Supplementary figure legends

Figure S1 RFI Population Distribution of the Trial Cohort (146 Cows). Cows with extreme low (n = 40) and extreme high (n = 38) RFI, are colored red and grey, respectively and represent the 25% most and 25% least efficient from a cohort of 146 cows. These 78 cows were chosen for rumen and fecal sampling.

Figure S2 Percentage of Mapped Reads From Efficient and Inefficient Cows' Samples to the Total Microbiome Genes.

Data are expressed as mean ± SEM.

Figure S3 Number of Genes in a Sample As a Function of the Number of Reads. Reads from each sample were aligned to the total genes. The number of aligned reads is plotted against the number of genes obtained for each sample. No correlation was found between the two variables.

Figure S4 Phyla Abundances. Microbiome composition of the two efficiency groups at the phylum level. Phyla with relative abundance above 0.001% are presented.

Data are expressed as mean ± SEM.

Figure S5 Shannon Diversity and Dominance of Efficient and Inefficient Microbiomes. (**a**) Shannon diversity at different phylogenetic levels. (**b**) Dominance at different phylogenetic levels.

Data are expressed as mean ± SEM. Wilcoxon rank sum, *$P < 0.05$.

Figure S6 Prediction of Physiological and Metabolic Traits According to Species. Species that differed in abundance between efficient and inefficient cows were sorted according to their *P*-values and grouped into bins of 100. The bins were used as predictive features for the different physiological parameters using the k-Nearest Neighbors (KNN) algorithm with k = 3. Each iteration used a different bin as predictive features, in ascending *P*-value order. (**a**) Conversion ratio (CR) prediction accuracy. (**b**) Milk fat prediction accuracy. (**c**) Dry matter intake (DMI) prediction accuracy. (**d**) Milk yield prediction accuracy. (**e**) Milk lactose prediction accuracy. (**f**) pH prediction accuracy. (**g**) Milk protein prediction accuracy. (**h**) Milk energy prediction accuracy. (**i**) Body conditioning score (BCS) change prediction accuracy.

Figure S7 Prediction of Physiological and Metabolic Traits According to Genes. Genes that differed in abundance between efficient and inefficient cows were sorted according to their *P*-values and grouped into bins of 100. The bins were used as predictive features for the different physiological parameters using the k-Nearest Neighbors (KNN) algorithm with k = 3. Each iteration used a different bin as predictive features, in ascending *P*-value order. Different graphs represent predictions of different physiological parameters. (**a**) CR prediction accuracy. (**b**) Milk fat prediction accuracy. (**c**) DMI prediction accuracy. (**d**) Milk yield prediction accuracy. (**e**) Milk lactose prediction accuracy. (**f**) pH prediction accuracy. (**g**) Milk protein prediction accuracy. (**h**) Milk energy prediction accuracy. (**i**) BCS change prediction accuracy.

Figure S8 Specificity and Sensitivity Evaluation of Predictions of Physiological and Metabolic Traits According to Species. Receiver Operation Characteristics (ROC) curves and Area Under Curve (AUC) measures were obtained for the first five prediction bins (see Fig. 2a, Supplementary Fig. 4) based on the average of 1,000 KNN cross-validation iterations. (**a**) CR ROC analysis. (**b**) Milk fat ROC analysis. (**c**) DMI ROC analysis. (**d**) Milk yield ROC analysis. (**e**) Milk lactose ROC analysis. (**f**) pH ROC analysis. (**g**) Milk protein ROC analysis. (**h**) Milk energy ROC analysis. (**i**) BCS change ROC analysis.

Figure S9 Specificity and Sensitivity Evaluation of Predictions of Physiological and Metabolic Traits According to Genes. Receiver Operation Characteristics (ROC) curves and Area Under Curve (AUC) measures were obtained for the first five prediction bins (see Fig. 2b, Supplementary Fig. 5) based on the average of 1,000 KNN cross-validation iterations. (**a**) CR ROC analysis. (**b**) Milk fat ROC analysis. (**c**) DMI ROC analysis. (**d**) Milk yield ROC analysis. (**e**) Milk lactose ROC analysis. (**f**) pH ROC analysis. (**g**) Milk protein ROC analysis. (**h**) Milk energy ROC analysis. (**i**) BCS change ROC analysis.

Figure S10 In-Vitro Digestibility and In Vivo Digestibility. (**a**) In-vitro dry matter (DM) digestibility of feed after 24 h incubation with rumen fluid of efficient and inefficient cows. (**b**) In-vitro neutral detergent fiber (NDF) digestibility of feed after 24 h incubation with rumen fluid of efficient and inefficient cows. (**c**) In-vivo DM digestibility of efficient and inefficient stool samples. (**d**) In-vivo NDF digestibility of efficient and inefficient stool samples.

Data are expressed as mean ± SEM.

Figure S11 Relative Abundance of Significantly Different Species. Relative abundance of the 18 species that were found to be significantly different between the two efficiency groups.

Data are expressed as mean ± SEM.

Figure S12 Principal Coordinates Analysis (PCoA) of Genes Enriched in the Two Efficiency Groups – PCoA was performed for the efficient and inefficient cows' microbiomes using the 34,166 genes that were significantly different between the two efficiency groups.

Figure S13 End Product Compounds of KEGG Annotated Genes Significantly Enriched in Efficient Cows' Microbiomes. End product compounds resulting from the overall annotation of the efficient cows' enriched genes are denoted in blue. Metabolites which were measured in the metabolomics analysis are denoted in green.

Figure S14 End Product Compounds of KEGG Annotated Genes Significantly Enriched in Inefficient Cows' Microbiomes. End product compounds resulting from the overall annotation of the inefficient cows' enriched genes are denoted in blue. Metabolites which were measured in the metabolomics analysis are denoted in green.

Figure S15 Read Alignment to Known Rumen Microbial Genomes. Reads from each sample were aligned to sequenced genomes of known rumen microorganisms using the BWA tool. The ratios between recruitment of efficient/inefficient samples to each genome are presented.

Data are expressed as mean ± SEM. Permutations t test, *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Figure S16 Acrylate Pathway Distribution in Organisms of the Rumen Microbiome. Genes from all samples were blasted against genes of lactoyl-CoA dehydratase subunits alpha, beta and gamma (Reichardt *et al*. 2014). Genes that passed a cutoff of 60% identity were gathered and annotated using the NR database. The percentage of each annotation in the overall genes above a threshold of 2% is presented.

Supplementary table legends

Table S1 EC Numbers of KEGG Metabolic Pathways Enzymes. EC numbers of KEGG metabolic pathways enzymes used for alignment of reads from the different feed efficiency microbiome samples.

Table S2 Shannon Diversity and Dominance of Species and Genes of Efficient and Inefficient Cows’ Microbiomes.

Data are expressed as mean ± SEM. Wilcoxon rank sum, *$P < 0.05$, **$P < 0.01$.

Table S3 SCFAs Concentrations. Concentrations of SCFAs before and after normalization of total OM. Data are expressed as mean ± SEM. $^{a, b}$ Means within a row with different superscript letters are significantly different ($P < 0.05$).

Table S4 ko Numbers of Amino Acid Metabolisn Category Enriched in Inefficient Cows' Microbiomes.

Supplementary Data legends

Data 1 Tukey's Test Analysis of RFI of the Trial Cohort (146 cows). Cows with extreme low (n = 40) and extreme high (n = 38) RFI values that are significantly different from each other ($P < 0.05$) are denoted in red and grey, respectively, and represent the 25% most and 25% least efficient animals of the cohort which were chosen for further analysis. Letters or signs represent groups of animals which are not significantly different in their RFI value.

Data 2 Relative Abundances at Different Phylogenetic Levels. Relative abundances of taxa in the two efficiency groups according to 16S data. Species were binned at different taxonomic levels to receive taxon abundances for each phylogenetic level. Data are expressed as mean ± SEM.

Data 3 Prediction Accuracy of Different Bin Sizes. Bins with different number of features ranging from 50 to 500 and covering the first 1,000 Fisher's exact test significant features, were used to predict the RFI phenotype. Kruskal-Wallis test was used to determine if bin size significantly affects prediction accuracy.

Data 4 Measured Physiological Parameters for Each Animal Sampled.

Data 5 Taxonomic Annotation of Species in the First Five RFI Predictive Bins.

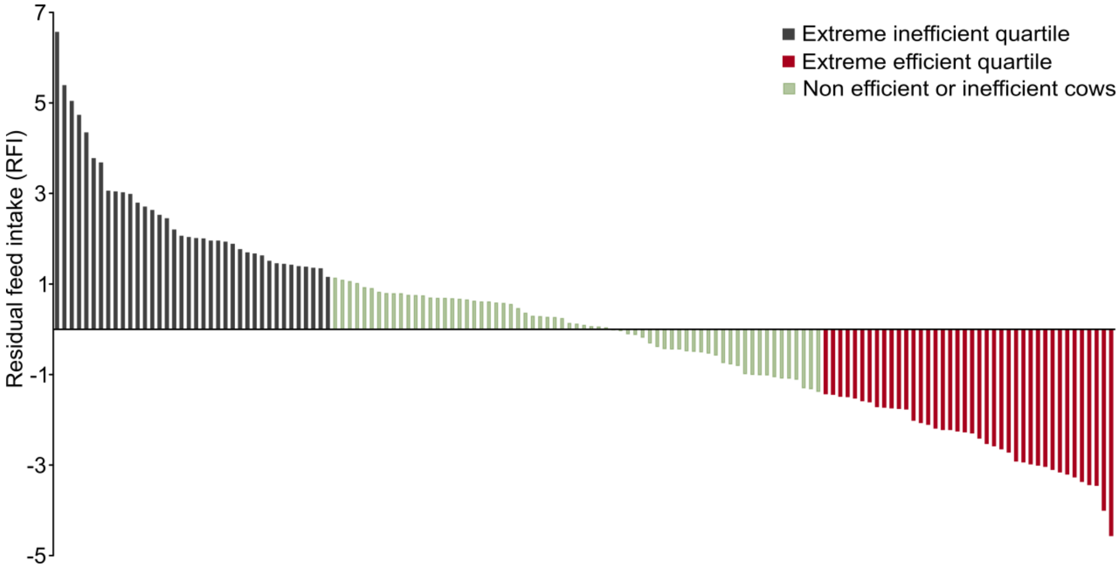
Data 6 Functional Annotation of Genes in the First Five RFI Predictive Bins. Genes in the first five RFI predictive bins were annotated using the NR and GO databases. Sequences description and Gene-Ontology-Terms are presented.

Data 7 Genes Significantly Enriched in Efficient or Inefficient Cows’ Microbiomes. Genes that were found to be significantly enriched in the microbiomes of one of the efficiency groups are listed, along with their abundances in each group (mean ± SEM), E.C. number and KEGG Orthology. The r and *P*-values of Spearman correlation of each gene to the RFI phenotype are also presented. *P*-values were corrected for false discovery rate.

Data 8 ko Annotations of Genes Significantly Enriched in Efficient or Inefficient Cows’ Microbiomes. The pathway names and ko numbers of each gene enriched in efficient or inefficient cows are listed.

Data 9 End Product Compounds of KEGG Annotated Genes Significantly Enriched in Efficient or Inefficient Cows’ Microbiomes. The end product compounds resulting from the overall annotation of the efficient cows' enriched genes or the inefficient cows' enriched genes are listed.

Figure S1



Residual feed intake (RFI)
7
5
3
1
-1
-3
-5
Extreme inefficient quartile
Extreme efficient quartile
Non efficient or inefficient cows

Figure S2

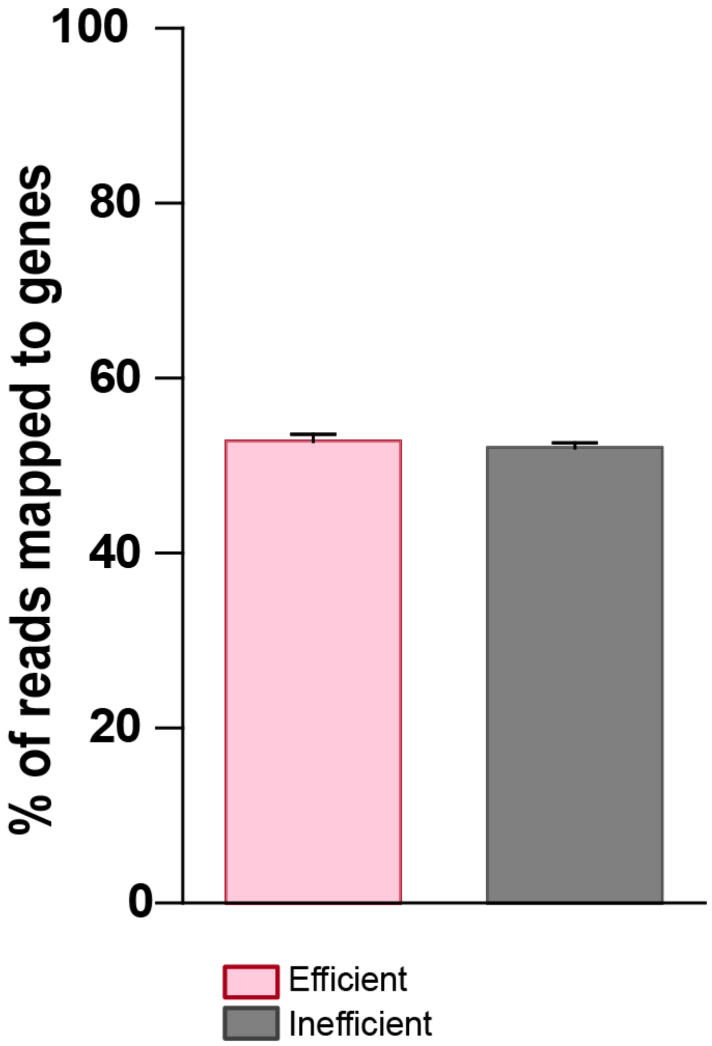
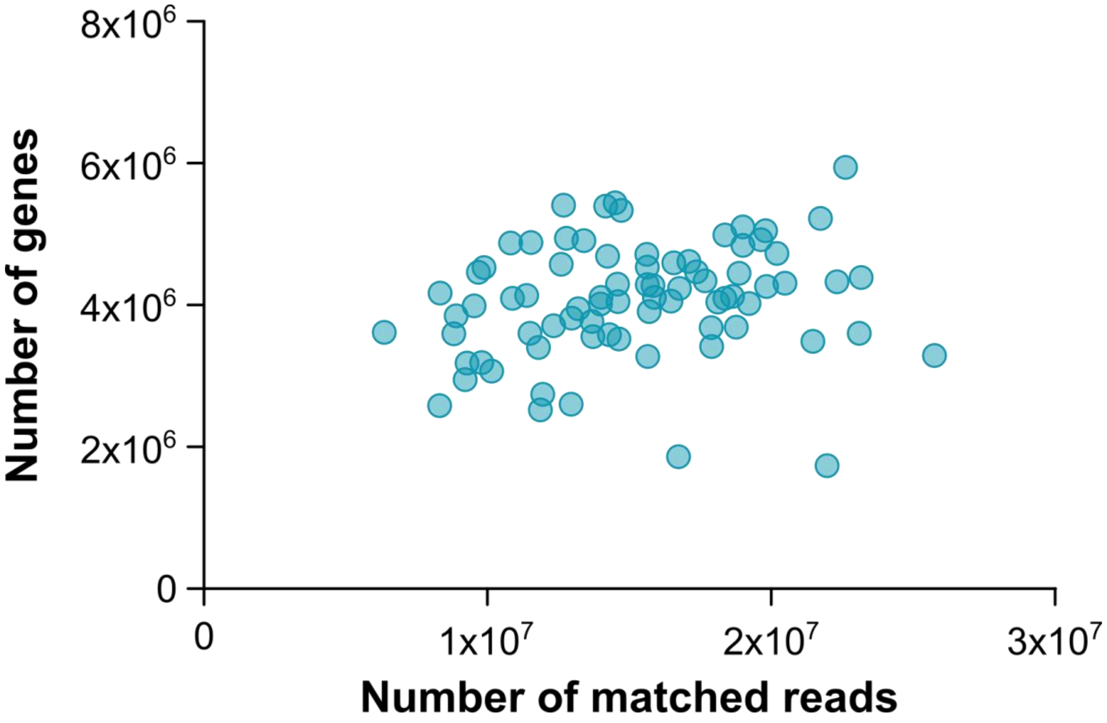
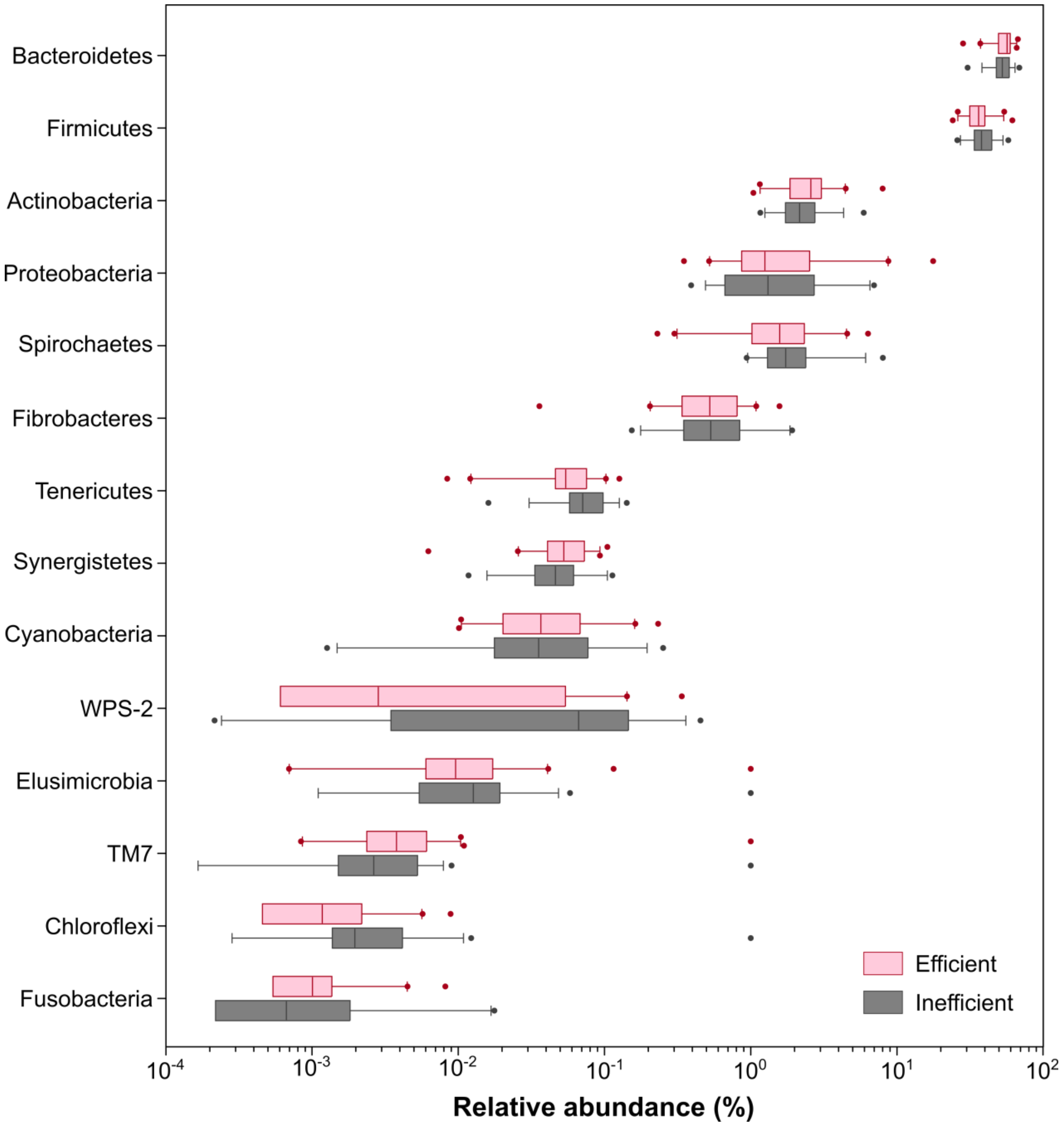

% of reads mapped to genes
100
80
60
40
20
0
Efficient
Inefficient

Figure S3



Number of genes
8x10^6
6x10^6
4x10^6
2x10^6
0
0
1x10^7
2x10^7
3x10^7
Number of matched reads

175 **Figure S4**



176

Figure S5

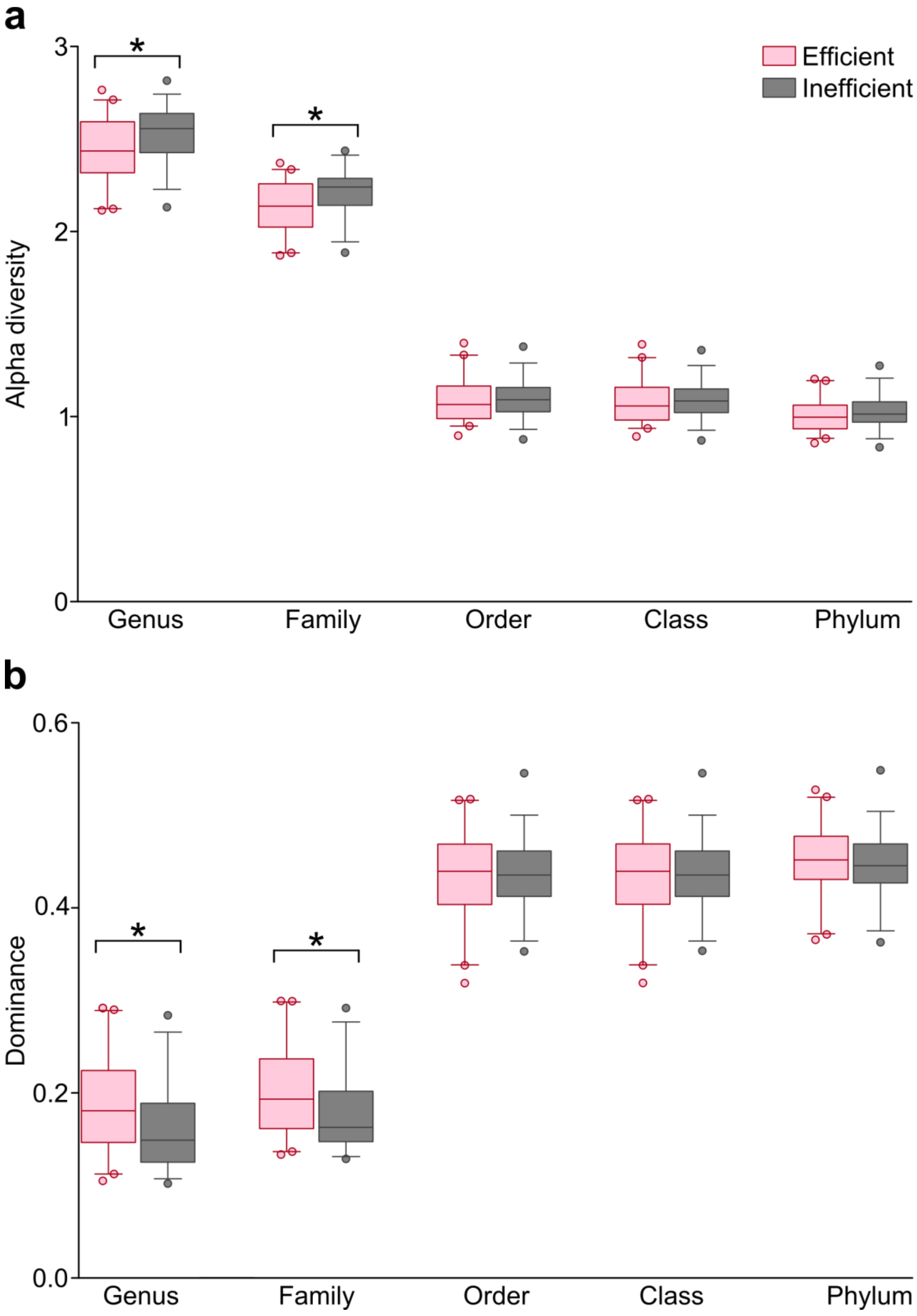

a
Alpha diversity
0
1
2
3
*
*
Efficient
Inefficient
Genus
Family
Order
Class
Phylum
b
Dominance
0.0
0.2
0.4
0.6
*
*
Genus
Family
Order
Class
Phylum

Figure S6

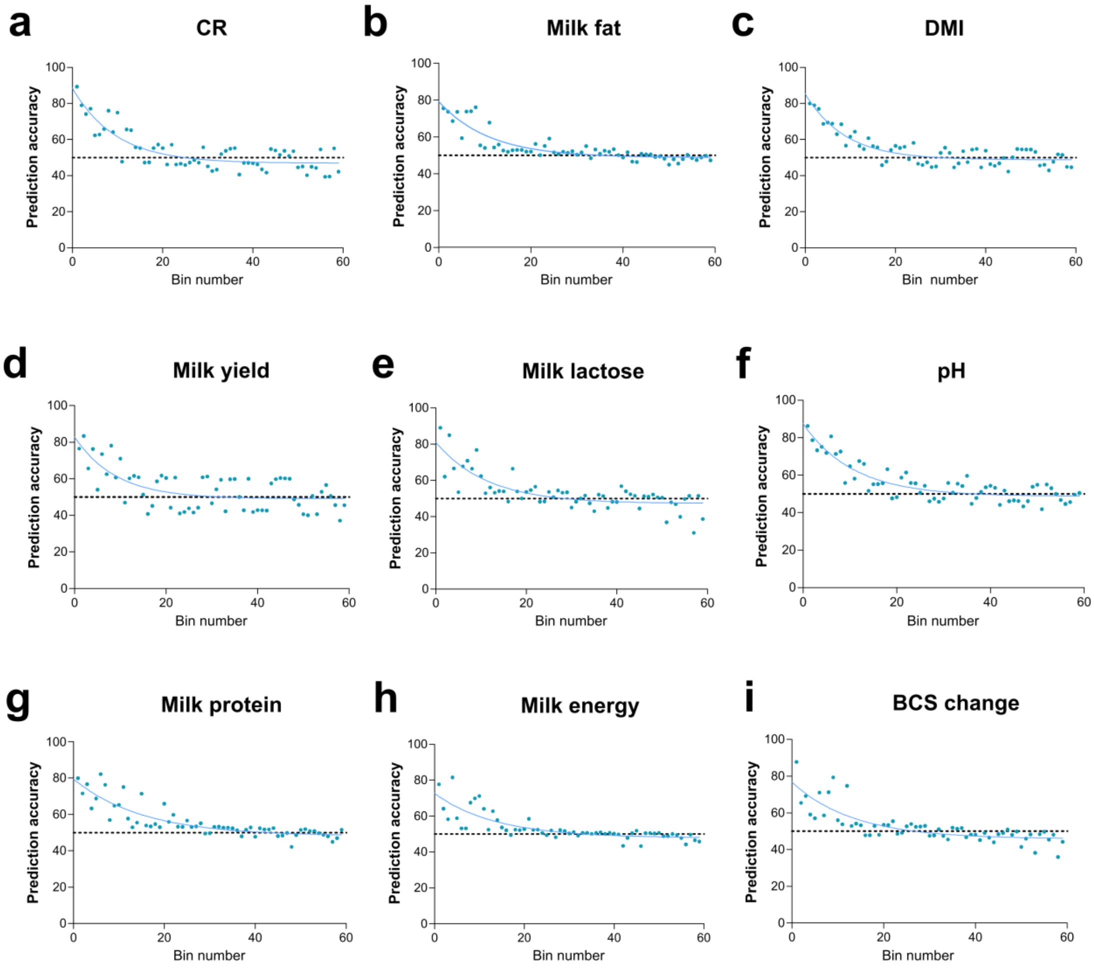

a
CR
b
Milk fat
c
DMI
d
Milk yield
e
Milk lactose
f
pH
g
Milk protein
h
Milk energy
i
BCS change
Prediction accuracy
Bin number
0
20
40
60
80
100

Figure S7

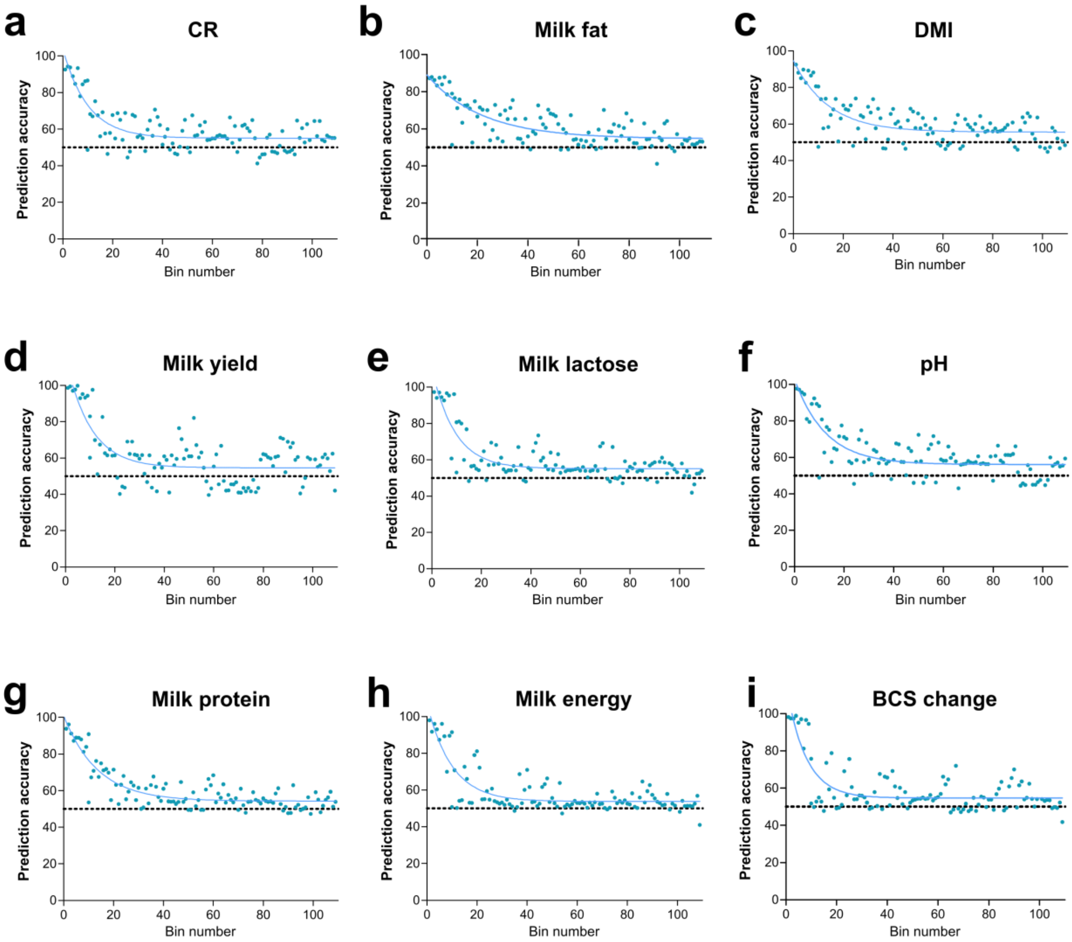

a
CR
b
Milk fat
c
DMI
d
Milk yield
e
Milk lactose
f
pH
g
Milk protein
h
Milk energy
i
BCS change
Prediction accuracy
Bin number
0
20
40
60
80
100

Figure S8

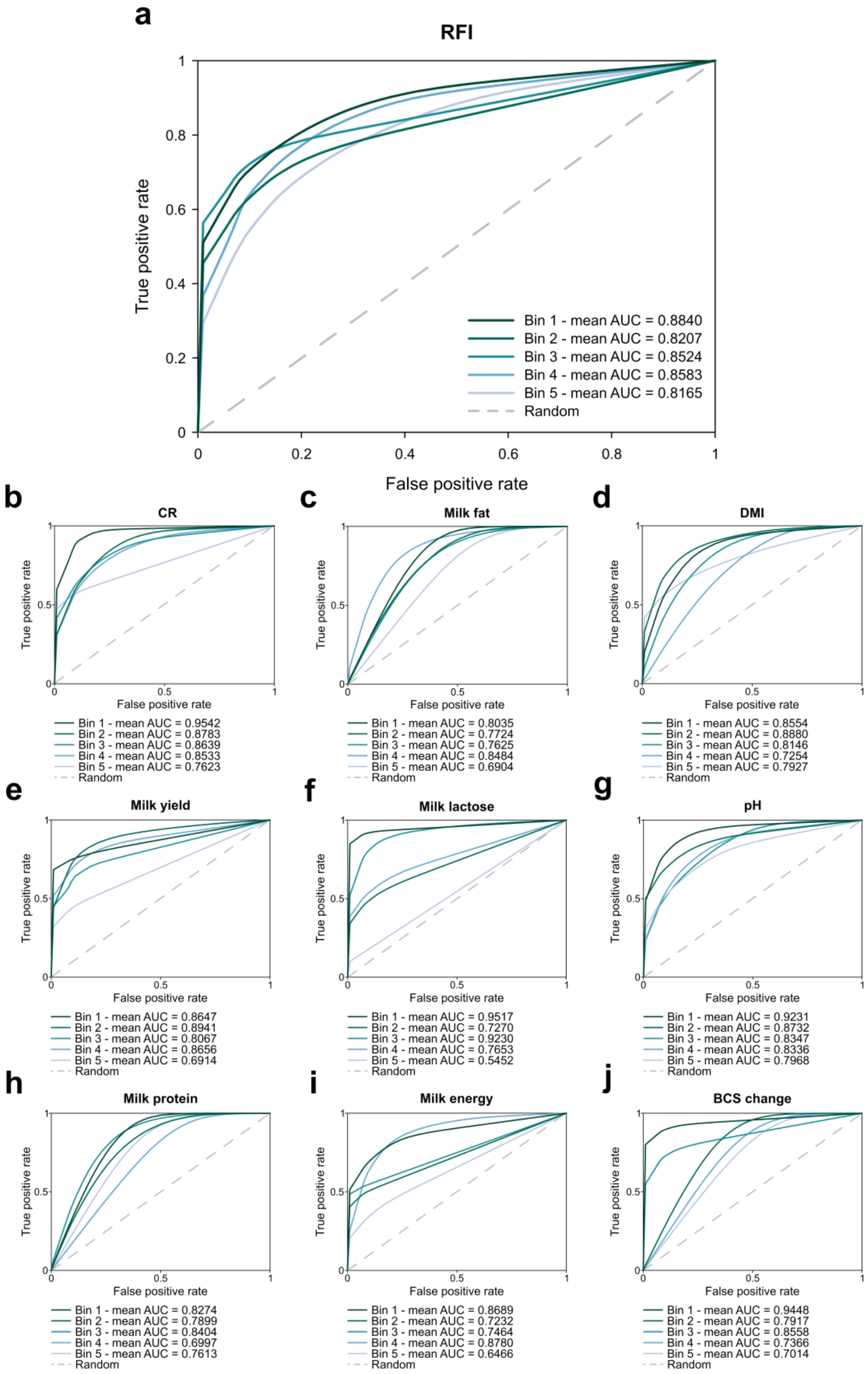

a
RFI
True positive rate
False positive rate
0
0.2
0.4
0.6
0.8
1
Bin 1 - mean AUC = 0.8840
Bin 2 - mean AUC = 0.8207
Bin 3 - mean AUC = 0.8524
Bin 4 - mean AUC = 0.8583
Bin 5 - mean AUC = 0.8165
Random
b
CR
Bin 1 - mean AUC = 0.9542
Bin 2 - mean AUC = 0.8783
Bin 3 - mean AUC = 0.8639
Bin 4 - mean AUC = 0.8533
Bin 5 - mean AUC = 0.7623
Random
c
Milk fat
Bin 1 - mean AUC = 0.8035
Bin 2 - mean AUC = 0.7724
Bin 3 - mean AUC = 0.7625
Bin 4 - mean AUC = 0.8484
Bin 5 - mean AUC = 0.6904
Random
d
DMI
Bin 1 - mean AUC = 0.8554
Bin 2 - mean AUC = 0.8880
Bin 3 - mean AUC = 0.8146
Bin 4 - mean AUC = 0.7254
Bin 5 - mean AUC = 0.7927
Random
e
Milk yield
Bin 1 - mean AUC = 0.8647
Bin 2 - mean AUC = 0.8941
Bin 3 - mean AUC = 0.8067
Bin 4 - mean AUC = 0.8656
Bin 5 - mean AUC = 0.6914
Random
f
Milk lactose
Bin 1 - mean AUC = 0.9517
Bin 2 - mean AUC = 0.7270
Bin 3 - mean AUC = 0.9230
Bin 4 - mean AUC = 0.7653
Bin 5 - mean AUC = 0.5452
Random
g
pH
Bin 1 - mean AUC = 0.9231
Bin 2 - mean AUC = 0.8732
Bin 3 - mean AUC = 0.8347
Bin 4 - mean AUC = 0.8336
Bin 5 - mean AUC = 0.7968
Random
h
Milk protein
Bin 1 - mean AUC = 0.8274
Bin 2 - mean AUC = 0.7899
Bin 3 - mean AUC = 0.8404
Bin 4 - mean AUC = 0.6997
Bin 5 - mean AUC = 0.7613
Random
i
Milk energy
Bin 1 - mean AUC = 0.8689
Bin 2 - mean AUC = 0.7232
Bin 3 - mean AUC = 0.7464
Bin 4 - mean AUC = 0.8780
Bin 5 - mean AUC = 0.6466
Random
j
BCS change
Bin 1 - mean AUC = 0.9448
Bin 2 - mean AUC = 0.7917
Bin 3 - mean AUC = 0.8558
Bin 4 - mean AUC = 0.7366
Bin 5 - mean AUC = 0.7014
Random
True positive rate
False positive rate
0
0.5
1

Figure S9

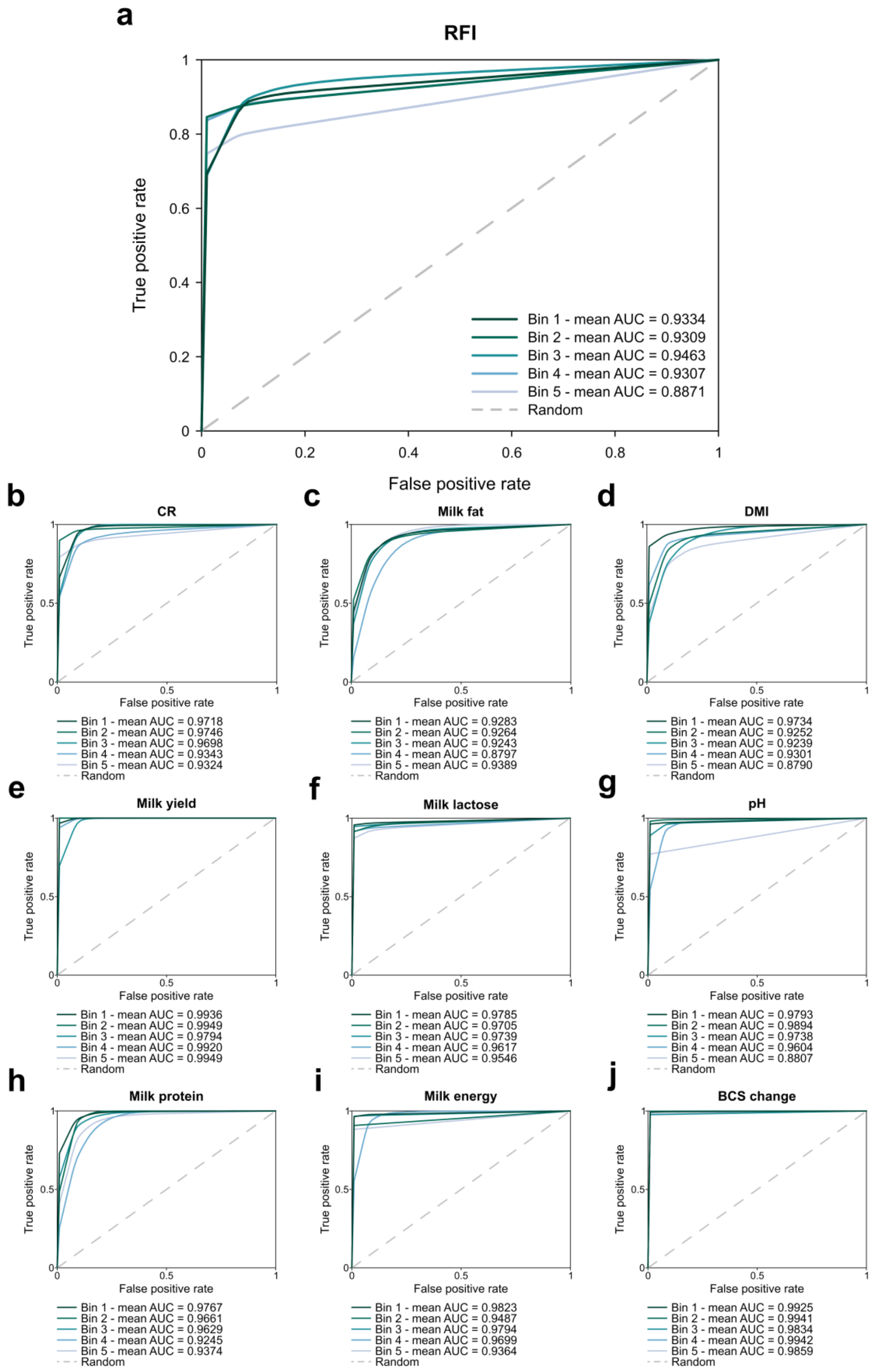

a
RFI
True positive rate
False positive rate
0
0.2
0.4
0.6
0.8
1
Bin 1 - mean AUC = 0.9334
Bin 2 - mean AUC = 0.9309
Bin 3 - mean AUC = 0.9463
Bin 4 - mean AUC = 0.9307
Bin 5 - mean AUC = 0.8871
Random
b
CR
Bin 1 - mean AUC = 0.9718
Bin 2 - mean AUC = 0.9746
Bin 3 - mean AUC = 0.9698
Bin 4 - mean AUC = 0.9343
Bin 5 - mean AUC = 0.9324
Random
c
Milk fat
Bin 1 - mean AUC = 0.9283
Bin 2 - mean AUC = 0.9264
Bin 3 - mean AUC = 0.9243
Bin 4 - mean AUC = 0.8797
Bin 5 - mean AUC = 0.9389
Random
d
DMI
Bin 1 - mean AUC = 0.9734
Bin 2 - mean AUC = 0.9252
Bin 3 - mean AUC = 0.9239
Bin 4 - mean AUC = 0.9301
Bin 5 - mean AUC = 0.8790
Random
e
Milk yield
Bin 1 - mean AUC = 0.9936
Bin 2 - mean AUC = 0.9949
Bin 3 - mean AUC = 0.9794
Bin 4 - mean AUC = 0.9920
Bin 5 - mean AUC = 0.9949
Random
f
Milk lactose
Bin 1 - mean AUC = 0.9785
Bin 2 - mean AUC = 0.9705
Bin 3 - mean AUC = 0.9739
Bin 4 - mean AUC = 0.9617
Bin 5 - mean AUC = 0.9546
Random
g
pH
Bin 1 - mean AUC = 0.9793
Bin 2 - mean AUC = 0.9894
Bin 3 - mean AUC = 0.9738
Bin 4 - mean AUC = 0.9604
Bin 5 - mean AUC = 0.8807
Random
h
Milk protein
Bin 1 - mean AUC = 0.9767
Bin 2 - mean AUC = 0.9661
Bin 3 - mean AUC = 0.9629
Bin 4 - mean AUC = 0.9245
Bin 5 - mean AUC = 0.9374
Random
i
Milk energy
Bin 1 - mean AUC = 0.9823
Bin 2 - mean AUC = 0.9487
Bin 3 - mean AUC = 0.9794
Bin 4 - mean AUC = 0.9699
Bin 5 - mean AUC = 0.9364
Random
j
BCS change
Bin 1 - mean AUC = 0.9925
Bin 2 - mean AUC = 0.9941
Bin 3 - mean AUC = 0.9834
Bin 4 - mean AUC = 0.9942
Bin 5 - mean AUC = 0.9859
Random

Figure S10

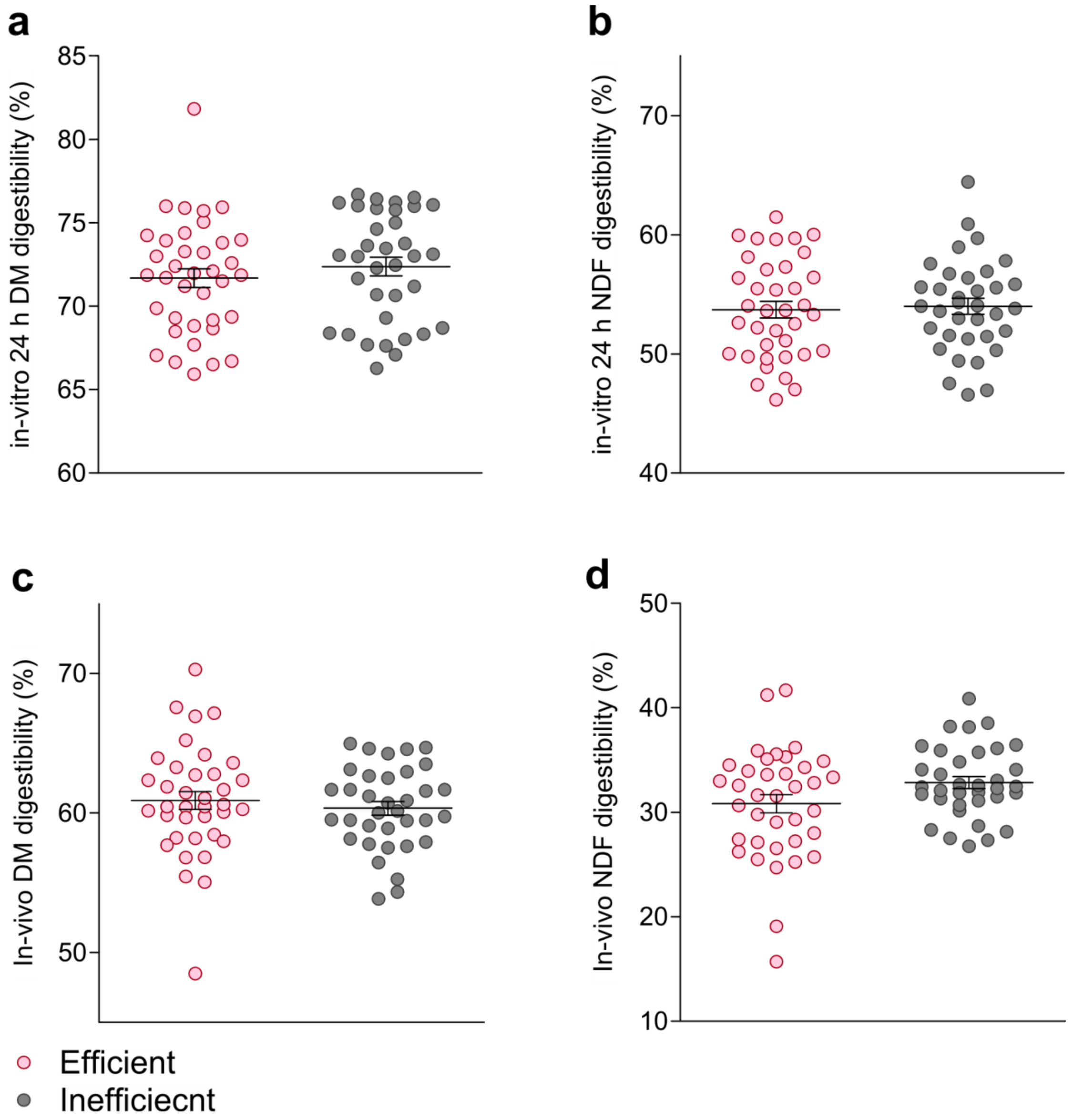

a
in-vitro 24 h DM digestibility (%)
85
80
75
70
65
60
b
in-vitro 24 h NDF digestibility (%)
70
60
50
40
c
In-vivo DM digestibility (%)
70
60
50
d
In-vivo NDF digestibility (%)
50
40
30
20
10
○ Efficient
● Inefficiecnt

Figure S11

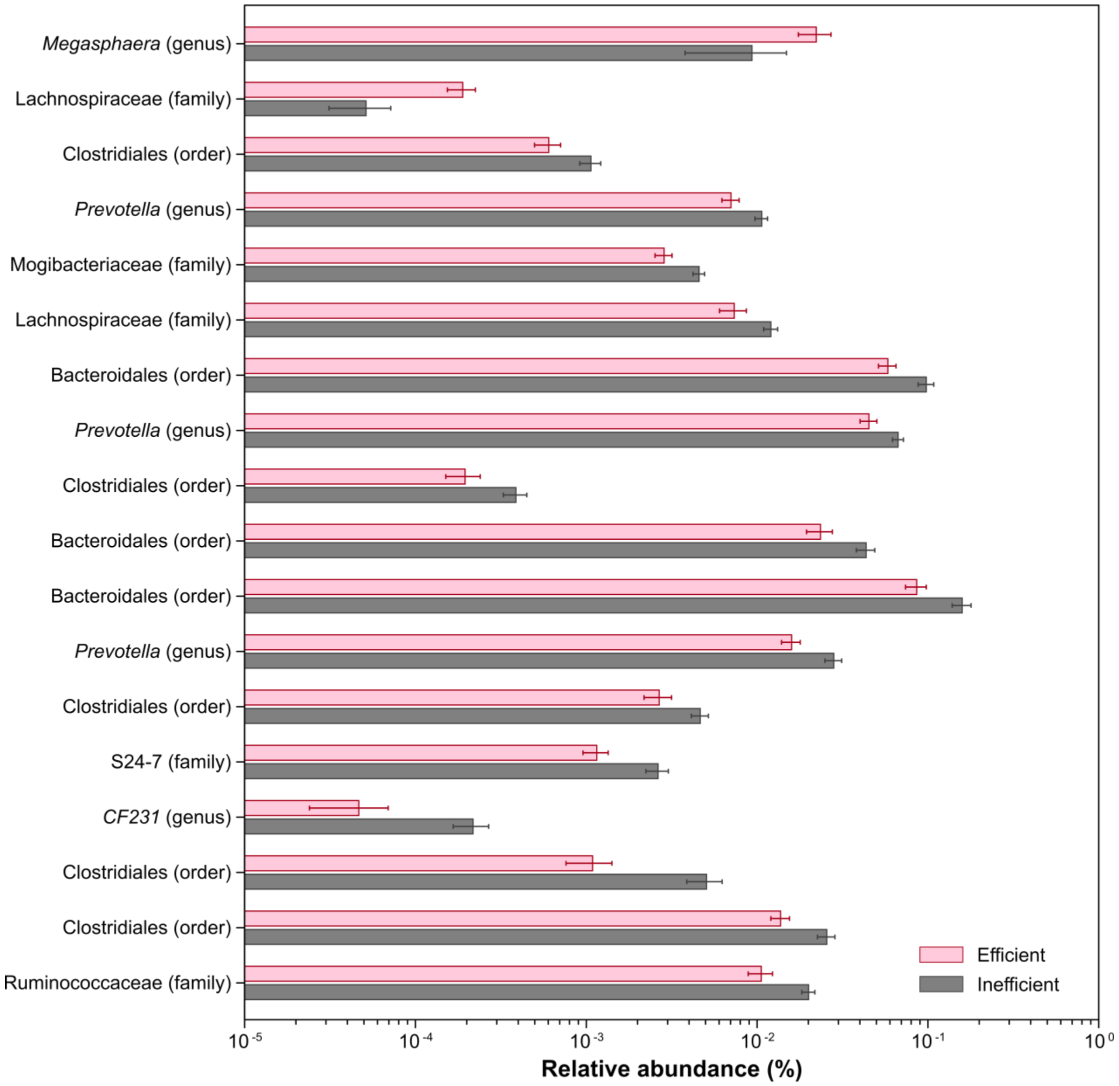

Megasphaera (genus)
Lachnospiraceae (family)
Clostridiales (order)
Prevotella (genus)
Mogibacteriaceae (family)
Lachnospiraceae (family)
Bacteroidales (order)
Prevotella (genus)
Clostridiales (order)
Bacteroidales (order)
Bacteroidales (order)
Prevotella (genus)
Clostridiales (order)
S24-7 (family)
CF231 (genus)
Clostridiales (order)
Clostridiales (order)
Ruminococcaceae (family)
Efficient
Inefficient
10^{-5}
10^{-4}
10^{-3}
10^{-2}
10^{-1}
10^{0}
Relative abundance (%)

Figure S12 PCoA

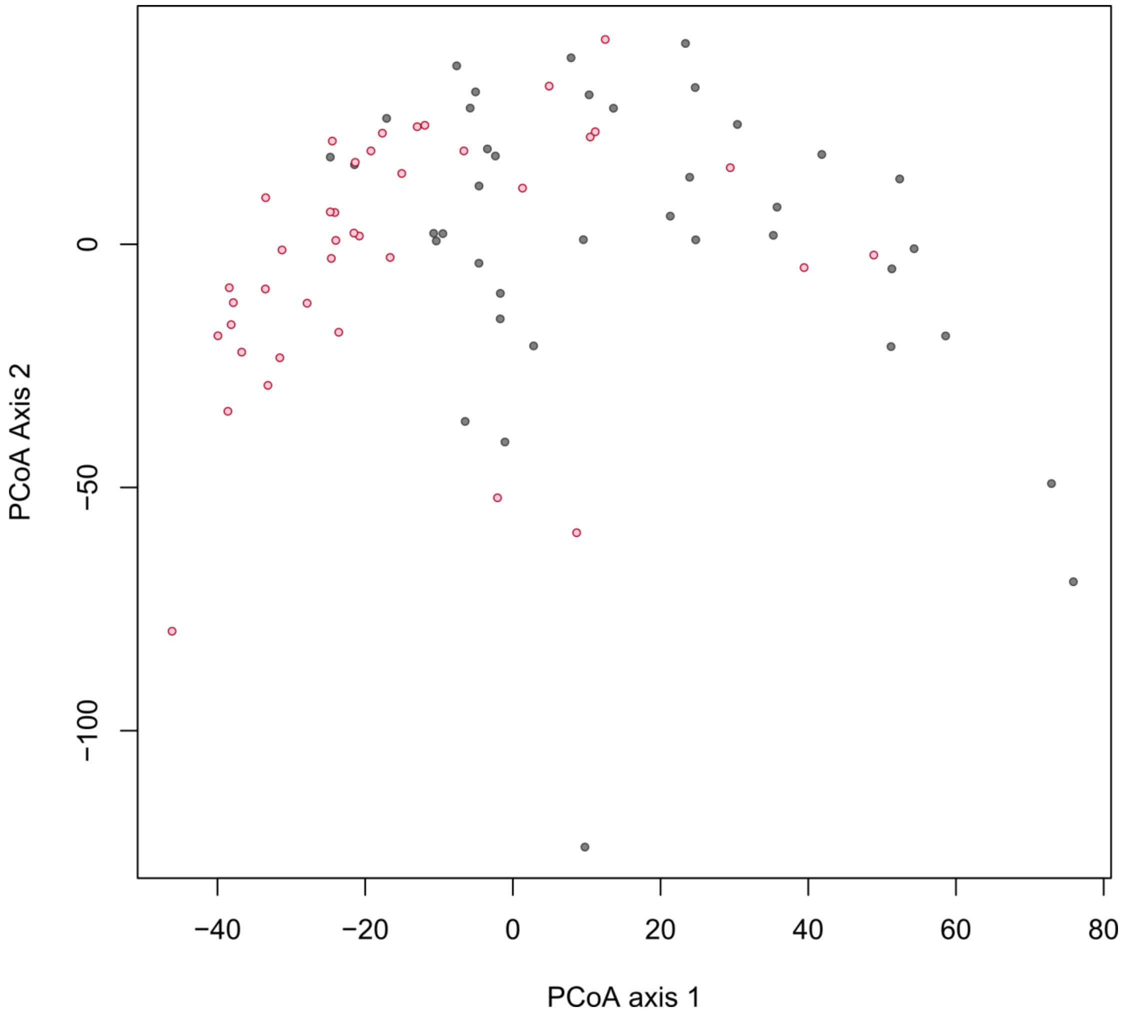


Figure S13

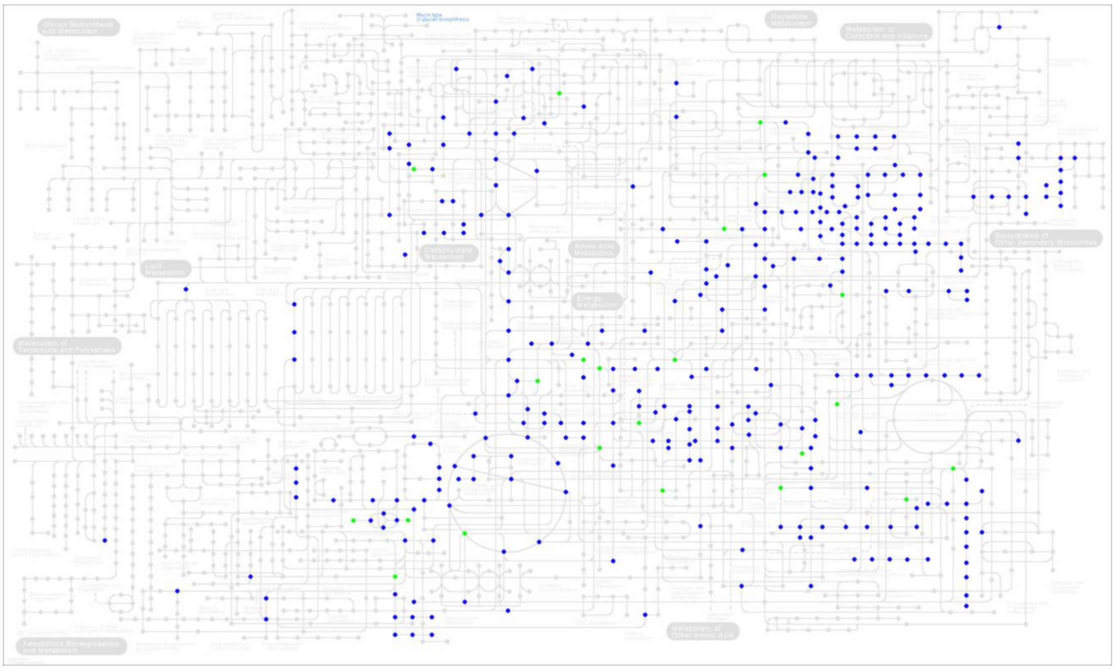

Glycan Biosynthesis and Metabolism
Mucin type O-glycan biosynthesis
Nucleotide Metabolism
Metabolism of Cofactors and Vitamins
Lipid Metabolism
Carbohydrate Metabolism
Amino Acid Metabolism
Energy Metabolism
Biosynthesis of Other Secondary Metabolites
Metabolism of Terpenoids and Polyketides
Xenobiotics Biodegradation and Metabolism
Metabolism of Other Amino Acids

Figure S14

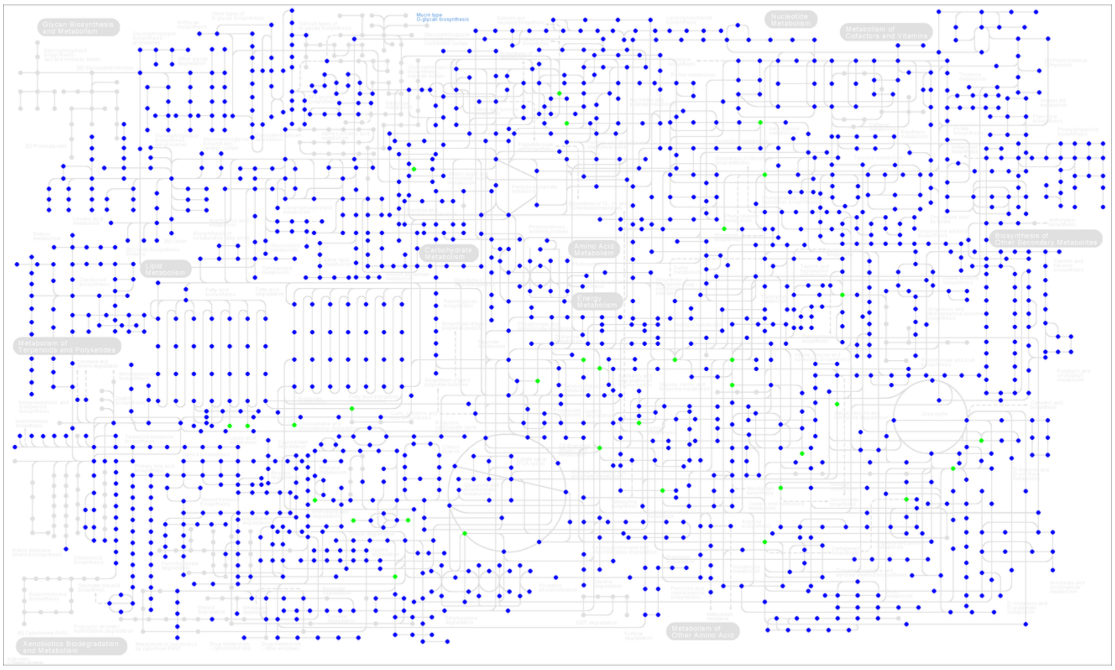

Glycan Biosynthesis and Metabolism
Mucin type O-glycan biosynthesis
Nucleotide Metabolism
Metabolism of Cofactors and Vitamins
Lipid Metabolism
Carbohydrate Metabolism
Amino Acid Metabolism
Energy Metabolism
Biosynthesis of Other Secondary Metabolites
Metabolism of Terpenoids and Polyketides
Xenobiotics Biodegradation and Metabolism
Metabolism of Other Amino Acids

Figure S15

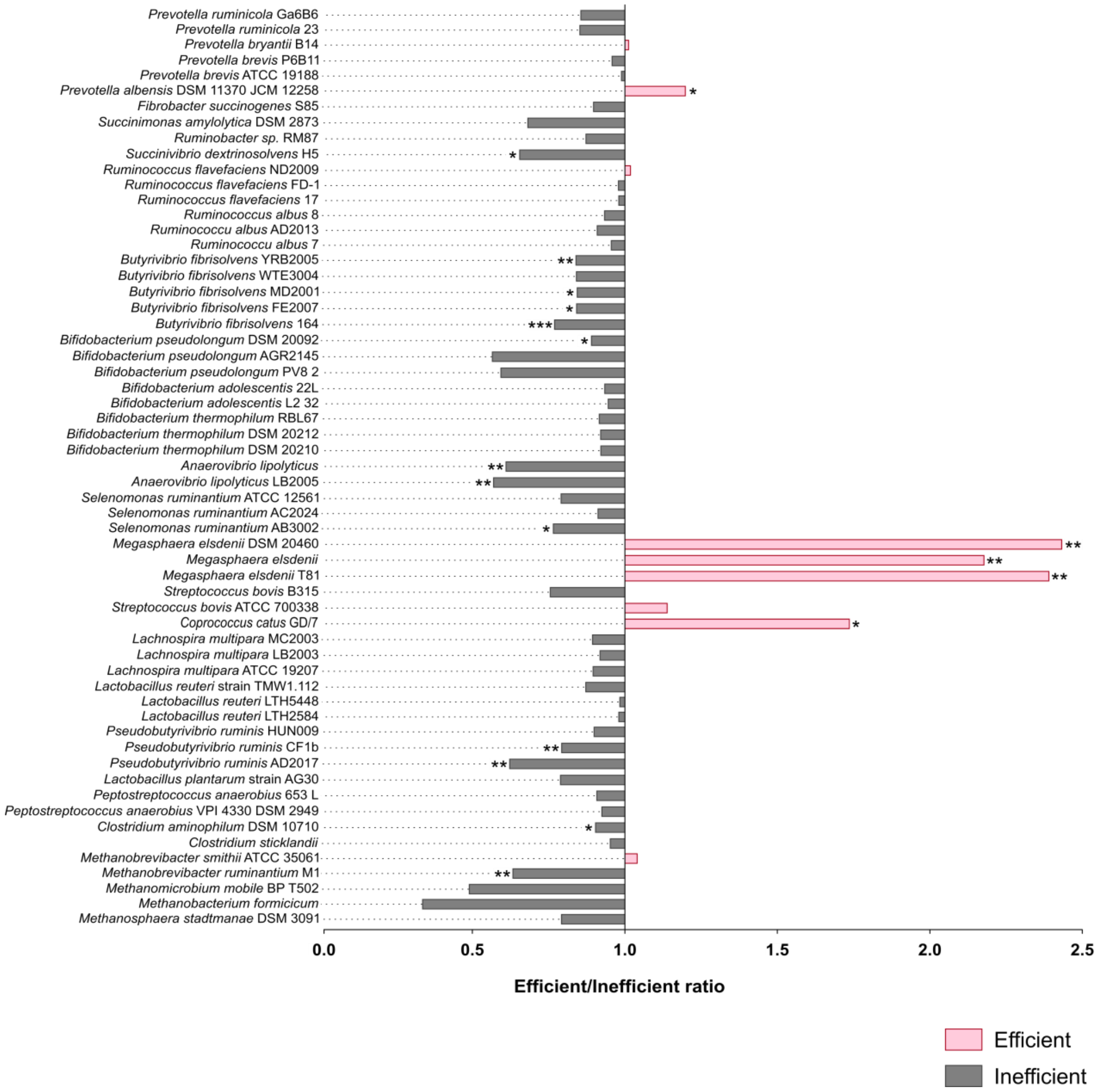

Prevotella ruminicola Ga6B6
Prevotella ruminicola 23
Prevotella bryantii B14
Prevotella brevis P6B11
Prevotella brevis ATCC 19188
Prevotella albensis DSM 11370 JCM 12258
Fibrobacter succinogenes S85
Succinimonas amylolytica DSM 2873
Ruminobacter sp. RM87
Succinivibrio dextrinosolvens H5
Ruminococcus flavefaciens ND2009
Ruminococcus flavefaciens FD-1
Ruminococcus flavefaciens 17
Ruminococcus albus 8
Ruminococcu albus AD2013
Ruminococcu albus 7
Butyrivibrio fibrisolvens YRB2005
Butyrivibrio fibrisolvens WTE3004
Butyrivibrio fibrisolvens MD2001
Butyrivibrio fibrisolvens FE2007
Butyrivibrio fibrisolvens 164
Bifidobacterium pseudolongum DSM 20092
Bifidobacterium pseudolongum AGR2145
Bifidobacterium pseudolongum PV8 2
Bifidobacterium adolescentis 22L
Bifidobacterium adolescentis L2 32
Bifidobacterium thermophilum RBL67
Bifidobacterium thermophilum DSM 20212
Bifidobacterium thermophilum DSM 20210
Anaerovibrio lipolyticus
Anaerovibrio lipolyticus LB2005
Selenomonas ruminantium ATCC 12561
Selenomonas ruminantium AC2024
Selenomonas ruminantium AB3002
Megasphaera elsdenii DSM 20460
Megasphaera elsdenii
Megasphaera elsdenii T81
Streptococcus bovis B315
Streptococcus bovis ATCC 700338
Coprococcus catus GD/7
Lachnospira multipara MC2003
Lachnospira multipara LB2003
Lachnospira multipara ATCC 19207
Lactobacillus reuteri strain TMW1.112
Lactobacillus reuteri LTH5448
Lactobacillus reuteri LTH2584
Pseudobutyrivibrio ruminis HUN009
Pseudobutyrivibrio ruminis CF1b
Pseudobutyrivibrio ruminis AD2017
Lactobacillus plantarum strain AG30
Peptostreptococcus anaerobius 653 L
Peptostreptococcus anaerobius VPI 4330 DSM 2949
Clostridium aminophilum DSM 10710
Clostridium sticklandii
Methanobrevibacter smithii ATCC 35061
Methanobrevibacter ruminantium M1
Methanomicrobium mobile BP T502
Methanobacterium formicicum
Methanosphaera stadtmanae DSM 3091
*
*
**
*
*
***
*
**
**
*
**
**
**
*
**
**
*
**
0.0
0.5
1.0
1.5
2.0
2.5
Efficient/Inefficient ratio
Efficient
Inefficient

Figure S16

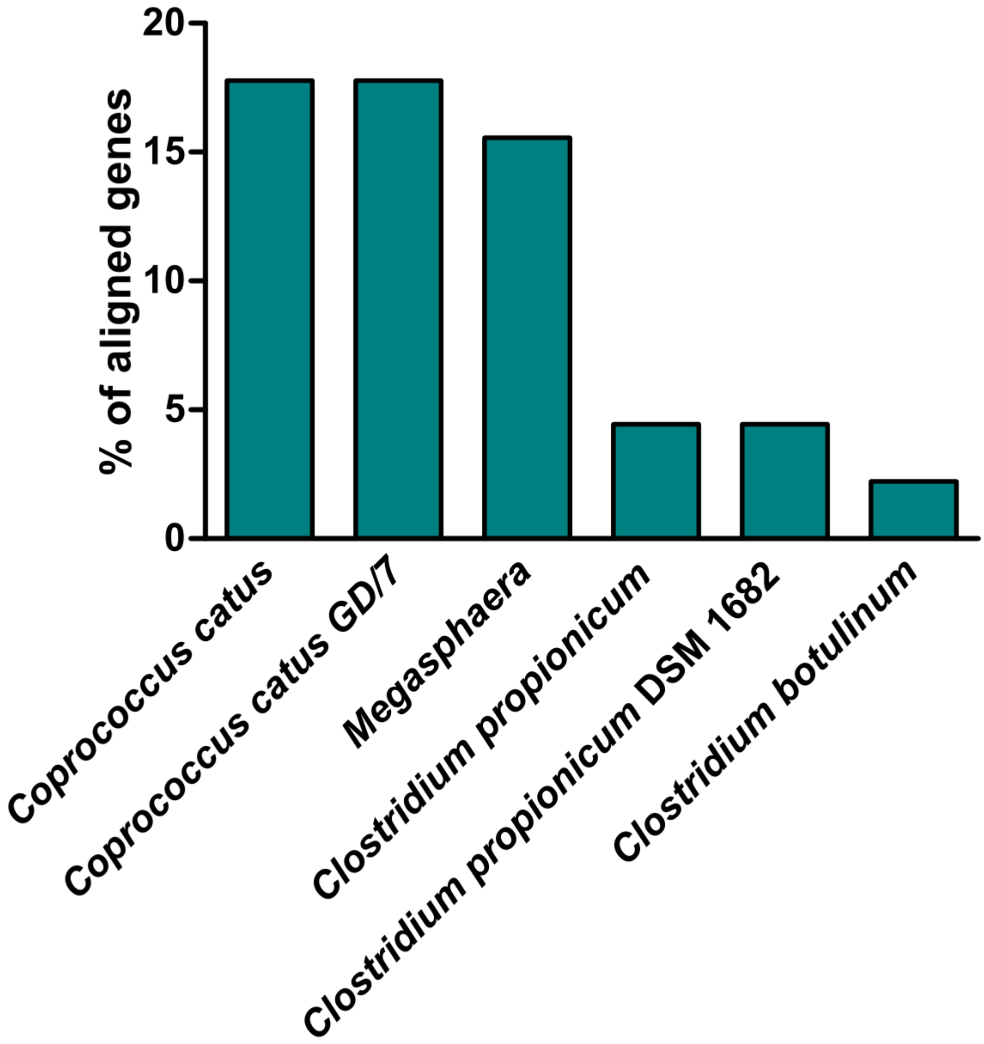

% of aligned genes
20
15
10
5
0
Coprococcus catus
Coprococcus catus GD/7
Megasphaera
Clostridium propionicum
Clostridium propionicum DSM 1682
Clostridium botulinum

Table S1

Acetate	Butyrate	Propionate	Valerate	Isovalerate	Lactate	Methane
2.8.3.8	2.7.2.7	1.3.8.7	1.2.1.19	1.1.1.26	1.1.1.27	1.12.98.1
	2.8.3.8	2.8.3.1	1.2.1.3	1.2.1.24	1.1.1.28	1.12.982
		4.1.1.41	1.2.1.5	1.2.1.3	1.13.12.4	1.2.99.5
		4.2.1.54	1.2.1.47	1.2.1.4	1.2.1.2	1.5.1
		5.4.99.2	1.2.3.1	1.2.1.5	3.1.2.6	1.5.98.1
		51.99.1	1.2.1.8	1.2.1.77	4.2.1.130	1.5.98.2
			2.8.3.8	1.2.1.8		1.8.98.1
			3.1.1.1	1.2.7.5		2.1.1.86
			3.1.1.22	3.1.1.8		2.8.4.1
			3.1.1.8	3.1.2.20		23.1.101
				3.5.5.5		3.5.4.27
				3.5.5.7		

Table S2

Diversity index	Efficient	Inefficient
Alpha diversity species**	5.87 ± 0.05	6.092 ± 0.03
Alpha diversity genes**	14.32 ± 0.06	14.53 ± 0.028
Dominance species*	0.01 ± 0.001	0.0079 ± 0.0003
Dominance genes*	$1.14 \times 10^{-21} \pm 9.72 \times 10^{-23}$	$0.86 \times 10^{-21} \pm 6.15 \times 10^{-23}$

Table S3

Metabolite	Efficient (mM)	Inefficient (mM)	Efficient (mM per g/L OM)	Inefficient (mM per g/L OM)
Acetate	35.61 ± 1.24	33.26 ± 1.55	11.84 ± 0.26	11.1 ± 0.42
Propionate	22.32 ± 0.82	19.45 ± 0.84	7.45 ± 0.23^{a}	6.49 ± 0.22^{b}
Isobutyrate	1.5 ± 0.05	1.35 ± 0.06	0.5 ± 0.01	0.45 ± 0.02
Butyrate	22.03 ± 1.16	19.6 ± 0.9	7.25 ± 0.27^{a}	6.52 ± 0.22^{b}
Isovalerate	2.94 ± 0.15	2.5 ± 0.13	0.97 ± 0.03^{a}	0.84 ± 0.04^{b}
Valerate	3.75 ± 0.14	3.15 ± 0.14	1.25 ± 0.04^{a}	1.05 ± 0.04^{b}
Total VFAs	88.14 ± 3.13	79.3 ± 3.28	29.26 ± 0.62^{a}	26.46 ± 0.82^{b}

Table S4

ko number	ko description
04974	Protein digestion and absorption
00250	Alanine, aspartate and glutamate metabolism
00260	Glycine, serine and threonine metabolism
00270	Cysteine and methionine metabolism
00290	Valine, leucine and isoleucine biosynthesis
00300	Lysine biosynthesis
00330	Arginine and proline metabolism
00340	Histidine metabolism
00350	Tyrosine metabolism
00360	Phenylalanine metabolism
00380	Tryptophan metabolism
00400	Phenylalanine, tyrosine and tryptophan biosynthesis
00680	Methane metabolism