Outlook: Further Trait Complexes for Hologenomic Breeding

The studies undertaken thus far in poultry and pigs have mainly considered feed or nutrient efficiency-related traits. Considering the effects of the microbiota on animal health, microbial colonization in the gut is closely linked to the immune system, e.g., the energy supply of the immune system through the production of short-chain fatty acids and the anti-infection barrier by inhibiting pathogens (Diaz Carrasco et al., 2019; Shang et al., 2018). Future breeding

strategies for improved immune systems could consider microbial colonization in the GIT and stabilize a natural barrier to unwanted microorganisms, especially in the gut habitat (Yegani and Korver, 2008). For example, in poultry production, *Salmonella* and *Campylobacter* are unwanted microorganisms with zoonotic potential and some antibiotic resistance (Shang et al., 2018). Other studies have considered the GIT microbiota in terms of complex behavioral characteristics such as feather pecking or aggressive pecking (e.g., van der Eijk et al., 2020, Borda-Molina et al., 2021, Mindus et al., 2021). The authors argued that certain microorganisms in the GITs of animals can influence behavior or stress response via hormonal changes through the gut-brain axis. A broader consideration of social or hormonally influenced behaviors in relation to microbiota composition may become more important in the future. However, such concepts still need substantial research.

Further development or increased research in the field of hologenomic selection for metabolic and digestive efficiency could possibly also enable a better understanding of the current welfare debate in poultry breeding. An example could be keel bone fractures in laying hens, since phosphorus utilization is influenced by the gut microbiota (Haas et al., 2022; Vollmar et al., 2020) and tibia and foot ash have significant microbiabilities (not yet published).

The collection of gastrointestinal microbiota data for trait prediction is not easy to capture, and for some phenotypic observations, the useability of simple fecal samples is restricted. The main absorption in the GIT takes place in the small intestine. However, microbiota sampling in this section on live animals is invasive or can only be cultured from dead animals, which is why fecal samples are used in many studies (e.g., Déru et al., 2022b). A potential way would be the prediction of microbial information of the desired gut section via microbiota from noninvasive sampling, e.g., cloaca samples. Andreani et al. (2020) found, for example, that cloaca samples are good for predicting cecal microbiota in chickens, which needs further investigation. Another solution would be to collect only standardized DNA samples from selection candidates and, together with a large reference population with known microbiota, host DNA, and phenotypic information, predict the phenotypes of the selection candidate. This restricts one to what is possible from a breeding point of view. A way to do so is the hologenomic selection approach discussed above. In the future, technical sampling solutions such as osmotic pills, which are inserted into the GIT by mouth and can take samples in the desired section using an external magnet, may play a role (Rezaei Nejad et al., 2019).

Conclusion

Part of the gastrointestinal microbiota is influenced by host genetics, which allows for breeding. The microbial composition in turn influences some quantitative traits, and the microbiota thus acts as a mediator between host genetics and quantitative traits. The consideration of the microbial composition could therefore be worthwhile in breeding approaches. Challenges include the expensive and laborious microbial phenotyping of high numbers of animals and the lack of understanding of the relationship between complex quantitative traits and the GIT microbiota.

References

- Aliakbari, A., O. Zemb, Y. Billon, C. Barilly, I. Ahn, J. Riquet, and H. Gilbert. 2021. Genetic relationships between feed efficiency and gut microbiome in pig lines selected for residual feed intake. *J. Anim. Breed. Genet.* 138(4):491–507.
- Aliakbari, A., O. Zemb, L. Cauquil, C. Barilly, Y. Billon, and H. Gilbert. 2022. Microbiability and microbiome-wide association analyses of feed efficiency and performance traits in pigs. *Genet. Sel. Evol.* 54(1):29.
- Andreani, N. A., C. J. Donaldson, and M. Goddard. 2020. A reasonable correlation between cloacal and cecal microbiomes in broiler chickens. *Poult. Sci.* 99(11):6062–6070.
- Apajalahti, J., A. Kettunen*, and H. Graham. 2004. Characteristics of the gastrointestinal microbial communities, with special reference to the chicken. *worlds poult sci j* 60(2):223–232.
- Bergamaschi, M., C. Maltecca, C. Schillebeeckx, N. P. McNulty, C. Schwab, C. Shull, J. Fix, and F. Tiezzi. 2020. Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. *Sci. Rep.* 10(1):10134.
- Borda-Molina, D., H. Iffland, M. Schmid, R. Müller, S. Schad, J. Seifert, J. Tetens, W. Bessei, J. Bennewitz, and A. Camarinha-Silva. 2021. Gut microbial composition and predicted functions are not associated with feather pecking and antagonistic behavior in laying hens. *Life (Basel, Switzerland)* 11(3).
- Borda-Molina, D., J. Seifert, and A. Camarinha-Silva. 2018. Current perspectives of the chicken gastrointestinal tract and its microbiome. *Comput. Struct. Biotechnol. J.* 16:131–139.
- Bordenstein, S. R., and K. R. Theis. 2015. Host biology in light of the microbiome: Ten principles of holobionts and hologenomes. *PLoS biology* 13(8):e1002226.
- Camarinha-Silva, A., M. Maushammer, R. Wellmann, M. Vital, S. Preuss, and J. Bennewitz. 2017. Host genome influence on gut microbial composition and microbial prediction of complex traits in pigs. *Genetics* 206(3):1637–1644.
- Chen, C., X. Huang, S. Fang, H. Yang, M. He, Y. Zhao, and L. Huang. 2018. Contribution of host genetics to the variation of microbial composition of cecum lumen and feces in pigs. *Front. Microbiol.* 9:2626.
- Christensen, O. F., V. Börner, L. Varona, and A. Legarra. 2021. Genetic evaluation including intermediate omics features. *Genetics* 219(2).

- Déru, V., A. Bouquet, O. Zemb, B. Blanchet, M. L. de Almeida, L. Cauquil, C. Carillier-Jacquín, and H. Gilbert. 2022a. Genetic relationships between efficiency traits and gut microbiota traits in growing pigs being fed with a conventional or a high-fiber diet. *J. Anim. Sci.* 100(6).
- Déru, V., F. Tiezzi, C. Carillier-Jacquín, B. Blanchet, L. Cauquil, O. Zemb, A. Bouquet, C. Maltecca, and H. Gilbert. 2022b. Gut microbiota and host genetics contribute to the phenotypic variation of digestive and feed efficiency traits in growing pigs fed a conventional and a high fiber diet. *Genet. Sel. Evol.* 54(1):55.
- Díaz Carrasco, J. M., N. A. Casanova, and M. E. Fernández Miyakawa. 2019. Microbiota, gut health and chicken productivity: What is the connection? *Microorganisms* 7(10).
- Difford, G., J. Lassen, and P. Lovendahl. 2016. Genes and microbes, the next step in dairy cattle breeding. In: Wageningen Academic Publishers, editor, *Book of Abstracts of the 67th Annual Meeting of the European Federation of Animal Science*. p. 285.
- Estellé, J. 2019. Benefits from the joint analysis of host genomes and metagenomes: Select the holobiont. *J. Anim. Breed. Genet.* 136(2):75–76.
- Gianola, D., and D. Sorensen. 2004. Quantitative genetic models for describing simultaneous and recursive relationships between phenotypes. *Genetics* 167(3):1407–1424.
- Guevarra, R. B., J. H. Lee, S. H. Lee, M.-J. Seok, D. W. Kim, B. N. Kang, T. J. Johnson, R. E. Isaacson, and H. B. Kim. 2019. Piglet gut microbial shifts early in life: causes and effects. *J. Anim. Sci. Biotechnol.* 10(1):1.
- Haas, V., M. Rodehutschord, A. Camarinha-Silva, and J. Bennewitz. 2023. Inferring causal structures of gut microbiota diversity and feed efficiency traits in poultry using Bayesian learning and genomic structural equation models. *J. Anim. Sci.* 101. skad044.
- Haas, V., S. Vollmar, S. Preuß, M. Rodehutschord, A. Camarinha-Silva, and J. Bennewitz. 2022. Composition of the ileum microbiota is a mediator between the host genome and phosphorus utilization and other efficiency traits in Japanese quail (*Coturnix japonica*). *Genet. Sel. Evol.* 54(1).
- He, Y., F. Tiezzi, J. Jiang, J. Howard, Y. Huang, K. Gray, J.-W. Choi, and C. Maltecca. 2022. Exploring methods to summarize gut microbiota composition for microbiability estimation and phenotypic prediction in swine. *J. Anim. Sci.* 100(9).
- Huber, K. 2018. Invited review: resource allocation mismatch as pathway to disproportionate growth in farm animals - prerequisite for a disturbed health. *Animal an international journal of animal bioscience* 12(3):528–536.

- Kers, J. G., F. C. Velkers, E. A. J. Fischer, G. D. A. Hermes, J. A. Stegeman, and H. Smidt. 2018. Host and environmental factors affecting the intestinal microbiota in chickens. *Front. Microbiol.* 9:235.
- Khanal, P., C. Maltecca, C. Schwab, J. Fix, M. Bergamaschi, and F. Tiezzi. 2020. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. *Genet. Sel. Evol.* 52(1):41.
- Lu, D., F. Tiezzi, C. Schillebeeckx, N. P. McNulty, C. Schwab, C. Shull, and C. Maltecca. 2018. Host contributes to longitudinal diversity of fecal microbiota in swine selected for lean growth. *Microbiome* 6(1):4.
- Maki, J. J., C. L. Klima, M. J. Sylte, and T. Looft. 2019. The microbial pecking order: Utilization of intestinal microbiota for poultry health. *Microorganisms* 7(10).
- Maltecca, C., M. Bergamaschi, and F. Tiezzi. 2020. The interaction between microbiome and pig efficiency: A review. *J. Anim. Breed. Genet.* 137(1):4–13.
- Maltecca, C., D. Lu, C. Schillebeeckx, N. P. McNulty, C. Schwab, C. Shull, and F. Tiezzi. 2019. Predicting growth and carcass traits in swine using microbiome data and machine learning algorithms. *Sci. Rep.* 9(1):6574.
- Martin, P., V. Ducrocq, P. Faverdin, and N. C. Friggens. 2021. Invited review: Disentangling residual feed intake-Insights and approaches to make it more fit for purpose in the modern context. *J. Dairy Sci.* 104(6):6329–6342.
- Meng, H., Y. Zhang, L. Zhao, W. Zhao, C. He, C. F. Honaker, Z. Zhai, Z. Sun, and P. B. Siegel. 2014. Body weight selection affects quantitative genetic correlated responses in gut microbiota. *PLoS One* 9(3):e89862.
- Meuwissen, T., B. Hayes, and M. Goddard. 2016. Genomic selection: A paradigm shift in animal breeding. *Anim. Fron.* 6(1):6–14.
- Mignon-Grasteau, S., A. Narcy, N. Rideau, C. Chantry-Darmon, M.-Y. Boscher, N. Sellier, M. Chabault, B. Konsak-Ilievski, E. Le Bihan-Duval, and I. Gabriel. 2015. Impact of selection for digestive efficiency on microbiota composition in the chicken. *PLoS One* 10(8):e0135488.
- Mindus, C., N. van Staaveren, D. Fuchs, J. M. Gostner, J. B. Kjaer, W. Kunze, M. F. Mian, A. K. Shoveller, P. Forsythe, and A. Harlander-Matauschek. 2021. *L. rhamnosus* improves the immune response and tryptophan catabolism in laying hen pullets. *Sci. Rep.* 11(1):19538.

- Momen, M., M. T. Campbell, H. Walia, and G. Morota. 2019. Utilizing trait networks and structural equation models as tools to interpret multi-trait genome-wide association studies. *Plant Methods* 15(1):107.
- Pegolo, S., M. Momen, G. Morota, G. J. M. Rosa, D. Gianola, G. Bittante, and A. Cecchinato. 2020. Structural equation modeling for investigating multi-trait genetic architecture of udder health in dairy cattle. *Sci. Rep.* 10(1):7751.
- Pérez-Cobas, A. E., L. Gomez-Valero, and C. Buchrieser. 2020. Metagenomic approaches in microbial ecology: an update on whole-genome and marker gene sequencing analyses. *Microb. Genom.* 6(8).
- Pérez-Enciso, M., L. M. Zingaretti, Y. Ramayo-Caldas, and G. de Los Campos. 2021. Opportunities and limits of combining microbiome and genome data for complex trait prediction. *Genet. Sel. Evol.* 53(1):65.
- Puillet, L., D. Réale, and N. C. Friggens. 2016. Disentangling the relative roles of resource acquisition and allocation on animal feed efficiency: insights from a dairy cow model. *Genet. Sel. Evol.* 48(1):72.
- Qadri, Q. R., Q. Zhao, X. Lai, Z. Zhang, W. Zhao, Y. Pan, and Q. Wang. 2022. Estimation of complex-trait prediction accuracy from the different holo-omics interaction models. *Genes* 13(9).
- Rezaei Nejad, H., B. C. M. Oliveira, A. Sadeqi, A. Dehkharghani, I. Kondova, J. A. M. Langermans, J. S. Guasto, S. Tzipori, G. Widmer, and S. R. Sonkusale. 2019. Ingestible osmotic pill for in vivo sampling of gut microbiomes. *Adv. Intell. Syst.* 1(5):1900053.
- Rodehutsord, M., V. Sommerfeld, I. Kühn, and M. R. Bedford. 2022. Phytases: Potential and limits of phytate destruction in the digestive tract of pigs and poultry. In: M. R. Bedford, G. Partridge, C. L. Walk, M. Hruby, C. Evans, H. Irving, J. Vehmanperä, K. Juntunen, J. Patience, Q. Li, A. Petry, J. Lee, K. Brown, A. Cowieson, D. Menezes-Blackburn, R. Greiner, and U. Konietzny, editors, *Enzymes in farm animal nutrition*. CABI, GB. p. 124–152.
- Rosa, G. J. M., B. D. Valente, G. de Los Campos, X.-L. Wu, D. Gianola, and M. A. Silva. 2011. Inferring causal phenotype networks using structural equation models. *Genet. Sel. Evol.* 43:6.
- Ross, E. M., and B. J. Hayes. 2022. Metagenomic predictions: A review 10 years on. *Front. Genet.* 13:865765.

- Saborío-Montero, A., M. Gutiérrez-Rivas, A. García-Rodríguez, R. Atxaerandio, I. Goiri, E. López de Maturana, J. A. Jiménez-Montero, R. Alenda, and O. González-Recio. 2020. Structural equation models to disentangle the biological relationship between microbiota and complex traits: Methane production in dairy cattle as a case of study. *J. Anim. Breed. Genet.* 137(1):36–48.
- Schmid, M., and J. Bennewitz. 2017. Invited review: Genome-wide association analysis for quantitative traits in livestock – a selective review of statistical models and experimental designs. *Arch. Anim. Breed.* 60(3):335–346.
- Shang, Y., S. Kumar, B. Oakley, and W. K. Kim. 2018. Chicken gut microbiota: importance and detection technology. *Front. Vet. Sci.* 5:254.
- Stanley, D., R. J. Hughes, and R. J. Moore. 2014. Microbiota of the chicken gastrointestinal tract: influence on health, productivity and disease. *Appl. Microbiol. Biotechnol.* 98(10):4301–4310.
- Tiezzi, F., J. Fix, C. Schwab, C. Shull, and C. Maltecca. 2021. Gut microbiome mediates host genomic effects on phenotypes: a case study with fat deposition in pigs. *Comput. Struct. Biotechnol. J.* 19(2-3):530–544.
- van der Eijk, J. A. J., T. B. Rodenburg, H. de Vries, J. B. Kjaer, H. Smidt, M. Naguib, B. Kemp, and A. Lammers. 2020. Early-life microbiota transplantation affects behavioural responses, serotonin and immune characteristics in chicken lines divergently selected on feather pecking. *Sci. Rep.* 10(1):2750.
- Verschuren, L. M. G., D. Schokker, R. Bergsma, A. J. M. Jansman, F. Molist, and M. P. L. Calus. 2020. Prediction of nutrient digestibility in grower-finisher pigs based on faecal microbiota composition. *J. Anim. Breed. Genet.* 137(1):23–35.
- Vollmar, S., R. Wellmann, D. Borda-Molina, M. Rodehutschord, A. Camarinha-Silva, and J. Bennewitz. 2020. The gut microbial architecture of efficiency traits in the domestic poultry model species Japanese quail (*Coturnix japonica*) assessed by mixed linear models. *G3 (Bethesda)* 10(7):2553–2562.
- Wang, Y., P. Zhou, X. Zhou, M. Fu, T. Wang, Z. Liu, X. Liu, Z. Wang, and B. Liu. 2022. Effect of host genetics and gut microbiome on fat deposition traits in pigs. *Front. Microbiol.* 13:925200.
- Weishaar, R., R. Wellmann, A. Camarinha-Silva, M. Rodehutschord, and J. Bennewitz. 2020. Selecting the hologenome to breed for an improved feed efficiency in pigs-A novel selection index. *J. Anim. Breed. Genet.* 137(1):14–22.

- Wen, C., W. Yan, C. Mai, Z. Duan, J. Zheng, C. Sun, and N. Yang. 2021. Joint contributions of the gut microbiota and host genetics to feed efficiency in chickens. *Microbiome* 9(1):126.
- Wen, C., W. Yan, C. Sun, C. Ji, Q. Zhou, D. Zhang, J. Zheng, and N. Yang. 2019. The gut microbiota is largely independent of host genetics in regulating fat deposition in chickens. *ISME J.* 13(6):1422–1436.
- Witzig, M., A. Camarinha-Silva, R. Green-Engert, K. Hoelzle, E. Zeller, J. Seifert, L. E. Hoelzle, and M. Rodehutschord. 2015. Correction: Spatial variation of the gut microbiota in broiler chickens as affected by dietary available phosphorus and assessed by T-RFLP analysis and 454 pyrosequencing. *PLoS One* 10(12):e0145588.
- Yadav, S., and R. Jha. 2019. Strategies to modulate the intestinal microbiota and their effects on nutrient utilization, performance, and health of poultry. *J. Anim. Sci. Biotechnol.* 10:2.
- Yegani, M., and D. R. Korver. 2008. Factors affecting intestinal health in poultry. *Poult. Sci.* 87(10):2052–2063.
- Zhao, L., G. Wang, P. Siegel, C. He, H. Wang, W. Zhao, Z. Zhai, F. Tian, J. Zhao, H. Zhang, Z. Sun, W. Chen, Y. Zhang, and H. Meng. 2013. Quantitative genetic background of the host influences gut microbiomes in chickens. *Sci. Rep.* 3:1163.

GENERAL DISCUSSION

The overriding reason for this work was the potential improved utilization of phosphorus (P) from feed components through genetic and microbial adaptation. This may result in savings of mineral P as supplemental feed and thus a lower environmental impact of livestock, especially poultry production. The experimental design was conducted using an F2 cross of Japanese quail (*Coturnix japonica*) as a model species of agriculturally important poultry species. All animals were genotyped with a 4k SNP chip, the phenotypes were recorded, and the ileal microbiota compositions were quantified. Previous studies confirmed the influence of host genetics and gastrointestinal microbiota composition on P utilization (PU) and related quantitative traits. Further consideration of the relationships between host genetics, gastrointestinal microbiota composition and quantitative traits was elaborated. This thesis has been divided into four main chapters.

Chapter one constructed a genetic linkage map based on genome-wide genotypes and examined quantitative traits by applying quantitative trait loci (QTL) linkage mapping. Several significant QTL regions with trait-associated markers could be detected, and potential candidate genes located within the identified QTL regions were found. In addition to performance traits, bone ash traits that were correlated with the focal trait PU were considered.

Chapter two incorporated gastrointestinal microbiota as a source of information and complemented a previous study that was part of the same project (Vollmar et al., 2020) to investigate the links between host genetics, gastrointestinal microbiota and quantitative traits. The host genetic influence on ileal genera was demonstrated by significant heritabilities and QTL linkage analyses. Using causal relationships between heritable genera and efficiency traits, as well as the confirmed usability of a hologenomic selection approach for the investigated traits, the mediator property of gastrointestinal microbiota between host genetics and efficiency traits was confirmed.

Chapter three used Bayesian networks to uncover causal relationships between microbial features and efficiency traits. With structural equation model association analyses, total SNP effects on a trait were divided into direct and indirect effects mediated by upstream traits.

Chapter four was an invited review of microbial genetic-statistical studies of poultry and pigs. In this review important analyses and approaches were discussed.

This **general discussion** provides additional results of the analyses of bone ash traits and genomic analyses of microbial features and debates the usability of microbial features in Bayesian network analyses and discusses the general findings across the chapters.

Microbial Analysis of Bone Ash Traits

In addition to PU and related efficiency traits, bone ash traits of the foot and tibia were investigated in **chapter one**. Bone ash consists mainly of calcium (Ca) and P (Darwish et al., 2017), and the genetic correlations between PU and bone ash traits reflect the genetic links (Künzel et al., 2019). The genetic correlations between PU and the four bone ash traits are all at a similar level approximately 0.5 (Künzel et al., 2019). Multiple significant QTL regions were found for these ash traits, suggesting a polygenic origin (**chapter one**). In the previous study of Vollmar et al. (2020), the microbiabilities of the efficiency traits used in **chapter one** were estimated. Microbial influence was confirmed for PU, feed intake, feed per gain, and body weight gain but not for Ca utilization (CaU). PU was found to have a significant microbiability of 0.15 (Vollmar et al., 2020). Microbial influence on bone ash traits was confirmed by estimating significant microbiabilities (Table 1) using the microbial linear mixed model (Model 5) from **chapter two**, i.e., the proportion of phenotypic variance explained by microbial composition (see **chapter four**). The significance of the microbiabilities was tested by conducting a likelihood ratio test on the random animal effects in the same way as it was done for the heritabilities in **chapter two** with Model 1.

For completeness, the microbiability of P retention (PR) is shown in Table 1, as this has not yet been published from our dataset and is a substantial part of **chapter three**.

Table 1 Microbiabilities (m^2) from the microbial linear mixed model (Model 5) in **chapter two** with standard errors (SE) and P values.

Trait	m^2	SE	P value
Phosphorus retention	0.11	0.04	< 0.001
Tibia ash (mg)	0.02	0.08	0.153
Tibia ash (%)	0.16	0.09	< 0.001
Foot ash (mg)	0.03	0.09	0.121
Foot ash (%)	0.20	0.06	< 0.001

These microbiabilities (Table 1) show that part of the phenotypic variance of the bone ash traits of total dry matter content (tibia ash % and foot ash %) can be explained by the composition of the gastrointestinal microbiota. The results of total tibia and foot ash in mg do not depend on the microbial composition of the animals' ileum from our dataset. One explanation might be that the bone ash proportion of the total dry matter contains other bone components that are potentially microbially affected. The correlations between tibia ash in mg and % are genetically 0.59 and phenotypically 0.57, and between foot ash in mg and % are genetically 0.69 and phenotypically 0.59 (Künzel et al., 2019).

Bone stability in poultry is dependent on bone mineralization (e.g., Jansen et al., 2020; Alfonso-Carrillo et al., 2021), which requires the major minerals Ca and P (Darwish et al., 2017). Low

mineralization for various reasons can decrease bone stability, which can frequently manifest in laying hens in the form of keel bone fractures (Toscano et al., 2020). The study of Sjögren et al. (2012) and Rodrigues et al. (2012) found that bone mineral density is regulated by the gut microbiota in mice and rats. Other studies reviewed the potential influence of gut microbiota on bone metabolism in chickens (Chen et al., 2022).

Considering the microbiabilities of PU and bone ash traits and the prevalence of fractures in poultry housing, it might be worth investigating the microbiota in the GIT of the animals. For more in-depth analyses on this animal welfare topic, the additional use of phenotypes such as egg performance, eggshell quality, classification of keel bone fractures as in Fleming et al. (2000), for example, or bone breaking strength as in Santos et al. (2022), for example, could be beneficial. Approaches such as the hologenomic selection approach developed by Weishaar et al. (2020) for efficiency traits could be interesting for further investigations of fractures in poultry.

Use of Microbial Data in Quantitative Analyses

Microbial data are compositional data with a high number of low relative abundances and many animal-specific zero values (Pawlowsky-Glahn et al., 2015). Considering the use of microbial data as quantitative traits, for example, in the study of causal relationships (**chapters two and three**), the causal structures are useable to explain the up to high correlations between some microbial features and efficiency traits. Since correlations are not causalities, the use of Bayesian networks and the modeling of structural equations allows a differentiated consideration. However, it should be noted that the use of compositional data as phenotypes may complicate the detection of causal relationships.

Thus, the use of superordinate microbial features as quantitative traits is limited but not impossible. It depends mainly on the number of animal-specific zero values. From our analyses, we were able to determine that the use of superordinate microbial features can require a higher computational capacity, the results can be more unsteady, and care should be taken to ensure that the statistical models converge. For example, a stable Bayesian network with two microbial genera is shown in Figure 1. The genera *Bacillus* and *Leuconostoc* were studied together with performance and bone ash traits. Model 1 from **chapter three** was used to estimate the model residuals of the characteristics. Because of convergence problems of Model 1, the network (Figure 1) should be interpreted with caution and serves as an illustration of the limitations. In addition, it should be noted that some priorities can be set in such Bayesian networks to facilitate network discovery or avoid trivial safe dependencies, especially when using small sample sizes. Examples are given in Figure 1 with the use of

compositional data of different microbial features (*Bacillus* and *Leuconostoc*) or biologically improbable interactions (tibia and foot ash).

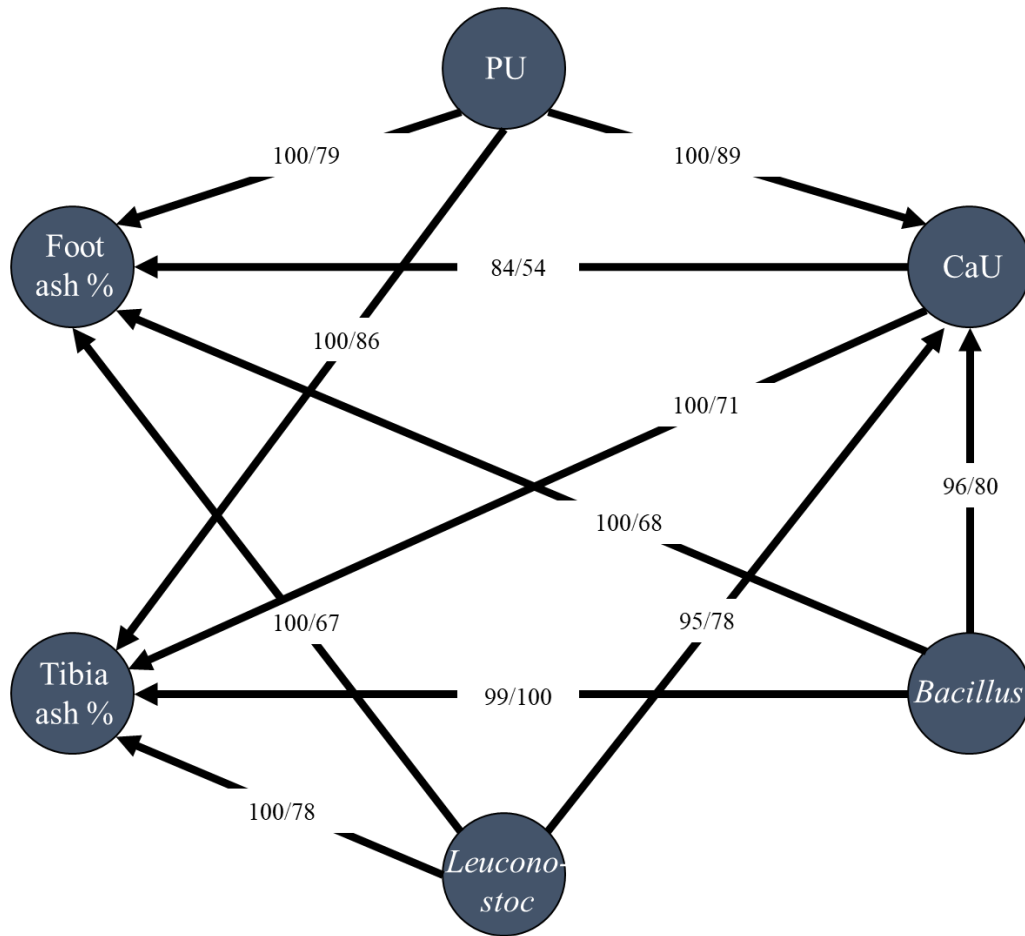


Figure 1 Bayesian network structure resulting from the Hill-Climbing learning algorithm (PU: Phosphorus utilization, CaU: Calcium utilization, Foot ash %, Tibia ash %, microbial genus *Bacillus* and *Leuconostoc*). The numbers are bootstrap sample proportions (50,000 samples), indicating an existing arc between the variables (left figures) and the proportion of this direction of the arc (right figures).

The convergence problems could be solved with lower classification levels, as was done in **chapter four** with the use of microbial features at the phylum level or with a microbial interaction feature such as the alpha diversity index (**chapter three**). From a phenotypic point of view, subordinate microbial features such as phyla are clustered superordinate ones due to taxonomic assignment (see **chapter four**). To find possible drivers of the observed phenotypes, it makes more sense from the point of view of causal agent identifiability to include no clustered microbial information in network analyses. However, the GIT microbiota occurs as a community with a high degree of interaction, which complicates the identifiability of single microbial drivers (Li, 2015). The alpha diversity index used in **chapter three** considers these interactions within the microbial community in the GIT (Pielou, 1966).

Apart from ethical aspects, well-thought-out animal designs with high sample sizes combined with high computational capacity would be beneficial to generate powerful results to further expand such Bayesian network studies.

Subordinate Microbial Features

For further classification of the results shown thus far, the heritability estimates at the phylum level are listed below. The microbial phyla partly used in the Bayesian network analysis in **chapter four** showed heritabilities between 0.05 and 0.11 (Table 2). They were calculated in the same way as the heritabilities of the microbial genera in **chapter two**. Only phyla with an average relative abundance of > 0.001 were examined. Each microbial phylum was transformed to an approximate Gaussian distribution using Box–Cox transformation with individual transformation parameters (Table 2); for more information, see **chapter two**. The results of the heritabilities at the phylum level are higher than those published in Zhou et al. (2022) for chicken fecal samples.

Table 2 Heritabilities of bacterial phyla, with average relative abundances (mean RA), standard errors (SE), P values, and the individual Box–Cox transformation parameters (Lambda).

Phylum ¹	Mean RA	h ²	SE	P value	Lambda (Box–Cox)
Firmicutes	83.25	0.11	0.06	< 0.001	2.97
Proteobacteria	14.30	0.09	0.05	< 0.001	0.39
Actinobacteria	1.65	0.09	0.05	< 0.001	0.01
Bacteroidetes	0.70	0.05	0.04	0.059	-0.11

¹ all microbial phyla with relative abundance (RA) > 0.001

QTL linkage mapping at the phylum level was performed in the same way as in **chapter two** at the genus level. The results of the QTL linkage mapping scans are illustrated in Figure 2. Two significant QTL regions were found for the four phyla studied, one each for Actinobacteria and Proteobacteria (Table 3).

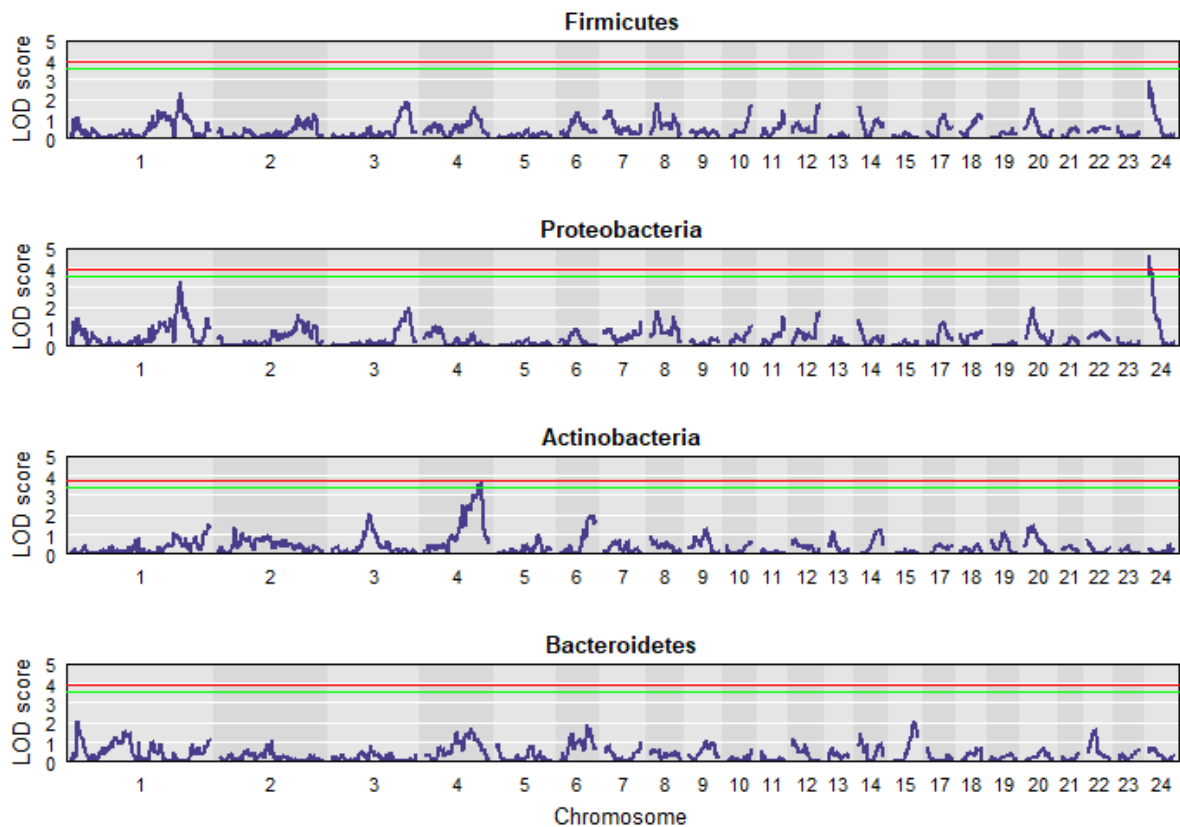


Figure 2 Plots of the QTL linkage mapping scan of microbial phyla, with LOD score as the test statistic. The red and green lines correspond to genome-wide significance levels of 5 and 10%, respectively.

The bacterial phylum Proteobacteria is the subordinate group of the genus *Escherichia/Shigella* investigated in **chapter two**. Both had a significant QTL region on *Coturnix japonica* chromosome (CJA) 24 with overlapping support intervals.

Table 3 QTL linkage mapping results at the phylum level.

Trait	CJA	Pos (cM)	LOD	Support interval borders (cM)	
				Low	High
Proteobacteria	24	0	4.65**	0	8
Actinobacteria	4	101	3.59*	69	106

Position (Pos) in cM of 5% (*) and 10% (**) genome-wide significant QTL on the *Coturnix japonica* chromosome (CJA), with LOD score test statistic (LOD) and the corresponding QTL support intervals in cM.

In **chapter four**, a stable Bayesian network is shown with two microbial features at the phylum level (Actinobacteria and Firmicutes). Both bacterial phyla affect tibia ash, CaU and BWG directly, and indirectly via the connection between these performance traits. PU affects the three performance traits directly but is not affected by the two phyla. The microbial effect on bone ash traits supposed in the **general discussion** above can be confirmed from the Bayesian network analyses from **chapter four**.

Hologenomic Modeling

In **chapter two**, we investigated the hologenomic approach developed in Weishaar et al. (2020) with our dataset for efficiency traits in Japanese quail. Three competing studies, which were mentioned in **chapter four**, considered both the microbial features and the host genetics together in one model because of interdependency.

As mentioned in **chapter four**, no exact interaction between markers and microbiota features can be calculated if the covariances are specified as independent. Therefore, Khanal et al. (2020) developed a model with a genomic (G) and microbial relationship matrix (M) and an additional interaction term GM that generates a covariance function. Statistically, this interaction means that the similarity between datasets is due to similarity of both the host genetics and the intestinal microbial composition. This interaction approach is similar to the models of genotype-by-environment interactions, where the intestinal microbial component corresponds to the environmental component. Translated into a biological explanation that allele substitution effects for each marker depend on the intestinal microbiota composition, and conversely, that microbial effects depend on the host genotype. Pérez-Enciso et al. (2021) and Qadri et al. (2022) also followed similar approaches for different forms of interaction between the host genome and microbiome in cattle. Qadri et al. (2022) for example, used Cholesky decomposition and the Hadamard product to model interactions.

However, the scientific community has not yet agreed on which models or interactions the relationships between the microbial and the genetic information of the animals can be correctly mapped together in a statistical model.

Gut Microbiota Development

The egg is the first potential microbiota source of hatching chicks. Before oviposition, the eggs are colonized with many microbes on their way through the reproductive tract and cloaca of hens (Lee et al., 2019). Therefore, the eggshell could play an important role in microbial colonization in the gut of chicks during hatching (Maki et al., 2020). It can be assumed that colonization from the shell to the embryo or chick takes place from brooding day 4 or 5 and during hatching (Reu et al., 2006; Gantois et al., 2009). Due to the disinfection of hatching eggs and the high hygiene status in hatcheries, vertical transmission of microbiota to chicks is limited (Stanley et al., 2013). If the eggs are disinfected as usual, the native maternal microbial colonization component is reduced. For the animals to develop their genetic potential of microbial colonization of the gastrointestinal tract, an early administration of probiotics (before hatching) would have a positive effect on the performance and utilization characteristics (Baldwin et al., 2018). The injection of native beneficial microbiota from healthy adults positively affected the immunity and performance of young chicks (Roto et al., 2016), and early

bacterial settlers can influence the entire development of the microbial gut flora (Nakphaichit et al., 2011).

To the best of our knowledge, in our study design, quail eggs were not disinfected before breeding. Vertical transmission was thus not reduced in any way. Additionally, we observed Japanese quail in a strong growth phase on the 10th to 15th day of life in the juvenile stage before puberty with some sexual changes, such as egg production (Balthazart et al., 2009; Narinc et al., 2013). Borda-Molina et al. (2020) and Zhao et al. (2013) found significant differences in the microbial colonization of the ileum of quail between sexes. We could not confirm this in our study, and the effect of sex effect was not significant. Therefore, sex was excluded from our analyses.

Microbial colonization in the GIT of animals is influenced by numerous external and internal factors (see **chapter four**). By using microbial information for trait predictions, for example, the predictors are not stable among different animal ages, GIT sections or environmental factors. Therefore, comparisons between different animals at different ages or in different environments should be performed with caution. In the case of our study, standardized procedures were advantageous in the experimental design and the subsequent interpretation of the results.

Concluding Remarks

In this thesis, PU and related traits were detected as typical quantitative traits, the gastrointestinal microbiota was confirmed to be partially heritable, and causal relationships between the gastrointestinal microbiota and quantitative traits were quantified. It became apparent that genes influence parts of the gastrointestinal microbiota and thus can be used for trait predictions and potential selective breeding via an adapted gastrointestinal microbiota composition. We can also state that parts of the gastrointestinal microbiota can be classified as mediators between host genetics and quantitative traits. Directional relationships between different performance traits and microbial features were decoded using Bayesian network analysis. Based on these Bayesian networks and structural equation models, the possibility of identifying and separating SNP effects into direct and indirect effects of mediating microbial features and other mediating traits was shown.

In conclusion, we can state that PU is influenced not only by host genetics but also by parts of the gastrointestinal microbiota. The thesis confirmed the relationships between host genetics, gastrointestinal microbiota and efficiency traits.

References

- Alfonso-Carrillo, C., C. Benavides-Reyes, J. de Los Mozos, N. Dominguez-Gasca, E. Sanchez-Rodríguez, A. I. Garcia-Ruiz, and A. B. Rodriguez-Navarro. 2021. Relationship between Bone Quality, Egg Production and Eggshell Quality in Laying Hens at the End of an Extended Production Cycle (105 Weeks). *Animals (Basel)* 11(3). doi:10.3390/ani11030623.
- Baldwin, S., R. J. Hughes, T. T. Hao Van, R. J. Moore, and D. Stanley. 2018. At-hatch administration of probiotic to chickens can introduce beneficial changes in gut microbiota. *PLoS One* 13(3):e0194825. doi:10.1371/journal.pone.0194825.
- Balthazart, J., C. A. Cornil, T. D. Charlier, M. Taziaux, and G. F. Ball. 2009. Estradiol, a key endocrine signal in the sexual differentiation and activation of reproductive behavior in quail. *J. Exp. Zool. A Ecol. Genet. Physiol.* 311(5):323–345. doi:10.1002/jez.464.
- Borda-Molina, D., C. Roth, A. Hernández-Arriaga, D. Rissi, S. Vollmar, M. Rodehutschord, J. Bennewitz, and A. Camarinha-Silva. 2020. Effects on the Ileal Microbiota of Phosphorus and Calcium Utilization, Bird Performance, and Gender in Japanese Quail. *Animals (Basel)* 10(5):885. doi:10.3390/ANI10050885.
- Chen, P., T. Xu, C. Zhang, X. Tong, A. Shaukat, Y. He, K. Liu, and S. Huang. 2022. Effects of Probiotics and Gut Microbiota on Bone Metabolism in Chickens: A Review. *Metabolites* 12(10). doi:10.3390/metabo12101000.
- Darwish, M., A. Aris, M. H. Puteh, M. N. H. Jusoh, and A. Abdul Kadir. 2017. Waste bones ash as an alternative source of P for struvite precipitation. *J. Environ. Manage.* 203(Pt 2):861–866. doi:10.1016/j.jenvman.2016.02.033.
- Fleming, R. H., H. A. McCormack, and C. C. Whitehead. 2000. Prediction of breaking strength in osteoporotic avian bone using digitized fluoroscopy, a low cost radiographic technique. *Calcif. Tissue Int.* 67(4):309–313. doi:10.1007/s002230001120.
- Gantois, I., R. Ducatelle, F. Pasmans, F. Haesebrouck, R. Gast, T. J. Humphrey, and F. van Immerseel. 2009. Mechanisms of egg contamination by *Salmonella* Enteritidis. *FEMS Microbiol. Rev.* 33(4):718–738. doi:10.1111/j.1574-6976.2008.00161.x.
- Jansen, S., U. Baulain, C. Habig, A. Weigend, I. Halle, A. M. Scholz, H. Simianer, A. R. Sharifi, and S. Weigend. 2020. Relationship between Bone Stability and Egg Production in Genetically Divergent Chicken Layer Lines. *Animals (Basel)* 10(5):850. doi:10.3390/ani10050850.

- Khanal, P., C. Maltecca, C. Schwab, J. Fix, M. Bergamaschi, and F. Tiezzi. 2020. Modeling host-microbiome interactions for the prediction of meat quality and carcass composition traits in swine. *Genet. Sel. Evol.* 52(1):41. doi:10.1186/s12711-020-00561-7.
- Künzel, S., J. Bennewitz, and M. Rodehutschord. 2019. Genetic parameters for bone ash and phosphorus utilization in an F2 cross of Japanese quail. *Poult. Sci.* 98(10):4369–4372. doi:10.3382/ps/pez398.
- Lee, S., T.-M. La, H.-J. Lee, I.-S. Choi, C.-S. Song, S.-Y. Park, J.-B. Lee, and S.-W. Lee. 2019. Characterization of microbial communities in the chicken oviduct and the origin of chicken embryo gut microbiota. *Sci. Rep.* 9(1):6838. doi:10.1038/s41598-019-43280-w.
- Li, H. 2015. Microbiome, Metagenomics, and High-Dimensional Compositional Data Analysis. *Annu. Rev. Stat. Appl.* 2(1):73–94. doi:10.1146/annurev-statistics-010814-020351.
- Maki, J. J., E. A. Bobeck, M. J. Sylte, and T. Looft. 2020. Eggshell and environmental bacteria contribute to the intestinal microbiota of growing chickens. *J. Anim. Sci. Biotechnol.* 11:60. doi:10.1186/s40104-020-00459-w.
- Nakphaichit, M., S. Thanomwongwattana, C. Phraephaisarn, N. Sakamoto, S. Keawsompong, J. Nakayama, and S. Nitisinprasert. 2011. The effect of including *Lactobacillus reuteri* KUB-AC5 during post-hatch feeding on the growth and ileum microbiota of broiler chickens. *Poult. Sci.* 90(12):2753–2765. doi:10.3382/ps.2011-01637.
- Narinc, D., E. Karaman, T. Aksoy, and M. Z. Firat. 2013. Investigation of nonlinear models to describe long-term egg production in Japanese quail. *Poult. Sci.* 92(6):1676–1682. doi:10.3382/ps.2012-02511.
- Pawlowsky-Glahn, V., J. J. Egozcue, and D. Lovell. 2015. Tools for compositional data with a total. *Stat. Modelling* 15(2):175–190. doi:10.1177/1471082X14535526.
- Pérez-Enciso, M., L. M. Zingaretti, Y. Ramayo-Caldas, and G. de Los Campos. 2021. Opportunities and limits of combining microbiome and genome data for complex trait prediction. *Genet. Sel. Evol.* 53(1):65. doi:10.1186/s12711-021-00658-7.
- Pielou, E. C. 1966. The measurement of diversity in different types of biological collections. *J. Theor. Biol.* 13:131–144. doi:10.1016/0022-5193(66)90013-0.
- Qadri, Q. R., Q. Zhao, X. Lai, Z. Zhang, W. Zhao, Y. Pan, and Q. Wang. 2022. Estimation of Complex-Trait Prediction Accuracy from the Different Holo-Omics Interaction Models. *Genes* 13(9). doi:10.3390/genes13091580.

GENERAL DISCUSSION

- Reu, K. de, K. Grijspeerd, W. Messens, M. Heyndrickx, M. Uyttendaele, J. Debevere, and L. Herman. 2006. Eggshell factors influencing eggshell penetration and whole egg contamination by different bacteria, including *Salmonella enteritidis*. *Int. J. Food Microbiol.* 112(3):253–260. doi:10.1016/j.ijfoodmicro.2006.04.011.
- Rodrigues, F. C., A. S. B. Castro, V. C. Rodrigues, S. A. Fernandes, E. A. F. Fontes, T. T. de Oliveira, H. S. D. Martino, and C. L. de Luces Fortes Ferreira. 2012. Yacon flour and *Bifidobacterium longum* modulate bone health in rats. *J. Med. Food* 15(7):664–670. doi:10.1089/jmf.2011.0296.
- Roto, S. M., Y. M. Kwon, and S. C. Ricke. 2016. Applications of In Ovo Technique for the Optimal Development of the Gastrointestinal Tract and the Potential Influence on the Establishment of Its Microbiome in Poultry. *Front. Vet. Sci.* 3:63. doi:10.3389/fvets.2016.00063.
- Santos, M. N., T. M. Widowski, E. G. Kiarie, M. T. Guerin, A. M. Edwards, and S. Torrey. 2022. In pursuit of a better broiler: tibial morphology, breaking strength, and ash content in conventional and slower-growing strains of broiler chickens. *Poult. Sci.* 101(4):101755. doi:10.1016/j.psj.2022.101755.
- Sjögren, K., C. Engdahl, P. Henning, U. H. Lerner, V. Tremaroli, M. K. Lagerquist, F. Bäckhed, and C. Ohlsson. 2012. The gut microbiota regulates bone mass in mice. *J. Bone Miner. Res.* 27(6):1357–1367. doi:10.1002/jbmr.1588.
- Stanley, D., M. S. Geier, R. J. Hughes, S. E. Denman, and R. J. Moore. 2013. Highly variable microbiota development in the chicken gastrointestinal tract. *PLoS One* 8(12):e84290. doi:10.1371/journal.pone.0084290.
- Toscano, M. J., I. C. Dunn, J.-P. Christensen, S. Petow, K. Kittelsen, and R. Ulrich. 2020. Explanations for keel bone fractures in laying hens: are there explanations in addition to elevated egg production? *Poult. Sci.* 99(9):4183–4194. doi:10.1016/j.psj.2020.05.035.
- Vollmar, S., R. Wellmann, D. Borda-Molina, M. Rodehutschord, A. Camarinha-Silva, and J. Bennewitz. 2020. The Gut Microbial Architecture of Efficiency Traits in the Domestic Poultry Model Species Japanese Quail (*Coturnix japonica*) Assessed by Mixed Linear Models. *G3 (Bethesda)* 10(7):2553–2562. doi:10.1534/g3.120.401424.
- Weishaar, R., R. Wellmann, A. Camarinha-Silva, M. Rodehutschord, and J. Bennewitz. 2020. Selecting the hologenome to breed for an improved feed efficiency in pigs-A novel selection index. *J. Anim. Breed. Genet.* 137(1):14–22. doi:10.1111/jbg.12447.

Zhao, L., G. Wang, P. Siegel, C. He, H. Wang, W. Zhao, Z. Zhai, F. Tian, J. Zhao, H. Zhang, Z. Sun, W. Chen, Y. Zhang, and H. Meng. 2013. Quantitative genetic background of the host influences gut microbiomes in chickens. *Sci. Rep.* 3:1163. doi:<https://doi.org/10.1038/srep01163>.

Zhou, Q., F. Lan, S. Gu, G. Li, G. Wu, Y. Yan, X. Li, J. Jin, C. Wen, C. Sun, and N. Yang. 2022. Genetic and microbiome analysis of feed efficiency in laying hens. *Poult. Sci.*:102393. doi:[10.1016/j.psj.2022.102393](https://doi.org/10.1016/j.psj.2022.102393).

ACKNOWLEDGEMENT

Nach einer großartigen Zeit voller neuen Erfahrungen, Errungenschaften, gemeisterten Zielen, aber auch anstrengenden Momenten, heißt es nun Danke zu sagen.

Mein größter Dank gilt meinem Doktorvater Prof. Dr. Jörn Bennewitz, der mir die Zeit im Fachgebiet für Tiergenetik und Züchtung ermöglichte, der immer ein offenes Ohr hatte und wo die flache Hierarchie verbunden mit seiner aufbauenden Führungsweise zu einem sehr angenehmen Arbeitsverhältnis führte. Lieber Jörn, vielen Dank für deine investierte Zeit, für die Antworten auf meine vielen Fragen und vor allem für das aufgebrachte Vertrauen in den letzten Jahren.

Ich möchte mich ebenso bei Jun.-Prof. Amélia Camarinha da Silva für die großartige mikrobielle Unterstützung bedanken. Liebe Amélia, ohne dich wäre meine Arbeit nicht möglich gewesen. Danke auch an Prof. Dr. Markus Rodehutschord für die Projektbegleitung, der konstruktiven Kritik und für die Hilfe von gebietsübergreifenden Fragestellungen.

Auch möchte ich mich herzlich bei Prof. Dr. Klaus Wimmers für die Übernahme des Zweitgutachtens bedanken und die dazu aufgebrachten Zeit und Mühen.

Ein großes Dankeschön gebührt auch Christina Schweizer. Christina, vielen Dank für deine tatkräftige Unterstützung in allen Bereichen. Vermutlich wären ohne dich so einige Fristen verstrichen.

Auch vielen Dank an Dr. Siegfried Preuß, Dr. Dr. Robin Wellmann und Dr. Markus Schmid. Vielen Dank euch Drei, für den fachlichen Beistand meiner Arbeit. Danke dir Sigi vor allem rund um den Themenbereich der Genotypisierung; Danke Robin für deine genetisch-statistische Expertise und lieber Markus, danke für deine Beratungen in allen Bereichen.

Vielen Dank gilt auch besonders meinen lieben Kollegen und Kolleginnen, die das Fachgebiet zu dem Ort gemacht haben, an dem man jeden Tag von neuem gerne zur Arbeit geht. Vielen Dank für die vielen gemeinsamen Stunden!

Abschließend möchte ich mich noch bei meiner Familie bedanken. Meinen Eltern, Geschwistern, meiner großartigen Frau Soli, aber auch meinen Großeltern, die mich immer unterstützt haben und mir den Weg zu dieser Arbeit ermöglicht haben. Ohne euch wäre diese Arbeit nie entstanden. Vielen Dank auch an meine Freunde, die mich auf dem Weg des Studiums und der Promotion begleitet haben.

CURRICULUM VITAE

Personal Data

Name: Valentin Peter Haas
Date of Birth: 02.08.1994
Place of Birth: Heidelberg, Germany

Education

Since 11/2019 PhD Student
Department of Animal Science, Animal Genetics and
Breeding University of Hohenheim, Stuttgart, Germany

2017 - 2019 M.Sc. Agricultural Science, Animal Science
University of Hohenheim, Stuttgart, Germany

2014 - 2017 B.Sc. Agricultural Science, Animal Science
University of Hohenheim, Stuttgart, Germany

2011 - 2014 Abitur, Technisches Gymnasium, Buchen, Germany

Practical Experience

2014 - 2019 Internship and work experience on the farm Kieser,
Buchen, Germany

EIDESSTÄTLICHE VERSICHERUNG

Eidesstattliche Versicherung über die eigenständig erbrachte Leistung

gemäß § 18 Absatz 3 Satz 5 der Promotionsordnung der Universität Hohenheim für die Fakultäten Agrar-, Natur- sowie Wirtschafts- und Sozialwissenschaften

1. Bei der eingereichten Dissertation zum Thema "*Genomic and microbial analyses of quantitative traits in poultry*" handelt es sich um meine eigenständig erbrachte Leistung.
2. Ich habe nur die angegebenen Quellen und Hilfsmittel benutzt und mich keiner unzulässigen Hilfe Dritter bedient. Insbesondere habe ich wörtlich oder sinngemäß aus anderen Werken übernommene Inhalte als solche kenntlich gemacht.
3. Ich habe nicht die Hilfe einer kommerziellen Promotionsvermittlung oder -beratung in Anspruch genommen.
4. Die Bedeutung der eidesstattlichen Versicherung und der strafrechtlichen Folgen einer unrichtigen oder unvollständigen eidesstattlichen Versicherung sind mir bekannt.

Die Richtigkeit der vorstehenden Erklärung bestätige ich. Ich versichere an Eides Statt, dass ich nach bestem Wissen die reine Wahrheit erkläre und nichts verschwiegen habe.

Stuttgart - Hohenheim, 27.07.2023

Valentin Haas

Belehrung über die Bedeutung und die strafrechtlichen Folgen der eidesstattlichen Versicherung**gemäß § 18 Absatz 3 Satz 6 der Promotionsordnung der Universität Hohenheim für die Fakultäten Agrar-, Natur- sowie Wirtschafts- und Sozialwissenschaften**

Die Universität Hohenheim verlangt eine Eidesstattliche Versicherung über die Eigenständigkeit der erbrachten wissenschaftlichen Leistungen, um sich glaubhaft zu versichern, dass die oder der Promovierende die wissenschaftlichen Leistungen eigenständig erbracht hat.

Weil der Gesetzgeber der Eidesstattlichen Versicherung eine besondere Bedeutung beimisst und sie erhebliche Folgen haben kann, hat der Gesetzgeber die Abgabe einer falschen eidesstattlichen Versicherung unter Strafe gestellt. Bei vorsätzlicher (also wissentlicher) Abgabe einer falschen Erklärung droht eine Freiheitsstrafe bis zu drei Jahren oder eine Geldstrafe.

Eine fahrlässige Abgabe (also Abgabe, obwohl Sie hätten erkennen müssen, dass die Erklärung nicht den Tatsachen entspricht) kann eine Freiheitsstrafe bis zu einem Jahr oder eine Geldstrafe nach sich ziehen.

Die entsprechenden Strafvorschriften sind in § 156 StGB (falsche Versicherung an Eides Statt) und in § 161 StGB (Fahrlässiger Falscheid, fahrlässige falsche Versicherung an Eides Statt) wiedergegeben.

§ 156 StGB: Falsche Versicherung an Eides Statt

Wer vor einer zur Abnahme einer Versicherung an Eides Statt zuständigen Behörde eine solche Versicherung falsch abgibt oder unter Berufung auf eine solche Versicherung falsch aussagt, wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

§ 161 StGB: Fahrlässiger Falscheid, fahrlässige falsche Versicherung an Eides Statt

Abs. 1: Wenn eine der in den §§ 154 und 156 bezeichneten Handlungen aus Fahrlässigkeit begangen worden ist, so tritt Freiheitsstrafe bis zu einem Jahr oder Geldstrafe ein.

Abs. 2: Straflosigkeit tritt ein, wenn der Täter die falsche Angabe rechtzeitig berichtigt. Die Vorschriften des § 158 Absätze 2 und 3 gelten entsprechend

Ich habe die Belehrung zur Eidesstattlichen Versicherung zur Kenntnis genommen.

Stuttgart - Hohenheim, 27.07.2023

Valentin Haas

