# Copy number variation on *ABCC2-DNMBP loci* impacts the diversity and composition of the gut microbiota in pigs

Yuliaxis Ramayo-Caldas[1+*], Daniel Crespo-Piazuelo[1+], Jordi Morata[2], Olga González-Rodríguez[1], Cristina Sebastià[3,4], Anna Castello[3,4], Antoni Dalmau[5], Sebastian Ramos-Onsins[3], Konstantinos G. Alexiou[3], Josep M. Folch[3,4], Raquel Quintanilla[1], Maria Ballester[1*]

[1] Animal Breeding and Genetics Program, Institute of Agrifood Research and Technology (IRTA), Torre Marimon, 08140, Caldes de Montbui, Spain

[2] CNAG-CRG, Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Baldiri i Reixac 4, 08028 Barcelona, Spain

[3] Plant and Animal Genomics Program, Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Campus UAB, Bellaterra 08193, Spain

[4]Animal and Food Science Department, University Autonomous of Barcelona (UAB), Bellaterra 08193, Spain

[5] Animal Welfare Program, IRTA, 17121, Monells, Girona, Spain

[+] equal contribution

[*] corresponding authors

## Abstract

### Background

Genetic variation in the pig genome partially modulates the composition of porcine gut microbial communities. Previous studies have been focused on the association between single nucleotide polymorphisms (SNPs) and the gut microbiota, but little is known about the relationship between structural variants and gut microbial traits.

### Results

The main goal of this study was to assess the effect of porcine genome copy number variants (CNVs) on the diversity and composition of pig gut microbiota. For this purpose, we used whole-genome sequencing data to undertake a comprehensive identification of CNVs followed by a genome-wide association analysis between the estimated CNV status and the gut bacterial diversity in a commercial Duroc pig population. A CNV predicted as gain (DUP) partially harboring *ABCC2-DNMBP loci* was associated with richness ($p$-value=$5.41\times10^{-5}$) and Shannon α-diversity ($p$-value=$1.42\times10^{-4}$). The *in-silico* predicted gain of copies was validated by real-time quantitative PCR (qPCR), and its segregation, and positive association with the richness and Shannon α-diversity of the porcine gut bacterial ecosystem was confirmed in an unrelated F1 (Duroc×Iberian) cross. Furthermore, despite genetic and environmental differences between both populations, the gut microbiota of DUP samples showed a significant over-abundance of the *Desulfovibrio, Blautia, Phascolarctobacterium, Faecalibacterium, Succinivibrio* and *Anaerovibrio* genera.

### Conclusions

In summary, this is the first study that evaluate the putative modulatory role of CNVs on pig gut microbiota. Our results advice the relevance of considering the role of host-genome structural variants as modulators of microbial ecosystems, and suggest the *ABCC2-DNMBP* CNV as a host-genetic factor for the modulation of the diversity and composition of the gut microbiota in pigs.

**Key words:** microbiota, diversity, structural variants, CNVs, modulators, porcine

## BACKGROUND

Gut microbiomes have a profound impact on many aspects of pig health, such as the modulation of metabolic functions, physiological processes, and relevant porcine traits like growth [1] , feed efficiency [2] [3] , and immunocompetence [4]. Host-microbiome interactions are mediated by both environmental and host factors. Among them, genetic variation in the pig genome can modulate, in a taxa-specific manner, the composition and function of the pig gut eukaryotic and prokaryotic communities. Pig gut microbiota is heritable to an extent, showing low to medium heritabilities [2] [5] [6]. Quantitative trait loci (QTLs), genetic variants, and candidate genes associated to pig gut microbiota have been reported [7] [8] [9] [10].

However, since previous studies were focused on the association between SNPs and microbial traits, little is known about the relationship between the gut microbiota and structural variants in the porcine genome. Copy-number variants (CNVs) are structural variants that produce a change in the number of copies (gain or loss) of a genomic region. Compared to SNPs, CNVs involve large DNA segments that span a significant proportion of the genome, and account for greater genomic variability than SNPs. Consequently, CNVs are a relevant source of genetic variation that contribute to evolutionary adaptations, variation in gene expression and phenotypic traits in human and domestic animals [11] [12]. In humans, gain of copies of salivary amylase (*AMY1*) gene was associated with oral and gut microbiome composition [13]. In this seminal study, Poole et al. found that individuals with greater number of copies of *AMY1* showed greater levels of salivary *Porphyromonas*, followed by an increased abundance of resistant starch-degrading microbes in the gut.

We hypothesized that alike humans, CNVs are likely to contribute to animal gut microbial variability. However, to the best of our knowledge, such associations have not been documented in livestock. Consequently, the putative modulatory role of CNVs on the diversity, composition and function of livestock gastrointestinal microbiota remains to be elucidated. The main goal of this study was to assess the effect of porcine CNVs on the diversity and composition of pig gut microbiota.

## MATERIAL AND METHODS

### Animal samples

Samples employed in this study are a subset of pigs reported in [14] and [9]. In brief, a total of 100 weaned piglets (50 males and 50 females) from a commercial Duroc pig line were used as a discovery dataset (Table 1). The pigs were distributed in three batches, all animals were raised on the same farm and had *ad libitum* access to the same commercial cereal-based diet. Furthermore, a subset of 24 unrelated F1 (Duroc×Iberian) crossbred pigs with phenotypically extreme gut microbial diversity index (12 high and 12 low) from [7] were employed as independent validation dataset (Table 1).

### Microbial DNA extraction, sequencing, and bioinformatics analysis

Fecal samples were collected from the Duroc piglets at 60 ± 8 days of age and microbial DNA was extracted with the DNeasy PowerSoil Kit (QIAGEN, Hilden, Germany). Extracted DNA was sent to the University of Illinois Keck Center for paired-end (2 × 250 bp) sequencing on an Illumina NovaSeq (Illumina, San Diego, CA, USA). The 16S rRNA gene fragment was amplified using the primers V3_F357_N: 5'-CCTACGGGNGGCWGCAG-3' and V4_R805: 5'-GACTACHVGGGTATCTAATCC-3'. Sequences were analysed with QIIME2 [15]; barcode sequences, primers, and low-quality reads (Phred score < 30) were removed. The quality control process also trimmed sequences based on expected amplicon length and removed chimeras. Afterwards, sequences were clustered into Amplicon Sequences Variants (ASVs) at 99% of identity. ASVs were classified to the lowest possible taxonomic level based on a primer-specific trained version of GreenGenes Database 13.8 [16]. Before the estimation of the diversity indices, to correct for the sequencing depth, samples were rarefied at 10,000 reads. Diversity metrics were estimated with the vegan R package v2.6-2 [17]. The α-diversity was evaluated with the Shannon index [18], and the β-diversity was assessed using the Whittaker index [19].

### Host-genome data analysis and CNV-calling

Simultaneously with fecal sampling, blood was collected at 60 ± 8 days of age via the external jugular vein. Host genomic DNA was extracted from blood using the NucleoSpin Blood (Macherey–Nagel). Whole genome was paired-end sequenced (2 × 150

bp) in an Illumina NovaSeq6000 platform (Illumina) at *Centro Nacional de Análisis Genómico* (CNAG-CRG; Barcelona, Spain). Reads were mapped to the porcine reference assembly Sscrofa.11.1 with BWA-MEM 0.7.17 [20]. Alignment files containing only properly paired, uniquely mapping reads without duplicates were processed using Picard [ http://broadinstitute.github.io/picard/ ] to add read groups and to remove duplicates. Variant calling was performed with the HaplotypeCaller tool from Genome Analysis Tool Kit (GATK 4.1.8.0) [21]. Applying GATK Best Practices, variants with minimum read depth of 5 on at least one sample were retained. Joint genotyping was conducted with combined gVCFs. Functional annotations were added using SnpEff v.5 [22] against the Sscrofa.11.1 reference database. CNV prediction was performed with ControlFREEC 11.5 [23], using a pool of samples as CNV baseline and using intervals of 20kb windows. CNV calls from all samples that were less than 10kb apart were merged with Survivor [24]. Individual CNV-calls were combined into copy number variant regions (CNVR) following the reciprocal overlap approach [11] with CNVRange [25]. Therefore, contiguous CNVs intervals with at least 50% of mutual overlap were merged into the same CNVR.

Nucleotide diversity pattern estimates were calculated considering the SNPs present in each individual separately. Here we tested two estimators: (i) Tajima's theta estimator (π or nucleotide diversity, called ***ATajima*** hereafter), that is simply the number of variants present in the individual; (ii) ***RTajima*** estimator, which considers the frequency observed in the entire sample but calculated given the SNPs present in each individual. This estimate is equivalent to ***ATajima*** but needs to be corrected by the probability that only a portion of the total SNPs from a sample are present in each of the samples (by using a hypergeometrical distribution). Specifically, the calculation for a single individual is:

$$\hat{\theta} = \frac{1}{\sum_{i=1}^{n-1}(n-i)} \sum_{i=1}^{n-1} (n-i) i \xi_i \left(1 - \frac{\binom{n-i}{2}}{\binom{n}{2}}\right)^{-1}$$

where $n$ is the number of samples in the population (2×number of individuals) and $\xi_i$ is the number of SNPs observed in this individual that are at frequency $i$ in the whole population.

The dataset in VCF format was converted to FASTA and from FASTA to transposed FASTA (tFASTA). This tFASTA file was read with the software *mstatspop* (https://github.com/CRAGENOMICA/mstatspop) to obtain the frequencies of SNPs from each of the pigs at the desired region. Finally, we calculated ***RTajima*** per fragment using

self-made R scripts. All these estimates were finally divided by the effective length size of the studied region to obtain comparative estimates per nucleotide.

**CNV-wide association analysis**

A genome-wide association analysis between the estimated CNV-status and gut bacterial diversity index was done using the following mixed model:

$$y_{ijk} = sex_j + b_k + u_i + cnv_{li} + e_{ijk}$$

where $y_{ijk}$ corresponds to the microbial index under scrutiny (richness or Shannon α-diversity) of the i-th individual animal of sex j in the k-th batch; $sex_j$ and $b_k$ correspond to the systematic effects of j-th sex (2 levels) and k-th batch (3 levels), respectively; $u_i$ is the random additive genetic effect of the i-th individual, collectively distributed as $\boldsymbol{u} \sim N(0, \mathbf{G}\sigma_u^2)$ where $\sigma_u^2$ is the additive genetic variance and **G** is the numerator of the genomic relationship matrix calculated using the autosomal SNPs; $cnv_{li}$ is the genotype (recoded as 11=loss, 12=diploid and 22=gain) for the l-th CNV of the i-th individual, and $e_{ijk}$ is the residual.

**Quantitative real time PCR**

Real-time quantitative PCR (qPCR) was used to validate the CNV on the *ABCC2* gene in a total of 72 samples including a subset of 48 Duroc samples (24 diploid and 24 *in silico* predicted as DUP), and 24 unrelated F1 Duroc×Iberian cross (Table 1). The CNV breakpoint was re-estimated with Manta: v1.6.0 [26]. All primers were designed using the Primer Express 2.0 software (Applied Biosystems). The pair of primers ABCC2_CNV_F 5'-*TGGCATCATTTATGTGGCTGTT-3'* and ABCC2_CNV_R 5'-*AGGAAGGAGCTTGGGCTTTTA-3'* amplify a specific region of the *ABCC2* gene containing the CNV, while the pair of primers ABCC2_F 5'-*TGGACAAGAACCAGAGTCAAAGC*-3' and ABCC2_R 5'-*ACATAGAGCGCATTTGAACGAA*-3' amplify a region outside of the estimated CNV breakpoint that was used as single copy control region. The 2-ΔΔCt method for relative quantification (RQ) of CNVs was used as previously described in [27]. qPCRs were carried out using SYBR Green chemistry (SYBRTM Select Master Mix, Applied Biosystems) and the instruments ABI PRISM® 7900HT and 7500 Real-Time PCR System (Applied Biosystems, Inc.; Foster City, CA). The reactions were carried out in a 20µl volume containing 10ng of genomic DNA. All primers were used at 900 nM. The thermal cycle was: 10 min at 95°C, 40 cycles of 15 sec at 95°C and 1 min at 60°C. Each sample was

analyzed in quadrupled. PCR efficiencies (<95%) were evaluated with standard curves and dissociation curves were drawn for each primer pair to assess for the specificity of the PCR reactions. Three samples without CNV were used as reference. Results were analyzed with Thermo Fisher Cloud software 1.0 (Applied Biosystems), and qBase Plus v3.2 (Biogazelle).

**Identification of microbial signatures**

The identification of ASVs that discriminate samples according to the number of copies of the *ABCC2-DNMBP loci* was performed based on the compositional kernel as implemented the function '*classify*' of *kernInt* R package [28]. The '*classify'* function run a supervised classification model based on Support Vector Machine. For that purpose, the available dataset was split at random into training set (80% of data) and validation set (20%). The *C* hyperparameter's optimal value was obtained by 10×10 cross-validation on the training set. To estimate the mean classification accuracy the 'classify' function was run ten times using different training/test splits of the dataset. Microbial signatures were obtained from the hyperplane vector w, and the importance of the ASV k was computed by *kernInt* as $(wk)^2$ [29]. Initially, the top 5% relevant taxa were retained, but a conservative approach was applied afterwards, keeping for subsequent analyses only the ASVs reported as relevant in at least the 50% of the replicates. Finally, to identify overrepresentation at genus level, the list of selected features was submitted to a taxa-set enrichment analysis [30].

## RESULTS

**Detection of copy number variants and association analysis**

In this study we used whole-genome sequencing data from 100 healthy 60-day-old Duroc pigs to undertake a comprehensive identification of CNVs. A total of 1,292 CNVs distributed across 531 CNVR on autosomal pig chromosomes were identified (Figure 1, Supplementary table 1). After quality control, 1,005 CNVs grouped into 291 CNVR, presented in at least the 5% of the samples, were used for the association analysis. Among them, a CNV predicted as gain (DUP) located on CNVR454 (SSC14:111000000-111075999) that partially contain the ATP Binding Cassette Subfamily C Member 2 (*ABCC2*) and the Dynamin Binding Protein (*DNMBP*) genes showed a significant association with richness (p-value= $5.41\times10^{-5}$) and the Shannon α-diversity (p-value=$1.42\times10^{-4}$) (Figure 2).

**Validation by quantitative PCR**

Since the *in silico* identification of CNVs may result in both false positive and negative results [27] [31], we conducted a qPCR assay with primers located on the *ABCC2* gene for the experimental validation of the CNV. The *in silico* predicted genotype was confirmed by qPCR in 46 out of the 48 Duroc samples, corresponding to an accuracy of 95.83%. To be noted, all the 24 samples *in silico* predicted as DUP presented the gain in number of copies. Thus, deviations from the diploid status were observed in two out of the 24 animals, where variation in the number of copies was not *in silico* predicted by ControlFREEC [23].

Because the existence of false negative sample assignation of CNV status could impact the GWAS results, and therefore, to avoid spurious associations and confirm our findings, we repeated the diversity index comparison using the subset of samples analyzed by qPCR (2N=22 vs DUP=26). In agreement with the CNV-GWAS, qPCR results corroborated that DUP pigs significantly had greater richness ($p=1.8\times10^{-3}$) and α-diversity values ($p=3.8\times10^{-3}$) (Table 1). Furthermore, RQ of the number of copies was positively correlated with the richness ($r=0.474$, p-value$=6.72\times10^{-4}$) and the α-diversity ($r=0.401$, p-value$=7.77\times10^{-3}$) (Figure 3). We also observed a positive relationship between the RQ of the number of copies and the nucleotide variability of the CNV genomic interval estimators ***ATajima*** (r=0.43) and ***RTajima*** (r=0.71). The high correlation observed between ***RTajima*** and the RQ values (Supplementary Fig 1) can be explained by the characteristic of RTajima statistic, which gives more importance to intermediate frequencies observed in the whole population, indicating that intermediate frequency diversity in this CNV is over-represented at greater number of copies.

Remarkably, variation in the number of copies of the *ABCC2-DNMBP loci* was also segregating in an unrelated commercial F1 Duroc×Iberian crossbred pigs, with 13 out of 24 pigs showing a gain of copies. Moreover, despite differences on genetic background, age, or other environmental factors such as diet of farm of origin, the association between the *ABCC2-DNMBP loci* and the gut microbial diversity was replicated in the F1 Duroc×Iberian cross (Table 1, Figure 4). Indeed, in both Duroc and F1 Duroc×Iberian cross datasets, the qPCR reaffirmed that a gain of copies of the *ABCC2-DNMBP loci* was positively associated to the richness and Shannon α-diversity of the pig gut microbiota (Table 1).

## Microbial signatures linked to variation in the number of copies

Results from the supervised classification model showed that the relative abundance of 122 ASVs allowed the classification between groups of DUP vs 2N samples (Supplementary table 2). The taxa-set enrichment analysis pointed out a higher overall discriminant importance of ASVs members of the *Desulfovibrio*, *Blautia*, *Phascolarctobacterium*, *Fibrobacter, Roseburia, Faecalibacterium, Megasphaera, Succinivibrio, Coprococcus, RFN20,* and *Anaerovibrio* genera (Figure 5A). Furthermore, supporting their discriminative role, we observed that compared to their diploid counterparts, the gut microbiota of DUP pigs exhibited a higher relative abundance (FDR<0.05) of the *Desulfovibrio, Blautia, Phascolarctobacterium, Faecalibacterium, Megasphaera, Succinivibrio* and *Anaerovibrio* genera, but lower relative abundance of the *Fibrobacter* and *RFN20* genera (Figure 5B). To be noted, the results obtained from the differential abundance analysis done on the unrelated F1 Duroc×Iberian crossbred population confirmed a higher relative abundance of the *Desulfovibrio, Blautia, Phascolarctobacterium, Faecalibacterium, Succinivibrio* and *Anaerovibrio* genera in the gut microbiota of samples with a gain of copies of the *ABCC2-DNMBP loci* (Figure 6).

## DISCUSSION

In this study we report, for the first time in a livestock species, a CNV partially containing the *ABCC2* and *DNMBP* genes associated to the diversity and composition of the pig gut microbiota. *ABCC2* encodes a multidrug resistance-associated protein 2 (MRP2) that plays a relevant role preserving hepatic and intestinal homeostasis [32]. ABCC2 is involved in the excretion of conjugated bile acids (BAs), bilirubin, xenobiotics, and the transport of other organic anions [33][34]. In pigs, *ABCC2* has reported as co-associated to the intramuscular profile of fatty acid composition in an Iberian×Landrace cross [35]. In addition, the genomic interval harboring the *ABCC2-DNMBP loci* overlapped with QTLs associated to muscle profile of palmitic (QTLId: 95385), stearic (QTLId: 95386) and palmitoleic (QTLId: 95387) fatty acids content in a Duroc×(Landrace×Yorkshire) cross [36]. In other species such as mice, rat or humans, mutations in *ABCC2* are related to hereditary liver diseases. *Mrp2*$^{-/-}$ mice are viable [37] [38], but like *ABCC2*-knockout rats, showed chronic hyperbilirubinemia followed by a reduction in biliary excretion of bilirubin glucuronides [39] [37]. Meanwhile, mutations in the human *ABCC2* gene results in the Dubin-Johnson syndrome, an autosomal recessive disorder characterized by a defect in the transport of endogenous and exogenous anionic conjugates from hepatocytes into the bile [40]. It is worth to highlight that a genomic duplication of 5,299 base pairs comprising

exons 24 and 25 of human *ABCC2* gene was predicted to result in the insertion of a premature stop codon [41].

Considering the key role of *ABCC2* on the excretion of bilirubin and conjugated BAs, we hypothesize that variation in the number of copies of *ABCC2* may influence gut levels of conjugated BAs and/or bilirubin. A bidirectional crosstalk between gut microbiota and the metabolism of bilirubin and conjugated BAs has been documented. Bilirubin can regulate the composition of gut microbiota by being potentially toxic towards Gram-positive bacteria, while promoting the proliferation of Gram-negative species [42]. In a similar way, a higher BA tolerance is evidenced by Gram-negative bacteria [43]. In agreement with these studies, the gut microbiota of DUP samples in both the discovery and validation datasets showed a higher relative abundance of Gram-negative bacteria, such as members of the *Desulfovibrio*, *Phascolarctobacterium, Faecalibacterium, Succinivibrio,* and *Anaerovibrio* genera. On another note, gut microbiota composition can regulate BA and bilirubin production and signaling. Evidence from germ-free (GF) rats reveals that gut microbiota is a key player in the reduction of bilirubin to urobilinoids with significant lower fecal urobilin levels in GF rats compared with conventional ones [44]. Regarding BAs, GF mice showed significant differences on enterohepatic circulation and BA composition compared with conventional mice [45] with a lower proportion of secondary and tertiary unconjugated and glycine-conjugated BA in tissues of GF rats [46]. In addition, conjugated BAs can have a protective role on gut barrier integrity [47]. The oral administration of two major conjugated BAs, tauro-cholic acid and β-tauro-murocholic acid, increased the richness of neonatal small intestinal microbiota with a positive effect on the postnatal microbiota maturation [48]. To be noted, among the top discriminant ASVs we observed butyrate producer species with a potential beneficial effect for the host, such as *Blautia obeum* (*ASV2433, ASV2171, ASV2278*), *Faecalibacterium prausnitzii* (*ASV2371*, *ASV2378*), *Butyricicoccus pullicaecorum* (*ASV2567*) and *Roseburia faecis (ASV1822).* Interestingly, the genome of all these species encodes bile salt hydrolases (*BSH*, EC 3.5.1.24) [49] [50], enzymes that mediate the primary BA deconjugation and successive conversion to secondary BAs. Therefore, partly determining the amount of secondary BAs in the colonic epithelium, which in turn acts as signaling molecules mediating different metabolic processes interconnected with health and diseases [51] [52].

The CNVR454 also included *DNMBP,* a gene that regulates the structure of apical junctions through F-actin organization in epithelial cells [53]. *DNMBP* is also involved in

luminal morphogenesis and enterocyte polarization [54] [55], thus potentially contributing to the function and homeostasis of intestinal epithelial barrier (IEB). In fact, the crosstalk between IEB and the gut microbiota is crucial for the maintenance of intestinal homeostasis. For example, enterocytes, which are the most abundant population among intestinal epithelial cells, express a range of pattern recognition receptors for sensing the microbe-associated molecular patterns. Further, enterocyte apex is covered by thousands of microvilli that are vital in colonic wound repair and the transport of molecules and nutrients such as bile salts, electrolytes and vitamins [56] [57] [58] [59]. Interestingly, depletion of microbiota in mice resulted in altered patterns of microvilli formation [60]. Likewise, compared with conventional piglets, germ-free (GM) pigs displayed aberrant intestinal morphology with longer villi and shorter crypts. Meanwhile, the oral administration of commensal bacteria increased crypt depth, and induced enterocyte brush border microvilli enzyme activities on these GM piglets [61] [62] [63] [64]. Therefore, considering the functional roles of *DNMBP* in the IEB, we cannot rule out the contribution of *DNMBP* to the modulation of the diversity and composition of the pig gut microbiota.

Altogether, our results pinpointed a positive association of the variation in the number of copies of the *ABCC2-DNMBP loci* with the richness, α-diversity, and composition of the pig gut microbial ecosystems. Such findings open the possibility to modulate the gut microbiota, which has emerged as a promising breeding or therapeutic tool to optimize livestock production efficiency, animal health and well-being. A greater gut microbial diversity is usually desired, and generally accepted as an indicator of a resilient microbial ecosystem, gut, and host health. Indeed, a diverse and healthy gut has a positive effect on the absorption of dietary nutrients, feed efficiency and animal well-being.

We are aware of some limitations of our study like the limited taxonomic resolution achieved by targeting the V3-V4 16S rRNA genomic region with short-read sequencing. We are also awake about the convenience of performing further analyses to confirm the raised hypotheses by assessing the metabolic profile of BAs as well as evaluating the role of the CNV on gene expression (at both microbial and host-level) of genes involved in BA metabolism. Despite these limitations, our findings contribute to the understanding of host-microbiome interactions. Moreover, our results open the possibility to breed the holobiont via the incorporation of this source of variation on custom-made arrays that can be used in routine genotyping tasks applied to breeding programs, and together with nutritional or

management strategies, will favor the simultaneous improvement of microbial traits, gut health, and host-performance.

## CONCLUSIONS

Here we report the first study exploring associations between porcine CNV and the diversity and composition of the pig gut microbiota. We identified, functionally validated, and replicated in an unrelated population a positive association between the gain of copies of *ABCC2-DNMBP loci* and the composition and diversity of the pig gut microbiota. These results suggest a role of the host-genome structural variants in the modulation of microbial ecosystems, and open the possibility of including CNVs in selection programs to simultaneously improve microbial traits, gut health and host-performance.

**Table 1.** Descriptive statistics, mean and standard deviation (SD), of the richness and α-diversity in the discovery and validation datasets.

| Population | Dataset (n) | Groups | Richness (SD) | α-diversity (SD) |
|---|---|---|---|---|
| Duroc (Discovery) | Complete (100) | | 604.92 (98.17) | 6.05 (0.22) |
| | Validate qPCR (48) | Diploid | 553.73 (92.71) | 5.95 (0.21) |
| | | DUP | 657.85 (89.70) | 6.16 (0.17) |
| Duroc x Iberic (Validation) | | | | |
| | Complete (285) | | 459.65 (104.33) | 5.82 (0.25) |
| | Validate qPCR (24) | Diploid | 292.36 (67.18) | 5.36 (0.27) |
| | | DUP | 543.19 (158) | 6.06 (0.33) |

## Figure legends

**Figure 1**. Graphical representation of the CNVRs detected. Green circles represent loss predicted status, gains are indicated in blue, and regions with either loss or gain status are represented in red. Chromosome sizes are represented in proportion to sequence length of the Sus scrofa 11.1 reference assembly.

**Figure 2**. Results from the association analyses of CNVs identified across the pig genome with gut bacterial richness (A) and Shannon -diversity (B). The x-axis represents the CNV position in the pig autosomal chromosomes (1–18), and the y-axis reflects the significance level represented as the –log10 (p-value).

**Figure 3**. Relationship between the CNV relative quantification of the number of copies (RQ) with the richness and Shannon α-diversity index.

**Figure 4**. Results from the replication analysis comparing mean RQ and Shannon α-diversity in the F1 Duroc×Iberian cross. Green represents diploid (2N) samples, and purples ones DUP samples.

**Figure 5**. Results from microbial signature analyses at Genus level. A) Taxa-set enrichment. B) Patterns of differential abundance analysis between DUP and 2N pigs in the purebred Duroc population.

**Figure 6**. Differential abundance patterns at genus level between DUP and 2N samples in the F1 Duroc×Iberian crossbred population.

**Supplementary Data**

**Table S1.** Description of the 531 Copy Number Variant Regions.

**Figure S1.** Correlation coefficients between CNV relative quantification (RQ) and the nucleotide variability of the CNV genomic interval estimators ATajima and RTajima.

**Table S2.** Taxonomic composition at family level of the 122 discriminant ASVs.

## ABBREVIATIONS

**QIIME2:** Quantitative insights into microbial ecology

**GWAS**: Genome-wide association studies

**SNP:** Single nucleotide polymorphisms

**CNV**: Copy number variants

**CNVR**: Copy number variant region

**DUP:** CNV predicted as gain

**BA:** Bile acids

**ATajima:** Tajima's theta estimator or nucleotide diversity

**RTajima:** Relative Tajima's theta estimator

## DECLARATIONS

### Ethics approval and consent to participate

Animal care and experimental procedures were carried out following the institutional guidelines for the Good Experimental Practices and the Spanish Policy for Animal Protection RD 53/2013, which meets the European Union Directive 2010/63/EU for protection of animals used in experimentation, and were approved by the IRTA Ethical Committee. Consent to participate is not applicable in this study.

### Consent for publication

Not applicable.

### Availability of supporting data

The raw sequencing data employed in this article has been submitted to the NCBI's sequence read archive (https://www.ncbi.nlm.nih.gov/sra); BioProject: PRJNA608629.

### Competing Interests

The authors declare no competing interests

### Funding

The project was funded by the Spanish Ministry of Science and Innovation-State Research Agency (AEI, Spain, 10.13039/501100011033) grant number PID2020-112677RB-C21 to MB, PID2021-126555OB-I00 to YRC, and the GENE-SWitCH project (https://www.gene-

switch.eu) funded by the European Union's Horizon 2020 research and innovation programme under the grant agreement n° 817998 to MB and DCP. SRO is supported by grant PID2020-119255GB-I00 (MICINN,Spain) and by the CERCA Programme/Generalitat de Catalunya. CRAG acknowledges financial support from the Spanish Ministry of Economy and Competitiveness, through the Severo Ochoa Programme for Centers of Excellence in R&D 2016-2019 and 2020-2023 (SEV-2015-0533, CEX2019-000917) and the European Regional Development Fund (ERDF). YRC is recipient of a Ramon y Cajal post-doctoral fellowship (RYC2019-027244-I) funded by the Spanish Ministry of Science and Innovation. CS is funded by AGUAR (2020FI_B 00225). DCP, MB, OGR, RQ, and YRC belonged to a Consolidated Research Group AGAUR, ref. 2017SGR-1719

## Author Contributions

YRC, DCP and MB designed the study. OGP and MB carried out the DNA extractions. YRC, MB, OGP and RQ, performed the sampling. YRC, DCP, MB, JM, CS, SRO, and KGA analyzed the data. YRC, DCP, RQ, JMF, SRO and MB interpreted the results and wrote the manuscript. All the authors read and approved the final version of the manuscript.

## Acknowledgements

The authors warmly thank all technical staff from *Selección Batallé* S.A for providing the animal material and their collaboration during the sampling.

## REFERENCES

1. Ramayo-Caldas Y, Mach N, Lepage P, Levenez F, Denis C, Lemonnier G, et al. Phylogenetic network analysis applied to pig gut microbiota identifies an ecosystem structure linked with growth traits. ISME J [Internet]. 2016;10:2973–7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27177190

2. Camarinha-Silva A, Maushammer M, Wellmann R, Vital M, Preuss S, Bennewitz J. Host Genome Influence on Gut Microbial Composition and Microbial Prediction of Complex Traits in Pigs. Genetics [Internet]. 2017;206:1637–44. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28468904

3. McCormack UM, Curião T, Metzler-Zebeli BU, Magowan E, Berry DP, Reyer H, et al. Porcine Feed Efficiency-Associated Intestinal Microbiota and Physiological Traits: Finding Consistent Cross-Locational Biomarkers for Residual Feed Intake. mSystems [Internet]. 2019;4. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31213524

4. Ramayo-Caldas Y, Zingaretti LM, Pérez-Pascual D, Alexandre PA, Reverter A, Dalmau A, et al. Leveraging host-genetics and gut microbiota to determine immunocompetence in pigs. Anim Microbiome [Internet]. 2021;3:74. Available from: https://doi.org/10.1186/s42523-021-00138-9

5. Bergamaschi M, Maltecca C, Schillebeeckx C, McNulty NP, Schwab C, Shull C, et al. Heritability and genome-wide association of swine gut microbiome features with growth and fatness parameters. Sci Rep [Internet]. 2020;10:10134. Available from: https://doi.org/10.1038/s41598-020-66791-3

6. Aliakbari A, Zemb O, Cauquil L, Barilly C, Billon Y, Gilbert H. Microbiability and microbiome-wide association analyses of feed efficiency and performance traits in pigs. Genetics Selection Evolution [Internet]. 2022;54:29. Available from: https://doi.org/10.1186/s12711-022-00717-7

7. Crespo-Piazuelo D, Migura-Garcia L, Estellé J, Criado-Mesas L, Revilla M, Castelló A, et al. Association between the pig genome and its gut microbiota composition. Sci Rep [Internet]. 2019;9:8791. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31217427

8. Ramayo-Caldas Y, Prenafeta-Boldú F, Zingaretti LM, Gonzalez-Rodriguez O, Dalmau A, Quintanilla R, et al. Gut eukaryotic communities in pigs: diversity, composition and host genetics contribution. Anim Microbiome [Internet]. 2020;2:18. Available from: https://doi.org/10.1186/s42523-020-00038-4

9. Reverter A, Ballester M, Alexandre PA, Mármol-Sánchez E, Dalmau A, Quintanilla R, et al. A gene co-association network regulating gut microbial communities in a Duroc pig population. Microbiome [Internet]. 2021;9:52. Available from: https://www.ncbi.nlm.nih.gov/pubmed/33612109

10. Ryu EP, Davenport ER. Host Genetic Determinants of the Microbiome Across Animals: From Caenorhabditis elegans to Cattle. Annu Rev Anim Biosci [Internet]. 2022;10:203–26. Available from: https://doi.org/10.1146/annurev-animal-020420-032054

11. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, et al. Origins and functional impact of copy number variation in the human genome. Nature [Internet]. 2010;464:704–12. Available from: https://www.ncbi.nlm.nih.gov/pubmed/19812545

12. Clop A, Vidal O, Amills M. Copy number variation in the genomes of domestic animals. Anim Genet [Internet]. 2012;43:503–17. Available from: https://doi.org/10.1111/j.1365-2052.2012.02317.x

13. Poole AC, Goodrich JK, Youngblut ND, Luque GG, Ruaud A, Sutter JL, et al. Human Salivary Amylase Gene Copy Number Impacts Oral and Gut Microbiomes. Cell Host Microbe [Internet]. 2019;25:553-564.e7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30974084

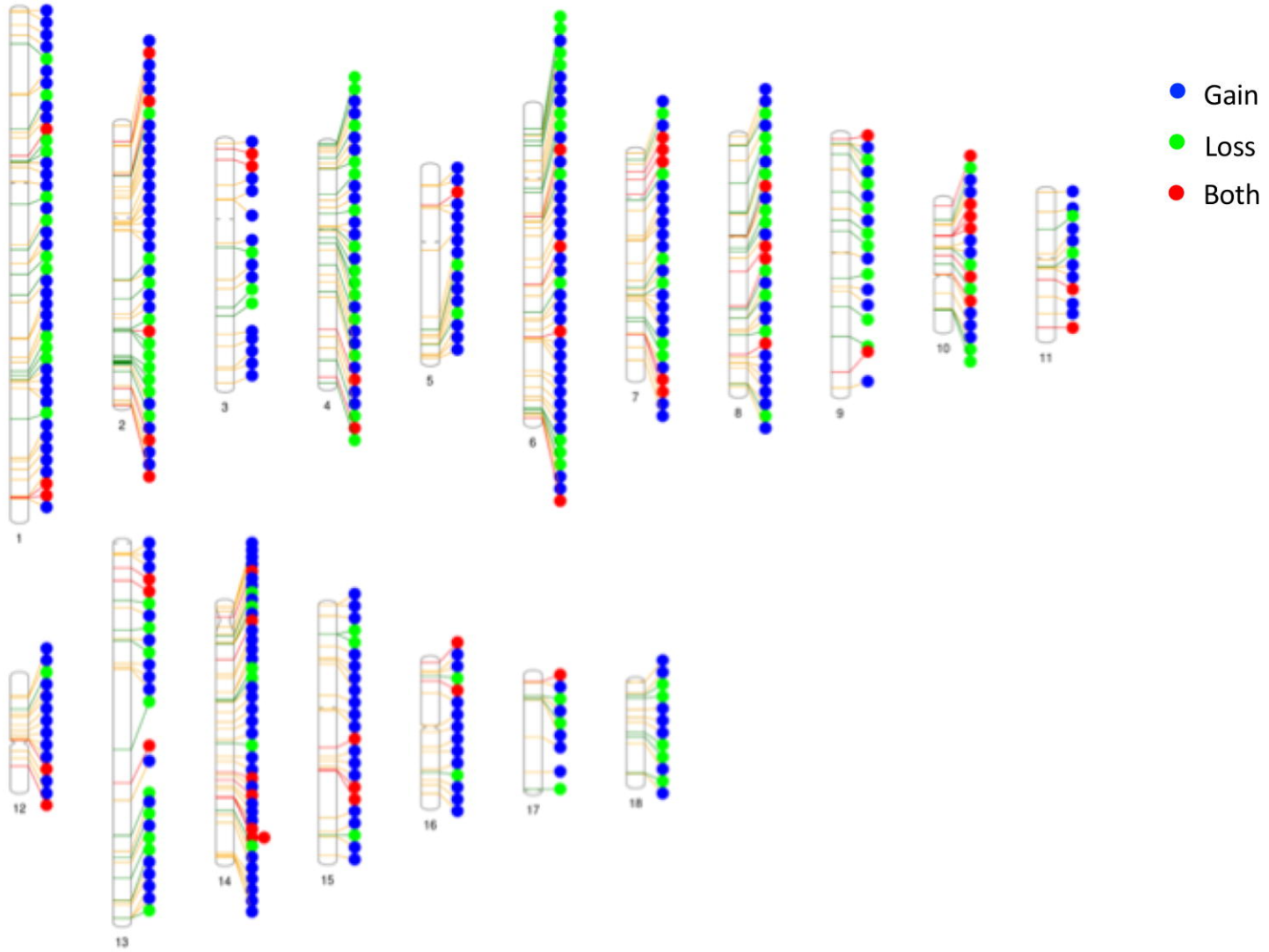
14. Ballester M, Ramayo-Caldas Y, González-Rodríguez O, Pascual M, Reixach J, Díaz M, et al. Genetic parameters and associated genomic regions for global immunocompetence and other health-related traits in pigs. Sci Rep [Internet]. 2020;10:18462. Available from: https://doi.org/10.1038/s41598-020-75417-7

15. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol [Internet]. 2019;37:852–7. Available from: https://www.ncbi.nlm.nih.gov/pubmed/31341288
16. DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. Appl Environ Microbiol [Internet]. 2006;72:5069–72. Available from: https://www.ncbi.nlm.nih.gov/pubmed/16820507
17. Oksanen J, Blanchet G, Kindt R, Legendre P, Minchin P, O'Hara RB, et al. vegan: Community Ecology Package. R package version 2.0-9 [Internet]. 2009. Available from: http://CRAN.R-project.org/package=vegan
18. Shannon C. A mathematical theory of communication. 1948.
19. Whittaker RH. EVOLUTION AND MEASUREMENT OF SPECIES DIVERSITY. Taxon [Internet]. 1972;21:213–51. Available from: https://doi.org/10.2307/1218190
20. H. L. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. 2013.
21. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res [Internet]. 2010;20:1297–303. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20644199
22. Cingolani P, Platts A, Wang le L, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin) [Internet]. 2012;6:80–92. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22728672
23. Boeva V, Popova T, Bleakley K, Chiche P, Cappo J, Schleiermacher G, et al. Control-FREEC: a tool for assessing copy number and allelic content using next-generation sequencing data. Bioinformatics [Internet]. 2012;28:423–5. Available from: https://www.ncbi.nlm.nih.gov/pubmed/22155870
24. Jeffares DC, Jolly C, Hoti M, Speed D, Shaw L, Rallis C, et al. Transient structural variations have strong effects on quantitative traits and reproductive isolation in fission yeast. Nat Commun [Internet]. 2017;8:14061. Available from: https://www.ncbi.nlm.nih.gov/pubmed/28117401
25. da Silva V, Ramos M, Groenen M, Crooijmans R, Johansson A, Regitano L, et al. CNVRanger: association analysis of CNVs with gene expression and quantitative phenotypes. Bioinformatics [Internet]. 2020;36:972–3. Available from: https://doi.org/10.1093/bioinformatics/btz632
26. Chen X, Schulz-Trieglaff O, Shaw R, Barnes B, Schlesinger F, Källberg M, et al. Manta: rapid detection of structural variants and indels for germline and cancer sequencing applications. Bioinformatics [Internet]. 2016;32:1220–2. Available from: https://www.ncbi.nlm.nih.gov/pubmed/26647377
27. Ramayo-Caldas Y, Castelló A, Pena RN, Alves E, Mercadé A, Souza CA, et al. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. BMC Genomics [Internet]. 2010;11:593. Available from: https://www.ncbi.nlm.nih.gov/pubmed/20969757
28. Ramon E, Belanche-Muñoz L, Molist F, Quintanilla R, Perez-Enciso M, Ramayo-Caldas Y. : A Kernel Framework for Integrating Supervised and Unsupervised Analyses in Spatio-Temporal Metagenomic Datasets. Front Microbiol [Internet]. 2021;12:609048. Available from: https://www.ncbi.nlm.nih.gov/pubmed/33584612
29. Guyon I, Weston J, Barnhill S, Vapnik V. Gene Selection for Cancer Classification using Support Vector Machines. Mach Learn [Internet]. 2002;46:389–422. Available from: https://doi.org/10.1023/A:1012487302797
30. Zingaretti LM, Renand G, Morgavi DP, Ramayo-Caldas Y. Link-HD: a versatile framework to explore and integrate heterogeneous microbial communities. Bioinformatics [Internet]. 2020;36:2298–9. Available from: https://doi.org/10.1093/bioinformatics/btz862

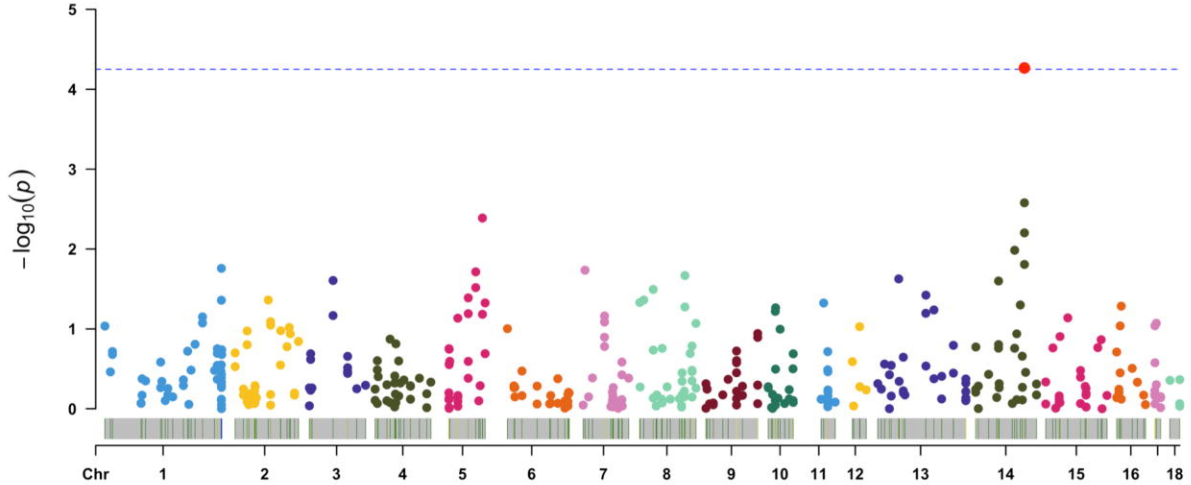
585 31. Yalcin B, Wong K, Agam A, Goodson M, Keane TM, Gan X, et al. Sequence-based
586 characterization of structural variation in the mouse genome. Nature [Internet]. 2011;477:326–9.
587 Available from: https://doi.org/10.1038/nature10432
588 32. Nies AT, Keppler D. The apical conjugate efflux pump ABCC2 (MRP2). Pflugers Arch
589 [Internet]. 2007;453:643–59. Available from: https://doi.org/10.1007/s00424-006-0109-y
590 33. Jemnitz K, Heredi-Szabo K, Janossy J, Ioja E, Vereczkey L, Krajcsi P. ABCC2/Abcc2: a
591 multispecific transporter with dominant excretory functions. Drug Metab Rev [Internet].
592 2010;42:402–36. Available from: https://doi.org/10.3109/03602530903491741
593 34. Klaassen CD, Aleksunes LM. Xenobiotic, bile acid, and cholesterol transporters: function and
594 regulation. Pharmacol Rev [Internet]. 2010;62:1–96. Available from:
595 https://www.ncbi.nlm.nih.gov/pubmed/20103563
596 35. Ramayo-Caldas Y, Ballester M, Fortes MR, Esteve-Codina A, Castelló A, Noguera JL, et al.
597 From SNP co-association to RNA co-expression: novel insights into gene networks for
598 intramuscular fatty acid composition in porcine. BMC Genomics [Internet]. 2014;15:232. Available
599 from: https://www.ncbi.nlm.nih.gov/pubmed/24666776
600 36. Zhang W, Zhang J, Cui L, Ma J, Chen C, Ai H, et al. Genetic architecture of fatty acid
601 composition in the longissimus dorsi muscle revealed by genome-wide association studies on
602 diverse pig populations. Genet Sel Evol [Internet]. 2016;48:5. Available from:
603 https://www.ncbi.nlm.nih.gov/pubmed/26796620
604 37. Chu XY, Strauss JR, Mariano MA, Li J, Newton DJ, Cai X, et al. Characterization of mice
605 lacking the multidrug resistance protein MRP2 (ABCC2). J Pharmacol Exp Ther [Internet].
606 2006;317:579–89. Available from: https://www.ncbi.nlm.nih.gov/pubmed/16421286
607 38. Vlaming ML, Mohrmann K, Wagenaar E, de Waart DR, Elferink RP, Lagas JS, et al.
608 Carcinogen and anticancer drug transport by Mrp2 in vivo: studies using Mrp2 (Abcc2) knockout
609 mice. J Pharmacol Exp Ther [Internet]. 2006;318:319–27. Available from:
610 https://www.ncbi.nlm.nih.gov/pubmed/16611851
611 39. Jansen PL, Peters WH, Lamers WH. Hereditary chronic conjugated hyperbilirubinemia in
612 mutant rats caused by defective hepatic anion transport. Hepatology [Internet]. 1985;5:573–9.
613 Available from: https://www.ncbi.nlm.nih.gov/pubmed/4018730
614 40. DUBIN IN, JOHNSON FB. Chronic idiopathic jaundice with unidentified pigment in liver
615 cells; a new clinicopathologic entity with a report of 12 cases. Medicine (Baltimore) [Internet].
616 1954;33:155–97. Available from: https://www.ncbi.nlm.nih.gov/pubmed/13193360
617 41. Dixon PH, Sambrotta M, Chambers J, Taylor-Harris P, Syngelaki A, Nicolaides K, et al. An
618 expanded role for heterozygous mutations of ABCB4, ABCB11, ATP8B1, ABCC2 and TJP2 in
619 intrahepatic cholestasis of pregnancy. Sci Rep [Internet]. 2017;7:11823. Available from:
620 https://www.ncbi.nlm.nih.gov/pubmed/28924228
621 42. Nobles CL, Green SI, Maresso AW. A product of heme catabolism modulates bacterial function
622 and survival. PLoS Pathog [Internet]. 2013;9:e1003507. Available from:
623 https://www.ncbi.nlm.nih.gov/pubmed/23935485
624 43. Ridlon JM, Harris SC, Bhowmik S, Kang D-J, Hylemon PB. Consequences of bile salt
625 biotransformations by intestinal bacteria. Gut Microbes [Internet]. Taylor & Francis; 2016;7:22–39.
626 Available from: https://doi.org/10.1080/19490976.2015.1127483
627 44. GUSTAFSSON BE, LANKE LS. Bilirubin and urobilins in germfree, ex-germfree, and
628 conventional rats. J Exp Med [Internet]. 1960;112:975–81. Available from:
629 https://www.ncbi.nlm.nih.gov/pubmed/13709962
630 45. Sayin SI, Wahlström A, Felin J, Jäntti S, Marschall HU, Bamberg K, et al. Gut microbiota
631 regulates bile acid metabolism by reducing the levels of tauro-beta-muricholic acid, a naturally
632 occurring FXR antagonist. Cell Metab [Internet]. 2013;17:225–35. Available from:
633 https://www.ncbi.nlm.nih.gov/pubmed/23395169
634 46. Swann JR, Want EJ, Geier FM, Spagou K, Wilson ID, Sidaway JE, et al. Systemic gut microbial
635 modulation of bile acid metabolism in host tissue compartments. Proceedings of the National

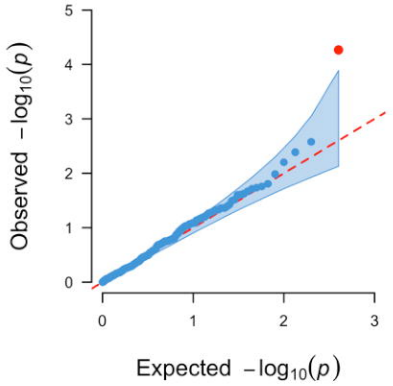
636 Academy of Sciences [Internet]. 2011;108:4523–30. Available from:
637 https://doi.org/10.1073/pnas.1006734107
638 47. Li DK, Chaudhari SN, Sojoodi M, Lee Y, Adhikari AA, Zukerberg L, et al. Inhibition of
639 microbial deconjugation of micellar bile acids protects against intestinal permeability and liver
640 injury. bioRxiv [Internet]. 2021;2021.03.24.436896. Available from:
641 http://biorxiv.org/content/early/2021/11/09/2021.03.24.436896.abstract
642 48. van Best N, Rolle-Kampczyk U, Schaap FG, Basic M, Olde Damink SWM, Bleich A, et al. Bile
643 acids drive the newborn's gut microbiota maturation. Nat Commun [Internet]. 2020;11:3692.
644 Available from: https://doi.org/10.1038/s41467-020-17183-8
645 49. Molinero N, Ruiz L, Sánchez B, Margolles A, Delgado S. Intestinal Bacteria Interplay With Bile
646 and Cholesterol Metabolism: Implications on Host Physiology. Front Physiol [Internet].
647 2019;10:185. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30923502
648 50. Song Z, Cai Y, Lao X, Wang X, Lin X, Cui Y, et al. Taxonomic profiling and populational
649 patterns of bacterial bile salt hydrolase (BSH) genes based on worldwide human gut microbiome.
650 Microbiome [Internet]. 2019;7:9. Available from: https://doi.org/10.1186/s40168-019-0628-3
651 51. Shapiro H, Kolodziejczyk AA, Halstuch D, Elinav E. Bile acids in glucose metabolism in health
652 and disease. Journal of Experimental Medicine [Internet]. 2018;215:383–96. Available from:
653 https://doi.org/10.1084/jem.20171965
654 52. Cipriani S, Mencarelli A, Chini MG, Distrutti E, Renga B, Bifulco G, et al. The Bile Acid
655 Receptor GPBAR-1 (TGR5) Modulates Integrity of Intestinal Barrier and Immune Response to
656 Experimental Colitis. PLoS One [Internet]. 2011;6:e25637. Available from:
657 https://doi.org/10.1371/journal.pone.0025637
658 53. Otani T, Ichii T, Aono S, Takeichi M. Cdc42 GEF Tuba regulates the junctional configuration of
659 simple epithelial cells. J Cell Biol [Internet]. 2006;175:135–46. Available from:
660 http://europepmc.org/abstract/MED/17015620
661 54. Qin Y, Meisen WH, Hao Y, Macara IG. Tuba, a Cdc42 GEF, is required for polarized spindle
662 orientation during epithelial cyst formation. J Cell Biol [Internet]. 2010;189:661–9. Available from:
663 http://europepmc.org/abstract/MED/20479467
664 55. Ubelmann F, Chamaillard M, El-Marjou F, Simon A, Netter J, Vignjevic D, et al. Enterocyte
665 loss of polarity and gut wound healing rely upon the F-actin–severing function of villin.
666 Proceedings of the National Academy of Sciences [Internet]. Proceedings of the National Academy
667 of Sciences; 2013;110:E1380–9. Available from: https://doi.org/10.1073/pnas.1218446110
668 56. M BK, L WS, D LD. Epithelial Microvilli Establish an Electrostatic Barrier to Microbial
669 Adhesion. Infect Immun [Internet]. American Society for Microbiology; 2014;82:2860–71.
670 Available from: https://doi.org/10.1128/IAI.01681-14
671 57. Ferrary E, Cohen-Tannoudji M, Pehau-Arnaudet G, Lapillonne A, Athman R, Ruiz T, et al. In
672 vivo, villin is required for Ca(2+)-dependent F-actin disruption in intestinal brush borders. J Cell
673 Biol [Internet]. 1999;146:819–30. Available from: http://europepmc.org/abstract/MED/10459016
674 58. Wells JM, Rossi O, Meijerink M, van Baarlen P. Epithelial crosstalk at the microbiota–mucosal
675 interface. Proceedings of the National Academy of Sciences [Internet]. Proceedings of the National
676 Academy of Sciences; 2011;108:4607–14. Available from:
677 https://doi.org/10.1073/pnas.1000092107
678 59. Gieryńska M, Szulc-Dąbrowska L, Struzik J, Mielcarska MB, Gregorczyk-Zboroch KP.
679 Integrity of the Intestinal Barrier: The Involvement of Epithelial Cells and Microbiota-A Mutual
680 Relationship. Animals (Basel). 2022;12.
681 60. Yu LC-H, Wang J-T, Wei S-C, Ni Y-H. Host-microbial interactions and regulation of intestinal
682 epithelial barrier function: From physiology to pathology. World J Gastrointest Pathophysiol.
683 2012;3:27–43.
684 61. Willing BP, van Kessel AG. Enterocyte proliferation and apoptosis in the caudal small intestine
685 is influenced by the composition of colonizing commensal bacteria in the neonatal gnotobiotic pig.
686 J Anim Sci [Internet]. 2007;85:3256–66. Available from: https://doi.org/10.2527/jas.2007-0320

687 62. Shirkey TW, Siggers RH, Goldade BG, Marshall JK, Drew MD, Laarveld B, et al. Effects of
688 Commensal Bacteria on Intestinal Morphology and Expression of Proinflammatory Cytokines in
689 the Gnotobiotic Pig. Exp Biol Med [Internet]. SAGE Publications; 2006;231:1333–45. Available
690 from: https://doi.org/10.1177/153537020623100807
691 63. Danielsen M, Hornshøj H, Siggers RH, Jensen BB, van Kessel AG, Bendixen E. Effects of
692 Bacterial Colonization on the Porcine Intestinal Proteome. J Proteome Res [Internet]. American
693 Chemical Society; 2007;6:2596–604. Available from: https://doi.org/10.1021/pr070038b
694 64. Kozakova H, Kolinska J, Lojda Z, Rehakova Z, Sinkora J, Zakostelecka M, et al. Effect of
695 bacterial monoassociation on brush-border enzyme activities in ex-germ-free piglets: comparison of
696 commensal and pathogenic Escherichia coli strains. Microbes Infect. 2006;8:2629–39.
697

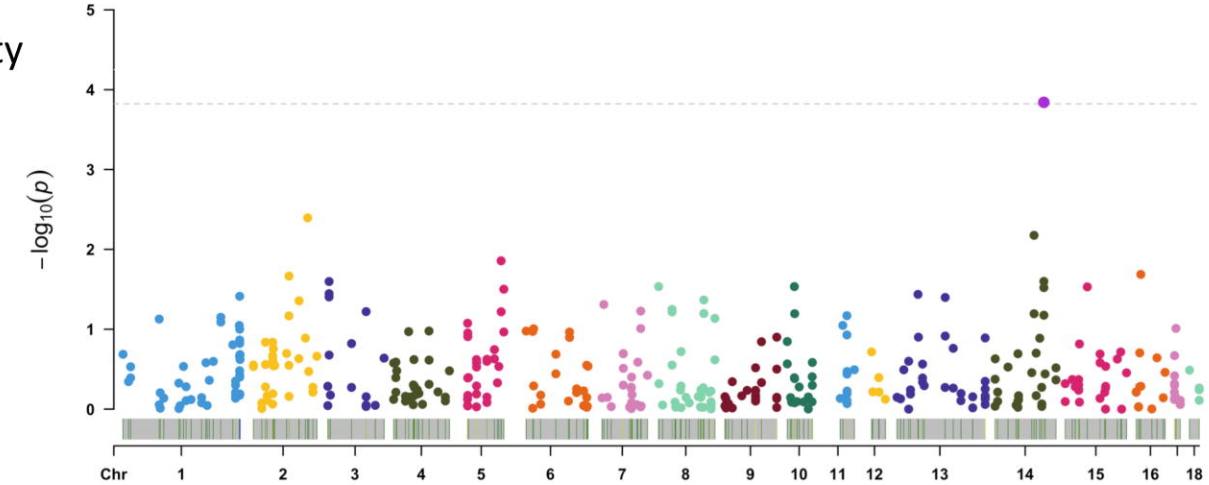

Gain
Loss
Both
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

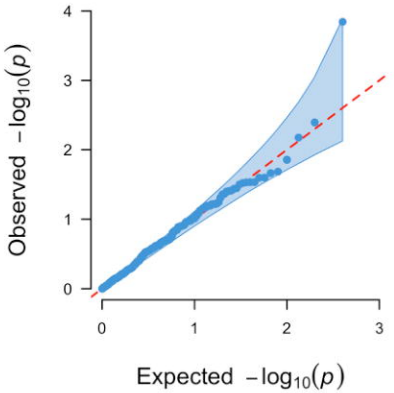
Richness



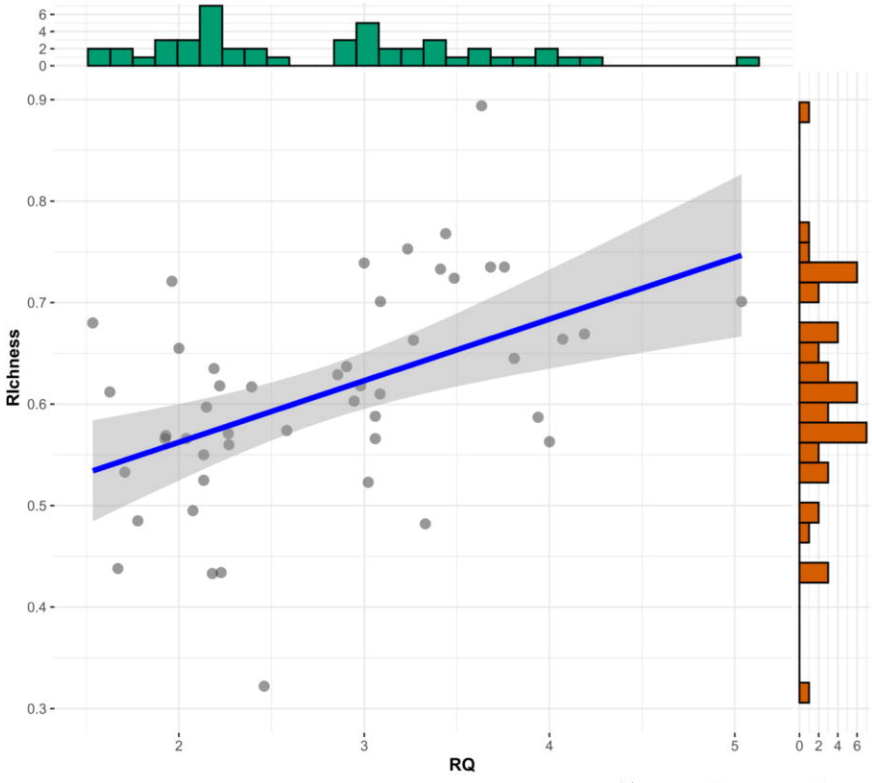

α-diversity



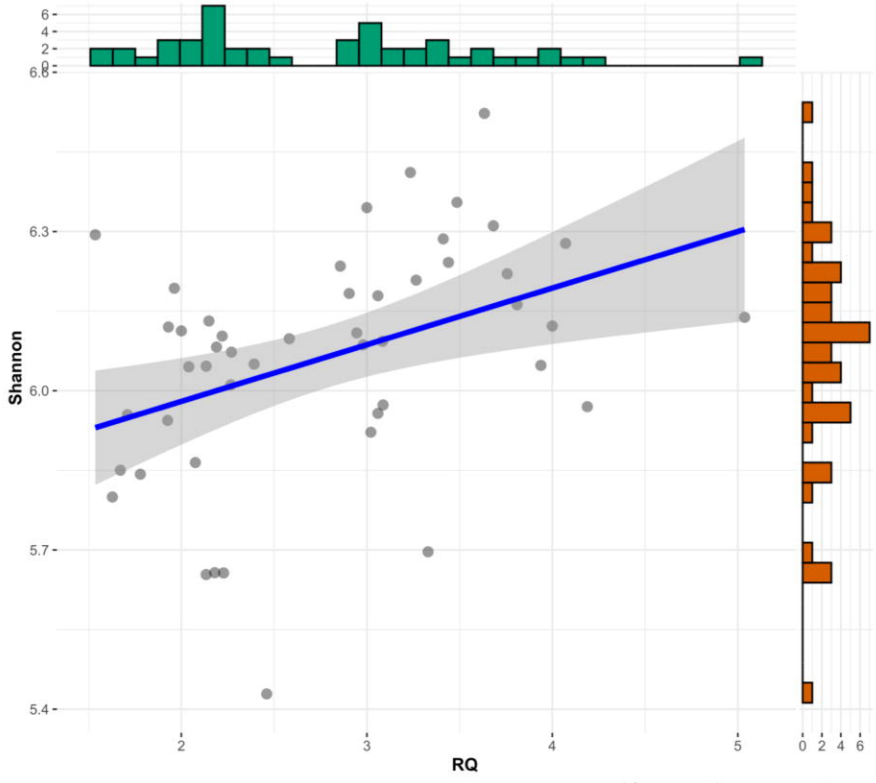
Richness

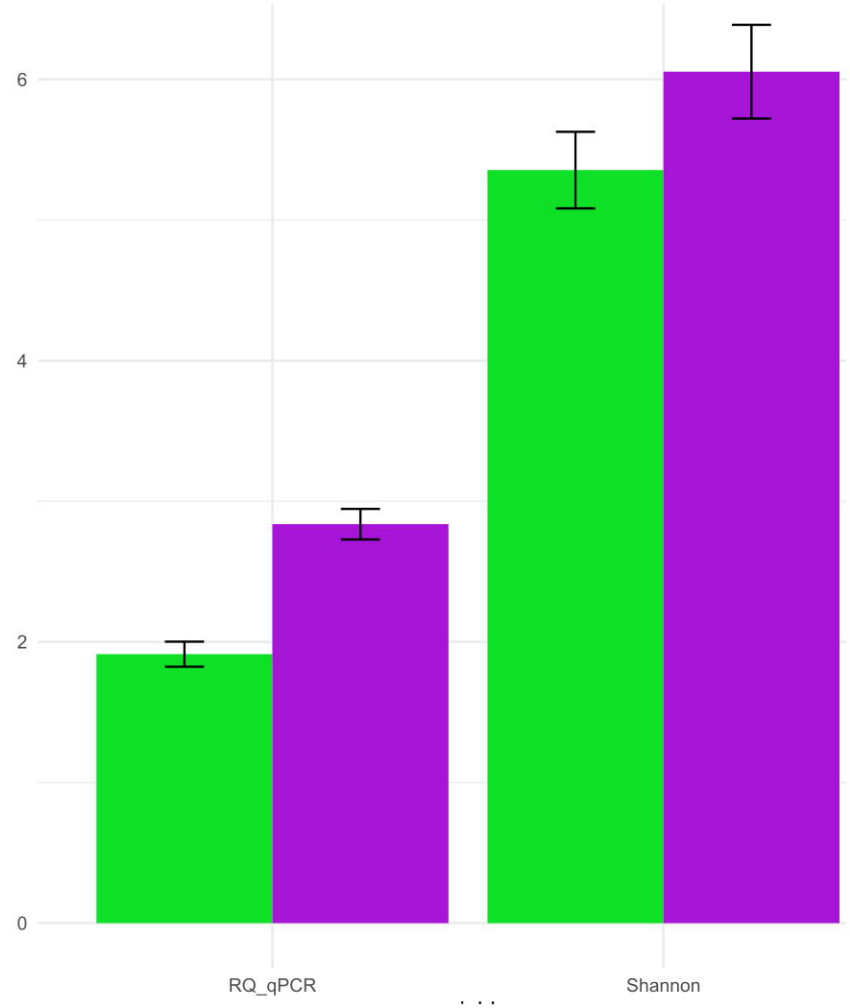
r=0.474, p-value=6.72×10$^{-4}$

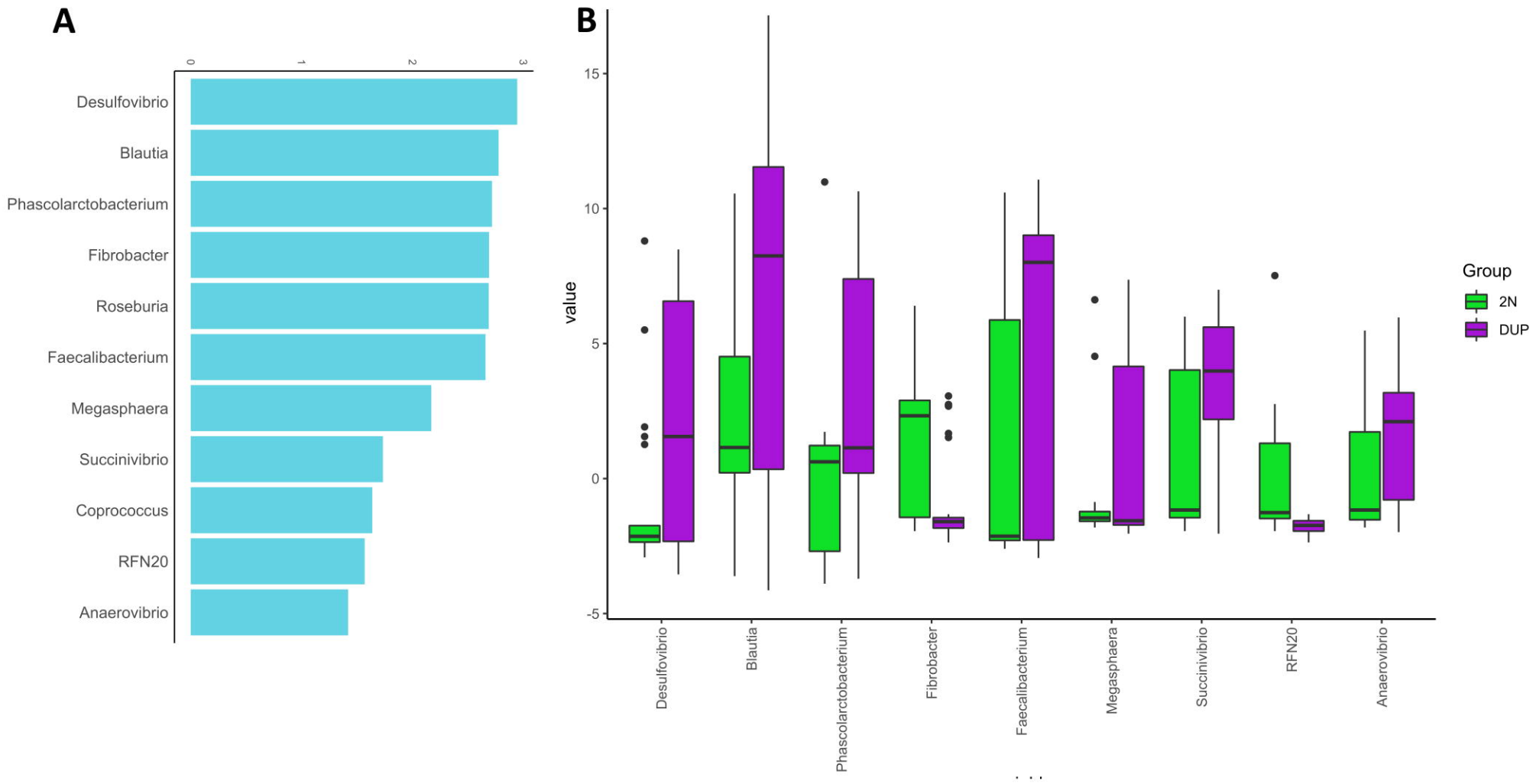


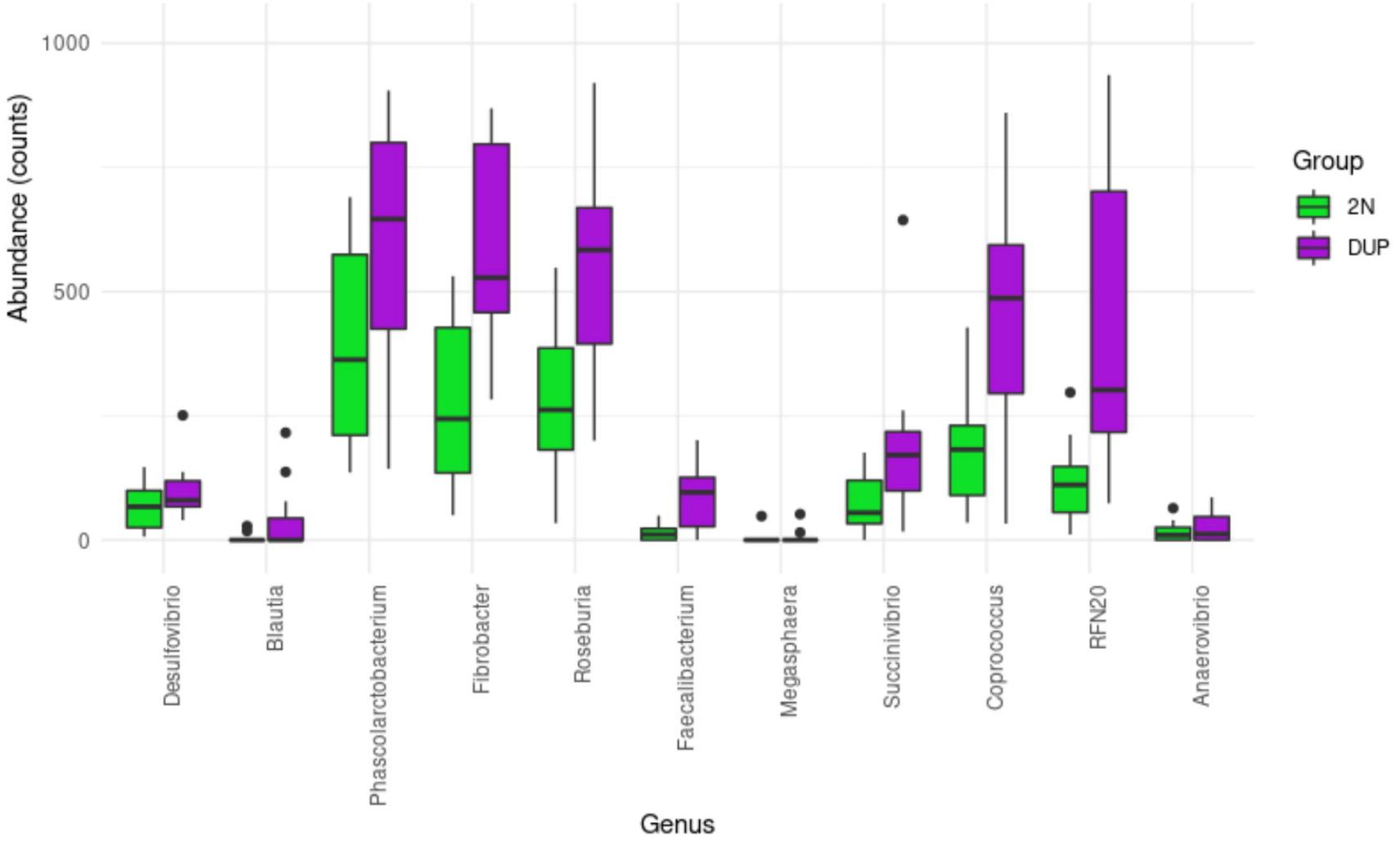

α-diversity

r=0.401, p-value=7.77×10$^{-3}$




6
4
2
0
RQ_qPCR
Shannon


A
0
1
2
3
Desulfovibrio
Blautia
Phascolarctobacterium
Fibrobacter
Roseburia
Faecalibacterium
Megasphaera
Succinivibrio
Coprococcus
RFN20
Anaerovibrio
B
value
15
10
5
0
-5
Desulfovibrio
Blautia
Phascolarctobacterium
Fibrobacter
Faecalibacterium
Megasphaera
Succinivibrio
RFN20
Anaerovibrio
Group
2N
DUP


Abundance (counts)
1000
500
0
Desulfovibrio
Blautia
Phascolarctobacterium
Fibrobacter
Roseburia
Faecalibacterium
Megasphaera
Succinivibrio
Coprococcus
RFN20
Anaerovibrio
Genus
Group
2N
DUP