Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

A multi-study analysis reveals ecological succession and functional reorganization across the developing pig gut microbiome

Natanieli Silva Máximo
Universidade Federal de Viçosa
Jansller Luiz Genova
Universidade Federal de Viçosa
Mateus Ferreira Santana
mateus.santana@ufv.br
Universidade Federal de Viçosa

Research Article

Keywords: Age-adjusted body weight, Core microbiota, Developmental stage, Meta-analysis, Microbial association networks, Precision livestock farming

Posted Date: November 23rd, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11131385/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

A multi-study analysis reveals ecological succession and functional reorganization across the developing pig gut microbiome

Natanieli Silva Máximo[12], Janslller Luiz Genova[3], Mateus Ferreira Santana[12*]

[*]Corresponding author, mateus.santana@ufv.br

[1] Laboratory of Microbial Eco-Evolutionary Genomics

[2] Department of Microbiology, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil.

[3] Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil.

ABSTRACT

Background The pig gut microbiome undergoes pronounced successional changes throughout host development, yet whether these changes primarily reflect community replacement or reorganization within a persistent ecological framework remains unclear. Here, we hypothesized that the pig gut microbiome is hierarchically organized into two complementary ecological layers: a persistent structural core maintained across heterogeneous host conditions and an adaptive layer characterized by performance-associated taxonomic, functional, and topological variation.

Results We conducted a global integrative multi-study analysis of 1,608 pig gut microbial profiles spanning 25 dataset strata across seven countries. Ordination and constrained analyses consistently identified developmental stage as the primary chronological axis of microbiome succession, accounting for the largest independent fraction of baseline

compositional variation (R^2 = 0.270; p = 0.049). Alpha-diversity dynamics revealed a disturbance–recovery trajectory centered on weaning, characterized by transient reductions in community evenness and increased ecological heterogeneity. Despite substantial taxonomic turnover, a persistent structural core including *Prevotella*, *Faecalibacterium*, *Roseburia*, *Ruminococcus*, and *Lactobacillus* was maintained across developmental stages, host breeds, sexes, and sampling sites. Although age-adjusted body weight did not significantly explain overall community composition, consensus prioritization integrating Random Forest, ANCOM-BC2, and MaAsLin3 identified 120 microbial signatures associated with this phenotype. Predicted functional analyses further identified 113 KEGG pathways significantly associated with age-adjusted body weight (q < 0.05), encompassing complex carbohydrate utilization, amino acid metabolism, energy metabolism, and SCFA-associated processes. Network analysis of high age-adjusted BW animals revealed a sparse and highly modular genus-level association topology characterized by distinct degree-defined hubs and high-betweenness taxa, while structural-core membership was only partially aligned with network centrality.

Conclusions These findings support a hierarchical ecological framework in which a persistent structural layer provides continuity across heterogeneous host conditions, whereas age-adjusted body weight is associated with more selective taxonomic, predicted functional, and topological patterns. By distinguishing ecological persistence from performance-associated microbial organization, this study shifts emphasis from isolated microbial biomarkers toward an ecosystem-level interpretation of the pig gut microbiome and provides a foundation for microbiome-informed precision livestock strategies.

Keywords Age-adjusted body weight; Core microbiota; Developmental stage; Meta-analysis; Microbial association networks; Precision livestock farming.

Background

The intestinal microbiome is increasingly recognized as a central regulator of host nutrition, immune maturation, epithelial homeostasis, and metabolic function. In pig production, variation in microbial community composition and functional potential has been associated with feed efficiency, growth performance, intestinal resilience, and susceptibility to inflammatory disorders [1, 2]. More recently, ecological network analyses have further demonstrated that the organization of microbial co-abundance patterns is associated with productive traits such as average daily gain, highlighting that microbial interactions may be as biologically relevant as taxonomic composition itself [3]. Collectively, these observations identify the intestinal microbiome as a promising target for microbiome-informed precision livestock strategies [4, 5].

Ecological succession theory predicts that host-associated microbial communities undergo predictable developmental trajectories driven by continuous environmental and physiological changes [6-8]. In pigs, microbiome maturation is characterized by progressive ecological transitions associated with dietary shifts, immune system development, oxygen depletion within the intestinal lumen, and increasing niche specialization. Among these transitions, weaning represents one of the most disruptive ecological events, during which the abrupt replacement of milk-derived nutrients by complex plant carbohydrates frequently leads to transient community instability and increased susceptibility to dysbiosis [4]. Nevertheless, the magnitude and trajectory of these successional changes vary across production systems owing to differences in weaning practices, diet, host genetics, production environment, sampling site, and analytical methodology.

Despite the rapid expansion of pig microbiome research, microbial biomarkers associated with productive performance remain inconsistent across independent studies. Differences in host genetics, management practices, nutritional strategies, sequencing platforms, and analytical pipelines often result in distinct microbial biomarkers, which can limit biological interpretation and the identification of broadly applicable microbial targets. Although previous studies have independently investigated microbiome succession, productive performance, core microbiota, and microbial community composition, it remains unclear whether these ecological dimensions represent independent biological processes or interconnected components of a common organizational framework operating across heterogeneous production systems [9, 10].

Here, we propose the hypothesis that the pig gut microbiome is hierarchically organized into two complementary ecological layers: a hypothesized conserved structural microbiome that persists across heterogeneous host-associated conditions and offers ecological stability, and a dynamic functional layer that undergoes coordinated ecological remodeling during host development and is associated with age-adjusted body weight. Within this conceptual framework, developmental succession is expected to reshape microbial functional organization while preserving the underlying structural architecture of the intestinal ecosystem [11, 12].

Accumulating macro-ecological evidence indicates that host-associated microbial ecosystems frequently display pronounced taxonomic elasticity across cohorts while preserving a highly canalized functional repertoire [13, 14]. This decoupling highlights a high degree of functional redundancy [15], which serves as a primary stabilizing mechanism that maintains core metabolic homeostasis and ecosystem stability despite taxonomic shifts induced by environmental or physiological perturbations [16]. Such

observations indicate that taxonomic turnover does not necessarily imply disruption of ecosystem organization, raising the possibility that conserved structural features may coexist with predicted metabolic shifts across host development [9, 16].

To evaluate this hypothesis, we conducted a global ecological meta-analysis that integrated publicly available gut microbiome sequencing datasets encompassing multiple developmental stages, breeds, and production systems. By combining multivariate ecological analyses, consensus biomarker discovery, functional inference, and ecological network reconstruction, we investigated whether microbiome maturation can be explained as a hierarchical ecological process in which structural persistence and functional adaptation coexist across heterogeneous production systems.

Methods

Study collection, selection, and harmonization

Publicly available pig gut microbiome studies were identified through systematic searches of Google Scholar, Web of Science, PubMed, the NCBI Sequence Read Archive (SRA), ENA Browser, and EMBL-EBI databases using combinations of the terms *Sus scrofa*, *pig*, *porcine*, *gut microbiome,* or *microbiota*. Retrieved studies were screened according to predefined eligibility criteria.

Studies or experimental groups were included when: (i) animals were clinically healthy and had not received therapeutic antibiotics, antimicrobial treatments, or other microbiome-modulating interventions before sampling; (ii) samples originated from untreated control groups or observational studies without intentional microbiome manipulation; (iii) individual-level metadata included host age and sampling site, with sex, breed, body weight, and animal identification retained whenever available; and (iv)

raw sequencing data and sufficient methodological information were publicly accessible. Studies were excluded when animals received antimicrobial treatments or growth-promoting additives, when essential metadata were unavailable, or when sampled animals were older than 240 days. For intervention studies, only untreated control groups were retained.

Raw data retrieval and bioinformatic processing

Raw sequencing datasets comprising whole-genome shotgun (WGS), 16S rRNA gene, and ITS amplicons were retrieved from the NCBI SRA using the SRA Toolkit (v3.1.1) according to their respective BioProject accession numbers. Following quality assessment, 25 analytical datasets derived from 15 independent studies were selected for retention (Table S1). Several analytical datasets originated from the study of Xiao *et al*. [17], which included multiple pig breeds sampled across distinct developmental stages. Because each breed represented a biologically distinct host population, these cohorts were analyzed as independent datasets while preserving their common study origin to account for potential center-specific effects.

To eliminate study-specific bioinformatic biases and batch effects from disparate primary processing, all raw sequencing reads were processed uniformly from scratch. Raw FASTQ files were quality-filtered and adapter-trimmed using fastp (v1.0.1) [18] with a minimum Phred score threshold of Q20, and final sequencing quality was comprehensively assessed using MultiQC (v1.30) [19].

Amplicon datasets (16S rRNA gene and ITS) were processed using the DADA2 pipeline (v1.26.0) [20] for denoising, quality filtering, chimera removal, and amplicon sequence variant (ASV) inference. Taxonomic assignments for bacteria and fungi were

generated using the SILVA (v138.2) [21] and UNITE [22] reference databases, respectively. For WGS datasets, taxonomic profiling was performed directly from the quality-filtered raw reads using MetaPhlAn4 [23] against its native reference database to ensure seamless taxonomic alignment with amplicon-derived profiles.

Data harmonization, quality filtering and normalization

To bridge the architectural differences between sequencing technologies, taxonomic abundance tables derived from ASVs, OTUs, and shotgun metagenomic profiling were consolidated into a unified matrix. Only taxa confidently assigned to the genus level were retained to improve taxonomic harmonization across platforms and sequencing strategies.

Before downstream ecological modeling, a strict prevalence and abundance filter was applied independently within each sequencing stratum: taxa detected in fewer than 2% of samples or presenting a cumulative relative abundance below 0.10% were excluded [24]. To account for the highly heterogeneous sequencing depths characteristic of multi-study designs, the filtered genus-level abundance matrix was normalized via rarefaction to a uniform depth of 10,000 reads per sample for alpha and beta diversity estimations, thereby mitigating artifacts driven by uneven library sizes while preserving cross-study compatibility.

Metadata standardization and developmental stages

Host metadata were manually curated, standardized, and cross-validated against the taxonomic profiles to ensure complete correspondence between microbial signatures and host phenotypes. Host age was classified into five biologically relevant developmental stages based on pig production milestones: neonatal-suckling (stage A, 0–

20 days), nursery (stage B, 21–63 days), growing I (stage C, 64–90 days), growing II (stage D, 91–180 days), and finishing (stage E, 181–240 days).

Confounding host-associated and environmental variables, including breed nomenclature, sex, sampling location, and body weight, were equally harmonized into uniform categorical or continuous vectors across all 15 studies.

Diversity and ecological structure analysis

Microbial community diversity was characterized using QIIME2 [25] and the phyloseq [26] package in R (v4.5.2) [27], utilizing the rarefied genus abundance matrix. Alpha diversity was quantified using Chao1 richness, Shannon diversity, Simpson diversity, and Pielou's evenness indices. Data distribution was assessed using the Shapiro–Wilk test. Because alpha-diversity metrics did not satisfy normality assumptions, differences among host-associated categories were evaluated using Kruskal–Wallis tests followed by Dunn's multiple comparisons with Benjamini–Hochberg (BH) false discovery rate (FDR) correction. Statistical significance was defined as adjusted $p < 0.05$.

Beta diversity was assessed using Bray–Curtis dissimilarity and Jaccard distance to capture abundance-based and presence–absence community variation, respectively, and visualized via Principal Coordinates Analysis (PCoA). Community variation was further investigated using canonical analysis of principal coordinates (CAP) and distance-based redundancy analysis (db-RDA) based on Bray–Curtis dissimilarities using Hellinger-transformed rarefied abundance data to identify ecological gradients associated with microbiome maturation and age-adjusted body weight.

The relative contribution of host-associated variables to microbiome composition was evaluated using permutational multivariate analysis of variance (PERMANOVA; adonis2, vegan package) with 999 permutations [28]. Developmental stage, breed, sex, sampling location, and age-adjusted body weight were included as explanatory variables. Marginal (Type III) effects were estimated using the by = "margin" option, allowing each variable to be evaluated while controlling for the remaining covariates. To account for the multi-study design, permutations were constrained within BioProject (strata = BioProject), controlling for study-specific dependence while avoiding confounding caused by the uneven distribution of host variables among studies. Homogeneity of multivariate dispersion was evaluated using PERMDISP (betadisper, vegan package), and distances to group centroids were extracted to quantify ecological heterogeneity within biological groups.

Constrained community variation was further investigated using canonical analysis of principal coordinates (CAP) and distance-based redundancy analysis (db-RDA) based on Bray–Curtis dissimilarities. These analyses were used to identify ecological gradients associated with microbiome maturation and age-adjusted body weight while accounting for host-associated explanatory variables.

Integrative inference of age-adjusted body weight-associated microbiota

To distinguish productive performance from chronological development, body weight (BW) was regressed against host age using a linear regression model (BW ~ Age) fitted across the entire dataset. The resulting residuals were used as an age-adjusted body weight metric, representing individual deviations from the body weight expected at a given age. Positive residuals indicated animals heavier than expected for their age, whereas negative residuals indicated animals lighter than expected.

Microbial taxa associated with age-adjusted body weight were identified using an integrative analytical framework combining compositional differential abundance analysis – ANCOM-BC2 (v2.12.0) [29], multivariable association modeling – MaAsLin3 (v1.2.0) [30], and machine learning (Random Forest) [31]. Differential abundance models included developmental stage, breed, sex, sampling location, and age-adjusted body weight as explanatory variables, with BioProject incorporated as a covariate. Since both ANCOM-BC2 and MaAsLin3 rely on specialized, non-rarefied counts to execute their native normalization and bias-correction frameworks (sampling fraction bias estimation via log-linear modeling and internal centered log-ratio – CLR normalization, respectively), non-rarefied, quality-filtered baseline counts were supplied to these frameworks to maximize statistical power and control Type I error rates.

Random Forest regression models were implemented using the ranger algorithm through the caret framework to predict age-adjusted body weight from microbial community composition. Data were randomly partitioned into training (80%) and testing (20%) subsets. Hyperparameters were optimized using repeated 10-fold cross-validation (five repeats), and the final model was trained with 1,000 trees. Predictive performance was evaluated on the independent test set using the coefficient of determination (R^2), root mean square error (RMSE) and mean absolute error (MAE). Variable importance was estimated by permutation, and taxa with importance values above the 70th percentile were retained as candidate predictive features. To prioritize robust ecological signals and minimize method-specific bias, only taxa independently supported by at least two analytical approaches were retained as consensus age-adjusted body weight-associated microbial signatures [32].

Identification of the structural core microbiome

The structural core microbiome was defined as the set of microbial taxa consistently detected across all major host-associated biological strata evaluated. Presence–absence matrices were generated independently for developmental stage, breed, sex, and sampling location, and taxa occurring in every category within each host variable were classified as members of the structural core microbiome.

Core membership was determined using the intersection of taxa shared among all categories of each biological factor in place of abundance-based criteria, prioritizing ecological persistence across host-associated conditions while minimizing the influence of transient or condition-specific taxa [32]. Overlap between the structural core microbiome and consensus age-adjusted body weight-associated microbial signatures was quantified and visualized using UpSetR and VennDiagram. Taxa were subsequently classified into three ecological categories: (i) structural core taxa, (ii) age-adjusted body weight-associated taxa, and (iii) taxa shared between both groups.

Ecological network reconstruction and topological analysis

Microbial association networks were reconstructed using the NetCoMi package [33] from genus-level abundance profiles. Associations were inferred using SPRING (Semi-parametric Rank-based Correlation and Partial Correlation Estimation) [34, 35], which estimates sparse conditional dependence networks while accounting for the compositional nature of microbiome data. Network sparsity was determined using the Stability Approach to Regularization Selection (StARS), and the resulting association matrix was represented as a weighted undirected network in which nodes corresponded to microbial genera and edges represented retained positive or negative associations.

Hub genera were operationally defined as nodes with degree centrality values at or above the network-specific 90th-percentile cutoff. High-betweenness taxa were similarly defined using the 90th percentile of the betweenness distribution and were interpreted as occupying potential bridging positions within the network. Betweenness was calculated from the unweighted network topology to avoid interpreting SPRING association strengths as graph distances. Network modules were identified using Louvain community detection based on the absolute magnitude of edge weights, while association signs were retained for visualization. Topological analyses were performed using igraph and network visualization using ggraph.

For network analysis, animals were stratified according to residual-based age-adjusted body weight tertiles, and the High-BW network was reconstructed using samples belonging to the upper tertile of the residual distribution.

Functional annotation, pathway enrichment and performance-associated functional profiling

To characterize the predicted functional potential of microbial taxa associated with age-adjusted body weight, eggNOG-derived annotations [36] were used to retrieve KEGG Orthology (KO) identifiers and KEGG pathway assignments [37]. KO identifiers were standardized by removing the ko: prefix, and multiple KO and pathway annotations assigned to the same sequence were expanded into individual records. A custom TERM2GENE mapping was subsequently generated by linking unique KO identifiers to their corresponding KEGG pathways.

For taxon-level functional enrichment, unique KO repertoires were compiled independently for each microbial species and analyzed using the compareCluster function

277 of the clusterProfiler package [38] with enricher and the custom TERM2GENE
278 annotation. Pathways with nominal p values < 0.05 were retained at this enrichment stage.

279 Productive performance was represented by age-adjusted body weight. Body
280 weight was regressed against host age using a linear regression model (BW ~ Age), and
281 the resulting residuals were divided into tertiles using the ntile function. Animals were
282 thereby classified into Low-BW, Mid-BW, and High-BW groups, corresponding to the
283 lower, intermediate, and upper thirds of the residual distribution, respectively.

284 To estimate sample-level predicted functional profiles, KEGG pathway
285 annotations were linked to the microbial abundance matrix. Species abundances were
286 converted to long format and matched to species-specific KEGG pathway assignments.
287 For each sample, the abundance contribution of species associated with each pathway was
288 aggregated to generate a sample-by-pathway abundance matrix. This matrix was aligned
289 with the corresponding host metadata before downstream statistical analysis.

290 Differential pathway abundance among age-adjusted body weight groups was
291 evaluated using MaAsLin3, with the Low-BW group specified as the reference category.
292 Pathway profiles were normalized using centered log-ratio (CLR) transformation within
293 MaAsLin3, with no additional transformation applied, and performance group was
294 included as the fixed effect. False discovery rate-adjusted q values were used to identify
295 performance-associated pathways.

296 KEGG pathway identifiers were subsequently annotated with pathway
297 descriptions using KEGGREST. Significantly associated pathways were linked back to
298 the consensus microbial species carrying the corresponding functional annotations. To
299 facilitate biological interpretation, pathway results underwent biologically guided post-

processing to exclude pathways primarily related to human diseases, eukaryotic signaling, plant metabolism, and other categories not directly relevant to microbial metabolism. Retained pathways were prioritized according to microbial metabolic functions, including carbohydrate and central carbon metabolism, SCFA-associated metabolism, amino acid metabolism, energy metabolism, vitamin and cofactor metabolism, microbial signaling, transport, and other ecologically relevant processes.

Results

Global ecological integration reveals host development as a conserved chronological axis of gut microbiome succession

A total of 1,608 pig gut microbiome samples from 15 independent BioProjects representing seven countries (Austria, Canada, China, Denmark, France, Russia, and the United States) were retained after quality control and metadata harmonization (Figure 1A; Table S1). The integrated dataset encompassed multiple production systems, developmental stages, host genetic backgrounds, sampling locations, and sequencing strategies, providing a comprehensive ecological framework for cross-study analyses (Figure 1B).

Principal coordinates analysis (PCoA) based on Bray–Curtis dissimilarities revealed a pronounced age-dependent succession of microbial communities across host development (Figure 2A). Microbial communities from early developmental stages occupied distinct ecological space, whereas samples from finishing animals formed a compact cluster with markedly reduced dispersion, indicating increased compositional similarity among mature microbiomes.

Constrained ordination analyses corroborate these patterns. Canonical analysis of principal coordinates (CAP) and distance-based redundancy analysis (db-RDA) consistently identified developmental stage as a major ecological gradient structuring the microbial community (Figure 2B–C). Intermediate developmental stages (B–D) exhibited greater ecological dispersion, whereas finishing animals showed significantly lower within-group variability. Beta-dispersion analysis (PERMDISP) further demonstrated a progressive reduction in distances to group centroids throughout host development ($p < 0.001$), indicating increasing ecological convergence during microbiome maturation.

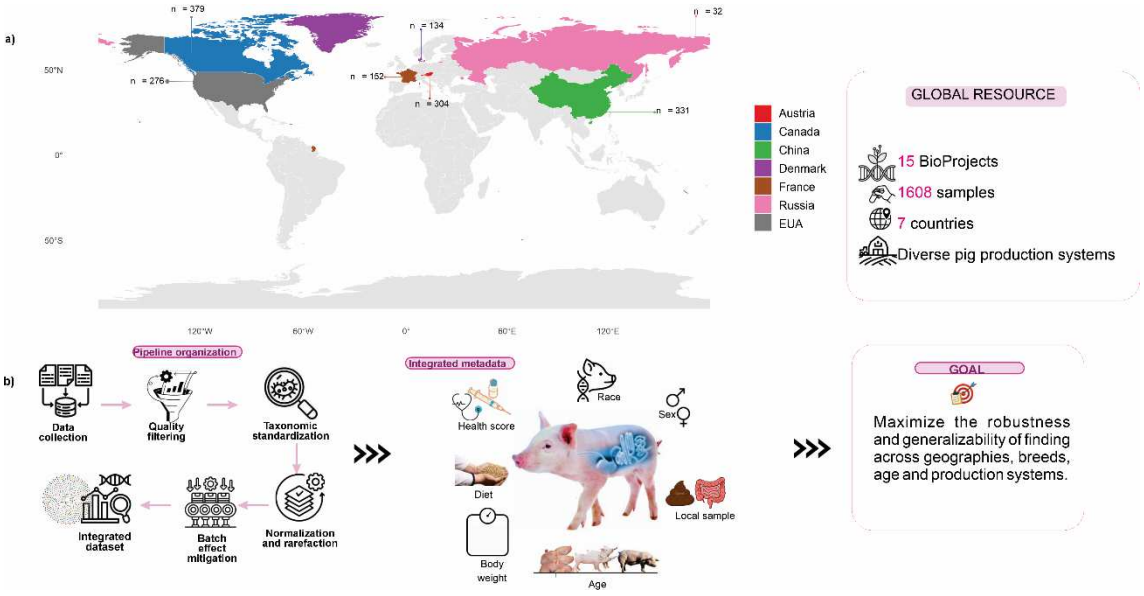
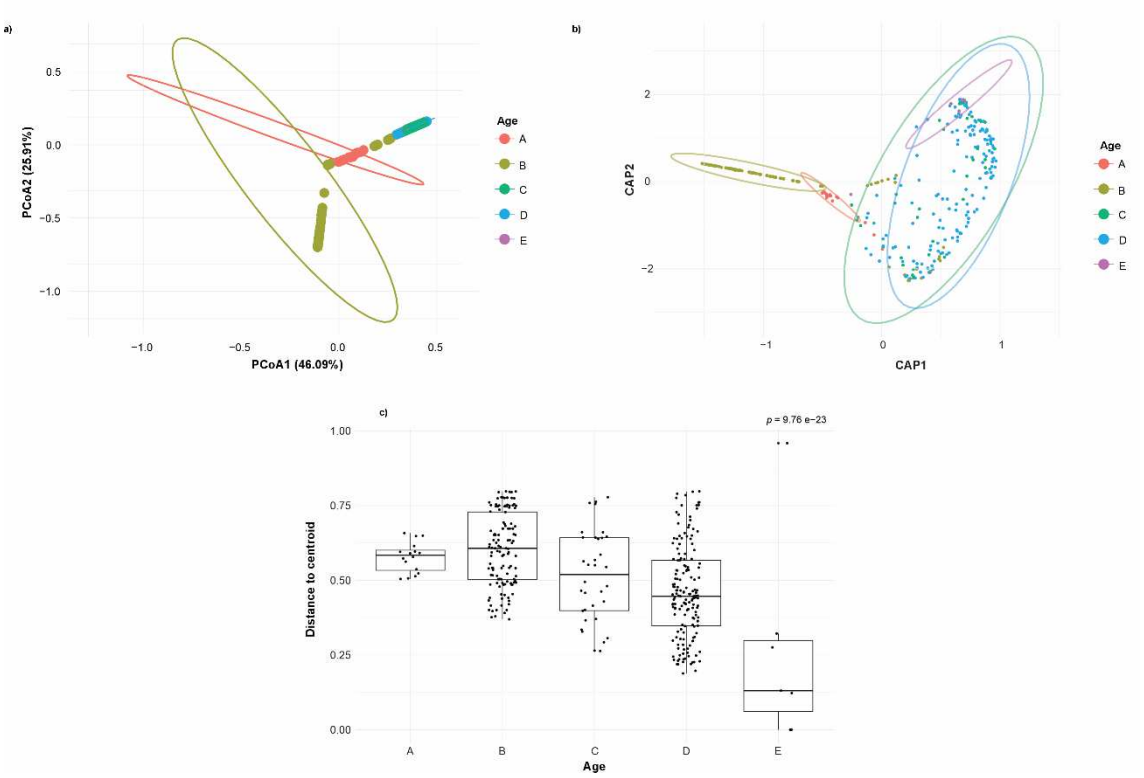


Figure 1. Global study design and ecological integration framework. (a) Geographic distribution of the 1,608 pig gut microbiome samples included in this meta-analysis, representing 15 independent BioProjects from seven countries. Numbers indicate the number of samples contributed by each country. (b) Overview of the analytical workflow used for dataset harmonization and ecological integration, including quality control, taxonomic standardization, metadata harmonization, normalization, and downstream ecological analyses.



337 **Figure 2. Host development drives the ecological succession of the pig gut**
338 **microbiome.** (a) Principal coordinates analysis (PCoA) based on Bray–Curtis
339 dissimilarities showing progressive ecological separation of microbial communities
340 across developmental stages. (b) Canonical analysis of principal coordinates (CAP)
341 constrained by developmental stage. (c) Beta-dispersion analysis showing distances to
342 group centroids across developmental stages. Developmental stages were defined as A
343 (0–20 days), B (21–63 days), C (64–90 days), D (91–180 days), and E (181–240 days).

344 **Microbiome maturation follows a disturbance–recovery trajectory centered on**
345 **weaning**

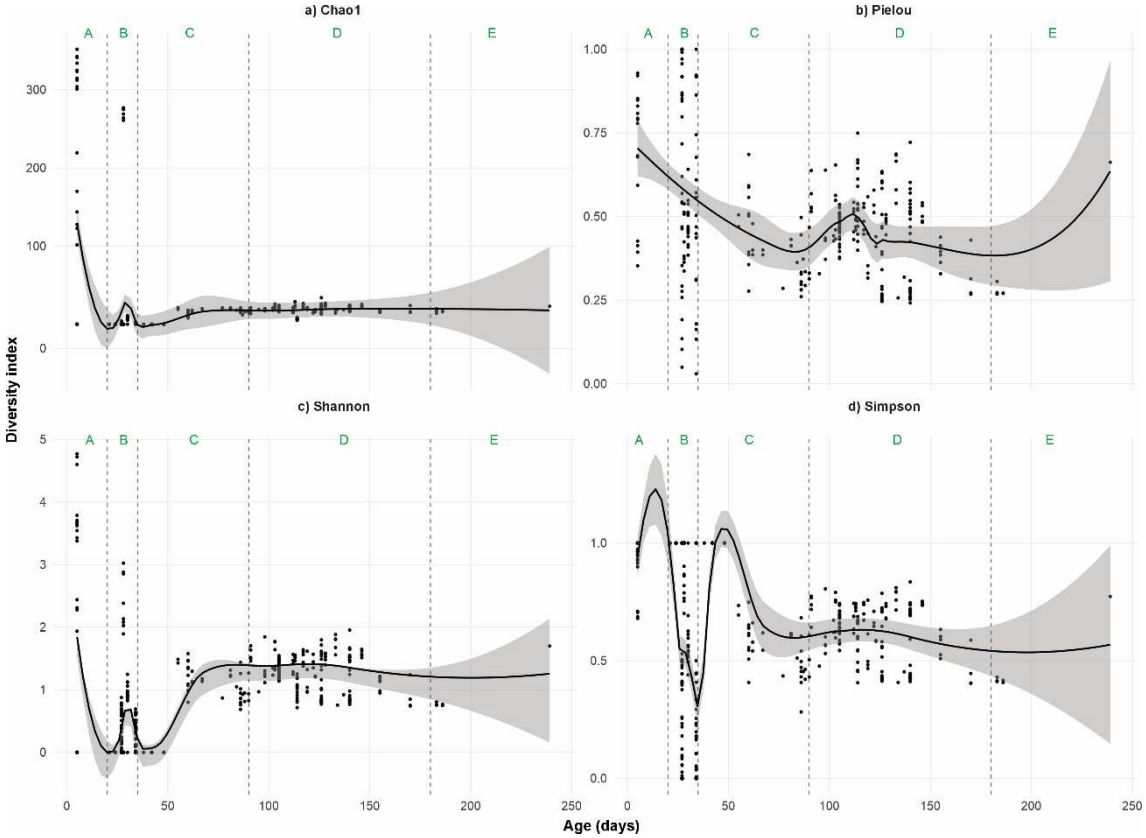
346 Following the identification of host development as the dominant ecological
347 driver, we investigated temporal changes in microbial diversity during microbiome
348 maturation. Alpha-diversity analyses revealed significant differences among

developmental stages (Figure 3; Figure S1). Chao1 richness increased progressively throughout host development, whereas Shannon diversity, Simpson diversity, and Pielou's evenness displayed non-linear trajectories characterized by transient reductions during the nursery stage.

The most pronounced changes occurred immediately after weaning, when microbial communities exhibited reduced evenness and increased dominance, followed by progressive recovery during subsequent developmental stages. Beta-dispersion analyses supported these observations by demonstrating elevated within-group ecological heterogeneity during early life and the post-weaning transition, followed by a gradual reduction in community variability toward adulthood.

These results indicate that microbiome maturation is characterized by a disturbance–recovery trajectory rather than by a continuous increase in diversity. Early developmental stages displayed greater ecological heterogeneity, whereas mature communities progressively converged toward more compositionally similar configurations.

Taxonomic succession accompanied these diversity patterns. Early-life microbiomes were dominated by taxa typically associated with milk-based nutrition, whereas later developmental stages showed increasing representation of anaerobic fermentative genera adapted to complex carbohydrate metabolism, reflecting progressive ecological succession during host maturation.



369 **Figure 3. Alpha-diversity dynamics during pig gut microbiome maturation.**
370 Scatterplots show individual samples and LOESS-smoothed trends (95% confidence
371 intervals) for (a) Chao1 richness, (b) Pielou's evenness, (c) Shannon diversity, and (d)
372 Simpson diversity across developmental stages. Developmental stages correspond to A
373 (0–20 days), B (21–63 days), C (64–90 days), D (91–180 days), and E (181–240 days).

374 **Host-associated biological variables modulate microbiome structure without**
375 **disrupting age-associated succession**

376 Although the developmental stage was identified as the primary driver of
377 microbiome maturation, additional host-associated biological variables also contributed
378 significantly to community organization. After accounting for study-specific variation by
379 restricting permutations within each BioProject, developmental stage explained the

largest independent component of baseline compositional variation (R^2 = 0.270; p = 0.049), followed by host genetic background (breed) (R^2 = 0.166; p = 0.001). In contrast, neither sex (R^2 = 0.068; p = 0.353) nor age-adjusted body weight (R^2 = 0.072; p = 0.633) significantly influenced overall community composition (Table S2).

Alpha-diversity analyses corroborated these findings by revealing significant breed-associated differences in microbial richness, diversity, dominance, and evenness (Figures S2–S3), indicating that host genetic background contributes to ecological modulation of microbial communities. Sampling location also affected diversity metrics, reflecting the distinct physiological environments represented by intestinal compartments and fecal samples. Nevertheless, the magnitude of these effects remained consistently smaller than the extensive ecological restructuring associated with microbiome maturation.

Multivariate homogeneity analyses (PERMDISP; Figure S4) further demonstrated significant differences in within-group ecological dispersion, particularly among host genetic backgrounds and age-adjusted body weight groups, indicating differences in community heterogeneity. However, these changes reflected secondary ecological modulation instead of alternative developmental trajectories, reinforcing that microbiome succession remained highly conserved across host-associated biological conditions.

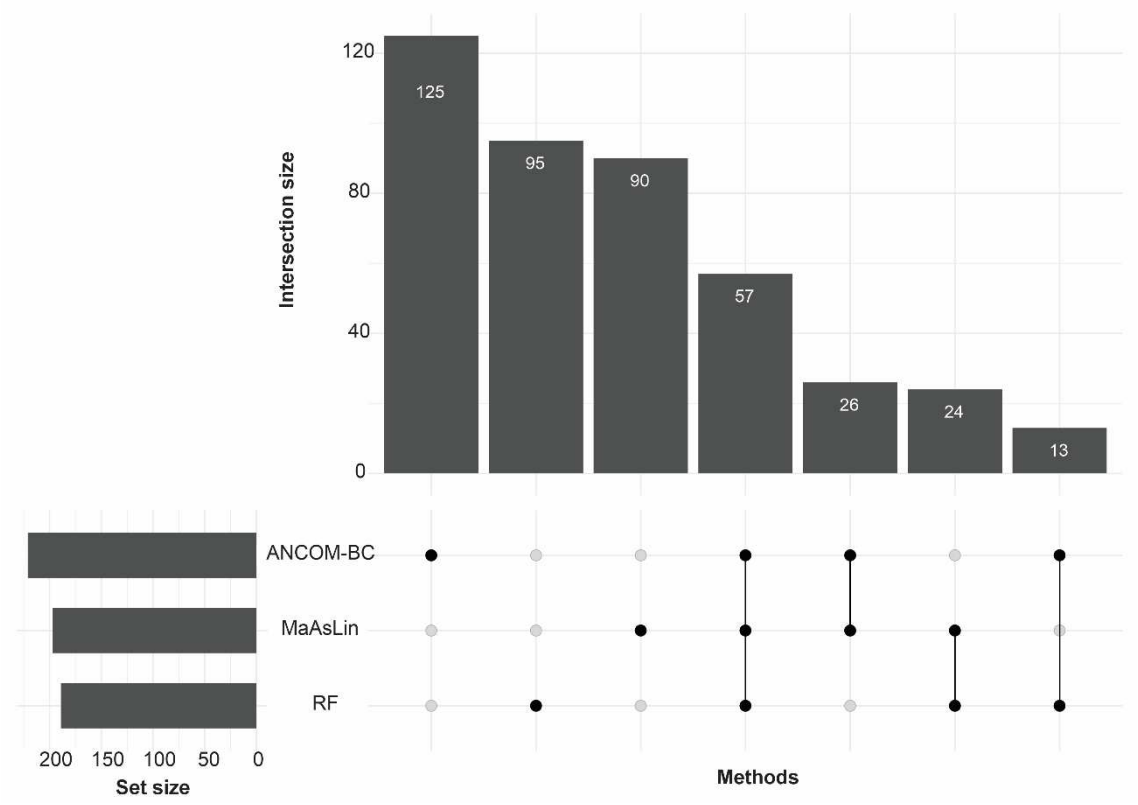
Collectively, these findings indicate that host-associated biological variables refine microbiome composition without disrupting the conserved developmental trajectory established during maturation. These observations support the hypothesis that microbiome organizations are maintained by an underlying structural component that persists across heterogeneous host contexts. We therefore investigated whether this

403 conserved structural layer could be identified across developmental stages and host-
404 associated biological variables.

405 **Identification of the conserved structural layer of the pig gut microbiome**

406 The predictive capacity of the Random Forest model (implemented via the *ranger*
407 package using an 80/20 train/test split) was evaluated on the independent, hold-out test
408 set. The model showed predictive performance in the held-out test subset, yielding R^2 =
409 0.612, RMSE = 0.556, and a correlation of r = 0.782 between observed and predicted
410 values. Ninety-five taxa exceeded the predefined permutation-importance threshold (70th
411 percentile) and were retained as predictive microbial signatures. In parallel, ANCOM-
412 BC2 and MaAsLin3 identified 221 and 197 taxa in total, respectively (125 and 90 of
413 which were method-exclusive; Figure 4). To increase robustness and reduce method-
414 specific false positives, taxa supported by at least two of the three independent analytical
415 approaches (Random Forest, ANCOM-BC2, and MaAsLin3) were retained as consensus
416 performance-associated biomarkers, yielding a final set of 120 taxa.

417 Because these consensus biomarkers could represent either persistent members of
418 the resident microbiome or transient microbial populations associated with host
419 performance, we next investigated whether they overlapped with a conserved structural
420 microbiome shared across heterogeneous biological conditions. This comparison enabled
421 us to distinguish microbial signatures associated with age-adjusted body weight from
422 those constituting the persistent ecological foundation of the pig gut microbiome.



423 **Figure 4. Consensus prioritization identifies robust performance-associated**
424 **microbial taxa**. The UpSet plot illustrates the overlap among microbial taxa identified
425 by independent analytical approaches, including ANCOM-BC2, MaAsLin3, and Random
426 Forest modeling. A total of 120 taxa were supported by at least two independent analytical
427 frameworks and retained as consensus performance-associated biomarkers.

428 **Structural core microbiome analyses reveal a conserved ecological mechanism**
429 **across heterogeneous host conditions**

430 Structural core microbiome analyses demonstrated that, despite extensive
431 ecological restructuring throughout host development and across host-associated
432 biological variables, a conserved subset of microbial taxa remained consistently
433 detectable under all evaluated biological conditions, supporting the existence of a resilient
434 structural core within the pig gut microbiome (Figure 5). Persistence analyses

incorporating developmental stage, sex, sampling location, and broad host genetic background (native versus commercial production systems) showed that large-scale compositional turnover occurred without disrupting this conserved ecological mechanism, indicating that microbiome maturation is largely characterized by community reorganization while preserving a substantial proportion of persistent microbial lineages.

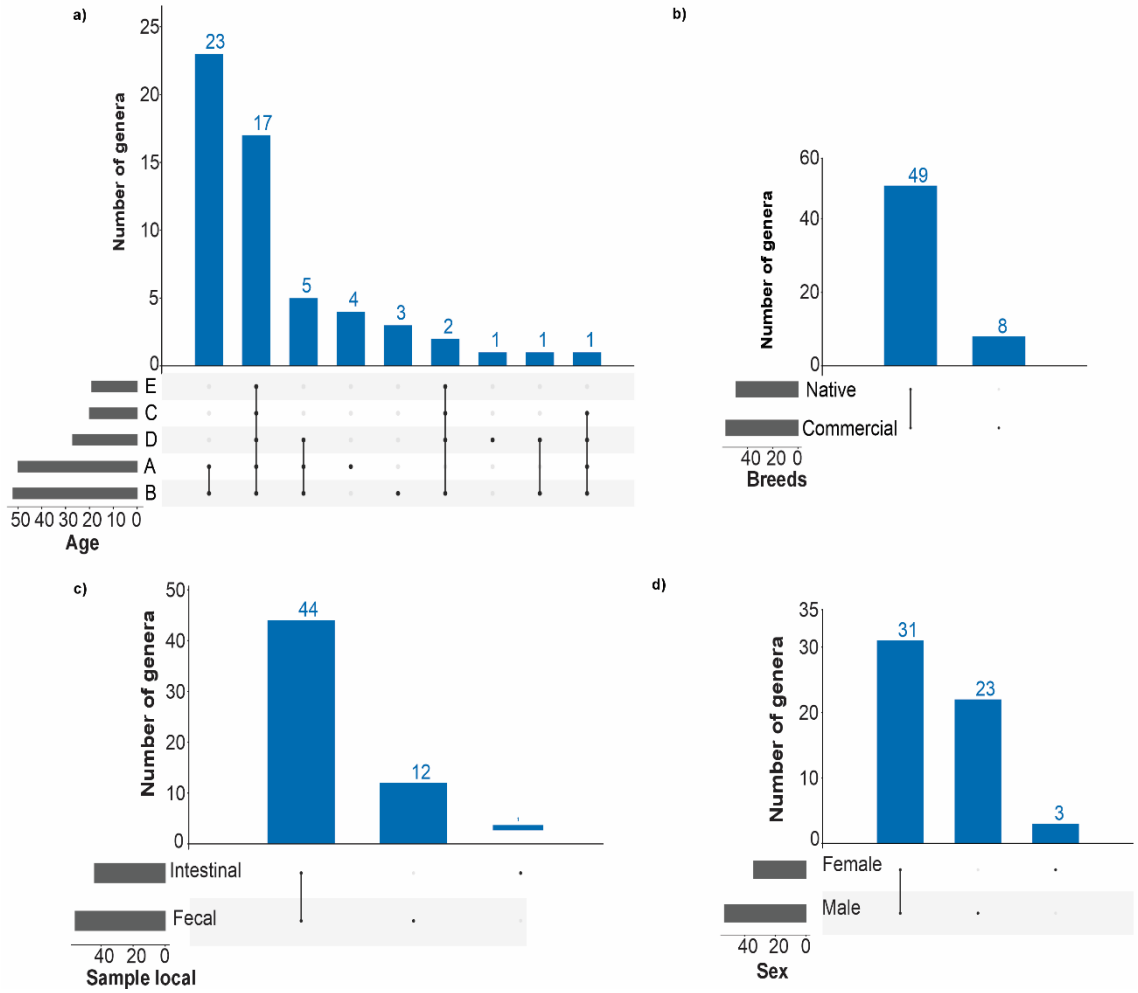
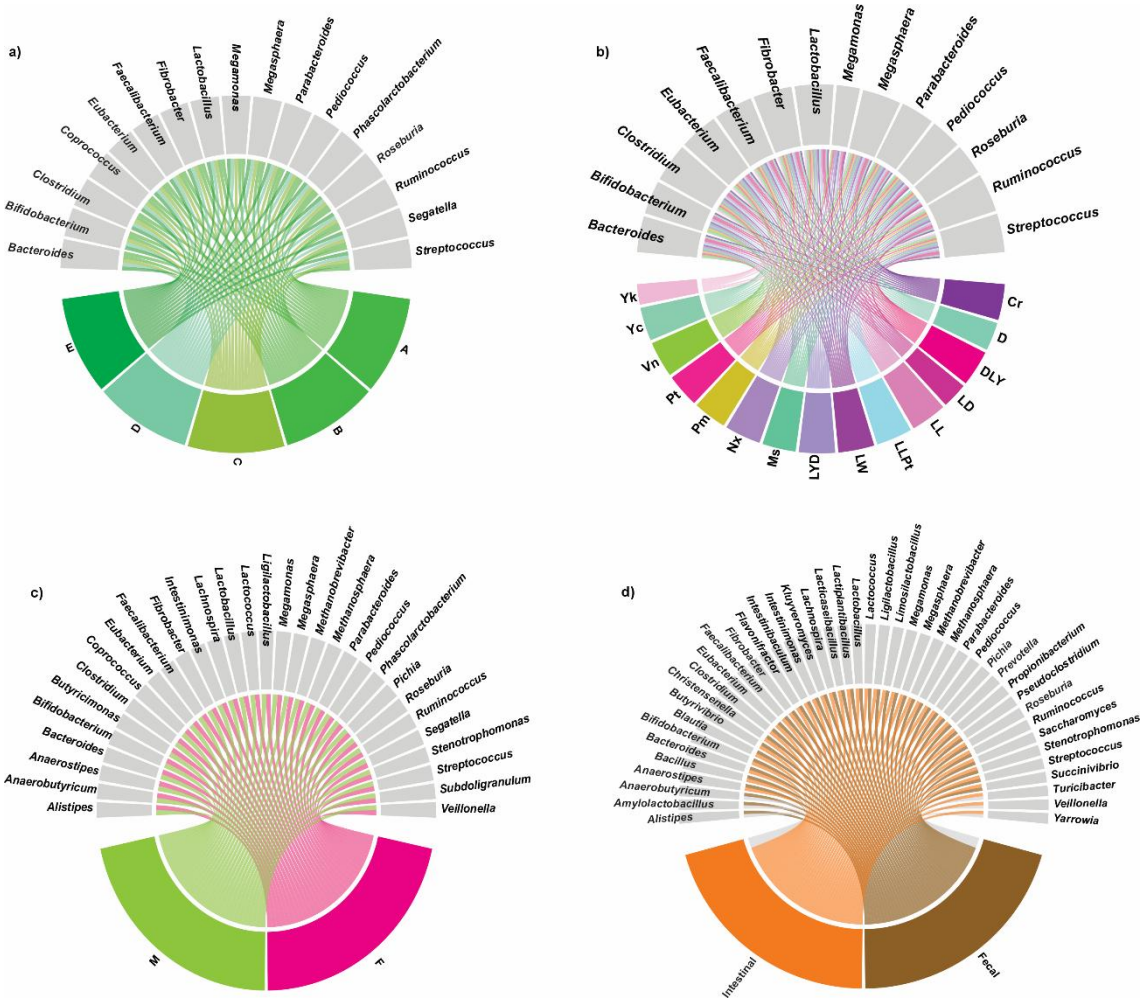


Figure 5. Identification of a conserved structural core microbiome across heterogeneous pig production systems. UpSet plot illustrating microbial taxa consistently detected across developmental stages, sex groups, sampling locations, and broad host genetic backgrounds (native vs. commercial production systems). Despite

445 substantial ecological succession throughout host development, a conserved core subset
446 remained persistently associated with the pig gut ecosystem. Developmental stages: A
447 (0–20 d), B (21–63 d), C (64–90 d), D (91–180 d), and E (181–240 d). Native breeds: Cr
448 (Creole), Ms (Meishan), Nx (Ningxiang), Pm (Pitman-Moore), Vn (Vietnamese), and Yc
449 (Yucatan). Commercial breeds: D (Duroc), DLY (Duroc × Landrace × Yorkshire), LD
450 (Landrace), LL (Large White × Landrace), LLPt (Large White × Landrace × Pietrain),
451 LW (Large White), LYD (Landrace × Yorkshire × Duroc), Pt (Pietrain), and Yk
452 (Yorkshire).

453 Having established the existence of this conserved structural core, we next
454 examined the microbial taxa responsible for maintaining ecological persistence across
455 individual host-associated biological conditions (Figure 6). Persistent taxa were
456 predominantly represented by metabolically important genera, including *Prevotella*,
457 *Faecalibacterium*, *Roseburia*, *Ruminococcus*, and *Lactobacillus*. Although their relative
458 abundances varied throughout host development and among breed groups, these genera
459 remained consistently detectable across developmental stages, host genetic backgrounds,
460 sampling locations, and sex groups, supporting their role as stable structural components
461 of the pig gut microbiome.



462 **Figure 6. Ecological distribution of conserved core microbial taxa across biological**
463 **conditions.** Chord diagrams illustrating the distribution of recurrent core genera across
464 a) developmental stages, b) host genetic backgrounds, c) sex groups, and d) sampling
465 locations. Connecting links represent the occurrence of each genus within the
466 corresponding biological category. Panel a highlights genus persistently detected
467 throughout host maturation, whereas panels b–d demonstrate conservation across host
468 genetic and environmental factors. Developmental stages: A (0–20 d), B (21–63 d), C
469 (64–90 d), D (91–180 d), and E (181–240 d). Breed abbreviations: Cr (Creole), D (Duroc),
470 DLY (Duroc × Landrace × Yorkshire), LD (Landrace), LL (Large White × Landrace),
471 LLPt (Large White × Landrace × Pietrain), LW (Large White), LYD (Landrace ×

Yorkshire × Duroc), Ms (Meishan), Nx (Ningxiang), Pm (Pitman-Moore), Pt (Pietrain), Vn (Vietnamese), Yc (Yucatan), and Yk (Yorkshire).

The identification of this persistent structural core also gives an ecological framework for interpreting performance-associated microbial signatures. If the structural microbiome remains conserved across host-associated biological conditions, microbial biomarkers associated with age-adjusted body weight are unlikely to arise from replacement of the structural community itself. Instead, these findings suggest that performance-associated microbial signatures emerge through ecological reconfiguration within a conserved structural mechanism. Together, these results provide empirical support for the structural layer proposed in our conceptual framework and establish the foundation for evaluating whether functional reorganization represents a complementary adaptive layer of the pig gut microbiome.

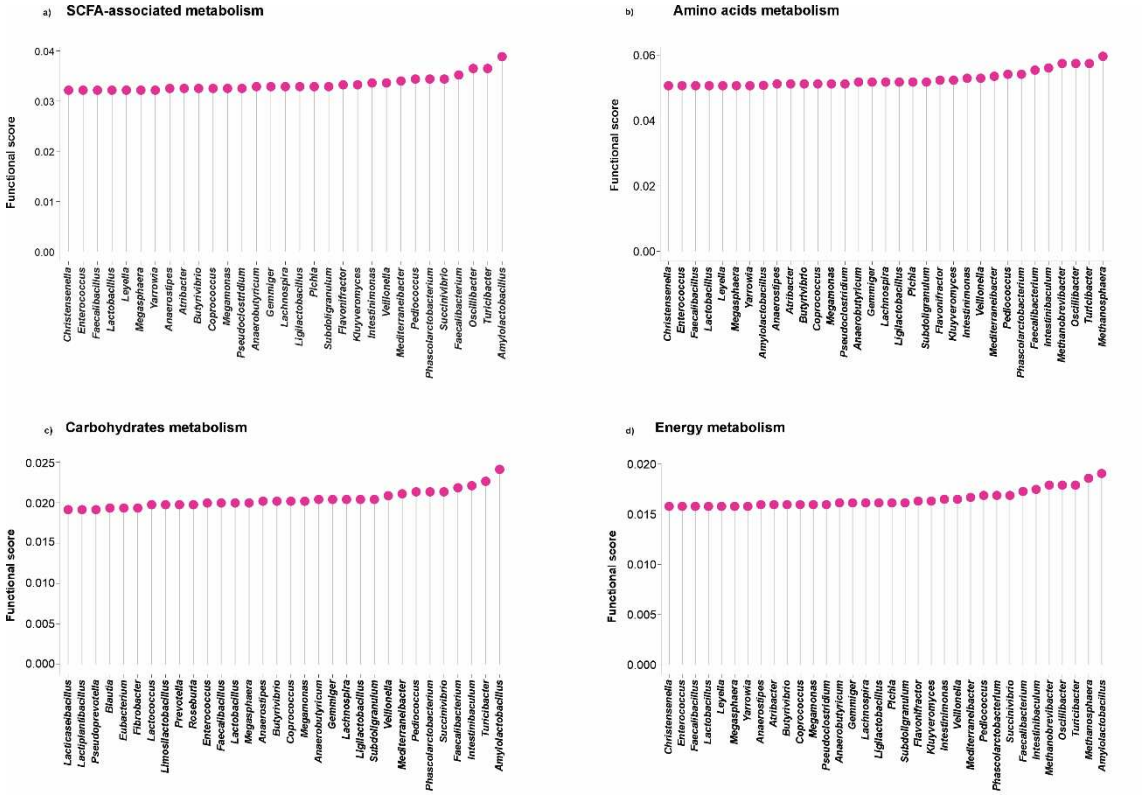
Predicted functional potential remodeling defines the adaptive layer associated with age-adjusted body weight

Having established the existence of a conserved structural layer, we next investigated whether variation in age-adjusted body weight was associated with a second, adaptive ecological layer characterized by coordinated predicted functional remodeling as opposed to structural replacement. Predicted KEGG pathways reconstructed from eggNOG annotations were evaluated using enrichment analysis. Functional enrichment was performed with compareCluster based on custom TERM2GENE mappings generated from eggNOG annotations, considering pathways with $p < 0.05$ as significantly enriched. This analysis identified coordinated metabolic functions shared among microbial

494 populations associated with age-adjusted body weight, providing the foundation for
495 subsequent differential abundance analyses.

496 To identify microbial functions specifically associated with body weight
497 independent of developmental stage, differential pathway abundance across residual-
498 based performance tertiles was evaluated using MaAsLin3. Among the 496 predicted
499 KEGG pathways, 113 were significantly associated with productive microbial profiles
500 after multiple-testing correction ($q < 0.05$), demonstrating extensive functional
501 remodeling beyond taxonomic restructuring.

502 For visualization purposes, significantly associated pathways were grouped into
503 four biologically meaningful functional categories according to their established
504 metabolic roles: (i) carbohydrate metabolism, including starch and sucrose metabolism,
505 glycolysis/gluconeogenesis, and pyruvate metabolism; (ii) amino acid metabolism; (iii)
506 energy metabolism; and (iv) short-chain fatty acid (SCFA)-associated metabolism,
507 comprising pathways involved in acetate, propionate, and butyrate biosynthesis (Figure
508 7; Table S3). Functional scores represent the normalized contribution of each microbial
509 genus to significantly enriched pathways within each category.



510 **Figure 7. Differentially enriched KEGG metabolic pathways associated with age-**
511 **adjusted body weight.** Functional potential was inferred from eggNOG annotations
512 mapped to KEGG pathways. Differential pathway abundance across residual-based
513 performance tertiles was evaluated using MaAsLin3 ($q < 0.05$, FDR corrected).
514 Significantly enriched pathways were grouped into four functional categories:
515 carbohydrate metabolism, amino acid metabolism, energy metabolism, and SCFA-
516 associated metabolism. Functional scores represent the normalized contribution of each
517 microbial genus to enriched pathways within each functional category.

518 The predicted functional profile associated with greater age-adjusted body weight
519 was characterized by pathways related to complex carbohydrate utilization, amino acid
520 metabolism, energy metabolism, and SCFA-associated processes. To characterize the
521 ecological organization of microbial genera within the high-performance group, a genus-

level SPRING association network was reconstructed for animals with high age-adjusted body weight (High BW; Figure 8a). Hub and high-betweenness genera were identified using the respective network-specific 90th-percentile thresholds for degree and betweenness centrality (Figure 8b).

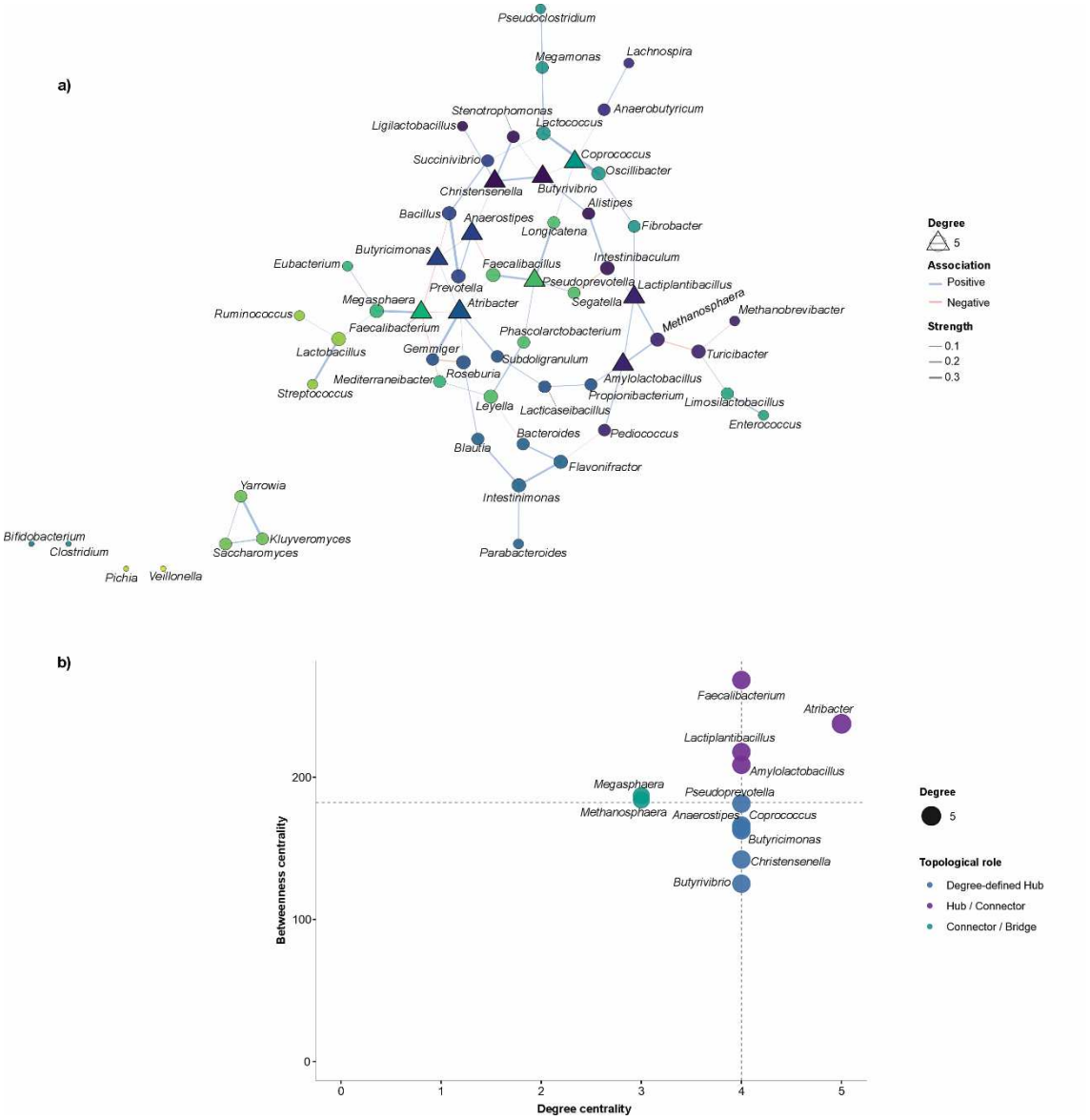


Figure 8. Topological organization of the high age-adjusted BW microbial association network. (a) SPRING network showing genus-level associations; node size represents degree, triangles indicate degree-defined hubs (≥90th percentile), colors

529 denote Louvain modules, and edge color and width represent association sign and
530 strength, respectively. (b) Relationship between degree and betweenness centrality;
531 dashed lines indicate the corresponding 90th-percentile thresholds for hubs and high-
532 betweenness taxa.

533 The High-BW network comprised 58 microbial genera connected by 67
534 associations, including 52 positive and 15 negative associations. The network exhibited
535 a sparse topology (density = 0.041; mean degree = 2.31) and was partitioned into multiple
536 modules by Louvain community detection (modularity = 0.831). The 90th-percentile
537 degree cutoff was 4, identifying ten degree-defined hubs, with *Atribacter* showing the
538 highest degree (degree = 5), followed by *Faecalibacterium*, *Lactiplantibacillus*,
539 *Amylolactobacillus*, *Pseudoprevotella*, *Coprococcus*, *Anaerostipes*, *Butyricimonas*,
540 *Christensenella*, and *Butyrivibrio* (degree = 4; Figure 8a).

541 Degree and betweenness captured partially distinct aspects of network topology
542 (Figure 8b). *Faecalibacterium*, *Atribacter*, *Lactiplantibacillus*, and *Amylolactobacillus*
543 exceeded both the degree and betweenness 90th-percentile thresholds, combining high
544 direct connectivity with high bridging centrality. In contrast, *Megasphaera* and
545 *Methanosphaera* exceeded the betweenness threshold despite having lower degree,
546 occupying positions consistent with potential connector roles within the inferred
547 association network.

548 Notably, structural-core genera occupied heterogeneous topological positions
549 within the High-BW network. Although *Faecalibacterium* simultaneously exhibited hub
550 and high-betweenness characteristics, other persistent core genera, including *Prevotella*,

551 *Roseburia*, *Lactococcus*, *Bacillus*, and *Subdoligranulum*, showed intermediate or low
552 degree centrality. Thus, persistence across host-associated conditions was not
553 synonymous with high network connectivity.

554 Together, these findings indicate that the High-BW microbial association network
555 contains distinct topological roles, ranging from highly connected hubs to high-
556 betweenness taxa with potential bridging positions. Importantly, these roles were only
557 partially aligned with structural-core membership, indicating that ecological persistence
558 and network centrality represent complementary properties of microbial community
559 organization.

560 **Discussion**

561 In this study, we leveraged a large-scale cohort of 1,608 pig gut microbial profiles
562 to resolve a problem in livestock microbiome research: the persistent cross-study
563 irreproducibility of single-taxon biomarkers associated with animal performance. As
564 opposed to viewing taxonomic variance as experimental noise, our findings demonstrate
565 that the pig gut microbiome operates through a two-layer hierarchical ecological model.

566 This architecture comprises a persistent structural component across
567 heterogeneous host conditions together with an adaptive layer characterized by
568 performance-associated taxonomic, predicted functional, and topological variation. By
569 shifting the analytical focus from individual microbial abundances to topological synergy,
570 this work offers a unifying ecological framework for host-microbiome interactions in
571 production systems.

572 Our multivariate ordination and variance partitioning revealed that while genetic
573 background and study cohort dictate baseline spatial coordinates ($R^2 = 0.166$, $p = 0.001$),

574 the temporal maturation trajectory follows a universal, highly canalized path. A critical
575 successional milestone in this trajectory occurs during stage B (days 21–63),
576 characterized by a marked, transient contraction in Pielou's evenness and Shannon
577 diversity index. This sharp decline captures the precise ecological footprint of the
578 weaning bottleneck, a severe, deterministic environmental disturbance driven by abrupt
579 dietary transitions from maternal milk to complex plant-derived polysaccharides,
580 compounded by social and physiological stress [2, 39, 40].

581 In place of representing stochastic drift, this transient collapse in evenness reflects
582 the rapid, opportunistic expansion of specialized primary fermenters outcompeting milk-
583 adapted pioneer taxa (*Lactobacillus* and *Bifidobacterium*) [41, 42]. As emphasized by
584 Saladrigas-García et al. [43], this window temporarily decreases maternal ecological
585 resilience and lowers the colonization resistance barrier before the ecosystem reorganizes
586 toward a mature climax state [44, 45]. During later finishing stages (stages D–E),
587 community dispersion progressively contracts, demonstrating that gut microbiome
588 assembly converges toward predictable, adult-like homeostatic configurations despite
589 environmental and management heterogeneity across cohorts [39, 46].

590 Our SPRING-derived microbial association network revealed a heterogeneous
591 topological organization within the High-BW microbiome, characterized by a sparse
592 overall structure in which degree-defined hubs and high-betweenness taxa occupied
593 partially distinct network positions. Instead of representing isolated phenomena,
594 microbiome succession, structural persistence, predicted functional remodeling, and
595 network topology operate as integrated dimensions of a unified hierarchical system. By
596 evaluating node metrics across the combined dataset, we identified a synergy between

597 local modularity and global connectivity operationalized through a two-layer ecological
598 model.

599 Importantly, membership in the persistent structural core did not necessarily
600 imply high topological centrality. Among recurrent core genera, *Faecalibacterium*
601 simultaneously exhibited hub and high-betweenness characteristics, whereas *Prevotella*,
602 *Roseburia*, *Blautia*, *Ruminococcus*, and *Subdoligranulum* occupied less-connected
603 positions. This distinction suggests that cross-condition persistence and topological
604 centrality capture different ecological properties: the former reflects recurrence across
605 heterogeneous host contexts, whereas the latter describes the position of individual genera
606 within a specific performance-associated association network.

607 The persistence of these genera across developmental stages, genetic
608 backgrounds, and sampling conditions suggests that the structural core may contribute to
609 ecological continuity despite substantial compositional variation. Functional redundancy
610 among gut microbial populations has been proposed as an important mechanism
611 supporting ecosystem stability under environmental perturbation [46, 47]; however, the
612 present analysis identifies taxonomic persistence rather than directly quantifying
613 functional redundancy or resilience.

614 In addition to this persistent structural baseline lies a second organizational tier
615 characterized by adaptive functional variation associated with age-adjusted body weight.
616 Previous studies have linked gut microbial functional organization to host energy
617 harvesting, SCFA-associated metabolism, and the coordination of metabolically
618 complementary microbial populations [48, 49]. Within the High-BW network,
619 *Faecalibacterium*, *Atribacter*, *Lactiplantibacillus*, and *Amylolactobacillus* combined
620 high degree with high betweenness centrality, whereas *Megasphaera* and

Methanosphaera exhibited high betweenness despite comparatively lower degree. These contrasting topological configurations are consistent with heterogeneous ecological roles, in which highly connected genera coexist with taxa occupying potential bridging positions between otherwise weakly connected regions of the network. Such connector-like organization is conceptually compatible with ecological models of microbial cross-feeding and metabolic complementarity among functional guilds [50, 51]. However, high betweenness in an association network does not directly demonstrate metabolite exchange, functional coupling, or causal interactions among taxa. Thus, these high-betweenness genera are more appropriately interpreted as topological candidates for bridging microbial subcommunities rather than as demonstrated metabolic bottlenecks.

This two-layer ecological model resolves decade-long contradictions in livestock microbiome literature. While individual microbial biomarkers may vary across studies because of environmental filtering and cohort-specific drivers, the recurrent detection of structural-core taxa across heterogeneous host conditions suggests that ecological persistence may be more reproducible than individual abundance-based biomarkers.

Age-adjusted body weight was not significantly associated with wholesale restructuring of microbial community composition in the PERMANOVA analysis, yet specific taxonomic, predicted functional, and topological patterns were associated with the High-BW phenotype. The sparse High-BW network contained distinct degree-defined hubs and high-betweenness configurations, suggesting that performance-associated microbial organization may involve specific network positions rather than a uniform increase in connectivity among structural-core taxa. Such organization is biologically compatible with previously described microbial cross-feeding relationships linking complex carbohydrate degradation, pyruvate fermentation, and the production of short-

chain fatty acids, including acetate, propionate, and butyrate [52–54]. However, because SPRING-derived associations do not directly demonstrate metabolite exchange or causal interactions, these patterns should be interpreted as topological signatures potentially consistent with metabolic complementarity rather than as direct evidence of enhanced cross-feeding or SCFA production.

These patterns are consistent with metabolic guild and cross-feeding models proposed in previous studies [48, 54], in which complementary microbial functions may contribute to substrate utilization and host energy harvesting. In the present study, however, these metabolic relationships remain hypotheses supported by predicted functional potential and association topology rather than direct measurements of metabolic exchange. Random Forest modeling further demonstrated that microbial taxonomic composition contained predictive information for age-adjusted body weight ($R^2 = 0.612$; RMSE = 0.556 in the held-out test subset), supporting the presence of distributed microbial signatures associated with performance. Moving beyond the reductionist search for linear, single-taxon "biomarkers" [49] is therefore essential for developing reproducible, network-targeted nutritional and biological interventions in livestock production.

Conclusions

This global integrative analysis reveals that pig gut microbiome maturation follows a conserved age-associated successional trajectory across heterogeneous production environments. Across 1,608 microbial profiles, host development emerged as the principal chronological axis of community variation, while a persistent subset of microbial genera was maintained across developmental stages, host genetic backgrounds, sexes, and sampling sites. Together, these findings support the existence of a structural

ecological layer that persists despite substantial variation in community composition across host and environmental contexts.

Beyond this persistent structural component, age-adjusted body weight was associated with a distinct set of taxonomic signatures and predicted functional features rather than with wholesale restructuring of the microbial community. Functional analyses identified coordinated variation in pathways related to carbohydrate utilization, amino acid metabolism, energy metabolism, and SCFA-associated processes. Within the High-BW microbial association network, degree-defined hubs and high-betweenness taxa occupied distinct topological positions, while structural-core membership was only partially associated with network centrality. These results indicate that ecological persistence, predicted functional potential, and network topology represent complementary dimensions of microbiome organization rather than interchangeable properties of the same microbial taxa.

Collectively, these findings support a hierarchical ecological framework in which a persistent structural layer provides continuity across heterogeneous host conditions, whereas performance-associated variation is reflected in more selective taxonomic, predicted functional, and topological patterns. This perspective shifts emphasis from the search for isolated microbial biomarkers toward an ecosystem-level understanding of microbial organization and provides a foundation for the development of microbiome-informed strategies for precision livestock production. Future experimental and longitudinal studies will be required to determine whether the identified functional and topological signatures contribute causally to host performance and whether they can be effectively targeted through nutritional or microbial interventions.

Acknowledgments

Acknowledgment to the Laboratory of Microbial Eco-evolutionary Genomics for the infrastructure and technical assistance essential to the development of this study. The authors also express their sincere appreciation to the Postgraduate program in agricultural microbiology at Universidade Federal de Viçosa for its academic and institutional support.

Authors' contributions

Conceptualization, N.S.M and M.F.S; bioinformatic analysis, N.S.M.; writing, N. S. M; reviewing and editing, M.F.S and J.L.G.

Funding

This research was financially supported by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Data availability

For the remaining data that may be relevant, the corresponding authors can be contacted.

Declarations

Ethics approval and consent to participate

Not applicable. This study involved the reanalysis of previously published, publicly available microbiome sequencing data and associated metadata. No new animal experiments were conducted by the authors. Ethical approval and animal welfare procedures for the original data collection were the responsibility of the respective primary studies, as referenced in Table S1.

Competing interests

The authors declare no competing interests.

Author details

*Corresponding author, mateus.santana@ufv.br. [1]Laboratory of Microbial Eco-Evolutionary Genomics. [2]Department of Microbiology, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil. [3]Department of Animal Science, Universidade Federal de Viçosa, Viçosa, MG, 36570-900, Brazil.

References

1. Wang X, Tsai T, Deng F, Wei X, Chai J, Knapp J, et al. Longitudinal investigation of the pig gut microbiome from birth to market reveals stage and growth performance associated bacteria. Microbiome. 2019; 7:109. https://doi.org/10.1186/s40168-019-0721-7

2. Arsenault M, Lillie B, Nadeem K, Khafipour E, Farzan A. Progression of pig fecal microbiota during early stages of life and its association with performance: a longitudinal study. BMC Microbiology. 2024; 24:182. https://doi.org/10.1186/s12866-024-03336-y

3. Liu Y, Chen C, Gao J. Topological characteristics and longitudinal dynamics of co-abundance networks involving beneficial commensal bacteria in the pig gut microbiome and its association with average daily gain. Frontiers Microbiology. 2026;

733 17:1818141. https://doi.org/10.3389/fmicb.2026.1818141
734 4. Wu W, Wang S, Xue C, Dong N. Role of gut microbiota metabolites in
735 maintaining intestinal health and preventing weaning-associated diarrhea in piglets.
736 Animal Nutrition. 2026; 25:46–61. https://doi.org/10.1016/j.aninu.2025.10.004
737 5. Wei G, Liu M, Huang L, Chen C. Metagenomic sequencing reveals the dynamic
738 changes of pig gut fungal composition following the ages and identifies fungal species
739 associated with diarrhea in piglets. Animal Microbiome. 2026; 8:60.
740 https://doi.org/10.1186/s42523-026-00551-y
741 6. Fukami T. Historical Contingency in Community Assembly: Integrating Niches,
742 Species Pools, and Priority Effects. Annual Review of Ecology, Evolution, and
743 Systematics. 2015; 46:1–23. https://doi.org/10.1146/annurev-ecolsys-110411-160340
744 7. Woelfel S, Silva MS, Stecher B. Intestinal colonization resistance in the context
745 of environmental, host, and microbial determinants. Cell Host & Microbe. 2024; 32:820–
746 36. https://doi.org/10.1016/j.chom.2024.05.002
747 8. Ning D, Wang Y, Fan Y, Wang J, Van Nostrand JD, Wu L, et al. Environmental
748 stress mediates groundwater microbial community assembly. Nature Microbiology.
749 2024; 9:490–501. https://doi.org/10.1038/s41564-023-01573-x
750 9. Coyte KZ, Schluter J, Foster KR. The ecology of the microbiome: Networks,
751 competition, and stability. Science. 2015; 350:663–6.
752 https://doi.org/10.1126/science.aad2602
753 10. Hervías-Parejo S, Cuevas-Blanco M, Lacasa L, Traveset A, Donoso I, Heleno R,
754 et al. On the structure of species-function participation in multilayer ecological networks.
755 Nature Communications. 2024; 15:8910. https://doi.org/10.1038/s41467-024-53001-1
756 11. Moya A, Ferrer M. Functional Redundancy-Induced Stability of Gut Microbiota
757 Subjected to Disturbance. Trends in Microbiology. 2016; 24:402–13.
758 https://doi.org/10.1016/j.tim.2016.02.002
759 12. Cross K, Beckman N, Jahnes B, Sabree ZL. Microbiome metabolic capacity is
760 buffered against phylotype losses by functional redundancy. Rudi K, organizador.
761 Applied and Environmental Microbiology. 2025; 91:e02368-24.
762 https://doi.org/10.1128/aem.02368-24
763 13. The Human Microbiome Project Consortium. Structure, function and diversity of
764 the healthy human microbiome. Nature. 2012; 486:207–14.
765 https://doi.org/10.1038/nature11234

14. Tremaroli V, Bäckhed F. Functional interactions between the gut microbiota and host metabolism. Nature. 2012; 489:242–9. https://doi.org/10.1038/nature11552

15. Louca S, Polz MF, Mazel F, Albright MBN, Huber JA, O'Connor MI, et al. Function and functional redundancy in microbial systems. Nature Ecology and Evolution. 2018; 2:936–43. https://doi.org/10.1038/s41559-018-0519-1

16. Chemla Y, Sweeney CJ, Wozniak CA, Voigt CA. Design and regulation of engineered bacteria for environmental release. Nature Microbiology. 2025; 10:281–300. https://doi.org/10.1038/s41564-024-01918-0

17. Xiao L, Estellé J, Kiilerich P, Ramayo-Caldas Y, Xia Z, Feng Q, et al. A reference gene catalogue of the pig gut microbiome. Nature Microbiology. 2016; 1:16161. https://doi.org/10.1038/nmicrobiol.2016.161

18. Chen S. fastp 1.0: An ultra-fast all-round tool for FASTQ data quality control and preprocessing. Imeta. 2025; 4:e70078. https://doi.org/10.1002/imt2.70078

19. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016; 32:3047–8. https://doi.org/10.1093/bioinformatics/btw354

20. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. Nature methods. 2016; 13:581–3. https://doi.org/10.1038/nmeth.3869

21. Chuvochina M, Gerken J, Frentrup M, Sandikci Y, Goldmann R, Freese HM, et al. SILVA in 2026: a global core biodata resource for rRNA within the DSMZ digital diversity. Nucleic Acids Research. 2026; 54:D334–41. https://doi.org/10.1093/nar/gkaf1247

22. Abarenkov K, Nilsson RH, Larsson K-H, Taylor AFS, May TW, Frøslev TG, et al. The UNITE database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. Nucleic Acids Research. 2024; 52:D791–7. https://doi.org/10.1093/nar/gkad1039

23. Blanco-Míguez A, Beghini F, Cumbo F, McIver LJ, Thompson KN, Zolfo M, et al. Extending and improving metagenomic taxonomic profiling with uncharacterized species using MetaPhlAn 4. Nature Biotechnology. 2023; 41:1633–44. https://doi.org/10.1038/s41587-023-01688-w

24. Weiss S, Xu ZZ, Peddada S, Amir A, Bittinger K, Gonzalez A, et al. Normalization and microbial differential abundance strategies depend upon data characteristics. Microbiome. 2017; 5:27. https://doi.org/10.1186/s40168-017-0237-y

800 25. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et
801 al. Reproducible, interactive, scalable and extensible microbiome data science using
802 QIIME 2. Nature Biotechnology. 2019; 37:852–7. https://doi.org/10.1038/s41587-019-
803 0209-9
804 26. McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive
805 Analysis and Graphics of Microbiome Census Data. PLOS ONE. 2013; 8:e61217.
806 https://doi.org/10.1371/journal.pone.0061217
807 27. R: The R Project for Statistical Computing. https://www.r-project.org/
808 28. Anderson MJ. A new method for non-parametric multivariate analysis of
809 variance. Austral Ecology. 2001; 26:32–46. https://doi.org/10.1111/j.1442-
810 9993.2001.01070.pp.x
811 29. Lin H, Peddada SD. Multigroup analysis of compositions of microbiomes with
812 covariate adjustments and repeated measures. Nature methods. 2024; 21:83–91.
813 https://doi.org/10.1038/s41592-023-02092-7
814 30. Nickols WA, Kuntz T, Shen J, Maharjan S, Mallick H, Franzosa EA, et al.
815 MaAsLin 3: refining and extending generalized multivariable linear models for meta-
816 omic association discovery. Nature Methods. 2026; 23:554–64.
817 https://doi.org/10.1038/s41592-025-02923-9
818 31. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al.
819 Microbiome differential abundance methods produce different results across 38 datasets.
820 Nature Communications. 2022; 13:342. https://doi.org/10.1038/s41467-022-28034-z
821 32. Risely A. Applying the core microbiome to understand host–microbe systems.
822 Tate A, organizador. Journal of Animal Ecology. 2020; 89:1549–58.
823 https://doi.org/10.1111/1365-2656.13229
824 33. Peschel S, Müller CL, Von Mutius E, Boulesteix A-L, Depner M. NetCoMi:
825 network construction and comparison for microbiome data in R. Briefings in
826 Bioinformatics. 2021; 22:bbaa290. https://doi.org/10.1093/bib/bbaa290
827 34. Kurtz ZD, Müller CL, Miraldi ER, Littman DR, Blaser MJ, Bonneau RA. Sparse
828 and Compositionally Robust Inference of Microbial Ecological Networks. PLOS
829 Computational Biology. 2015; 11:e1004226.
830 https://doi.org/10.1371/journal.pcbi.1004226
831 35. Yoon G, Gaynanova I, Müller CL. Microbial Networks in SPRING - Semi-
832 parametric Rank-Based Correlation and Partial Correlation Estimation for Quantitative

833 Microbiome Data. Frontiers in Genetics. 2019; 10:516.
834 https://doi.org/10.3389/fgene.2019.00516
835 36. Cantalapiedra CP, Hernández-Plaza A, Letunic I, Bork P, Huerta-Cepas J.
836 eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain
837 Prediction at the Metagenomic Scale. Molecular Biology Evolution. 2021; 38:5825–9.
838 https://doi.org/10.1093/molbev/msab293
839 37. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG
840 for taxonomy-based analysis of pathways and genomes. Nucleic Acids Research. 2023;
841 51:D587–92. https://doi.org/10.1093/nar/gkac963
842 38. Xu S, Hu E, Cai Y, Xie Z, Luo X, Zhan L, et al. Using clusterProfiler to
843 characterize multiomics data. Nature Protocols. 2024; 19:3292–320.
844 https://doi.org/10.1038/s41596-024-01020-z
845 39. Tang X, Xiong K, Fang R, Li M. Weaning stress and intestinal health of piglets:
846 A review. Frontiers in Immunology. 2022; 13.
847 https://doi.org/10.3389/fimmu.2022.1042778
848 40. St-Pierre B, Perez Palencia JY, Samuel RS. Impact of Early Weaning on
849 Development of the Pig Gut Microbiome. Microorganisms. 2023; 11:1753.
850 https://doi.org/10.3390/microorganisms11071753
851 41. Van De Vliet M, Joossens M. The Resemblance between Bacterial Gut
852 Colonization in Pigs and Humans. Microorganisms. 2022; 10:1831.
853 https://doi.org/10.3390/microorganisms10091831
854 42. Lv Y, Zhou Y, Lu H, Dong H, He Z. Determinants of piglet gut microbiome
855 colonization: roles of genetics, nutrition, therapeutics, and the impact of enteric pathogens
856 like PEDV. Frontiers in Cellular and Infection Microbiology. 2025; 15.
857 https://doi.org/10.3389/fcimb.2025.1626239
858 43. Saladrigas-García M, D'Angelo M, Ko HL, Nolis P, Ramayo-Caldas Y, Folch
859 JM, et al. Understanding host-microbiota interactions in the commercial piglet around
860 weaning. Scientific Reports. 2021; 11:23488. https://doi.org/10.1038/s41598-021-02754-
861 6
862 44. Li Y, Guo Y, Wen Z, Jiang X, Ma X, Han X. Weaning Stress Perturbs Gut
863 Microbiome and Its Metabolic Profile in Piglets. Scientific Reports. 2018; 8:18068.
864 https://doi.org/10.1038/s41598-018-33649-8
865 45. Mancin E, Maltecca C, Huang YJ, Mantovani R, Tiezzi F. A first characterization
866 of the microbiota-resilience link in pig. Microbiome. 2024; 12:53.
867 https://doi.org/10.1186/s40168-024-01771-7

868 46. Tian L, Wang X-W, Wu A-K, Fan Y, Friedman J, Dahlin A, et al. Deciphering
869 functional redundancy in the human microbiome. Nature Communications. 2020;
870 11:6217. https://doi.org/10.1038/s41467-020-19940-1
871 47. Fässler D, Heinken A, Hertel J. Characterising functional redundancy in
872 microbiome communities via relative entropy. Computational and Structural
873 Biotechnology Journal. 2025; 27:1482–97. https://doi.org/10.1016/j.csbj.2025.03.012
874 48. Vourlaki I-T, Rio-Lopez R, Clavell-Sansalvador A, Ramírez-Ayala LC, Ballester
875 M, Sanchez JP, et al. Co-occurring microbial guilds in pig fecal microbiota: key drivers
876 and effects on host performance. Genetics Selection Evolution. 2025; 57:27.
877 https://doi.org/10.1186/s12711-025-00979-x
878 49. Bergamaschi M, Tiezzi F, Howard J, Huang YJ, Gray KA, Schillebeeckx C, et al.
879 Gut microbiome composition differences among breeds impact feed efficiency in pig.
880 Microbiome. 2020; 8:110. https://doi.org/10.1186/s40168-020-00888-9
881 50. Shetty SA, Kostopoulos I, Geerlings SY, Smidt H, de Vos WM, Belzer C.
882 Dynamic metabolic interactions and trophic roles of human gut microbes identified using
883 a minimal microbiome exhibiting ecological properties. ISME Journal. 2022; 16:2144–
884 59. https://doi.org/10.1038/s41396-022-01255-2
885 51. Culp EJ, Goodman AL. Cross-feeding in the gut microbiome: Ecology and
886 mechanisms. Cell Host Microbe. 2023; 31:485–99.
887 https://doi.org/10.1016/j.chom.2023.03.016
888 52. Beaumont M, Cauquil L, Bertide A, Ahn I, Barilly C, Gil L, et al. Gut Microbiota-
889 Derived Metabolite Signature in Suckling and Weaned Piglets. Journal of Proteome
890 Research. 2021; 20:982–94. https://doi.org/10.1021/acs.jproteome.0c00745
891 53. Hou L, Yang Y, Sun B, Jing Y, Deng W. Dietary Fiber, Gut Microbiota, Short-
892 Chain Fatty Acids, and Host Metabolism. Am. J Life Sci. 2021; 9:162.
893 https://doi.org/10.11648/j.ajls.20210906.12
894 54. Hu J, Chen J, Ma L, Hou Q, Zhang Y, Kong X, et al. Characterizing core
895 microbiota and regulatory functions of the pig gut microbiome. The ISME Journal. 2024;
896 18:wrad037. https://doi.org/10.1093/ismejo/wrad037

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInformation.docx