Microbial Co-presence and Mutual-exclusion Networks in the Bovine Rumen Microbiome

Haiying Wang, Huiru Zheng*
School of Computing and Mathematics
Ulster University
Northern Ireland, United Kingdom
*h.zheng@ulster.ac.uk

Richard J. Dewhurst, Rainer Roehe
Future Farming Systems,
Scotland's Rural College,
Edinburgh, United Kingdom

Abstract— **The recognized significance of rumen microbiome has inspired efforts to examine the composition of rumen microbial communities in a large scale. One of the key research areas is to infer association and dependencies between members of rumen microbial communities through correlation analysis. However, it has been found that due to the compositional nature of data, simply applying correlation-based techniques to the analysis of relative abundance of microbial genes may produce artefactual correlation and loss of information. In an attempt to mitigate the compositional effect on the analysis of rumen microbiome data, this study applied a framework including a compendium of two correlation measures and three dissimilarity metrics that are intrinsically robust to compositionality. Based on the inference of significant positive and negative associations, co-presence and mutual-exclusion networks were constructed. The corresponding modules associated with methane production were identified. The modules are highly enriched with microbial genes associated with methane emissions and encoding enzymes involved in the methane methanogensis pathway. In comparisons to previous studies, our analysis demonstrates that deriving microbial associations based on the correlations between relative abundances may not only lead to missing information but also produce spurious associations.**

Keywords—rumen microbiome; compositional data; methane emission; co-occurrence networks; correlation analysis

I. INTRODUCTION

Thanks to their unique digestive system, ruminant livestock play an important role in human food production, producing high-quality milk and meat products from otherwise indigestible food components. Nevertheless, as a direct result and an inevitable outcome of rumen fermentation, methane production from ruminants contributes significantly to global anthropogenic greenhouse gas emissions [1], [2]. It has been estimated that ruminant livestock produce around 100 million tonnes of methane to the atmosphere each year representing the biggest man-made methane source after rice agriculture [3].

The fermentation process taking place in the rumen is a result of fine-tuned cooperation between the host animal and rumen microorganisms predominantly consisting of bacteria, archaea, protozoa and fungi [4]. These microorganisms play a vital role in their host's physiology and without a healthy microbial population in the rumen, ruminants cannot function properly [5].

Various studies have demonstrated the influence of rumen microbial communities on animal phenotypes [6], [7].

The recognized significance of rumen microbiome has inspired efforts to examine the composition of rumen microbial communities on a large scale. Due to the ability to reveal the full spectrum of microbial diversity, next-generation sequencing (NGS)-based metagenomics analysis has been widely applied. Examples include the study conducted by Henderson et al. [8], which performed metagenomics analysis of 742 samples collected from 32 animal species and 35 countries and found that rumen microbial community composition varies with diet and host. However, similar bacteria and archaea are found to dominate in nearly all samples across a wide geographical range. Lengowski et al. [9] examined ruminal microbial community composition alterations during adaption to a rumen simulation system and to the forages. It has been shown that the ruminal microbial community can be influenced significantly by forage source and sampling time. Nevertheless, a dynamic stable microbial community composition was achieved after 48h under the given incubation conditions on the domain level. Using young ruminants subjected to different microbial-modulating interventions, Morgavi et al. [7] examined rumen bacterial and archaeal communities and the interactions between microbial populations and the association with the host. The recent study carried out by Roehe et al. provides a comprehensive insight into host-microbe interaction in the rumen and highlighted that the host animal controls its own microbiota to a significant extent [6].

One of the key research areas in NGS-based metagenomics data analysis is to infer association and dependencies between members of microbial communities through correlation analysis [10]. For example, Williams et al. [11] introduced a framework to explore biological interactions occurring within microbial communities, in which the strength of correlation is derived from the calculation of the Spearman's correlation. The co-occurrence analysis can be performed at multiple scales ranging from the community level down to pairwise interactions between microbial taxa. Based on the relative abundance of 1570 KEGG genes across 8 samples, Roehe et al. [6] constructed a co-abundance network where nodes represent microbial genes and edges reflect the correlation in their abundance. They have successfully identified a close sub-network of the microbial genes associated with feed conversion efficiency and methane

emission respectively. Wang et al. [12] applied a random matrix theory-based approach for determination of the correlation threshold used to construct the co-abundance microbial network.

Despite encouraging results have been obtained, the correlation-based approaches to the inference of associations between microbial genes exhibit some limitations [10], [13]. Due to the nature of data generation and the normalization process involved, the abundance derived from NGS is a relative measurement associated with each microbial gene. As a such, abundances of microbial genes estimated under certain condition are not completely independent of each other. It has been shown that simply applying correlation-based techniques to the analysis of such compositional data may produce misleading results [13].

This study aims to apply a new framework including a compendium of correlation and dissimilarity measures to mitigate the effect of compositionality on the analysis of rumen microbiome data. The main objective is to infer both co-presence and mutual exclusion networks associated with methane emission. The remainder of this paper is organized as follows. Section II briefly describes the framework and methodology used in this study, including dataset and an ensemble of correlation of dissimilarity metrics. Section III presents the results and discussion. The paper concludes with a summary of contributions and limitations of this study followed by the direction of future research.

II. METHODOLOGY

We followed the approach introduced in [13] by using a compendium of similarity/dissimilarity measures for the analysis of rumen metagenomics data which include the relative abundance of 1570 microbial genes [6]. Without loss of generality, a network including 1000 top-ranking and 1000 bottom-ranking edges was constructed for each measure. To assess the significance of scores associated with each edge, we applied the Permutation-Renormalization and Bootstrap (ReBoot) method [13], which can construct a null distribution that reflects the compositional nature of the data. After merging p-values and multiple testing corrections, a final network consisting of significant co-presence (positive interaction) and mutual-exclusion (negative interaction) patterns was extracted. The resulting network was further examined in terms of topological analysis, biological relevance and pathway involvement. The key steps involved in the study are illustrated in Fig. 1.

A. Rumen Meteganomics Data

The data applied in this research was released by Roehe and his colleagues in a study [6] in which a 2×2 factorial design experiment of breed types and diets was performed using 72 steers from a two-breed rotational cross between Aberdeen-Angus (AA) or Limousin cattle (LIM). Based on genomic analysis of rumen contents taken from 8 extreme animals balanced for breed type and diet, a total of 3970 KEGG gene orthologues were identified, of which 1570 genes showed a relative abundance of more than 0.001%. The detailed description of data generation can be found in [2] and [6].

B. An Emsemble of Similarity and Disimilarity Measures

In order to mitigate the effect of compositionality on the analysis of rumen microbiome data, a compendium of two correlation measures, i.e. Spearman and Pearson correlations, and three dissimilarity metrics that are intrinsically robust to compositionality [13], i.e. Bray-Curtis dissimilarity (BC), Kullback-Leibler dissimilarity (KL), and Jensen-Shannon dissimilarity (JS) were utilized.

Let x and y be two vectors containing relative abundances across samples for two microbial genes. The three dissimilarities are defined as follows.

$$BC(x,y) = 1 - \frac{2\sum_k |x_k - y_k|}{\sum_k x_k + \sum_k y_k} \quad (1)$$

$$KL(x,y) = \sum_k (x_k \times \log\frac{x_k}{y_k} + y_k \times \log\frac{y_k}{x_k}) \quad (2)$$

$$JS(x,y) = \sum_k (x_k \log\frac{2x_k}{(x_k + y_k)} + y_k \log\frac{2y_k}{(x_k + y_k)}) \quad (3)$$

C. Statistical Significance of Emsemble Scores

To evaluate the significance of the association accounting for compositionality, we applied a nonparametric test based on the ReBoot method introduced in [13]. Unlike a standard procedure based on permutation test that essentially removes compositional effects and thus fails to identify spurious compositional correlations, the ReBoot method introduces sample-wise renormalization after permuting the abundance across samples. Such an approach leads to the construction of compositionality-aware null distribution. Comparing this null distribution to a standard bootstrap confidence interval, an appropriate significance level of the observed correlation can be established.

In this study, both permutation and bootstrap score distributions were computed with 100 iterations. Any edge with a score that falls outside of the bootstrapped confidence interval was removed.

D. Network Merging

After constructing a measure-specific network in which a node stands for a microbial gene and a score associated with each edge represents the strength of the association between two genes, we combined all the networks using Brown's method [14] which is an extension of Fisher's method for combining dependent p-values. The merged p-values on each final edge were adjusted using the Benjamini-Hochberg false discovery rate (FDR) correction and the final network was thresholded at a q-value less than 0.05.
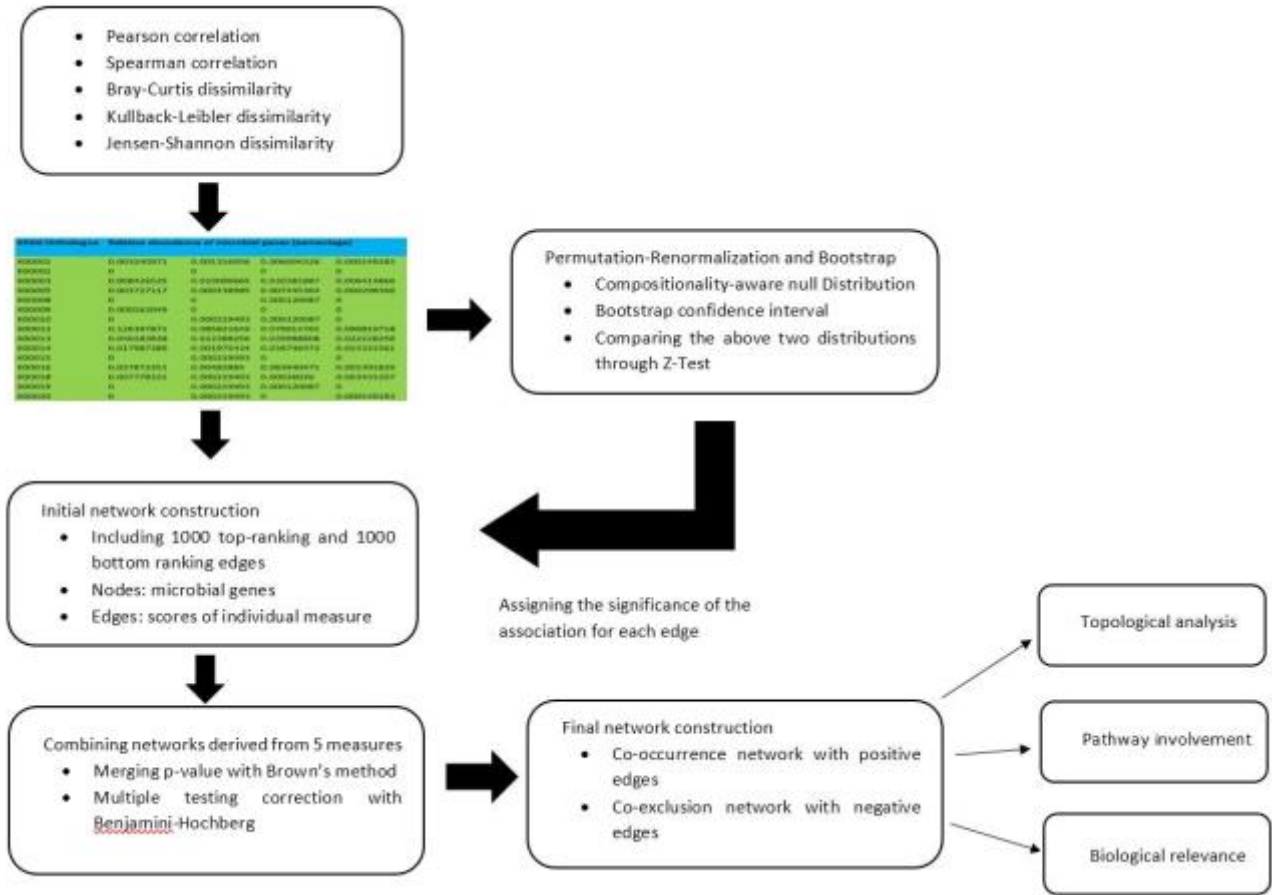
Fig. 1 A diagram to illustrate the key steps involved in the study.

III. RESULTS

A. Co-presence and Mutual-exclusion networks

The final association networks were constructed using the CoNet app [15] which offers a variety of approaches for inference of biological meaning using Cytoscape [16]. Only the interactions with an FDR corrected p value (q-value) less than 0.05 were kept. The co-presence network (Fig. 3) consists of 790 nodes (microbial genes) and 2106 edges with positive scores while the mutual-exclusion network is composed of 763 negative interactions between 473 microbial genes as shown in Fig. 4. The top 5 hub nodes are shown in Tables 1 and 2 respectively.

TABLE I TOP5 HUB NODES IN THE CO-PRESENCE NETWORK

Node degree	KEGG orthologue	Description of microbial genes
30	K07161	Uncharacterized protein
29	K00805	EC:2.5.1.30
25	K00169	EC:1.2.7.1
25	K06310	Spore germination protein
24	K02007	Cobalt/nickel transport system permease protein
24	K04070	EC:1.97.1.4
24	K00111	EC:1.1.5.3

TABLE II TOP5 HUB NODES IN THE MUTUAL-EXCLUSION NETWORK

Node degree	KEGG Orthologue	Description of microbial genes
77	K02986	Small subunit ribosomal protein S4
59	K06013	EC:3.4.24.84
57	K03780	EC:4.2.1.32
35	K02874	Large subunit ribosomal protein L14
35	K02931	Large subunit ribosomal protein L5
35	K02837	Peptide chain release factor RF-3

The distribution of the number of the interactions supported by the metrics is shown in TABLE III. There are 537 and 382 edges in co-presence and mutual-exclusion networks, respectively, which are supported by at least two metrics with a q value below 0.05. Interestingly all the pairs supported by the Spearson's correlation in the co-presence network exhibit a perfect monotonic relationship. None of links in the mutual-exclusion network is supported by all the 5 metrics used while in the co-presence network a total of 10 pairs are significantly supported by all 5 metrics as shown in TABLE IV.

Metrics	Co-presence network		Mutual-exclusion network	
	Threshold	Number of pairs	Threshold	Number of pairs
Pearson correlation	0.996	775	-0.913	397
Spearman correlation	1.0	975	-0.952	335
Bray-Curtis	0.049	490	0.917	135
Kullback-Leibler	0.018	441	9.17	136
Jensen-Shannon	0.0017	36	0.563	140

TABLE IV INTERACTIONS SUPPORTED BY BOTH CORRELATION METRICS AND THREE DISSIMILARITY MEASURES

| Interaction type	Interactor A	Interactor B	Corrected p value
co-presence	K13812	K00577	0.000
co-presence	K13812	K00400	0.000
co-presence	co-presence	K14105	0.000
co-presence	K00577	K00400	0.000
co-presence	K00320	K14127	0.000
co-presence	K07388	K01623	0.000
co-presence	K00123	K14128	0.000
co-presence	K00123	K03388	0.000
co-presence	K09154	K03042	0.000
co-presence	K03832	K03303	0.000
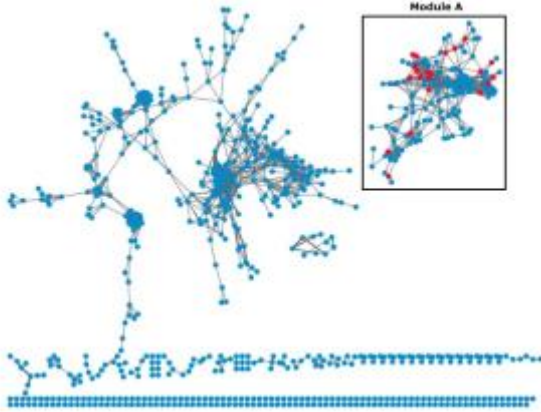
Fig. 3 Significant co-presence relationships among the abundances of KEGG microbial genes in the rumen microbiome. The width of edges is proportional to the level of significance of supporting evidence. Red nodes represent genes encoding enzymes that are directly involved in the methane production pathway.

B. Biological relevance

Given that the extreme animals selected in the data collection carried out by SRUC [2], [6] were based on methane emissions, we first checked the distribution of methane specific-microbial genes in both networks. The level of the enrichment of trait-
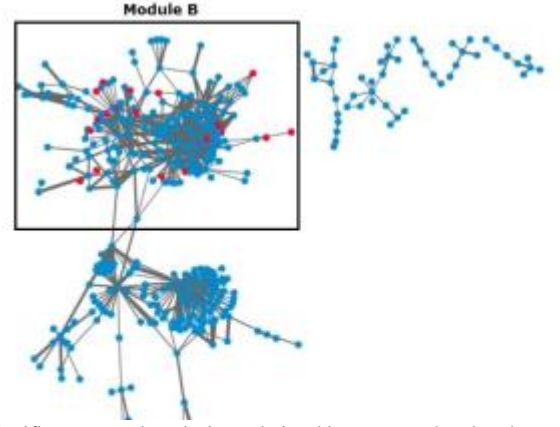

Fig. 2 Significant mutual-exclusion relationships among the abundances of KEGG microbial genes in the rumen microbiome. The width of edges is proportional to the level of significance of supporting evidence. Red nodes represent genes encoding enzymes that are directly involved in the methane production pathway.

specific genes can be quantitatively expressed by the hypergeometric distribution probability calculated as follows.

$$p = 1 - \sum_{i=0}^{m-1} \binom{m}{i}\binom{N-m}{n-i}/\binom{N}{n} \qquad (4)$$

where m is the number of microbial genes found in a module, i is the number of genes in the module associated with certain trait, N is the total number of microbial genes contained in the network and n is the number of trait-specific genes associated found in the network.

We found that, out of 31 genes that are directly involved in the methane production pathway studied in Wallace et al. [2], twenty-two and nineteen were found in the co-presence and mutual-exclusion networks respectively and all of them are grouped in Module A and B respectively ($p < 10^{-15}$). Furthermore, nineteen out of 20 methane emission specific genes identified by Roehe et al. [6] are contained in the co-presence network and grouped together in Module A ($p < 10^{-11}$). Based on these figures, one may confidently assume that Modules A and B are co-occurrence networks significantly associated with methane production.

We then turned to the topological analysis of Modules A and B. As depicted in TABLE V, both modules have a low average path length of less than four in comparison to 6 found in random networks on average [13]. Surprisingly, the clustering coefficient of Module B is equal to 0, indicating that none of the neighbours of nodes in Module B are connected. Moreover, Module B is more heterogeneous than Module A as indicated by the metric of network heterogeneity which reflects the tendency of a network containing hub nodes.

TABLE V THE CHARACTERISTICS OF MODULES A AND B

Parameters	Module A	Module B
Nodes	143	186
Edges	694	444
Clustering coefficient	0.572	0.000
Characteristics path length	3.628	3.777
Average number of neighbours	9.706	4.774
Network heterogeneity	0.676	1.517
Network density	0.068	0.026
Network centralization	0.145	0.296

There are two hub nodes in Module B having a degree more than 50. The top node (K06013, STE24 endopeptidase [EC:3.4.24.84]) exhibits significant mutual exclusion patterns over samples with 59 microbial genes supported by all three dissimilarity measures (BC, KL and JS) with a q-value less than 0.05. Similarly, K03780 is linked to 57 microbial genes in the form of strong mutual exclusions ($q < 0.05$).

In Module A, the most connected node is an uncharacterized protein (K07161), which shows significant co-presence patterns with 30 microbial genes across samples with a corrected p value less than 0.00001. In particular, it exhibits a similar abundance pattern (Fig. 4) across 8 samples with five genes (K00581, K00125, K00202, K00402 and K00401) encoding enzymes involved in the methane production pathway and four microbial genes associated with methane emission (K00581, K00125, K01499 and K00169). As shown in Fig. 4, K07161 has a relative high level of abundance in the samples in the high methane emission group, suggesting this uncharacterized protein might be involved in the methane production pathway.
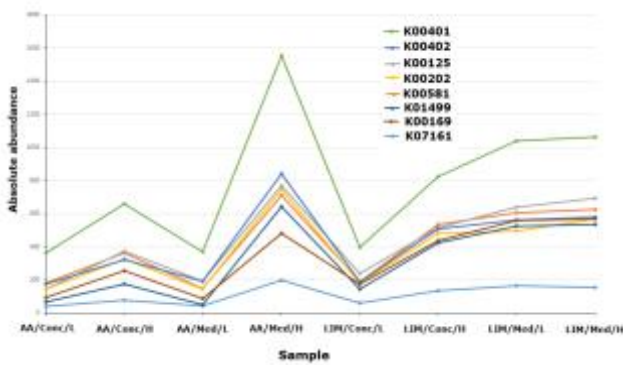

Fig. 4 The absolute abundance profile of the co-presence pattern observed in K07161 and seven microbial genes relevant to methane emission. AA: Aberdeen Angus, LIM: Limousin, Conc: concentrate diet, Med: medium concentrate diet, L: Low methane emission, H: High methane emission.

C. Pathway analysis

The interaction partners in the co-presence network for genes encoding enzymes involved in methanogenesis are listed in TABLE VI. As expected, there are a number of strong positive interactions among methane specific microbial genes. For example, Fig. 5 shows the interactions among 19 methane emission specific microbial genes in the form of co-presence. The significant co-presence patterns were also observed among genes encoding interacting enzymes. Examples include significant positive associations between K00125 encoding formate dehydrogenase (EC:1.2.1.2) and K00201 encoding formylmethanofaran dehydrogenase (EC:1.2.99.5). Similar observation is made between genes K00443 and K03388 encoding interacting enzymes, heterodisulfide reductase (EC:1.8.98.1) and coenzyme F420 hydrogenase (EC:1.12.98.1) respectively. However, no mutual exclusion patterns have been found among genes either associated with methane emissions or involved in the methane production pathway.

TABLE VI INTERACTION PARTNERS IN THE CO-PRESENCE NETWORK FOR KEY UNITS INVOLVED IN METHANOGENESIS

Enzymes	Enzyme encoding genes	Degree	Methane specific interaction partners
EC:1.2.1.2	K00123	14	K00125, K00201, K00399, K00401, K0051, K03388, K14128
	K00125	21	K03388, K00402, K00202, K00443, K00123, K00401, K00581, K14128, K14128
EC:1.2.99.5	K00200	7	K00203, K00170
	K00201	6	K03388, K00123, K14128, K00399
	K00202	18	K03388, K00581, K00125, K14128, K00401, K00402
	K00203	10	K14128, K00169, K00584, K00200
	K00205	3	K00672
EC:2.3.1.101	K00672	9	K00205, K00401, K00577, K00400, K12812
EC:3.5.4.27	K01499	18	K00169
EC:1.5.99.9/ EC:1.5.98.1	K00319	11	-
EC:1.5.99.11 /EC:1.5.98.2	K00320	8	-
EC:2.1.1.86	K00577	17	K00441, K00672, K03390, K13812, K00400, K14123
	K00581	20	K00125, K00202, K00402, K14128, K00399, K00123, K02288
	K00584	5	K00169, K00203, K00169
EC:2.8.4.1	K00399	6	K03388, K00201, K00123, K14128, K00581
	K00401	21	K00125, K14128, K00202, K00402, K00123, K03388, K00443, K00440
	K00402	19	K00125, K00581, K00202, K03388, K00401
EC:1.12.98.1	K00440	7	K00401, K00443
	K00441	7	K00672, K00577, K00400, K13812
	K00443	7	K00125, K00401, K03388, K00440
EC:1.8.98.1	K03388	18	K00399, K00201, K00123, K14128, K00125, K00581, K00202, K00401, K00402, K00443
	K03389	3	-
	K03390	14	K00577, K13823, K00400, K14123

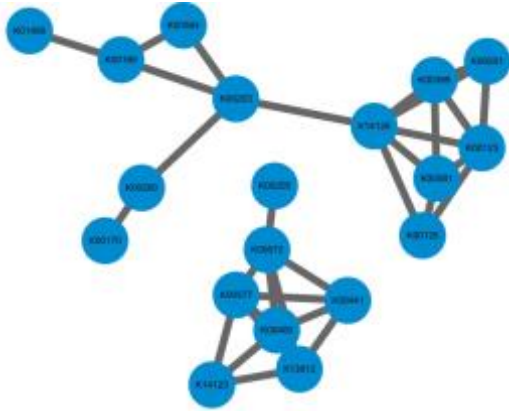Fig. 5 Positive interactions among 19 microbial genes associated with methane emission found in Module A.

D. Comparisons with previous studies

In comparison to our previous studies [6], [12] in which a co-abundance network was constructed using Pearson correlation coefficient to measure the similarity between two genes based on their relative abundances, the current study introduces two major improvements: (1) the system is able to construct a network containing either co-presence or mutual-exclusion patterns; and (2) the compositional effect in the analysis of rumen microbial communities based on relative abundance data is mitigated through an ensemble approach [13] containing two correlation measurements and 3 disisimilarity metrics.

It has been shown that assessing relationships between relative abundance profiles purely based on correlation-based metrics may lead to spurious correlation. For example, the actual counts and relative abundances which sum to one of K02986 and K00790 are shown in Fig. 6. Two microbial genes only have a weak negative correlation with Pearson correlation coefficient equal to -0.291 in Fig. 6(a) while they exhibit a strong negative correlation based on their relative abundance (-0.995) in Fig. 6(b). Another example is the correlation between K07636 and K03742 which show a strong positive correlation well above the threshold (0.99) used in our previous study [12] to construct the co-abundance network. However, if we look at their actual abundance profile, they have a correlation of 0.948 which is below the threshold identified (0.99).

Our results also provide the evidence that analysis of relative abundance profiles purely based on correlation-based metrics may lead to loss of information. For example, out of 31 genes encoding enzyme directly involved in the methane production pathway, 22 were found in Module A which is strongly associated with methane emission while only 18 were included in the module found in our previous study [12]. Out of 2106 positive interactions included in the co-presence network and 763 negative associations in the mutual-exclusion network, only 775 are found to have an absolute value of the Pearson correlation greater than the threshold identified in [12]. In particular, there is only one pair of microbial genes, i.e. K00790 and K02986, having a negative correlation less than -0.99. On

the other hand, the interactions supported by the Pearson correlation measure found in the co-presence network have a positive value higher than 0.995, suggesting that inferring a microbial association network solely based on a correlation measure may not only lead to missing information but also cause artefactual associations. A close examination of the interaction partners of K00123 (formate dehydrogenase major unit [EC:1.2.1.2]) in the co-expression network confirms our analysis. K00123 is found to be associated with methane emission [6] and involved in the methane production pathway [2]. It has 14 significant positive interactions with a corrected p value less than 0.05. However, more than half of interactions have a Pearson correlation coefficient less than 0.99 and thus were not included in our previous study including the interactions with another subunit of formate dehydrogenase (K00125) and K00201(formylmethanofuran dehydrogenase subunit B [EC:1.2.99.5]).
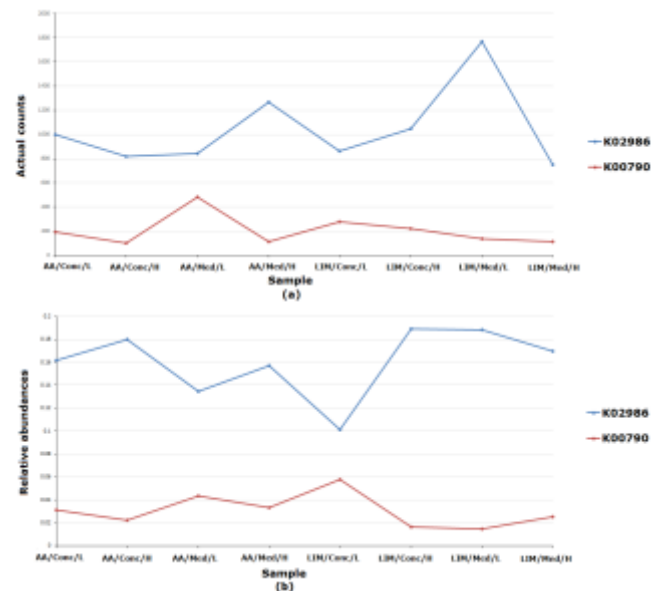

Fig. 6 The abundance profiles of two microbial genes, i.e. K02986 and K00790 across 6 samples: (a) actual counts; and (b) relative abundance

IV. CONCLUSIONS

Advances in NGS-based approaches have opened up new avenues in rumen microbial ecology studies. One of the key research areas is to infer association and dependencies between members of rumen microbial communities through correlation analysis. However, it has been found that due to the nature of data generation and the normalization process involved, traditional correlation-based analysis exhibits some significant limitations [10], [13]. Using a compendium of 2 correlation and 3 dissimilarity measures, this paper applied a new framework for the analysis of rumen metagenomics data which include the relative abundance of 1570 microbial genes. Robust co-presence and mutual exclusion networks were constructed which contains 1000 top-ranking and 1000 bottom-ranking edges with an FDR corrected value less than 0.05. Based on the assessment of level of enrichment of trait-specific microbial

genes, co-presence and mutual-exclusion modules associated with methane production, i.e. Modules A and B, were identified. While there exist strong positive correlations between methane specific genes in Module A which includes143 nodes and 694 positive edges, no mutual-exclusion patterns were observed among genes associated with methane emissions and encoding enzymes included in the methane production pathway in Module B which consists of 186 genes and 444 negative associations. The results demonstrate that deriving microbial associations based on the correlations between relative abundances may not only lead to loss of information but also produce spurious associations.

In this study, we adopted the parameters used [13] and the network construction was based on the analysis of the 2000 edges with extreme scores, i.e. 1000 top-scores representing strong positive interactions and 1000 bottom scoring associated with negative association. An important direction for our future research is to develop an advanced approach for the automatic determination of the optimal number of edges to be included for the inference of microbial association networks.

Currently the analyses of network topology and biological relevance were carried out by treating co-presence and mutual-exclusion networks as two independent networks. We are now working on the development of a multiplex-network based approach (Fig. 7) in an attempt to bridge together different co-presence and mutual-exclusion relations. As the first attempt, we applied the PageRank centrality developed for interconnected multilayer networks [17], [18] to rank the nodes. Different rankings were obtained as depicted in TABLE VII, however, the biological relevance of the results deserves further investigation.

Fig. 7 An illustration of a multiplex network based approach for combining co-presence and mutual-exclusion networks.

TABLE VII THE TOP 5 MICROBIAL GENES IN THE CO-PRESENCE NETWORK AND IN BOTH CO-PRESENCE AND MUTUAL-EXCLUSION NETWORKS BASED ON USING A MULTIPLEX NETWORK-BASED PAGERANK CENTRALITY

Rank	Ranking in two networks	Ranking in the co-presence network
1	K06013	K07161
2	K03780	K00169
3	K02837	K00401
4	K01468	K02007
5	K09824	K01622

REFERENCES

[1] J. Broucek, "Production of Methane Emissions from Ruminant Husbandry: A Review," Journal of Environmental Protection, 2014, 5, pp. 1482-1493.

[2] R. J. Wallace, J.A. Rooke, N. McKain, C-A. Duthie, J. J. Hyslop, D. W. Ross, et al. "The rumen microbial metagenome associated with high methane production in cattle," BMC Genomics. 2015;16: 839. doi: 10.1186/s12864-015-2032-0. pmid:26494241

[3] GreenGas House online, 2017. [Online]. Avaialble: http://www.ghgonline.org/methaneruminants.htm, [accessed: 04-07-2017].

[4] A. Kumar, P. Rameshwar, S. Devki, and N Kamra, Rumen Microbiology: From Evolution to Revolution, Springer, July 11, 2015

[5] E. Jami, B.A. White, and I. Mizrahi I, "Potential Role of the Bovine Rumen Microbiome in Modulating Milk Composition and Feed Efficiency," PLoS ONE 9(1): e85423. https://doi.org/10.1371/journal.pone.0085423

[6] R. Roehe, R.J. Dewhurst, C-A. Duthie, J.A. Rooke, N. McKain, et al., "Bovine host genetic variation influences rumen microbial methane production with best selection criterion for low methane emitting and efficiently feed converting hosts based on metagenomic gene abundance," PLoS Genet., 2016, 12: e1005846. doi:10.1371/journal.pgen.1005846.

[7] D.P. Morgavi, E. Rathahao-Paris, M. Popova, J. Boccard, K.F. Nielsen, and H. Boudra, "Rumen microbial communities influence metabolic phenotypes in lambs," Front. Microbiol. 6:1060. doi: 10.3389/fmicb.2015.01060

[8] G. Henderson, F. Cox, S. Ganesh, A. Jonker, Y. Wayne, et al. "Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range," Scientific Reports. 2015, 5, 14567 (http://dx.doi: 10.1038/srep14567).

[9] M.B. Lengowski, K.H.R. Zuber, M. Witzig, J. Möhring, J. Boguhn, and M. Rodehutscord, "Changes in Rumen Microbial Community Composition during Adaption to an In Vitro System and the Impact of Different Forages," PLoS ONE, 2016, 11(2): e0150115. https://doi.org/10.1371/journal.pone.0150115

[10] J. Friedman J and E. J.Alm EJ, "Inferring Correlation Networks from Genomic Survey Data," PLoS Comput Biol, 2012, 8(9): e1002687.

[11] R.J. Williams, A. Howe, and K. S. Hofmockel, "Demonstrating microbial co-occurrence pattern analyses within and between ecosystems," Front. Microbiol. 2014, 5:358. doi: 10.3389/fmicb.2014.00358

[12] Wang et al, "Integrated metagenomic analysis of the rumen microbiome of cattle reveals key biological mechanisms associated with methane traits," Methods, 2017. https://doi.org/10.1016/j.ymeth.2017.05.029.

[13] K. Faust, J.F. Sathirapongsasuti, J. Izard, N. Segata, D. Gevers, J. Raes, et al. "Microbial Co-occurrence Relationships in the Human Microbiome," PLoS Comput Biol., 2012, 8(7): e1002606.

[14] M.B. Brown, "A Method for Combining Non-Independent, One-Sided Tests of Significance", Biometrics, 1975, 31, pp. 987-992

[15] K. Faust and J.Raes, "CoNet app: inference of biological association networks using Cytoscape," F1000Research, 2016, 5:1519 (doi: 10.12688/f1000research.9050.2)

[16] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," Genome Research 2003 Nov; 13(11), pp.2498-504

[17] M. De Domenico, A. Solé-Ribalta, E. Omodei, S. Gómez, and A. Arenas, "Ranking in interconnected multilayer networks reveals versatile nodes," Nat. Commun., 2015, 6:6868 doi: 10.1038/ncomms7868

[18] M. De Domenico, M. Porter and A. Arenas, "MuxViz: a tool for multilayer analysis and visualization of networks," Journal of Complex Networks, 2015, 3(2), pp.159-17