

Title:

Multi-omics analysis reveals regime shifts in the gastrointestinal ecosystem in chickens following anticoccidial vaccination and *Eimeria tenella* challenge

Authors:

Po-Yu Liu (ORCID ID: 0000-0003-1290-0850)^{1,2,3}, Janie Liaw⁴, Francesca Soutter⁵, José Jaramillo Ortiz^{1, 6}, Fiona M. Tomley (ORCID ID: 0000-0003-2188-8013)¹, Dirk Werling (ORCID ID: 0000-0001-5411-4044)^{1, 6}, Ozan Gundogdu (ORCID ID: 0000-0002-3550-0545)⁴, Damer P. Blake (ORCID ID: 0000-0003-1077-2306)^{1, 6}, Dong Xia (ORCID ID: 0000-0003-4571-2776)^{1,*}

¹Pathobiology and Population Sciences, Royal Veterinary College, Hawkshead Lane, North Mymms AL9 7TA, UK

²School of Medicine, College of Medicine, National Sun Yat-sen University, Kaohsiung 804201, Taiwan

³Department of Biomedical Science and Environmental Biology, Kaohsiung Medical University, Kaohsiung 807378, Taiwan

⁴Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK

⁵Scotland's Rural College, Edinburgh, EH25 9RG, UK

⁶Centre for Vaccinology and Regenerative Medicine, Royal Veterinary College, Hawkshead Lane, North Mymms, AL9 7TA, UK

***Corresponding author:**

Dong Xia

Pathobiology and Population Sciences, Royal Veterinary College, Hawkshead Lane, North Mymms AL9 7TA, UK

E-mail: dxia@rvc.ac.uk

ABSTRACT

Coccidiosis, caused by *Eimeria* parasites, poses significant economic and welfare challenges in poultry farming. Beyond its direct impact on health, *Eimeria* infection disrupts enteric microbial populations leading to dysbiosis and increases vulnerability to secondary diseases such as necrotic enteritis, caused by *Clostridium perfringens*. The impact of *Eimeria* infection or anticoccidial vaccination on host gastrointestinal phenotypes and enteric microbiota remains understudied. In this study, the metabolomic profiles and microbiota composition of chicken caecal tissue and contents were evaluated concurrently during a controlled experimental vaccination and challenge trial. Cobb500 broilers were vaccinated with a *Saccharomyces cerevisiae*-vectored anticoccidial vaccine and challenged with 15,000 *Eimeria tenella* oocysts. Assessment of caecal pathology and quantification of parasite load revealed correlations with alterations to caecal microbiota and host metabolome linked to infection and vaccination status. Infection heightened microbiota richness with increases in potentially pathogenic species, while vaccination elevated beneficial *Bifidobacterium*. Using a multi-omics factor analysis (MOFA) machine learning model, data on caecal microbiota and host metabolome were integrated and distinct profiles for healthy, infected, and recovering chickens were identified. Healthy and recovering chickens exhibited higher vitamin B metabolism linked to short-chain fatty acid-producing bacteria, whereas essential amino acid and cell membrane lipid metabolisms were prominent in infected and vaccinated chickens. Notably, vaccinated chickens showed distinct metabolites related to the enrichment of sphingolipids, important components of nerve cells and cell membranes. Our integrated multi-omics model revealed latent biomarkers indicative of vaccination and infection status, offering potential tools for diagnosing infection, monitoring vaccination efficacy, and guiding the development of novel treatments or controls.

Keywords: *Eimeria*, yeast-based anticoccidial vaccine, gut microbiota, metabolome, multi-omics

INTRODUCTION

Protozoan parasites of the genus *Eimeria* cause coccidiosis in poultry and costs to the industry have been estimated to exceed £10 billion annually [1]. Clinical coccidiosis manifests as poor body weight gain and feed conversion with diarrhoea, bloody droppings, and mortality in severe cases. Infection induces strong pro- and anti-inflammatory cytokine responses that may exacerbate pathology [2-5]. Clinical coccidiosis is commonly avoided through a combination of good husbandry, parasite chemoprophylaxis with anticoccidial drugs and/or vaccination using varied formulations of live parasites [6, 7]. In some countries, public concern related to pathogen drug resistance and widespread use of antimicrobials in animal production are driving legislative and commercial changes, including increased use of anticoccidial vaccination [8]. Although current live parasite vaccines are effective, considerable efforts are also being made to develop recombinant anticoccidial vaccines [9]. In a previous study, a novel prototype inactivated yeast-based recombinant oral vaccine for *Eimeria tenella* was shown to result in reduced parasite replication, reduced caecal pathology and improved chicken performance compared to controls [10]. Using *Saccharomyces cerevisiae* to express and deliver *E. tenella* antigens apical membrane antigen 1 (EtAMA1) [11], immune mapped protein 1 (EtIMP1) [12] and repeat 3 from microneme protein 3 (EtMIC3) [13] induced significant protection against high-level challenge in vaccinated Cobb500 broiler chickens [10]. However, the impact of vaccination and subsequent parasite challenge on the host gut and its enteric microbiota was not evaluated. Oral administration of heat-inactivated and freeze-dried *S. cerevisiae* has previously been shown to ameliorate the effects of coccidiosis in broiler chickens while modulating the host immune response and microbiota [14, 15]. Understanding the influence of a yeast-vectored anticoccidial vaccine on host metabolome and microbiomes could therefore be used to inform future vaccine development.

Enteric microbiomes play crucial roles in shaping host physiological functions including provision of nutrients [16, 17], immune system maturation and regulation [18, 19]. *Eimeria* infection can cause imbalance in gastrointestinal ecosystems [20, 21], commonly referred to as dysbiosis, and raises the risk of enteric comorbidities such as necrotic enteritis caused by *Clostridium perfringens* [22]. Variation in the severity of damage caused by *Eimeria* infection has also been shown to associate with differences in enteric microbiomes. For example, high level caecal lesion scores recorded during *E. tenella* infection correlated with increased *Enterobacteriaceae* occurrence but decreased *Bacillales* and *Lactobacillales* [21]. However, little is known about physiological responses in gastrointestinal molecular and biochemical mechanisms, or variation in microbiota between immunologically naïve, infected and

vaccinated chickens. Few studies have provided insight into chickens' metabolic responses to infection or vaccination. Using an untargeted metabolomic profile assessment, Aggrey et al. (2019) found that carnitine-derived metabolites involved in fatty acid metabolism, and thromboxane B2, 12-HHTrE and itaconate involved in inflammatory responses, were influenced by *Eimeria acervulina* infection [23]. In the same way, a human shingles vaccine trial revealed that key metabolites such as sterol class metabolites, arachidonic acids, phosphoinositide, and diacylglycerol, were essential to immune signalling [24]. Here, we have created a multi-omics dataset defining caecal microbial populations (lumen contents and tissue-associated) and caecal tissue metabolomes using high-throughput sequencing of the 16S rRNA gene and liquid chromatography-mass spectrometry (LC-MS), respectively. We have used a Multi-Omics Factor Analysis (MOFA) [25, 26] machine learning model to systematically integrate data on caecal microbiota and the host metabolome sampled during an anticoccidial vaccine trial, investigating host microbe-associated signatures that can predict chicken health status and vaccine efficacy.

RESULTS

Caecal pathology and parasite load *post-Eimeria* challenge demonstrates efficacy of a candidate yeast-vectored anticoccidial vaccine

We previously evaluated the efficacy of an experimental *S. cerevisiae*-vectored anticoccidial vaccine using readouts of gut pathology (caecal lesion scores: 0-4), parasite replication (quantitative PCR of caecal tissue) and chicken performance (body weight gain, BWG) following oral challenge with 15,000 sporulated oocysts of *E. tenella* [10]. Briefly, lesion scores at 6 dpi were lower in vaccinated chickens compared to unvaccinated controls (V-C vs UV-C; $p < 0.001$; Figure 1A). Parasite replication measured by qPCR as parasite genomes per host genome was also lower in vaccinated chickens at 6 dpi ($p < 0.001$; Figure 1B). In contrast, BWG was not significantly different at 6 dpi (Figure 1C) although it was by 10 dpi [10].

In the present study, the level of *E. tenella* replication at 6 dpi was confirmed by quantification of *Eimeria* mitochondrial 16S rRNA amplicon reads in NGS microbiome data from caecal tissue and contents (Figures 1D and 1E). Comparison of all three *E. tenella* replication measures revealed a significant association with lesion score severity (qPCR ratio: $r = 0.89$, NGS reads of caecal contents: $r = 0.8$, NGS reads of caecal tissue: $r = 0.63$; all $p < 0.001$; Figure 1F). For comparison, 10 dpi unvaccinated and challenged chickens (UV-C10) considered to be recovering from infection also showed a significant reduction in gut pathology and *Eimeria* load compared to all infected subjects at 6 dpi ($p < 0.001$; Figures 1A and 1E).

Gut pathology and parasite load correlate with changes in gut microbiota

The composition of enteric microbial populations can reflect the health status of micro-ecosystems in the gastrointestinal (GI) tract. We performed 16S amplicon sequencing from caecal contents and tissues collected from the same individuals to characterize gut microbiota composition, with no significant differences in beta diversity detected between sample types (caecal tissue compared to caecal contents; PERMANOVA test $R^2 = 0.026$, $p = 0.052$) (Figure S1A). Comparison between caecal contents and tissue found 62.7-73.6% of microbiota composition to be shared (Figures S1B). Microbial populations enriched in caecal contents included *Lactobacillus mucosae*, *Lactobacillus salivarius*, *Paludicola psychrotolerans*, *Kineothrix alysoides*, *Anaerostipes butyraticus*, and [*Clostridium*] *polysaccharolyticum*; while microbial populations of *Anaerotruncus colihominis* (KTU 13) and *Flavonifractor plautii* (KTU 14) were enriched in caecal tissues (i.e. UV-C, MV-C, and V-C) (Figure S1C).

Principal coordinates analysis (PCoA) based on Bray-Curtis dissimilarity

measurements showed that the caecal contents microbiota composition of unchallenged versus all challenged groups were distinct from each other (6 dpi) along the PCoA1 axis (31.15% of observed variation) (Figure 2). A PERMANOVA test confirmed significant differences in microbiota ($R^2=0.33$, $p=0.001$) (Figure 2A) and there were significant correlations with caecal lesion scores ($|r|=0.73$, $p<0.001$), parasite load in caecal tissues (qPCR ratio: $|r|=0.76$, $p<0.001$) and caecal contents (NGS reads: $|r|=0.67$, $p<0.001$) (Figure 2B-D).

On average a low alpha diversity index of microbial richness was found in all chickens across all groups (71.64 ± 14.21) compared to a previous study by Hay et al. (2023) (493.13 ± 201.60 , re-analysed using the same pipeline used in the present study)[27]. This disparity may be due to requirement for broad-spectrum enrofloxacin treatment during this trial (Figure 2E). Comparison between the groups revealed a higher richness index in all challenged groups 6 dpi compared to the unvaccinated, unchallenged group (UV-UC), although the difference was not statistically significant. The dominant phyla were *Firmicutes*, followed by *Proteobacteria* in all chickens (combined, accounting for more than 98%) (Figure 2F); however, *Proteobacteria* were reduced in UV-UC chickens (4.83% comparing to 13.5%/26.65%/22.36% in other groups). *Actinobacteria* were enriched in both mock and true vaccinated groups (1.55% and 1.35%, respectively), dominated by genus *Bifidobacterium* (1.50% and 1.30%, respectively). Since the lesion scores and *Eimeria* loads were significantly correlated with the PCoA1 axis of beta diversity, 36 associated taxa enriched in challenged chickens were identified by Pearson's correlation analysis ($|r|\geq 0.4$, $FDR < 0.1$), including *Escherichia coli*, *Clostridium difficile*, *C. innocuum*, and *Proteus mirabilis* (Figure 2G).

Metabolomes reflect the molecular alterations of host physiology responses in health, infection, and recovery

Caecal tissue metabolomic profiling was performed for the same chickens as described above using samples collected in parallel with those used for microbiome sequencing analysis to characterize host physiological responses. An untargeted metabolomics approach was applied for screening metabolites within the tissues. Based on Euclidean distance measurements, PCoA of caecal tissue metabolome profiles showed a similar pattern to the caecal microbiota with unchallenged and challenged individuals differentiated along the PCoA1 axis (52.73% of observed variation) (Figure 3). The recovering (UV-C10) group displayed a broad but intermediate metabolome profile to that of 6 dpi challenged chickens and uninfected chickens and this group was also differentiated along the PCoA2 axis (9.57%). Host

metabolome profiles correlated with caecal lesion scores ($|r|=0.83$, $p<0.001$) and *Eimeria* loads (qPCR ratio: $|r|=0.73$, $p<0.001$; NGS reads of caecal contents: $|r|=0.83$, $p<0.001$; NGS reads of caecal tissues: $|r|=0.68$, $p<0.001$) (Figure 3B-C). Among 1,180 metabolites belonging to the 10 categories that were detected from all chickens (including partially characterized and uncharacterized; Figure 3D), 954 metabolites were either negatively (606, non-infection-associated) or positively (348, infection-associated) correlated with pathophysiology changes (lesion scores and *Eimeria* loads; significant negative correlation with PCo1 in Figure 3A by Pearson's correlation analysis, $FDR < 0.1$; Figure 3E). In more detail, xenobiotics, cofactors and vitamins, especially vitamin Bs, were characterized as non-infection-associated metabolites (Figure S2A and Table S2); while lipids, especially the sphingolipids, nucleotides, and carbohydrates, were characterized as infection-associated metabolites (Figure S2B and Table S2).

Multi-omics factor analysis reveals covariation patterns of disease status

Using multi-omics factor analysis (MOFA), integration of parallel caecal tissue and content microbiomes with host tissue metabolome data showed concordant responses that associated with gut pathology and parasite load. Host-microbe intercorrelated features were assessed between microbial and metabolite features using Spearman's correlation. A total of 151 KTUs and 767 metabolites were significantly associated ($FDR<0.05$), resulting in a MOFA model that contained 15 representative factors. The factors were decomposed and ordered by the fraction of significant associations they contributed to the major variances (Figure 4A). The first two MOFA factors explained the most variance that differentiated the unchallenged, challenged, and recovering groups on the MOFA scatter plot (Figure 4B). In addition, covariate (phenotype) correlation analysis demonstrated that the first two MOFA factors were associated with the majority of the covariates (Figure 4C) where factor 1 (FA1) was particularly associated with covariates related to infection ($r < -0.6$) and factor 2 (FA2) was associated with BWG ($r=0.56$); associations not identified in correlations of single omics analyses.

Multi-omics networks can contextualize the multiple types of microbiome disruption associated with various biological molecules found in different health statuses [28]. Additionally, a network's hotspot molecular features (hubs and clusters) can highlight targets to be followed-up. Here, we conducted network analyses downstream of MOFA to explore biomarkers that might associate with anticoccidial vaccination. Network analyses for the MOFA factors showed sub-structures (clusters of intercorrelated features) that were enriched in each MOFA

factor (Figure 4D-E). Three clusters were identified from FA1 components; two associated with *Eimeria* challenged chickens (including unvaccinated, vaccinated and recovering groups) (FA1-C1 and C3 in Figure 4D), whilst the third associated exclusively with unchallenged chickens (FA1-C2 in Figure 4D). Additionally, cluster 1 in the FA2 network demonstrated associations between unchallenged/recovering groups and the 6 dpi-challenged group (FA2-C1 in Figure 4E). Clustered components from the FA1 and FA2 networks associated with non-challenge and recovery were enriched by vitamin B and derivatives (e.g., pyridoxine, riboflavin, nicotinate derivatives), short-chain fatty acids (e.g., butyrate/isobutyrate and valerate), and short-chain fatty acid-producing bacteria (e.g., *Caproicibacter fermentans* and *Ruminococcoides bili*). Itaconate, an antipathogenic organic acid was enriched in recovering chickens. In contrast, uremic toxin (e.g., p-cresol sulfate), the long-chain fatty acids and derivatives (e.g., 14–18C fatty acids and glycerophospholipids (GPs), glycerophosphocholine (GPC), phosphoethanolamine (PE) derivatives), metabolites of fatty acid metabolism (eicosenoylcarnitine and docosadienoylcarnitine), and gut pathogens (e.g., *Clostridium difficile* and *C. innocuum*) and commensal bacteria (e.g., *Escherichia coli*, *Clostridium bolteae*, and *Fecalibacterium prausnitzii*), were enriched in post-*Eimeria* challenged associated clusters of both networks (Figure S3).

MOFA models discover potential signature markers of host response to challenge after vaccination

While highlighting the covariation patterns of disease status, the MOFA model constructed using data from all samples did not reveal factors specifically associated with unvaccinated-challenged and vaccinated-challenged (mock and true vaccines) chickens. A more focused MOFA model was performed on all 6 dpi challenged groups to identify signature markers after vaccination. In the second model, the first four MOFA factors contributed to the major variation of the data and the fraction of significant associations (Figure 5A). Interestingly, the phenotypic and pathological covariates were more closely associated with FA4 and FA11 (e.g., lesion score severity was more associated with FA4 than other FA; $r=-0.57$). Vaccine treatment conditions (Yeast: treating with yeast vectors or not; Vaccination: treating with the true vaccine or not) were negatively associated with FA4 and FA11, and the parasite load (qPCR ratio) was associated with both FAs ($r=-0.59$ and -0.37) (Figure 5C). Comparison of FA4 and FA11 using a scatter plot demonstrated that FA4 clearly distinguished the treatment condition of yeast vectors between unvaccinated (UV-C) and vaccinated groups (MV-C and V-C). FA11 showed a different trend between the mock vaccine group (MV-C) and the true vaccine group (V-C) (Figure 5B). Using network analysis, the signature features of various sphingolipids (e.g., sphingosine,

269 sphingomyelin) and *Ruminococcus lactaris* were clustered from both FAs and
270 enriched in most vaccinated subjects; whereas the long-chain fatty acids (e.g.,
271 linoleoyl-arachidonoyl-glycerol and oleoyl-oleoyl-glycerol) were enriched in
272 unvaccinated-unchallenged chickens (Figure 5D-E, and Figure S4).
273

DISCUSSION

An experimental yeast-vectored anticoccidial vaccine has recently been described as a step towards improved control of *Eimeria* species such as *E. tenella*, which cause coccidiosis in chickens [10]. Small-scale studies under commercial conditions found that vaccination could partially control the direct consequences of live parasite challenge, reducing parasite replication and its associated enteric pathology, while protecting performance (BWG and feed conversion ratio). In the present study, we have assessed the impact of vaccination on indirect consequences of *Eimeria* infection including microbial dysbiosis and metabolic disruption.

Using 16S rDNA amplicon sequencing from caecal contents and caecal tissue of experimentally vaccinated and challenged chickens (specific pathogen-free sourced), the microbiota richness was lower than in a previous farm study (Hay et al. 2023), likely due to the necessary enrofloxacin medication in the early rearing period to control an outbreak of colibacillosis. Indeed, it is well described that antibiotic treatment severely impacts on microbiome composition and richness, but recovering to its normal composition after stopping the treatment. While this was unexpected, such treatments are common under field conditions. Microbiota richness is also expected to be higher in populations of mixed breed chickens reared in the field under varied husbandry regimes than under the controlled conditions used in the present study. Comparison between caecal contents and tissues found no significant differences in alpha diversity ($p=0.87$) and beta diversity ($p=0.052$) (Figure S1). Only *Anaerotruncus (A.) colihominis* and *Flavonifractor (F.) plautii* were consistently enriched in caecal tissue samples across multiple groups (UV-UC, MV-C, and V-C), indicating their association with the intestinal mucosal environment. *A. colihominis*, originally isolated from mouse colonic mucosa by the Leibniz Institute DSMZ, has been detected in the intestinal lumen and stool samples of patients with bacteraemia and colorectal cancer, suggesting a potential broader role in gut dysbiosis and pathology [29, 30]. Similarly, *F. plautii*, known for its ability to degrade flavonoids and potentially mucins, was isolated by Levine *et al.* from mammalian intestinal mucosa [31]. Its presence in these tissue samples underscores its importance in gut health and disease [31, 32]. Based on our findings, investigation of caecal contents alone appears to be sufficient to investigate total gut microbiota because these reflect the primary condition of the intestinal ecosystem.

Eimeria infection is known to predispose chickens to diseases such as necrotic enteritis, caused by *C. perfringens* [33], and can disrupt enteric microbial populations leading to dysbiosis [21]. We anticipated that beta diversity, but not alpha diversity,

would change following *Eimeria* challenge. However, although average richness was lower in unchallenged chickens, the difference was not statistically significant (Figure 2). Comparison of bacterial abundance between infected and non-infected chickens revealed increased Gammaproteobacteria and pathogenic *Clostridia* in *Eimeria*-challenged chickens. Common gastrointestinal pathogens, including *Escherichia* (*E.*) *coli*, *Clostridium* (*C.*) *difficile*, *Enterococcus* (*E.*) *caecorum*, *Proteus mirabilis*, and *Clostridium* (*C.*) *innocuum*, were also in higher abundance (Figure 2G) suggesting significant dysbiosis occurred following *Eimeria* challenge infection. It is notable that some strains of *E. caecorum* have been reported to cause high morbidity and mortality in broiler chickens [34]. Additionally, *C. difficile* and *C. innocuum* can cause antibiotic-associated diarrhea and have shown vancomycin resistance [35, 36], suggesting that a compromised gut environment may facilitate colonization by antibiotic-resistant strains; this condition mirrors the mechanism of human pseudomembranous colitis, which arises due to the overgrowth of *C. difficile* following extensive antibiotic usage. *Eimeria* infection can alter the gut microenvironment by increasing intestinal permeability and inflammation [37], thereby interacting bidirectionally with the gut microbiota. Consequences of enteric dysbiosis include immune dysregulation causing gut-related disorders such as allergies, inflammatory bowel disease (IBD) and autoimmune disorders [38-40]. Thus, *Eimeria* challenge is likely to activate a synergistic response between the host's physiology and the commensal gut microbiota. Intestinal infections can decrease oxygen levels and lead to chronic tissue and mucosal hypoxia with dysregulation of activation of hypoxia-inducible factors (HIFs) and NF- κ B, exacerbating inflammation and injury of intestinal tissues [41, 42]. The metabolic environment of the mucosa is also altered during inflammation since the Enterobacteriaceae require terminal electron acceptors from the mucosa for anaerobic respiration and blooming [43, 44]. Inflammatory cells release ROS and RNS, forming NO₃⁻ as a terminal electron acceptor for Gammaproteobacteria growth via denitrification [43, 45-51].

In all *Eimeria*-challenged groups (UV-C, MV-C, V-C) the gut microbiota composition was similar (Figure 2A). However, yeast treatment groups (MV-C and V-C) showed a significant increase in *E. coli* abundance, nearly double in UV-C and over six times higher in UV-UC. While harmful *E. coli* may increase due to infection, it is possible that some protective *E. coli* strains that can stimulate an innate immune mechanism [52] and produce vitamins [53, 54] colonize after the reversion of dysbiosis. Notably, *Bifidobacterium*, a common lactic acid-producing probiotic, was present in both yeast treatment groups, irrespective of *Eimeria* antigen expression. In addition, Lactobacillales family bacteria (*Enterococcus*, *Lactobacillus*, *Pediococcus*) were

enriched in both non-infected and yeast treatment groups, with *Lactobacillus* and *Pediococcus* being particularly higher in non-infected groups. This enrichment suggests a beneficial modulation of the gut microbiota. Yeasts and lactic acid-producing bacteria, often found together in nature [55], decrease pH value during fermentation creating an unfavourable environment for some pathogens [56, 57].

We used a multi-omics integrative tool, MOFA, to infer how the host metabolome interacts with gut microbes under a range of vaccination and *Eimeria* infection conditions. MOFA modelling confirmed that metabolites involved in fatty acid metabolism and β -oxidation pathways were altered by *Eimeria* infection [23]. Inflammation and oxidative stress induced by *Eimeria* invasion and subsequent pathology increase demand for metabolites involved in fatty acid metabolism [58]. The model found that carnitine derivatives such as eicosenolcarnitine and docosadienolcarnitine, intermediate metabolites involved in fatty acid metabolism, were enriched in the *Eimeria* challenged groups (challenge groups compared to non-challenge and recovering groups; factor 2 of MOFA model 1). In addition, p-cresol sulfate (pCS), a uremic toxin formed by gut microbial fermentation of tyrosine [59, 60], was also enriched in all challenged groups, especially in unvaccinated, challenged chickens (factor 1 of MOFA model 1). The main producer of pCS, *C. difficile*, a significant cause of diarrhoea during microbial ecosystem collapse, was also identified (factor 1 of MOFA model 1) [61, 62]. These findings link both layers of omics and prove evidence that *Eimeria* infection causes dysbiosis.

Since the first MOFA model (the full model with all groups of the trial) could not distinguish an effect of vaccination among the challenged, non-challenged and recovering groups, a second MOFA model was used to explore latent grouping among vaccinated and non-vaccinated chickens. We found sphingolipids, including sphingosine, sphingomyelin, and sphingoinositol, were significant factors associated with vaccination. Sphingolipids are required in cell membrane structures of eukaryotes (especially the Schwann's cell, which surrounds the neuron axon) and some prokaryotes [63], as well as essential signalling molecules of inflammatory, immunity, cell autophagy, growth, and survival regulations [63-67]. Brown et al. (2019) indicated that the microbe-derived sphingolipids (especially from *Bacteroides*) are negatively correlated with gastrointestinal inflammation (i.e., inflammatory bowel disease) and maintaining homeostasis and symbiosis of gut microbiota [68]. This finding supports the efficacy of the yeast-based oral anti-coccidiosis vaccine and indicates that the vaccine can alter the symbiosis status of gut microbiota. However, only a few reads of *Bacteroides* were detected from yeast-based vaccine-treated

samples and non-*Eimeria*-challenged samples (including from caecal tissues and contents), possibly due to the early antibiotic treatment of all study subjects. It implies that the microbial anti-inflammatory sphingolipids could be produced via other microbial species in the chicken gut microbiota, then act as a signal of anti-coccidiosis for the further applications.

In conclusion, using MOFA machine learning to integrate evaluation of potential interactions between the enteric microbiome and host metabolism provided a mechanistic insight into effects of anticoccidial vaccination and *Eimeria* challenge. In the present study, we identified Gamma-proteobacteria, *p*-cresol sulfate, *Bifidobacterium*, carnitine-derived metabolites and sphingolipids as host-microbe-associated biomarkers that vary between healthy, infected, vaccinated and/or recovering chickens, providing insights into potential strategies for controlling, treating and preventing coccidiosis. As we look to the future, the findings of this study are poised to contribute to the advancement of precision agriculture, particularly in enhancing poultry health management and the development of novel interventions against coccidiosis.

MATERIALS AND METHODS

Ethics statement

The animal experiments in this study were approved by the Royal Veterinary College (RVC) Animal Welfare Ethical Review Body (AWERB) and performed under the Animals in Scientific Procedures Act 1986 (ASPA) with a UK Home Office Licence.

Study animals, metadata measurement and study design

Cobb500 broiler chickens were purchased from P. D. Hook (Hatcheries) Ltd. (Cote, UK) at day of hatch. All chickens received enrofloxacin (Baytril®, Bayer, Leverkusen, Germany, 10 mg Kg⁻¹) from days 16 to 18 of the trial due to an outbreak of colibacillosis. Feeding and vaccination treatments were as described in a previous study (Study 4 in [10]). Briefly, four groups of ten chickens were sampled from a larger vaccination study six days post *E. tenella* challenge including (1) unvaccinated, challenged (UV-C), (2) unvaccinated, unchallenged (UV-UC), (3) mock vaccinated, challenged (MV-C), and (4) vaccinated, challenged (V-C) groups. A fifth group of eight unvaccinated, challenged chickens were sampled ten days post challenge (UV-C10; **Table S1**). Mock and experimental vaccines were administered by oral inoculation in 100 µl phosphate buffer saline (PBS) every 3-4 days from day 7 of age (five doses per chicken in total). Group 3 (MV-C) was vaccinated using a mock vaccine including *S. cerevisiae* EBY100 strain (Invitrogen, Thermofisher Scientific, Waltham, MA, USA) containing the empty yeast display plasmid vector pYD1 (Invitrogen). Group 4 (V-C) was vaccinated at the same timepoints by oral inoculation of an experimental trivalent formulation of *S. cerevisiae*-vectored recombinant vaccine using pYD1 to separately express each of three *E. tenella* antigens including EtAMA1 ectodomain [11], EtIMP1 [12] and EtMIC3 [13]. The vaccine design and administration procedures were as described previously [10]. Groups 1, 3, 4 and 5 were challenged by oral inoculation with 15,000 sporulated *E. tenella* Houghton strain oocysts at 21 days of age. Challenge oocysts were prepared and inoculated following established protocols [69]. Caeca (paired) were collected immediately post-mortem at six or ten days post infection (dpi, Groups 1-4, and 5, respectively). The severity of infection was assessed using the Johnson and Reid scoring system [70]. Overall production performance was defined by Body Weight Gain (BWG) between 0 and 6 dpi. Parasite replication was measured using quantitative PCR for parasite genomes per host genome [10].

DNA extraction and 16S amplicon sequencing

Bacterial genomic DNA was extracted separately from caecal tissue (~100 mg) and caecal contents (~200 mg) using a QIAamp Fast DNA Stool Mini kit (QIAGEN, Valencia,

CA, USA) following the manufacturer's pathogen detection protocol. 16S amplicon library preparation followed the Illumina 16S Metagenomic Sequencing Library Preparation guidelines [71]. The 16S ribosomal RNA (rRNA) gene V3–V4 hypervariable regions were amplified by PCR with the adapter overhang primers 341F (5' -TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGC CCTACGGGNGGCWGCAG-3') and 805R (5' -GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG GACTACHVGGGTATCTAATCC-3') for 25 cycles. Indices and Illumina sequencing adapters were attached using the Nextera XT Index Kit with 8 cycles of a second amplification reaction. The final PCR products were purified using AMPure XP beads (Beckman Coulter, Brea, CA, USA). The amplicon DNA concentration was measured using Qubit dsDNA HS and BR Assay Kits (Thermo Fisher Scientific, Waltham, MA, USA). Library quality was determined using the Agilent Technologies 2100 Bioanalyzer system with a DNA-1000 chip. Eighty-eight samples representing caecal tissues from all chickens in Groups 1-5 (n=48) and caecal contents from all chickens in Groups 1-4 (n=40) were pooled with equal molality. The 16S amplicon libraries were sequenced using a 301 bp paired-end (301bp x 2) approach on an Illumina MiSeq platform using V3 chemistry.

Bioinformatic processing and microbiota analyses

The Illumina MiSeq platform generated a total of 22,525,182 paired-end sequences. Sequences were cleaned by sequence length ≥ 300 bp using Trimmomatic [72]. The 16S amplicon sequences were processed using the Quantitative Insights Into Microbial Ecology 2 (QIIME 2) pipeline (version 2019.10) [73]. Primer sequences were removed by Cutadapt (version 1.15) [74]. Trimmed sequences were truncated at 240 bp (forward) and 210 (reverse) and denoised using the DADA2 algorithm [75]. Amplicon sequence variants (ASVs) were obtained via the denoising process with quality filtering and chimera removal. A k-mer based re-clustering algorithm 'KTU' [76] was subsequently applied to assemble ASVs into optimal biological taxonomic units (KTUs). KTUs taxonomy was assigned by comparison with the SILVA SSU reference nr99 (v138) [77, 78] and NCBI 16S RefSeq (retrieved 10th Feb. 2022) databases using the kaxonomy function of the KTU R-package. Eukaryotic organelle 16S sequences (identified as *Eimeria*) were extracted and used for supplementary parasite load quantification; non-prokaryotic and unassigned KTUs were removed from the microbiota dataset. The 309 KTU microbiota dataset was rarefied at the minimum read counts among samples (10,034 reads) after removing twelve samples with shallow sequence depth (< 10,000 reads).

Microbiota analyses were conducted and visualized using the Microbiome Analysis R

code (MARco) [79], Community Ecology ‘vegan’ [80], and Pretty Heatmap (pheatmap) [81] packages in R (version 4.0.1) [82]. The ANOVA test with Tukey HSD *post-hoc* multiple comparison test or Kruskal-Wallis test with Dunn’s *post-hoc* multiple comparison test were used for parametric and non-parametric statistical analyses of group comparisons with a significance level of $\alpha = 0.05$, and the P values were adjusted with a false discovery rate (FDR). Alpha diversity indices were estimated by richness. Beta diversity of microbial communities was measured by Bray-Curtis dissimilarity using principal coordinates analysis (PCoA), and heterogeneity was tested using ADONIS and ANOSIM tests.

Metabolome profiling

Untargeted metabolome profiling of caecal tissues was performed by Metabolon (NC, USA) using their vendor protocol. Briefly, all samples were deproteinized by dissociating small molecules bound to protein or trapped in the precipitated protein matrix. To recover chemically diverse metabolites, methanol was used for protein precipitation under vigorous shaking for 2 min (Glen Mills GenoGrinder 2000) followed by centrifugation. The extract was aliquoted into five fractions: two for analysis by separate reverse phase (RP)/UPLC-MS/MS methods with positive ion mode electrospray ionization (ESI), one for analysis by RP/UPLC-MS/MS with negative ion mode ESI, one for analysis by HILIC/UPLC-MS/MS with negative ion mode ESI, and one sample was reserved as backup. Samples were placed briefly on a TurboVap® (Zymark) to remove the organic solvent. The sample extracts were stored overnight under nitrogen before preparation for analysis.

All methods used Waters ACQUITY ultra-performance liquid chromatography (UPLC) and a Thermo Scientific Q-Exactive high resolution/accurate mass spectrometer interfaced with a heated electrospray ionization (HESI-II) source and Orbitrap mass analyzer operated at 35,000 mass resolution. Each sample extract was dried then reconstituted in solvents compatible to each of the four methods. Each reconstitution solvent contained a series of standards at fixed concentrations to ensure injection and chromatographic consistency. One aliquot was analyzed using acidic positive ion conditions, chromatographically optimized for more hydrophilic compounds. The extract was gradient eluted from a C18 column (Waters UPLC BEH C18-2.1x100 mm, 1.7 μ m) using water and methanol, containing 0.05% perfluoropentanoic acid (PFPA) and 0.1% formic acid (FA). Another aliquot was also analyzed using acidic positive ion conditions; however, it was chromatographically optimized for more hydrophobic compounds. The extract was gradient eluted from the same afore mentioned C18 column using methanol, acetonitrile, water, 0.05%

PFFA and 0.01% FA and was operated at an overall higher organic content. Another aliquot was analyzed using basic negative ion optimized conditions using a separate dedicated C18 column. The basic extracts were gradient eluted from the column using methanol and water, however with 6.5mM ammonium bicarbonate at pH 8. The fourth aliquot was analyzed via negative ionization following elution from a HILIC column (Waters UPLC BEH Amide 2.1x150 mm, 1.7 μ m) using a gradient consisting of water and acetonitrile with 10mM ammonium formate, pH 10.8. The MS analysis alternated between MS and data-dependent MSⁿ scans using dynamic exclusion. The scan range varied slightly between methods, but covered 70-1000 m/z.

Raw data were extracted, peak-identified and QC processed by Metabolon's in-house systems. Compounds were identified by comparison to library entries of purified standards or recurrent unknown entities. The in-house library was built and maintained by Metabolon, and contained more than 3,300 commercially available purified standard compounds with the information of retention time/index (RI), mass to charge ratio (m/z), and chromatographic data (including MS/MS spectral data). Compound identification was based on the following criteria: retention index within a narrow RI window of the proposed identification, accurate mass match to the library +/- 10 ppm, and the MS/MS forward and reverse scores between the experimental data and authentic standards.

A subset of 1,180 metabolites was detected from the untargeted metabolomics screen. Each metabolite's peak area (i.e. total ion counts, integrated area-under-the-curve) was median-scaled to normalize. The missing values were then imputed with the observed minimum of each metabolite. Since the metabolomic data were typically close to log-normal distribution, the normalized-imputed data were transformed using the natural log for subsequent analyses.

Multi-Omics Factor Analysis (MOFA) model for microbiota and metabolome integrative analysis

MOFA model fittings were performed to integrate multi-omics data modalities based on an unsupervised machine learning model formulated in a probabilistic Bayesian framework. The 16S rRNA amplicons of caecal tissue and content, and host caecal metabolome were the separate data modalities in this study. In order to make all omics data comparable, the amplicon abundance was centered log-ratio transformed using the 'clr' function of the compositions R-package. Spearman's correlation (FDR <0.05) was implemented to select associated features from the omics datasets [83]. Downstream characterization was performed by variance decomposition, detecting

the fraction of significant associations between the features and each factor using Pearson's correlation ($FDR < 0.1$), and correlation of phenotype covariates. A subgrouped MOFA model fitting was performed on all 6 dpi challenged groups. A network analysis for identifying sub-structures of MOFA factors was performed with the R package igraph⁴⁷ [84]. An adjacency matrix based on Spearman's correlation coefficients of intercorrelated features was constructed from a MOFA factor of interest; these coefficients were also used for assessing length of edges on the network. The latter was conducted with the fast greedy modularity optimization algorithm [85] to identify clusters in the network.

SUPPORTING INFORMATION

Figure S1. Comparisons of microbiota profiles between chicken caecal contents and caecal tissues. (A) Beta diversity analysis with the PERMANOVA test demonstrated no significant differences between caecal contents and caecal tissues ($p=0.052$). (B) Venn diagrams show 62.7-73.6% microbiota composition shared between caecal contents and tissues in four study groups (UV-C, UV-UC, MV-C, and V-C). (C) Differential abundance ($p<0.05$ by DESeq2 test) of microbes between caecal contents and caecal tissues in four study groups, respectively.

Figure S2. Lesion score and *Eimeria* load correlated metabolites. (A) Categories of non-infection-associated metabolites. (B) Categories of infection-associated metabolites.

Figure S3. Details of subnetworks from the MOFA model. Microbial and metabolomic features of each cluster (C1-C3) in FA1(Figure 4D) and FA2(Figure 4E)—the abundance of microbial and metabolomic features among groups were presented by heatmap. The left columns were annotated by sample sources (C: caecal content microbiome; T: caecal tissue microbiome; M: caecal tissue metabolite).

Figure S4. Details of subnetworks from 6dpi -challenged MOFA model. Microbial and metabolomic features of each cluster (C1-C2) in FA4(Figure 5D) and FA11(Figure 5E)—the abundance of microbial and metabolomic features among groups were presented by heatmap. The left columns were annotated by sample sources (C: caecal content microbiome; T: caecal tissue microbiome; M: caecal tissue metabolite).

Table S1. Sample information of studying groups

Table S2. Infection-associated metabolites

DATA AVAILABILITY.

The original data set presented in the study is publicly available. These data can be found at NCBI under BioProject accession number: PRJNA990995.

ACKNOWLEDGEMENTS

We thank Dr Meiyeh Jade Lu from the NGS Core, Academia Sinica, and Ms Yu-Tang Yang from the National Taiwan University College of Medicine for the technical consultancy of the NGS library construction. We gratefully acknowledge the support of the Houghton Trust Research Grants for funding this project.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

AUTHORS' CONTRIBUTION

Conceptualization: P-Y.L., D.X., F.M.T., and D.P.B. Performed experiment: P-Y.L., J.L., F.S., J.J.O, and D.P.B. Data analysis: P-Y.L. Contributed data or analysis tools: F.S., D.W., and O.G. Writing—original draft: P-Y.L. and D.X. Writing—review and editing: P-Y.L., D.X., and D.P.B.

FIGURES & TABLES

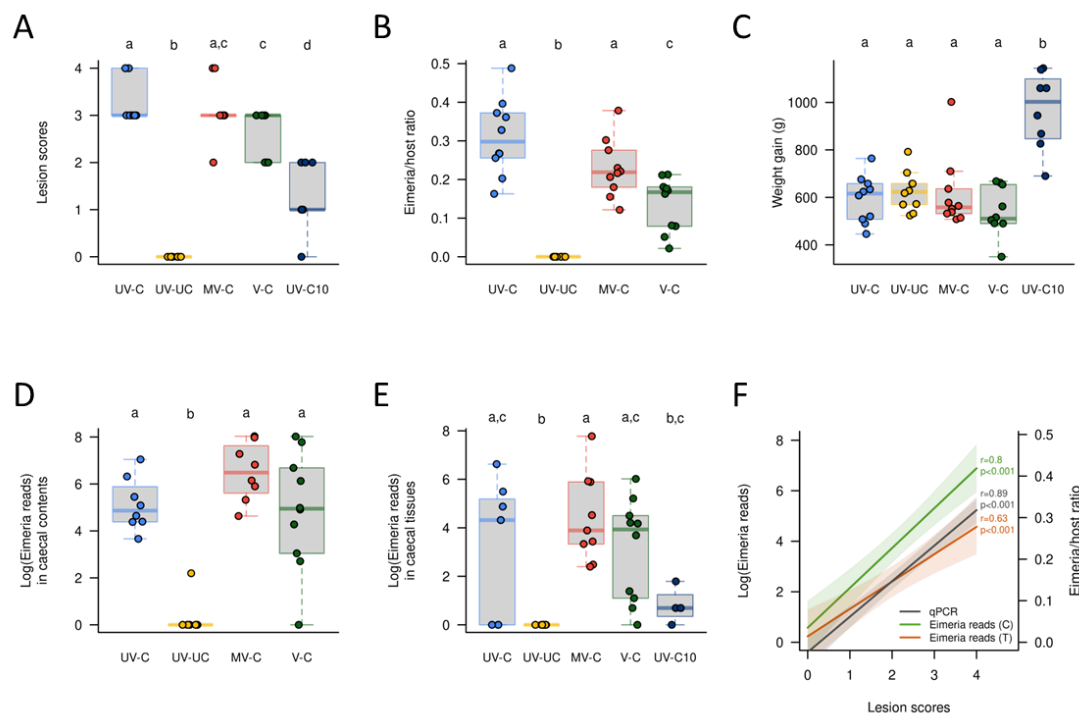


Figure 1. Summary of vaccine trial phenotypes assessed six days post infection (dpi) with 15,000 sporulated *Eimeria tenella* oocysts. (A) Caecal lesion scores, (B) parasite load represented as parasite genomes per host genome, determined using qPCR, (C) bodyweight gain from 0 to 6 dpi, (D and E) parasite load represented by *Eimeria* mitochondrial 16S rRNA sequence reads in caecal contents and tissue, (F) association between caecal lesion score and parasite load measures. Panels A-C reanalysed from Soutter et al., 2022.

Groups UV-C: unvaccinated, challenged, UV-UC: unvaccinated, unchallenged, MV-C: mock vaccinated, challenged, V-C: vaccinated, challenged.

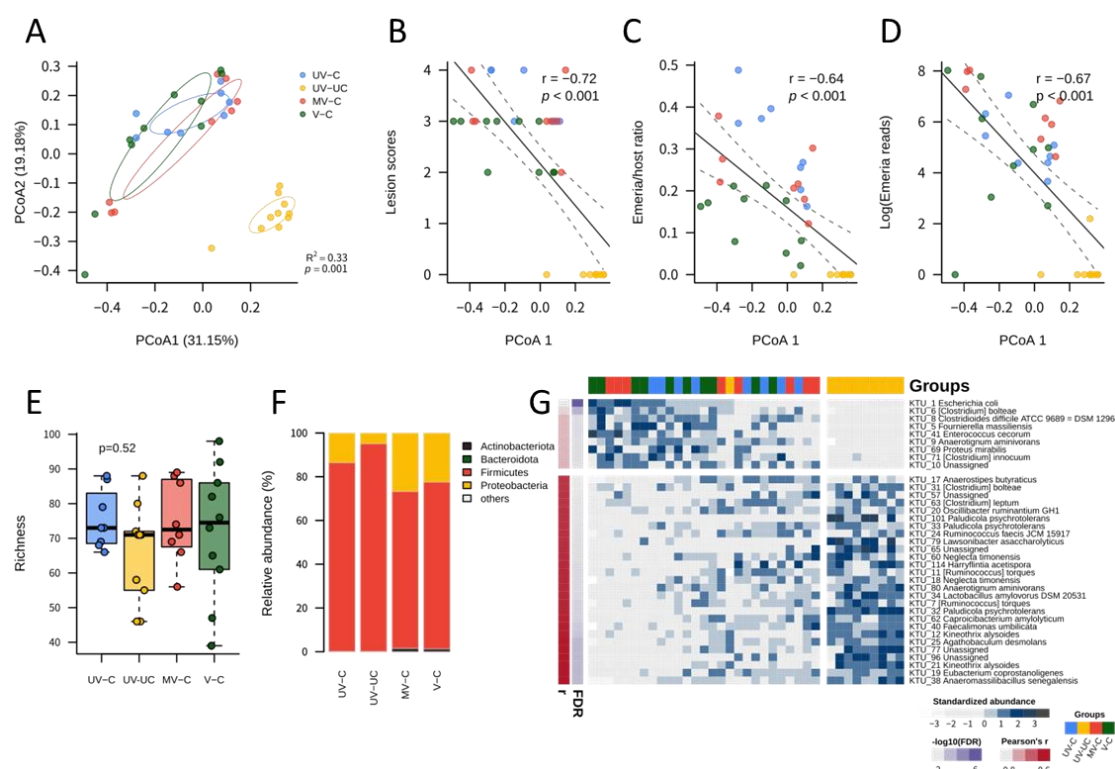


Figure 2. Gut microbiota profiling and associations of gut pathology and parasite loads. (A) Principal coordinates analysis (based on Bray-Curtis distance) for beta diversity of gut microbiota composition (caecal contents) among four groups of chickens. (B) The correlation between microbiota composition and lesion scores. (C)-(D) The correlations between microbiota composition and parasite loads, (C) based on qPCR quantification of the ratio of *Eimeria* and host genes, (D) based on NGS reads of *Eimeria* apicoplast 16S rRNA gene. (E) Alpha diversity (richness) of four groups of chickens. (F) Relative abundance of gut microbiota composition at the phylum level. (G) Gut pathology and parasite load associated microbes (Pearson's $r > 0.4$ or < -0.4 , FDR-adjusted $p < 0.05$) extracted from PCoA1 of panel (A). Groups UV-C: unvaccinated, challenged, UV-UC: unvaccinated, unchallenged, MV-C: mock vaccinated, challenged, V-C vaccinated, challenged.

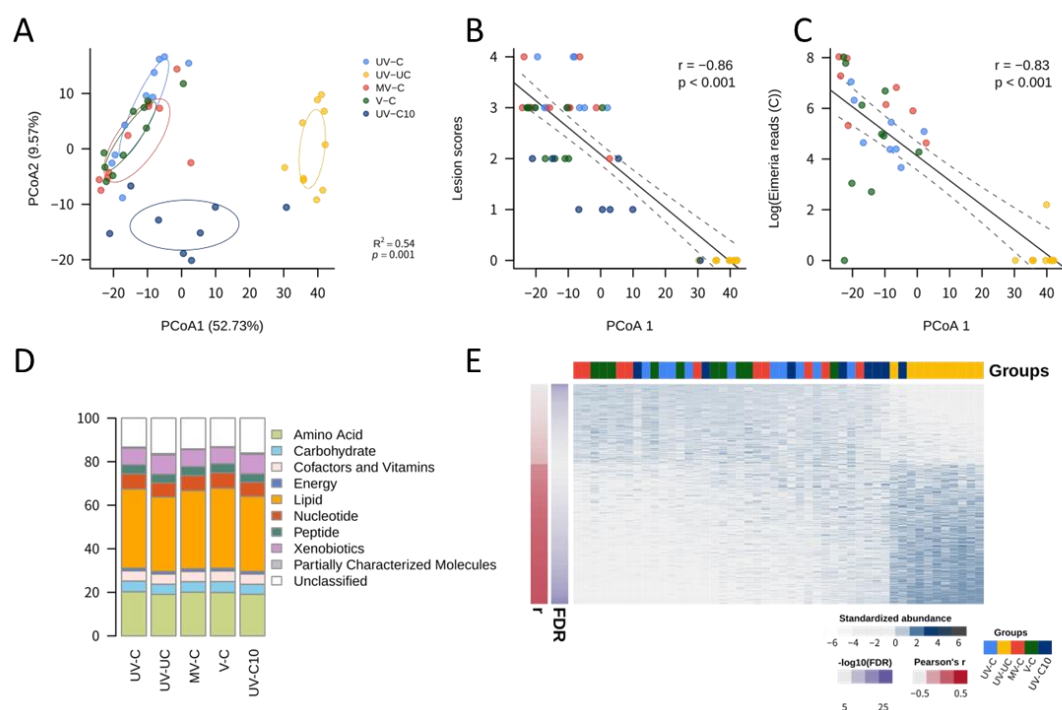


Figure 3. Chicken caecal tissue metabolome profiling and associations of gut pathology and parasite loads. (A) Principal coordinates analysis (based on Euclidean distance) for chicken caecal tissue metabolome composition among five groups of chickens. (B) The correlation between metabolome composition and lesion scores. (C) The correlation between metabolome composition and parasite loads, based on NGS reads of *Eimeria* apicoplast 16S rRNA gene. (D) Compositions and categories of the metabolome of five groups of chickens. (E) Gut pathology and parasite load associated metabolites (Pearson's $r > 0.4$ or < -0.4 , FDR-adjusted $p < 0.05$) extracted from PCoA1 of panel (A).

Groups UV-C: unvaccinated, challenged, UV-UC: unvaccinated, unchallenged, MV-C: mock vaccinated, challenged, V-C vaccinated, challenged, UV-C10: unvaccinated, challenged, 10 days post-infection.

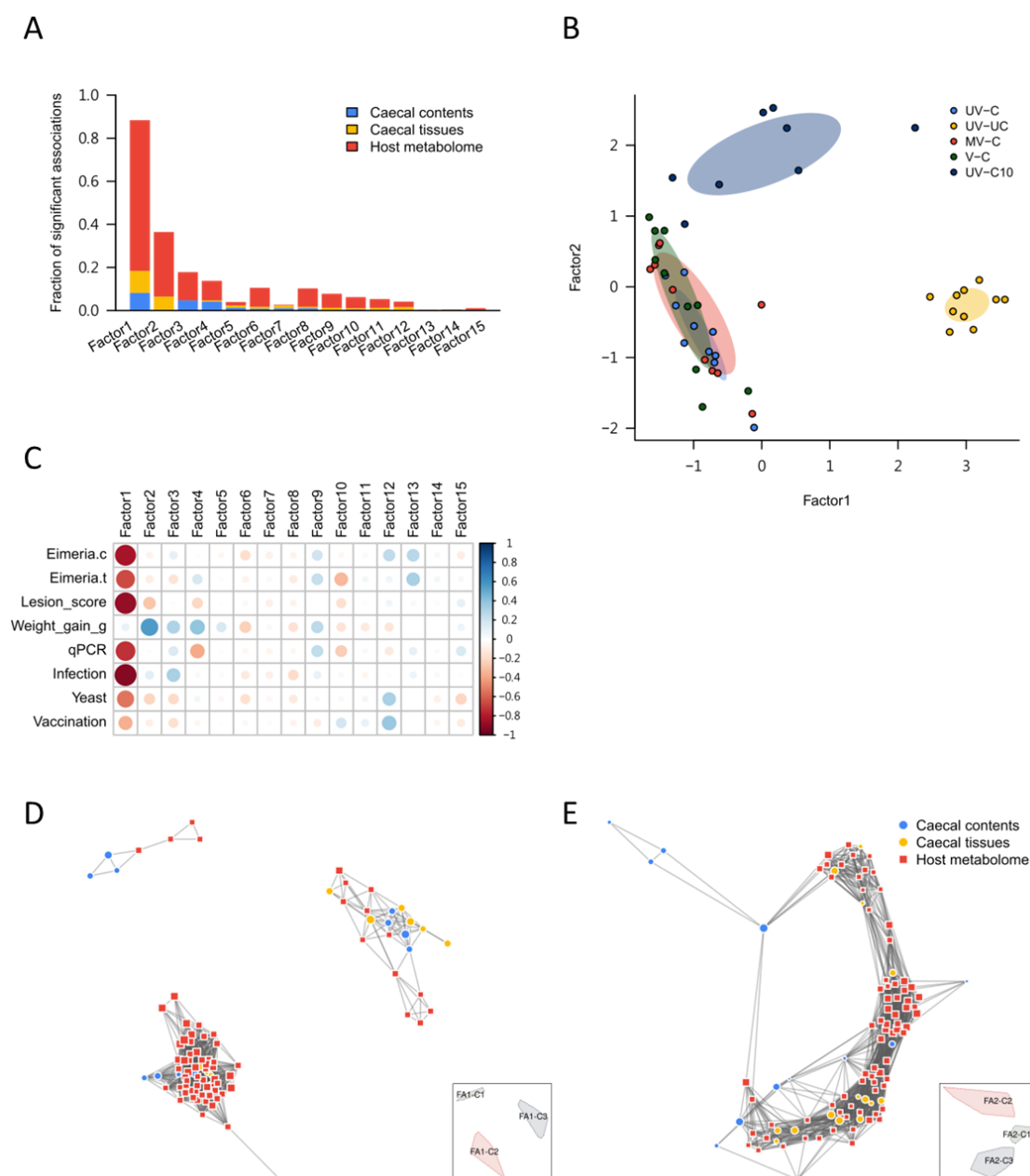


Figure 4. MOFA model for all trial groups and downstream signature marker identification by network analysis. (A) Bar plots showing the fraction of significant associations between the features of each microbiome or metabolome modality and each factor. The stacked bars interpret whether the variance-explained values are driven by a strong change in a small number of features or by a moderate effect across a large range of features. (B) Scatterplot of factor 1 (x axis) versus factor 2 (y axis). Each dot represents a sample, colored by the trial group. (C) The correlation heatmap of MOFA factors and phenotypes (Eimeria.c: NGS read-based *Eimeria* load in caecal contents; Eimeria.t: NGS read-based *Eimeria* load in caecal tissues; qPCR: qPCR-based *Eimeria* load in caecal tissues; Infection: infection condition- infected or non-infected; Yeast: yeast vector exposure or not; Vaccination: vaccination condition-

vaccinated or non-vaccinated). (D)-(E) Network analysis and visualization for the features from (D) factor 1 and (E) factor 2. Thumbnail legends present the regions of subnetworks.

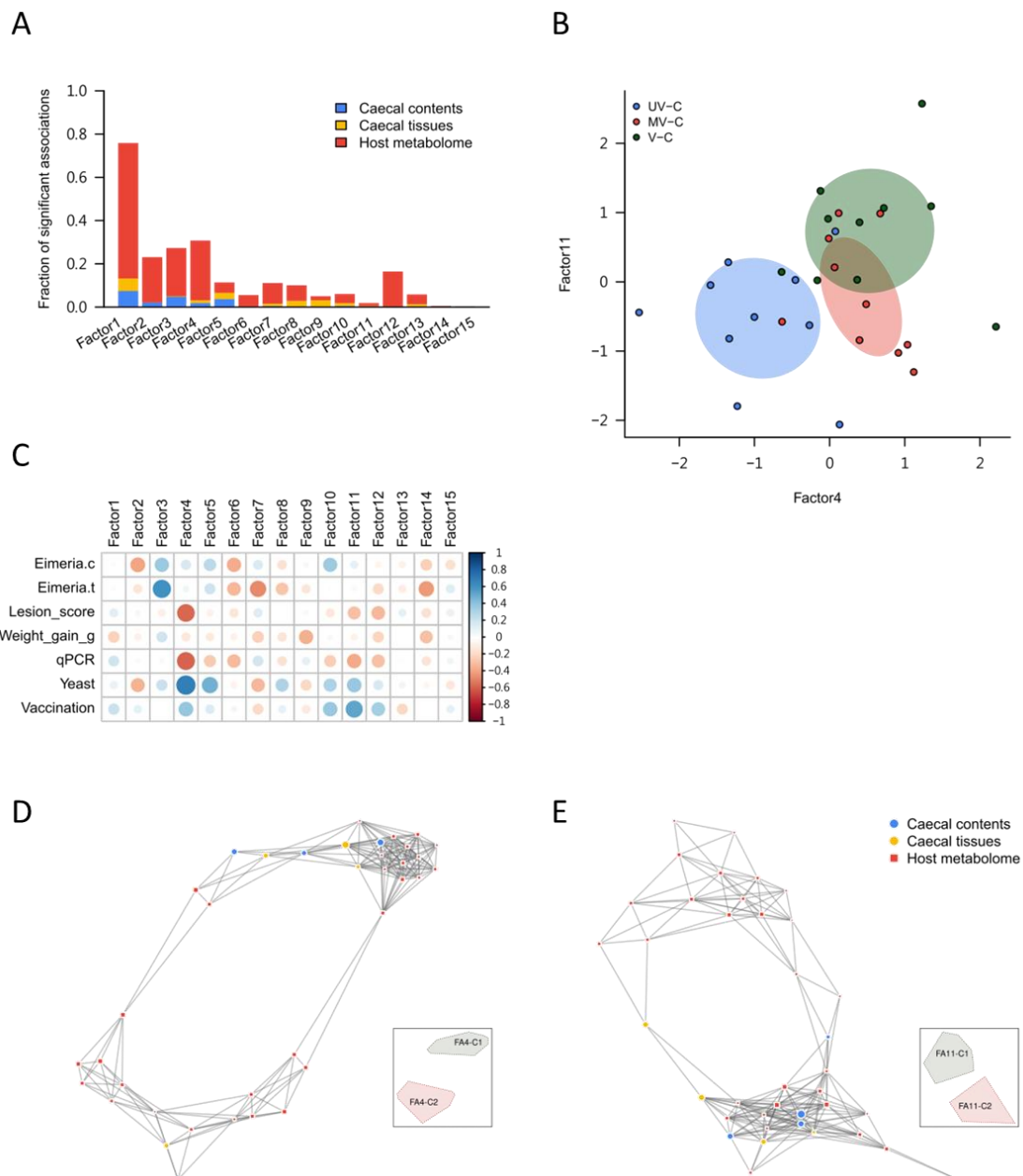
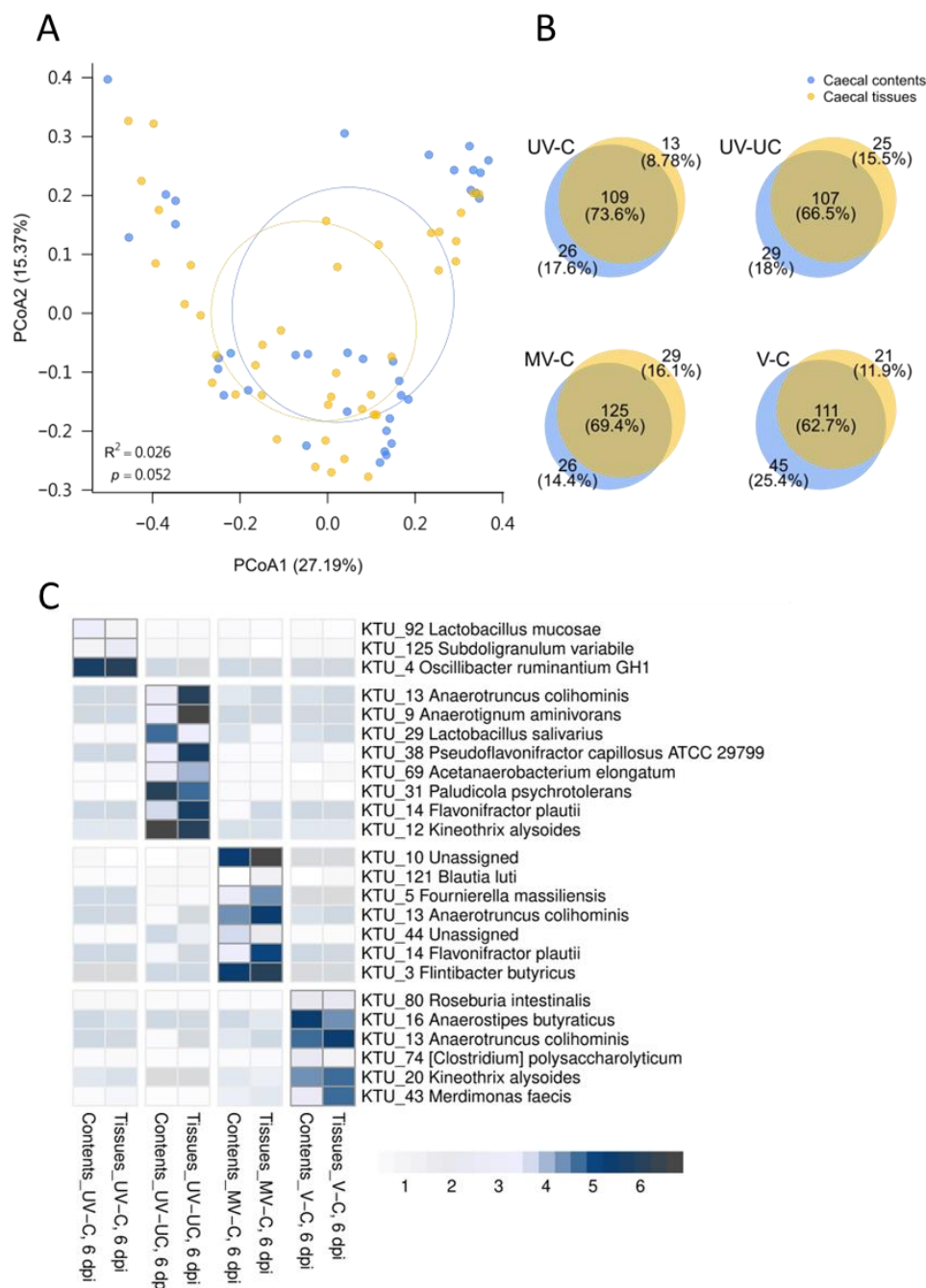


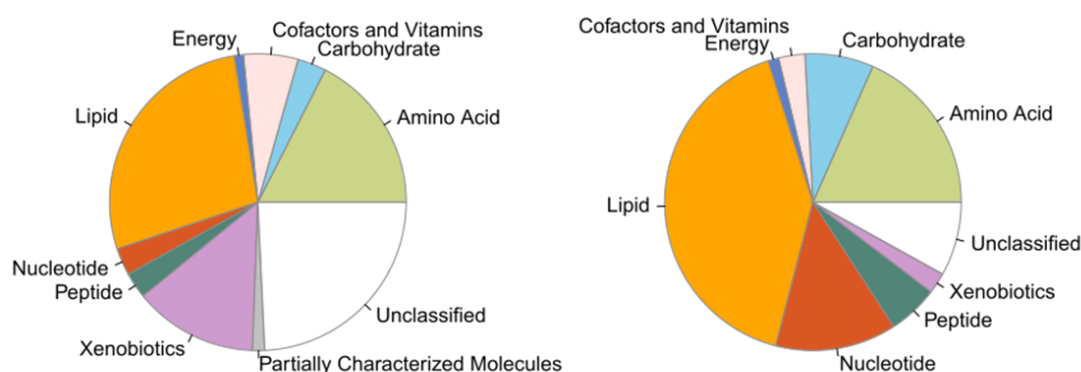
Figure 5. MOFA model for challenged 6dpi chickens and downstream signature marker identification by network analysis. (A) Bar plots showing the fraction of significant associations between the features of each microbiome or metabolome modality and each factor. The stacked bars interpret whether the variance-explained values are driven by a strong change in a small number of features or by a moderate effect across a large range of features. (B) Scatterplot of factor 4 (x axis) versus factor 11 (y axis). Each dot represents a sample, colored by the trial group. (C) The

683 correlation heatmap of MOFA factors and phenotypes (*Eimeria*.c: NGS read-based
684 *Eimeria* load in caecal contents; *Eimeria*.t: NGS read-based *Eimeria* load in caecal
685 tissues; qPCR: qPCR-based *Eimeria* load in caecal tissues; Yeast: yeast vector
686 exposure or not; Vaccination: vaccination condition- vaccinated or non-vaccinated).
687 (D)-(E) Network analysis and visualization for the features from (D) factor 4 and (E)
688 factor 11. Thumbnail legends present the regions of subnetworks.
689



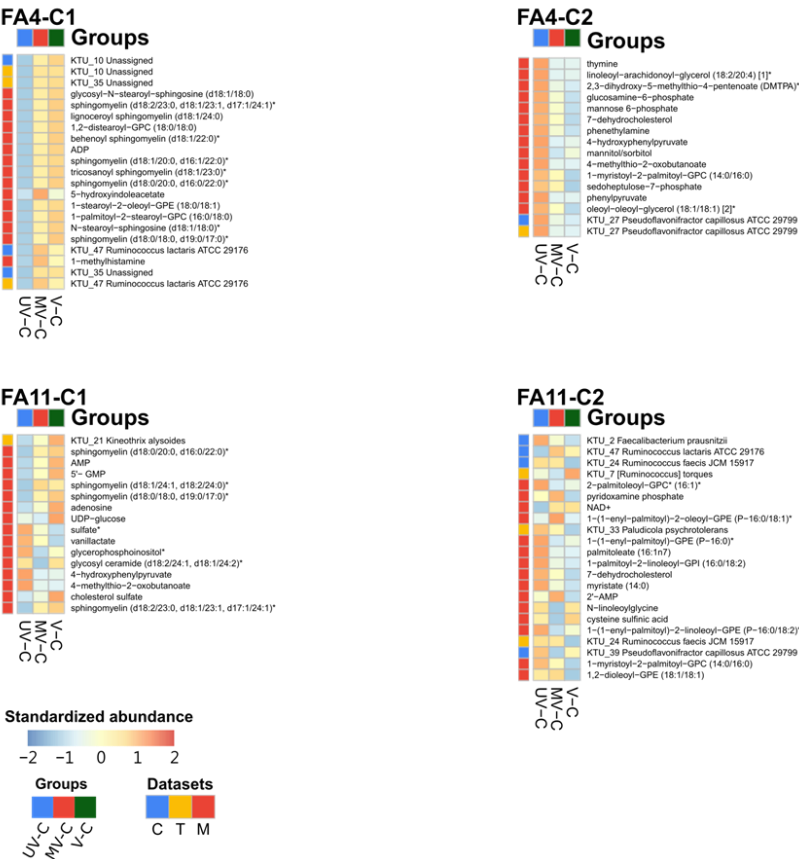
690
691 **Figure S1. Comparisons of microbiota profiles between chicken caecal contents and**
692 **caecal tissues. (A) Beta diversity analysis with the PERMANOVA test demonstrated**

693 no significant differences between caecal contents and caecal tissues ($p=0.052$). (B)
 694 Venn diagrams show 62.7-73.6% microbiota composition shared between caecal
 695 contents and tissues in four study groups (UV-C, UV-UC, MV-C, and V-C). (C)
 696 Differential abundance ($p<0.05$ by DESeq2 test) of microbes between caecal contents
 697 and caecal tissues in four study groups, respectively.
 698



699
 700 **Figure S2. Lesion score and *Eimeria* load correlated metabolites.** (A) Categories of
 701 non-infection-associated metabolites. (B) Categories of infection-associated
 702 metabolites.
 703

709 caecal content microbiome; T: caecal tissue microbiome; M: caecal tissue
710 metabolite).
711



712
713 **Figure S4. Details of subnetworks from 6dpi -challenged MOFA model.** Microbial
714 and metabolomic features of each cluster (C1-C2) in FA4(Figure 5D) and FA11(Figure
715 5E)—the abundance of microbial and metabolomic features among groups were
716 presented by heatmap. The left columns were annotated by sample sources (C:
717 caecal content microbiome; T: caecal tissue microbiome; M: caecal tissue
718 metabolite).
719

Table S1. Sample information of studying groups

Groups	# of samples	Description	Content ID	Tissue ID
UV-C	10	Unvaccinated, challenged, 6 dpi	C1	T1

UV-UC	10	Unvaccinated, unchallenged, 6 dpi	C2	T2
MV-C	10	Mock vaccinated, challenged, 6 dpi	C3	T3
V-C	10	Vaccinated, challenged, 6 dpi	C4	T4
UV-C10	8	Unvaccinated, challenged, 10 dpi	-	T5

720

721

REFERENCES

1. Blake DP, Knox J, Dehaeck B, Huntington B, Rathinam T, Ravipati V, et al. Re-calculating the cost of coccidiosis in chickens. *Vet Res.* 2020;51(1):115. Epub 2020/09/16. doi: 10.1186/s13567-020-00837-2. PubMed PMID: 32928271; PubMed Central PMCID: PMC7488756.
2. Rothwell L, Young JR, Zoorob R, Whittaker CA, Hesketh P, Archer A, et al. Cloning and characterization of chicken IL-10 and its role in the immune response to *Eimeria maxima*. *J Immunol.* 2004;173(4):2675-82. Epub 2004/08/06. doi: 10.4049/jimmunol.173.4.2675. PubMed PMID: 15294985.
3. Laurent F, Mancassola R, Lacroix S, Menezes R, Naciri M. Analysis of chicken mucosal immune response to *Eimeria tenella* and *Eimeria maxima* infection by quantitative reverse transcription-PCR. *Infect Immun.* 2001;69(4):2527-34. Epub 2001/03/20. doi: 10.1128/IAI.69.4.2527-2534.2001. PubMed PMID: 11254616; PubMed Central PMCID: PMC98188.
4. Hong YH, Lillehoj HS, Lee SH, Dalloul RA, Lillehoj EP. Analysis of chicken cytokine and chemokine gene expression following *Eimeria acervulina* and *Eimeria tenella* infections. *Vet Immunol Immunopathol.* 2006;114(3-4):209-23. Epub 2006/09/26. doi: 10.1016/j.vetimm.2006.07.007. PubMed PMID: 16996141.
5. Wu Z, Hu T, Rothwell L, Vervelde L, Kaiser P, Boulton K, et al. Analysis of the function of IL-10 in chickens using specific neutralising antibodies and a sensitive capture ELISA. *Dev Comp Immunol.* 2016;63:206-12. Epub 2016/04/25. doi: 10.1016/j.dci.2016.04.016. PubMed PMID: 27108075; PubMed Central PMCID: PMC4947970.
6. Peek HW, Landman WJ. Coccidiosis in poultry: anticoccidial products, vaccines and other prevention strategies. *Vet Q.* 2011;31(3):143-61. Epub 2011/10/28. doi: 10.1080/01652176.2011.605247. PubMed PMID: 22029884.
7. Blake DP, Tomley FM. Securing poultry production from the ever-present *Eimeria* challenge. *Trends Parasitol.* 2014;30(1):12-9. Epub 2013/11/19. doi: 10.1016/j.pt.2013.10.003. PubMed PMID: 24238797.
8. Chapman HD, Jeffers TK. Vaccination of chickens against coccidiosis ameliorates drug resistance in commercial poultry production. *Int J Parasitol Drugs Drug Resist.* 2014;4(3):214-7. Epub 2014/10/25. doi: 10.1016/j.ijpddr.2014.10.002. PubMed PMID: 25516830; PubMed Central PMCID: PMC4266793.
9. Venkatas J, Adeleke MA. A review of *Eimeria* antigen identification for the development of novel anticoccidial vaccines. *Parasitol Res.* 2019;118(6):1701-10. Epub 2019/05/08. doi: 10.1007/s00436-019-06338-2. PubMed PMID: 31065831.
10. Soutter F, Werling D, Nolan M, Kuster T, Attree E, Marugan-Hernandez V, et al. A Novel Whole Yeast-Based Subunit Oral Vaccine Against *Eimeria tenella* in Chickens.

- 760 Front Immunol. 2022;13:809711. Epub 2022/02/22. doi:
761 10.3389/fimmu.2022.809711. PubMed PMID: 35185896; PubMed Central PMCID:
762 PMCPMC8848252.
- 763 11. Blake DP, Clark EL, Macdonald SE, Thenmozhi V, Kundu K, Garg R, et al.
764 Population, genetic, and antigenic diversity of the apicomplexan *Eimeria tenella* and
765 their relevance to vaccine development. Proc Natl Acad Sci U S A.
766 2015;112(38):E5343-50. Epub 20150909. doi: 10.1073/pnas.1506468112. PubMed
767 PMID: 26354122; PubMed Central PMCID: PMCPMC4586875.
- 768 12. Kundu K, Garg R, Kumar S, Mandal M, Tomley FM, Blake DP, et al. Humoral and
769 cytokine response elicited during immunisation with recombinant Immune Mapped
770 protein-1 (EtIMP-1) and oocysts of *Eimeria tenella*. Vet Parasitol. 2017;244:44-53.
771 Epub 20170724. doi: 10.1016/j.vetpar.2017.07.025. PubMed PMID: 28917316.
- 772 13. Labbe M, de Venevelles P, Girard-Misguich F, Bourdieu C, Guillaume A, Pery P.
773 *Eimeria tenella* microneme protein EtMIC3: identification, localisation and role in
774 host cell infection. Mol Biochem Parasitol. 2005;140(1):43-53. doi:
775 10.1016/j.molbiopara.2004.12.002. PubMed PMID: 15694485.
- 776 14. Bortoluzzi C, Barbosa JGM, Pereira R, Fagundes NS, Rafael JM, Menten JFM.
777 Autolyzed Yeast (*Saccharomyces cerevisiae*) Supplementation Improves Performance
778 While Modulating the Intestinal Immune-System and Microbiology of Broiler
779 Chickens. Frontiers in Sustainable Food Systems. 2018;2. doi:
780 10.3389/fsufs.2018.00085.
- 781 15. Luquetti BC, Furlan RL, Alarcon MFF, Macari M. *Saccharomyces Cerevisiae* cell
782 wall dietary supplementation on the performance and intestinal mucosa
783 development and integrity of broiler chickens vaccinated against coccidiosis. Revista
784 Brasileira de Ciência Avícola. 2012;14(2):89-95. doi: 10.1590/s1516-
785 635x2012000200002.
- 786 16. LeBlanc JG, Milani C, de Giori GS, Sesma F, van Sinderen D, Ventura M. Bacteria
787 as vitamin suppliers to their host: a gut microbiota perspective. Curr Opin Biotechnol.
788 2013;24(2):160-8. Epub 2012/09/04. doi: 10.1016/j.copbio.2012.08.005. PubMed
789 PMID: 22940212.
- 790 17. Sannino DR, Dobson AJ, Edwards K, Angert ER, Buchon N. The *Drosophila*
791 *melanogaster* Gut Microbiota Provisions Thiamine to Its Host. mBio. 2018;9(2). Epub
792 2018/03/08. doi: 10.1128/mBio.00155-18. PubMed PMID: 29511074; PubMed
793 Central PMCID: PMCPMC5845000.
- 794 18. Thaïss CA, Zmora N, Levy M, Elinav E. The microbiome and innate immunity.
795 Nature. 2016;535(7610):65-74. Epub 2016/07/08. doi: 10.1038/nature18847.
796 PubMed PMID: 27383981.
- 797 19. Ximenez C, Torres J. Development of Microbiota in Infants and its Role in

798 Maturation of Gut Mucosa and Immune System. Arch Med Res. 2017;48(8):666-80.
799 Epub 2017/12/05. doi: 10.1016/j.arcmed.2017.11.007. PubMed PMID: 29198451.
800 20. Jebessa E, Guo L, Chen X, Bello SF, Cai B, Girma M, et al. Influence of Eimeria
801 maxima coccidia infection on gut microbiome diversity and composition of the
802 jejunum and cecum of indigenous chicken. Front Immunol. 2022;13:994224. Epub
803 20220905. doi: 10.3389/fimmu.2022.994224. PubMed PMID: 36131927; PubMed
804 Central PMCID: PMCPMC9483182.
805 21. Macdonald SE, Nolan MJ, Harman K, Boulton K, Hume DA, Tomley FM, et al.
806 Effects of Eimeria tenella infection on chicken caecal microbiome diversity, exploring
807 variation associated with severity of pathology. PLoS One. 2017;12(9):e0184890.
808 Epub 2017/09/22. doi: 10.1371/journal.pone.0184890. PubMed PMID: 28934262;
809 PubMed Central PMCID: PMCPMC5608234.
810 22. Moore RJ. Necrotic enteritis predisposing factors in broiler chickens. Avian
811 Pathol. 2016;45(3):275-81. Epub 2016/03/02. doi: 10.1080/03079457.2016.1150587.
812 PubMed PMID: 26926926.
813 23. Aggrey SE, Milfort MC, Fuller AL, Yuan J, Rekaya R. Effect of host genotype and
814 Eimeria acervulina infection on the metabolome of meat-type chickens. PLoS One.
815 2019;14(10):e0223417. Epub 2019/10/17. doi: 10.1371/journal.pone.0223417.
816 PubMed PMID: 31618222; PubMed Central PMCID: PMCPMC6795442.
817 24. Li S, Sullivan NL, Roupheal N, Yu T, Banton S, Maddur MS, et al. Metabolic
818 Phenotypes of Response to Vaccination in Humans. Cell. 2017;169(5):862-77 e17.
819 Epub 2017/05/16. doi: 10.1016/j.cell.2017.04.026. PubMed PMID: 28502771;
820 PubMed Central PMCID: PMCPMC5711477.
821 25. Haak BW, Argelaguet R, Kinsella CM, Kullberg RFJ, Lankelma JM, Deijs M, et al.
822 Integrative Transkingdom Analysis of the Gut Microbiome in Antibiotic Perturbation
823 and Critical Illness. mSystems. 2021;6(2). Epub 2021/03/18. doi:
824 10.1128/mSystems.01148-20. PubMed PMID: 33727397; PubMed Central PMCID:
825 PMCPMC8546997.
826 26. Argelaguet R, Velten B, Arnol D, Dietrich S, Zenz T, Marioni JC, et al. Multi-Omics
827 Factor Analysis-a framework for unsupervised integration of multi-omics data sets.
828 Mol Syst Biol. 2018;14(6):e8124. Epub 2018/06/22. doi: 10.15252/msb.20178124.
829 PubMed PMID: 29925568; PubMed Central PMCID: PMCPMC6010767.
830 27. Hay MC, Hinsu AT, Koringa PG, Pandit RJ, Liu P-Y, Parekh MJ, et al. Chicken caecal
831 enterotypes in indigenous Kadaknath and commercial Cobb chicken lines are
832 associated with Campylobacter abundance and influenced by farming practices.
833 Frontiers in Microbiomes. 2023;2. doi: 10.3389/frmbi.2023.1301609.
834 28. Lloyd-Price J, Arze C, Ananthakrishnan AN, Schirmer M, Avila-Pacheco J, Poon
835 TW, et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel

diseases. *Nature*. 2019;569(7758):655-62. Epub 2019/05/31. doi: 10.1038/s41586-019-1237-9. PubMed PMID: 31142855; PubMed Central PMCID: PMC6650278.

29. Lau SK, Woo PC, Woo GK, Fung AM, Ngan AH, Song Y, et al. Bacteraemia caused by *Anaerotruncus colihominis* and emended description of the species. *J Clin Pathol*. 2006;59(7):748-52. Epub 20060207. doi: 10.1136/jcp.2005.031773. PubMed PMID: 16467163; PubMed Central PMCID: PMC660411.

30. Chen W, Liu F, Ling Z, Tong X, Xiang C. Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer. *PLoS One*. 2012;7(6):e39743. Epub 20120628. doi: 10.1371/journal.pone.0039743. PubMed PMID: 22761885; PubMed Central PMCID: PMC3386193.

31. Levine UY, Looft T, Allen HK, Stanton TB. Butyrate-producing bacteria, including mucin degraders, from the swine intestinal tract. *Appl Environ Microbiol*. 2013;79(12):3879-81. Epub 20130412. doi: 10.1128/AEM.00589-13. PubMed PMID: 23584773; PubMed Central PMCID: PMC3675948.

32. Rodriguez-Castano GP, Rey FE, Caro-Quintero A, Acosta-Gonzalez A. Gut-derived Flavonifractor species variants are differentially enriched during in vitro incubation with quercetin. *PLoS One*. 2020;15(12):e0227724. Epub 20201202. doi: 10.1371/journal.pone.0227724. PubMed PMID: 33264299; PubMed Central PMCID: PMC7710108.

33. Stanley D, Wu SB, Rodgers N, Swick RA, Moore RJ. Differential responses of cecal microbiota to fishmeal, *Eimeria* and *Clostridium perfringens* in a necrotic enteritis challenge model in chickens. *PLoS One*. 2014;9(8):e104739. Epub 2014/08/29. doi: 10.1371/journal.pone.0104739. PubMed PMID: 25167074; PubMed Central PMCID: PMC4148237.

34. Jung A, Chen LR, Suyemoto MM, Barnes HJ, Borst LB. A Review of *Enterococcus cecorum* Infection in Poultry. *Avian Dis*. 2018;62(3):261-71. doi: 10.1637/11825-030618-Review.1. PubMed PMID: 30339512.

35. Chen YC, Kuo YC, Chen MC, Zhang YD, Chen CL, Le PH, et al. Case-Control Study of *Clostridium innocuum* Infection, Taiwan. *Emerg Infect Dis*. 2022;28(3):599-607. doi: 10.3201/eid2803.204421. PubMed PMID: 35195517; PubMed Central PMCID: PMC888209.

36. Chia JH, Wu TS, Wu TL, Chen CL, Chuang CH, Su LH, et al. *Clostridium innocuum* is a vancomycin-resistant pathogen that may cause antibiotic-associated diarrhoea. *Clin Microbiol Infect*. 2018;24(11):1195-9. Epub 20180217. doi: 10.1016/j.cmi.2018.02.015. PubMed PMID: 29458157.

37. Pham HHS, Matsubayashi M, Tsuji N, Hatabu T. Relationship between *Eimeria tenella* associated-early clinical signs and molecular changes in the intestinal barrier function. *Vet Immunol Immunopathol*. 2021;240:110321. Epub 20210828. doi:

874 10.1016/j.vetimm.2021.110321. PubMed PMID: 34520968.

875 38. Arrieta MC, Finlay BB. The commensal microbiota drives immune homeostasis.

876 Front Immunol. 2012;3:33. Epub 2012/05/09. doi: 10.3389/fimmu.2012.00033.

877 PubMed PMID: 22566917; PubMed Central PMCID: PMC3341987.

878 39. Matsuoka K, Kanai T. The gut microbiota and inflammatory bowel disease.

879 Semin Immunopathol. 2015;37(1):47-55. Epub 2014/11/26. doi: 10.1007/s00281-

880 014-0454-4. PubMed PMID: 25420450; PubMed Central PMCID: PMC34281375.

881 40. Sartor RB. Microbial influences in inflammatory bowel diseases.

882 Gastroenterology. 2008;134(2):577-94. Epub 2008/02/05. doi:

883 10.1053/j.gastro.2007.11.059. PubMed PMID: 18242222.

884 41. Singhal R, Shah YM. Oxygen battle in the gut: Hypoxia and hypoxia-inducible

885 factors in metabolic and inflammatory responses in the intestine. J Biol Chem.

886 2020;295(30):10493-505. Epub 2020/06/07. doi: 10.1074/jbc.REV120.011188.

887 PubMed PMID: 32503843; PubMed Central PMCID: PMC7383395.

888 42. Van Welden S, Selfridge AC, Hindryckx P. Intestinal hypoxia and hypoxia-induced

889 signalling as therapeutic targets for IBD. Nat Rev Gastroenterol Hepatol.

890 2017;14(10):596-611. Epub 2017/08/31. doi: 10.1038/nrgastro.2017.101. PubMed

891 PMID: 28853446.

892 43. Winter SE, Baumler AJ. Dysbiosis in the inflamed intestine: chance favors the

893 prepared microbe. Gut Microbes. 2014;5(1):71-3. Epub 2014/03/19. doi:

894 10.4161/gmic.27129. PubMed PMID: 24637596; PubMed Central PMCID:

895 PMC34049941.

896 44. Winter SE, Baumler AJ. Why related bacterial species bloom simultaneously in

897 the gut: principles underlying the 'Like will to like' concept. Cell Microbiol.

898 2014;16(2):179-84. Epub 2013/11/30. doi: 10.1111/cmi.12245. PubMed PMID:

899 24286560; PubMed Central PMCID: PMC34013256.

900 45. Bliska JB, van der Velden AW. Salmonella "sops" up a preferred electron receptor

901 in the inflamed intestine. mBio. 2012;3(4):e00226-12. Epub 2012/08/16. doi:

902 10.1128/mBio.00226-12. PubMed PMID: 22893385; PubMed Central PMCID:

903 PMC3419527.

904 46. Lopez CA, Rivera-Chavez F, Byndloss MX, Baumler AJ. The Periplasmic Nitrate

905 Reductase NapABC Supports Luminal Growth of Salmonella enterica Serovar

906 Typhimurium during Colitis. Infect Immun. 2015;83(9):3470-8. Epub 2015/06/24. doi:

907 10.1128/IAI.00351-15. PubMed PMID: 26099579; PubMed Central PMCID:

908 PMC34534643.

909 47. Lopez CA, Winter SE, Rivera-Chavez F, Xavier MN, Poon V, Nuccio SP, et al.

910 Phage-mediated acquisition of a type III secreted effector protein boosts growth of

911 salmonella by nitrate respiration. mBio. 2012;3(3). Epub 2012/06/14. doi:

912 10.1128/mBio.00143-12. PubMed PMID: 22691391; PubMed Central PMCID:
913 PMCPMC3374392.

914 48. Rivera-Chavez F, Winter SE, Lopez CA, Xavier MN, Winter MG, Nuccio SP, et al.
915 Salmonella uses energy taxis to benefit from intestinal inflammation. PLoS Pathog.
916 2013;9(4):e1003267. Epub 2013/05/03. doi: 10.1371/journal.ppat.1003267. PubMed
917 PMID: 23637594; PubMed Central PMCID: PMCPMC3630101.

918 49. Spees AM, Wangdi T, Lopez CA, Kingsbury DD, Xavier MN, Winter SE, et al.
919 Streptomycin-induced inflammation enhances Escherichia coli gut colonization
920 through nitrate respiration. mBio. 2013;4(4). Epub 2013/07/04. doi:
921 10.1128/mBio.00430-13. PubMed PMID: 23820397; PubMed Central PMCID:
922 PMCPMC3705454.

923 50. Vazquez-Torres A, Baumler AJ. Nitrate, nitrite and nitric oxide reductases: from
924 the last universal common ancestor to modern bacterial pathogens. Curr Opin
925 Microbiol. 2016;29:1-8. Epub 2015/10/02. doi: 10.1016/j.mib.2015.09.002. PubMed
926 PMID: 26426528; PubMed Central PMCID: PMCPMC4930242.

927 51. Winter SE, Winter MG, Xavier MN, Thiennimitr P, Poon V, Keestra AM, et al.
928 Host-derived nitrate boosts growth of E. coli in the inflamed gut. Science.
929 2013;339(6120):708-11. Epub 2013/02/09. doi: 10.1126/science.1232467. PubMed
930 PMID: 23393266; PubMed Central PMCID: PMCPMC4004111.

931 52. Ash C. The benefits of Escherichia coli. Science Signaling. 2015;8(401). doi:
932 10.1126/scisignal.aad7728.

933 53. Lawrence JG, Roth JR. Evolution of coenzyme B12 synthesis among enteric
934 bacteria: evidence for loss and reacquisition of a multigene complex. Genetics.
935 1996;142(1):11-24. doi: 10.1093/genetics/142.1.11. PubMed PMID: 8770581;
936 PubMed Central PMCID: PMCPMC1206939.

937 54. Bentley R, Meganathan R. Biosynthesis of vitamin K (menaquinone) in bacteria.
938 Microbiol Rev. 1982;46(3):241-80. doi: 10.1128/mr.46.3.241-280.1982. PubMed
939 PMID: 6127606; PubMed Central PMCID: PMCPMC281544.

940 55. Alexander M. Microbial Ecology. London, United Kingdom: John Wiley & Sons,
941 Inc.; 1971.

942 56. Narendranath NV, Power R. Relationship between pH and medium dissolved
943 solids in terms of growth and metabolism of lactobacilli and Saccharomyces
944 cerevisiae during ethanol production. Appl Environ Microbiol. 2005;71(5):2239-43.
945 Epub 2005/05/05. doi: 10.1128/AEM.71.5.2239-2243.2005. PubMed PMID:
946 15870306; PubMed Central PMCID: PMCPMC1087585.

947 57. Lund PA, De Biase D, Liran O, Scheler O, Mira NP, Cetecioglu Z, et al.
948 Understanding How Microorganisms Respond to Acid pH Is Central to Their Control
949 and Successful Exploitation. Front Microbiol. 2020;11:556140. Epub 2020/10/30. doi:

950 10.3389/fmicb.2020.556140. PubMed PMID: 33117305; PubMed Central PMCID:
951 PMCPMC7553086.

952 58. Esterbauer H, Schaur RJ, Zollner H. Chemistry and biochemistry of 4-
953 hydroxynonenal, malonaldehyde and related aldehydes. *Free Radic Biol Med*.
954 1991;11(1):81-128. Epub 1991/01/01. doi: 10.1016/0891-5849(91)90192-6. PubMed
955 PMID: 1937131.

956 59. Meijers BK, Evenepoel P. The gut-kidney axis: indoxyl sulfate, p-cresyl sulfate and
957 CKD progression. *Nephrol Dial Transplant*. 2011;26(3):759-61. Epub 2011/02/24. doi:
958 10.1093/ndt/gfq818. PubMed PMID: 21343587.

959 60. Smith EA, Macfarlane GT. Enumeration of human colonic bacteria producing
960 phenolic and indolic compounds: effects of pH, carbohydrate availability and
961 retention time on dissimilatory aromatic amino acid metabolism. *J Appl Bacteriol*.
962 1996;81(3):288-302. Epub 1996/09/01. doi: 10.1111/j.1365-2672.1996.tb04331.x.
963 PubMed PMID: 8810056.

964 61. Passmore IJ, Letertre MPM, Preston MD, Bianconi I, Harrison MA, Nasher F, et
965 al. Para-cresol production by *Clostridium difficile* affects microbial diversity and
966 membrane integrity of Gram-negative bacteria. *PLoS Pathog*. 2018;14(9):e1007191.
967 Epub 2018/09/13. doi: 10.1371/journal.ppat.1007191. PubMed PMID: 30208103;
968 PubMed Central PMCID: PMC6135563.

969 62. Paroni R, Casati S, Dei Cas M, Bignotto M, Rubino FM, Ciuffreda P. Unambiguous
970 Characterization of p-Cresyl Sulfate, a Protein-Bound Uremic Toxin, as Biomarker of
971 Heart and Kidney Disease. *Molecules*. 2019;24(20). Epub 2019/10/18. doi:
972 10.3390/molecules24203704. PubMed PMID: 31618977; PubMed Central PMCID:
973 PMC6832250.

974 63. Hannun YA, Obeid LM. Principles of bioactive lipid signalling: lessons from
975 sphingolipids. *Nat Rev Mol Cell Biol*. 2008;9(2):139-50. Epub 2008/01/25. doi:
976 10.1038/nrm2329. PubMed PMID: 18216770.

977 64. Koberlin MS, Snijder B, Heinz LX, Baumann CL, Fauster A, Vladimer GI, et al. A
978 Conserved Circular Network of Coregulated Lipids Modulates Innate Immune
979 Responses. *Cell*. 2015;162(1):170-83. Epub 2015/06/23. doi:
980 10.1016/j.cell.2015.05.051. PubMed PMID: 26095250; PubMed Central PMCID:
981 PMC4523684.

982 65. Merrill AH, Jr., Carman GM. Introduction to Thematic Minireview Series: Novel
983 Bioactive Sphingolipids. *J Biol Chem*. 2015;290(25):15362-4. Epub 2015/05/08. doi:
984 10.1074/jbc.R115.663708. PubMed PMID: 25947376; PubMed Central PMCID:
985 PMC4505448.

986 66. Spiegel S, Milstien S. The outs and the ins of sphingosine-1-phosphate in
987 immunity. *Nat Rev Immunol*. 2011;11(6):403-15. Epub 2011/05/07. doi:

988 10.1038/nri2974. PubMed PMID: 21546914; PubMed Central PMCID:
989 PMCPMC3368251.

990 67. Maceyka M, Spiegel S. Sphingolipid metabolites in inflammatory disease.
991 Nature. 2014;510(7503):58-67. Epub 2014/06/06. doi: 10.1038/nature13475.
992 PubMed PMID: 24899305; PubMed Central PMCID: PMCPMC4320971.

993 68. Brown EM, Ke X, Hitchcock D, Jeanfavre S, Avila-Pacheco J, Nakata T, et al.
994 Bacteroides-Derived Sphingolipids Are Critical for Maintaining Intestinal Homeostasis
995 and Symbiosis. Cell Host Microbe. 2019;25(5):668-80 e7. Epub 2019/05/10. doi:
996 10.1016/j.chom.2019.04.002. PubMed PMID: 31071294; PubMed Central PMCID:
997 PMCPMC6544385.

998 69. Pastor-Fernandez I, Pegg E, Macdonald SE, Tomley FM, Blake DP, Marugan-
999 Hernandez V. Laboratory Growth and Genetic Manipulation of Eimeria tenella. Curr
1000 Protoc Microbiol. 2019;53(1):e81. Epub 20190227. doi: 10.1002/cpmc.81. PubMed
1001 PMID: 30811108.

1002 70. Johnson J, Reid WM. Anticoccidial drugs: lesion scoring techniques in battery
1003 and floor-pen experiments with chickens. Exp Parasitol. 1970;28(1):30-6. doi:
1004 10.1016/0014-4894(70)90063-9. PubMed PMID: 5459870.

1005 71. Illumina. Illumina 16S metagenomic sequencing library preparation (Illumina
1006 Technical Note 15044223 Rev. A) 2013 [cited 2018 2018/10]. Available from:
1007 [http://support.illumina.com/content/dam/illumina-](http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf)
1008 [support/documents/documentation/chemistry_documentation/16s/16s-](http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf)
1009 [metagenomic-library-prep-guide-15044223-b.pdf](http://support.illumina.com/content/dam/illumina-support/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf).

1010 72. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
1011 sequence data. Bioinformatics. 2014;30(15):2114-20. Epub 20140401. doi:
1012 10.1093/bioinformatics/btu170. PubMed PMID: 24695404; PubMed Central PMCID:
1013 PMCPMC4103590.

1014 73. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, et al.
1015 Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME
1016 2's q2-feature-classifier plugin. Microbiome. 2018;6(1):90. Epub 2018/05/19. doi:
1017 10.1186/s40168-018-0470-z. PubMed PMID: 29773078; PubMed Central PMCID:
1018 PMCPMC5956843.

1019 74. Martin M. Cutadapt removes adapter sequences from high-throughput
1020 sequencing reads. EMBnetjournal. 2011;17(1):10. doi: 10.14806/ej.17.1.200.

1021 75. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2:
1022 High-resolution sample inference from Illumina amplicon data. Nat Methods.
1023 2016;13(7):581-3. Epub 2016/05/24. doi: 10.1038/nmeth.3869. PubMed PMID:
1024 27214047; PubMed Central PMCID: PMCPMC4927377.

1025 76. Liu PY, Yang SH, Yang SY. KTU: K-mer Taxonomic Units improve the biological

1026 relevance of amplicon sequence variant microbiota data. *Methods in Ecology and*
1027 *Evolution*. 2022;13(3):560-8. doi: 10.1111/2041-210x.13758. PubMed PMID:
1028 WOS:000717949300001.

1029 77. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA
1030 ribosomal RNA gene database project: improved data processing and web-based
1031 tools. *Nucleic Acids Res*. 2013;41(Database issue):D590-6. Epub 2012/11/30. doi:
1032 10.1093/nar/gks1219. PubMed PMID: 23193283; PubMed Central PMCID:
1033 PMCPMC3531112.

1034 78. Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, Quast C, et al. The SILVA and
1035 "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Research*.
1036 2014;42(D1):D643-D8. doi: 10.1093/nar/gkt1209. PubMed PMID:
1037 WOS:000331139800094.

1038 79. Liu P-Y. poyuliu/MARco: MARco: Microbiome Analysis RcodeDB (Version v1.0):
1039 Zenodo; 2021. Available from: <http://doi.org/10.5281/zenodo.4589898>.

1040 80. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al.
1041 vegan: community ecology package. R package version 2.3-1. Oulu, Finland 2015.

1042 81. Kolde R. pheatmap: Pretty heatmaps. R package version 1.0.8. 2015.

1043 82. R Core Team. R: A language and environment for statistical computing. R
1044 Foundation for Statistical Computing, Vienna, Austria; 2015.

1045 83. Dekkers KF, Sayols-Baixeras S, Baldanzi G, Nowak C, Hammar U, Nguyen D, et al.
1046 An online atlas of human plasma metabolite signatures of gut microbiome
1047 composition. medRxiv. 2021.

1048 84. Csardi G, Nepusz T. The igraph software package for complex network research.
1049 InterJournal 2006.

1050 85. Clauset A, Newman ME, Moore C. Finding community structure in very large
1051 networks. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;70(6 Pt 2):066111. doi:
1052 10.1103/PhysRevE.70.066111. PubMed PMID: 15697438.

1053

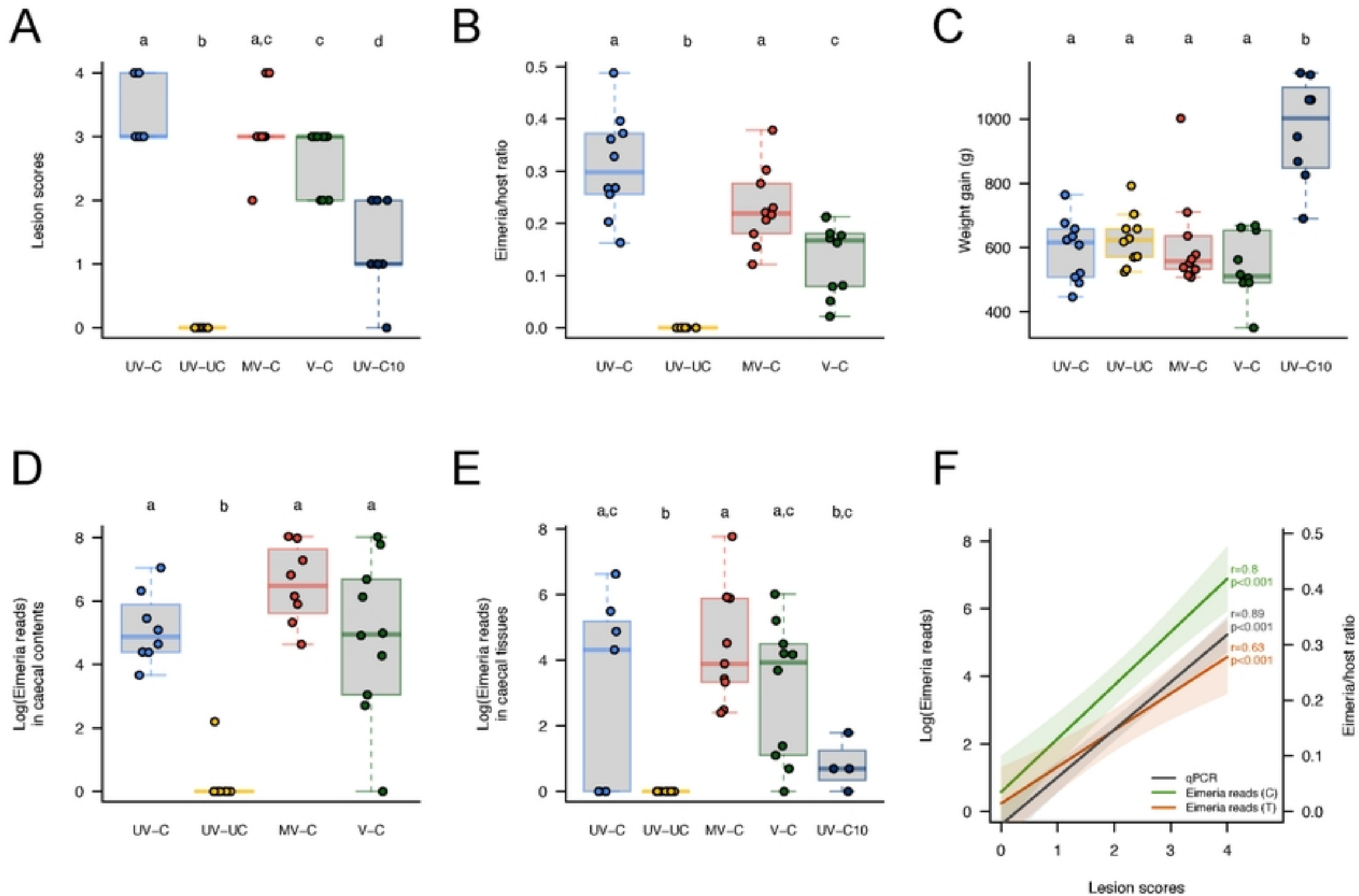


Figure 1

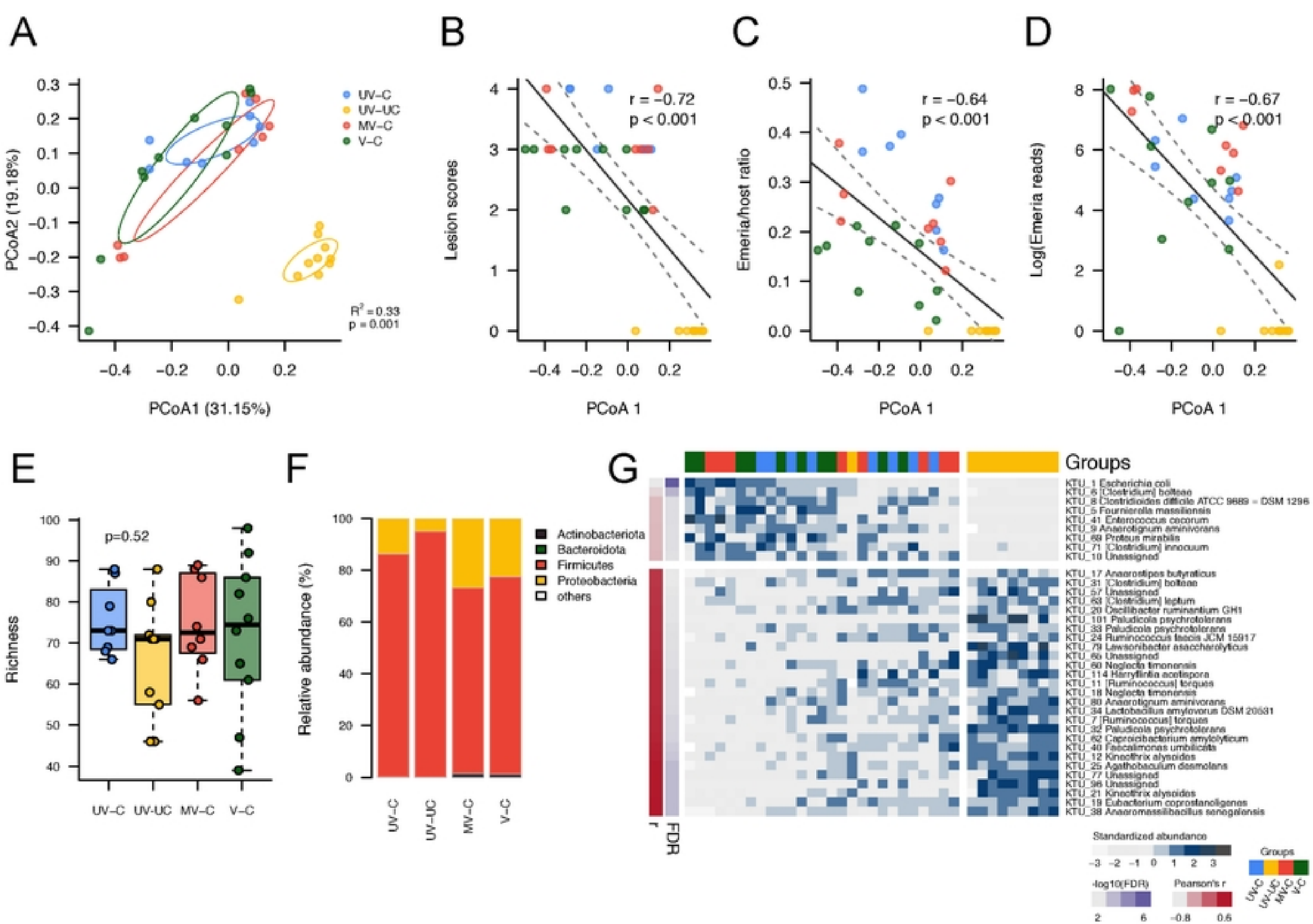


Figure 2

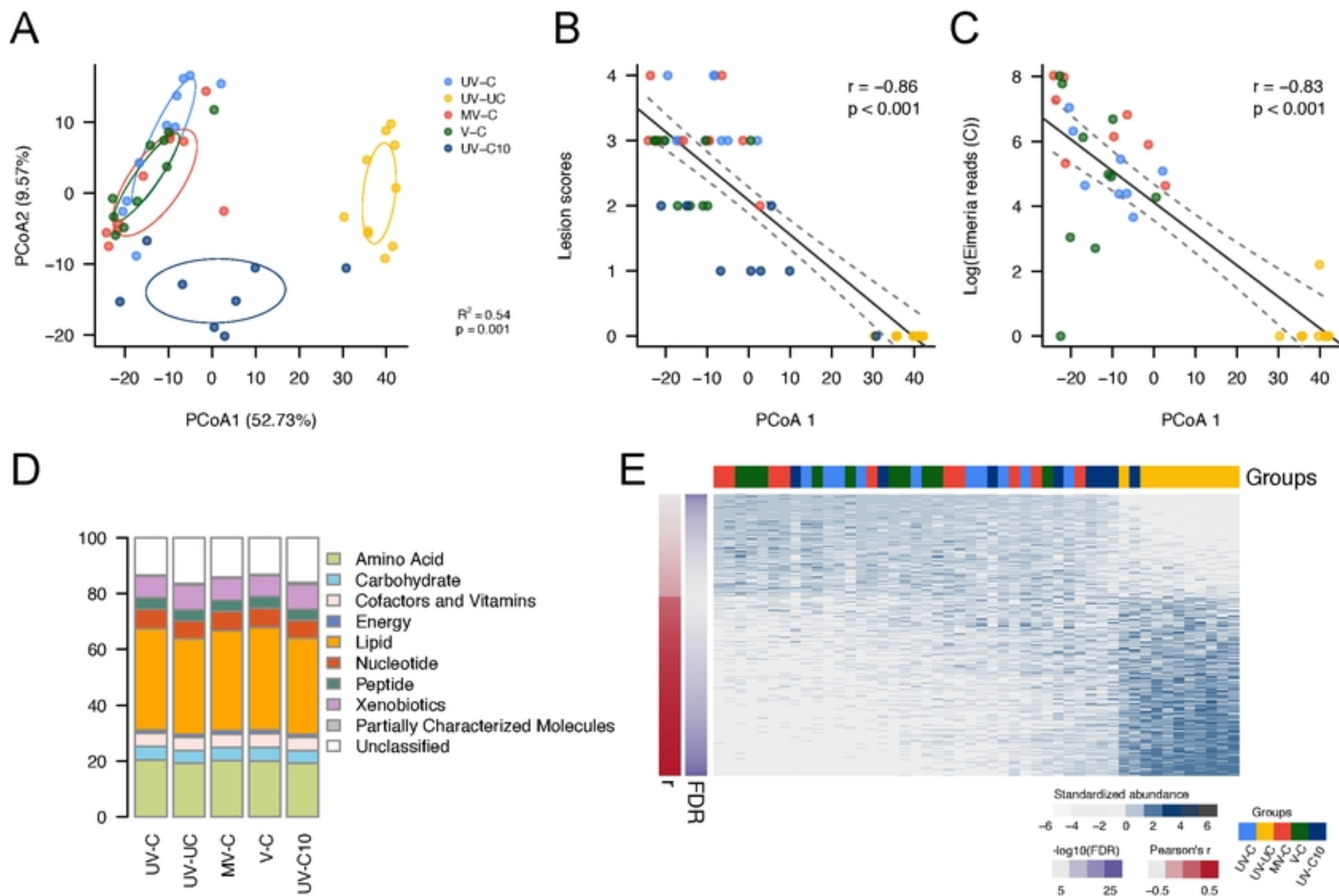


Figure 3

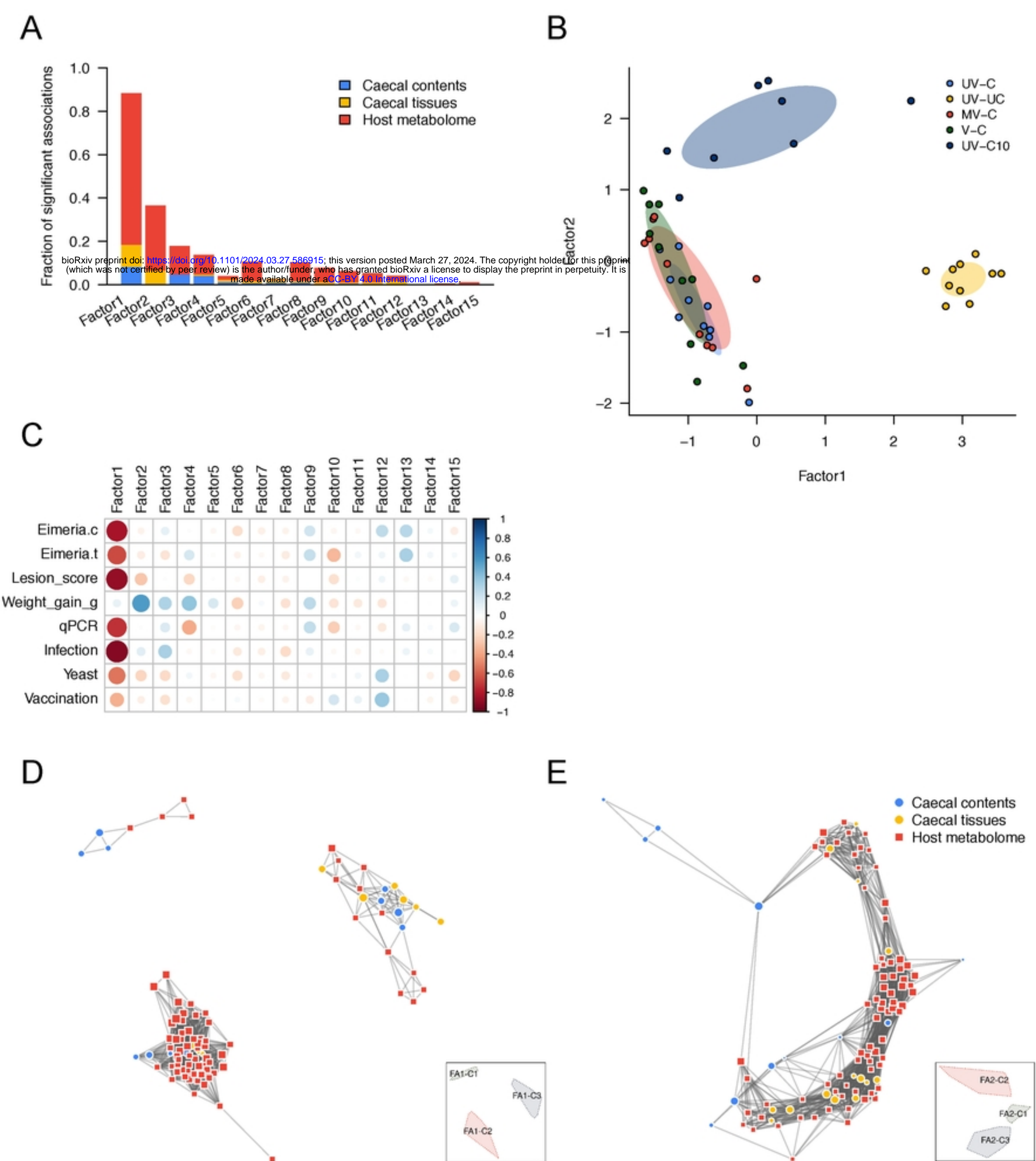


Figure 4

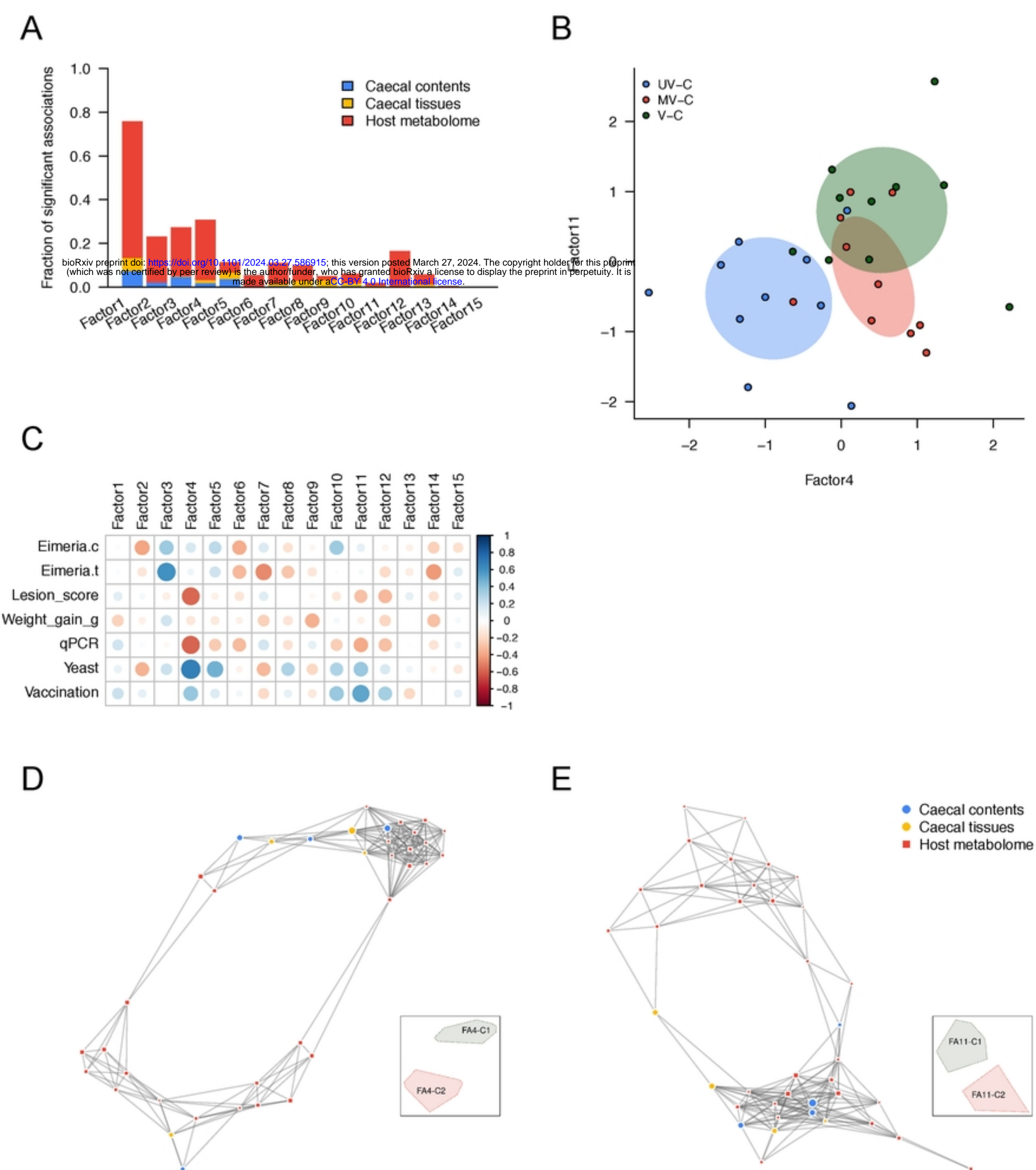


Figure 5