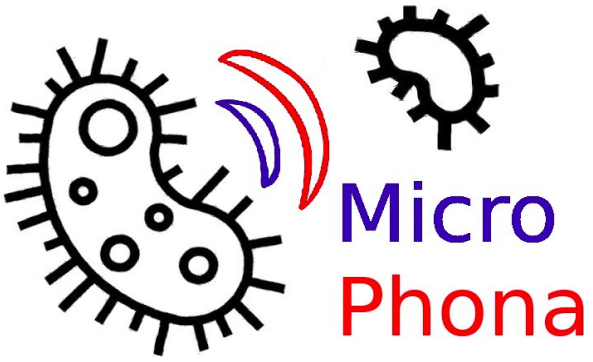
DABIN Aimeric, GAUDIN Marinna, POMMELLEC Jolann Décembre 2019

MicroPhona :

A Tool dedicated to Microbiota Metabolic Conversation



Key-words :
microbiota, *microbiome*, *co-occurrence graphs*, *anorexia nervosa*, *metabolism*, *gut-brain axis*, *python*

Abstract

The study of the microbiota remains a challenge and can lead to the understanding of many diseases. In the case of Anorexia Nervosa (AN), we suspect that the microbiota is involved contributing to mental illness. One way to study the microbiota is to establish co-occurrence graphs based on correlation data, but it does not provide an explanation for these observations. We propose a tool, called MicroPhona, that can annotate such graphs with metabolic data extracted the embl and Recon3D databases. A user-friendly python script enables users to annotate correlation graphs, with different options and outputs (.gml and/or.png files). We applied it to a data set of controls and AN patients to identify some potentially impaired biological pathways between the microbiota, or between us and the microbiota. We focused on the tryptophan case, as it has been suggested in previous studies that the serotonin metabolism is altered in AN patients through the potential sequestration of its precursor of synthesis : the tryptophan.

TABLE OF CONTENTS

1 - Introduction

1.1 - Context

The collection of microbes (bacteria, archea, fungi, virus and microbial eukaryotes) that compose us is called the microbiota and its genome the microbiome. There are as many cells in our microbiota as in our body, making it an important actor in our health [1]. It has been shown that dysbiosis (loss or gain of bacteria promoting either health or disease) in one's microbiota can lead to numerous diseases notably chronic diseases such as Inflammatory Bowel Disease, obesity, chronic inflammation [2].

But, studying microbiota remains a challenge. This is due in part to the large number of bacteria, and interactions, but also a high diversity (some bacteria species as abundant as 5 % in a person could be no more abundant than 0,01% in another, which is why it is not possible to define a standard healthy microbiome ecology [1] making experimental investigation very difficult to impossible, and justifying the need for computational exploration through microbial correlation network [3]. However those correlation networks consists in the mere observation of presence or absence of certain organisms together, which does not supply a sufficient explanation for most disease studies. Thus, we try here to offer an application that introduces causality in those correlation graphs through the addition of one other studied parameters: metabolites. Indeed in order to better understand how microorganisms interact with each other, we decided to look at the content of their conversation.

1.2 - Application to Anorexia Nervosa

The brain and the microbiota communicate both ways: this is called the gut-brain axis. This communication can induce a positive or negative impact, depending on the microbiota composition. In the case of Anorexia Nervosa (AN), a deadly mental illness consisting in diet restriction and fear of gaining weight, this could for instance impact its onset or progress. It has thus been proposed to include microbiota analysis and personalized treatment (pre and probiotics) for recovery protocol in AN patient [4].This recovery protocol could in particular help to avoid relapse, very common in AN.

We obtained data from a study, published in 2016 entitled "***Weight gain in anorexia nervosa does not ameliorate the faecal microbiota, branched chain fatty acid profiles, and gastrointestinal complaints***", conducted by Mack et Al. This study analyzed the gut microbiota (stool samples) and metabolomic (chained fatty acids in stool samples) profiles of three groups: Normal-Weight (NW), patients before (ANT1) and after weight gain (ANT2) following a high energy diet. However we only had access to the microbiota data (abundance table following 16S sequencing) and the group to which they belonged [5].

They found that even if NW and AN groups shared a similar diversity and the same dominant phyla (*Bacteroidetes*, *Firmicutes*, *Actinobacteria*, *Proteobacteria* and *Verrucomicrobia*) they had profound differences in their relative abundances (more *Firmicutes* and *Actinobacteria*, less *Bacteroidetes* in AN patients and more *Verrucomicrobia* in ANT1 but decreasing in ANT2).

They also caracterized the core microbiome (90% of samples containing a specific OTU). 13 OTUS (among these *Coprococcus*, *Clostridium cluster XIVa* and *Dorea*) were present in both ANT2 and NW but not ANT1, maybe markers of recovery and *Clostridium Cluster XI* (*Clostridium Difficile* notably) only present in ANT1 could be a sign of recovery when absent. Finally 4 OTUS seemed to be specific to AN patients in general (2 of them being *Bifidobacterium*). They also pointed out the reduced levels of butyrate-producing *Roseburia spp.* That is why we tried to see in our annotated graphs and analysis of networks such bacteria.

Concerning the metabolomic data, they found that the total Short Chain Fatty Acids (SCFA : acetate, propionate, butyrate) were normal in AN patients but Branched Chain Fatty Acids (BCFA) were increased in ANT1 and ANT2. Also a well-known phenomenon is the use of glycans, notably mucins that constitute the protective layer of the instestine, as a substitute for non digestible carbohydrates, and such mucin-degraders bacteria were more present in AN patients.

Moreover, we did not only focus on the product of fermentation but also some hormones. Indeed, the gut-brain axis is capable to communicate in three ways [4] :

-via the enteric nervous system were motor neurons in the ganglia act as effector cell, allowing a direct influence from the brain to the microbiota as it change the gut composition (secretion, movement..), and connected to the central nervous system via the vagus nerve (microbiota to the brain).

-via the immune system: damaged blood brain barrier (maybe through the action of mucin degraders for instance), neuroinflammation induced by the microglia or directly through the control of immune system cells by the microbiota

-via the endocrine system through a regulation of hormones or peptides involved in the body mass energy balance, and certain microbiota seemed to be able to have receptors to hormons or to be able to produce such hormons (ex: serotonin for Escherichia, Candida, Streptococcus and Enterococcus, production of butyrate known for good effect on the brain). One aspect that we found particularly interesting though not present in this article was the sequestration of tryptophane influencing the regulation of serotonin, that is why we presented it as an example of use of our tool in this report [6].

2 - METHODS

2.1 - Description of data

We collected the dataset of the interest study from **Mgnify database**, and downloaded the corresponding relative abundance tables.

Data are composed of three subgroups. The first one is composed of **Normal Weight people** (**NW**) (n = 55) as Healthy controls. The second one includes **diagnosed ANorexic people at admission** (at Time 1, forming **ANT1** subgroup) (n=55, paired with NW) with BMI < 18kg/m² in adults or <10% of the expected weight in adolescents. Among them, 44 **ANorexic people upon discharge** form the third subgroup (at Time 2, forming **ANT2** subgroup) (n=44), after they spent an hospital stay, followed a strict enriched diet between 2 and 4 months, and had a sufficient weight gain according to their BMI score.

Faecal stool samples of AN patients were collected as soon as possible after the beginning of their inpatient stay. The V4 16S rDNA bacterial sequences have been submitted to the EMBL databases. They were filtered, denoised, and clustered in Operational Taxonomic Units (OTUs) with at least 97% of similarity. Differences in the relative proportion of bacterial phyla between AN patients before and after weight gain and between AN patients and NW participants were compared using the Wilcoxon signed-rank test and Mann-Whitney-U-test respectively [5].

From this general relative abundances table that originally regrouped all the samples whatever their subgroup, we split it in three separated abundances tables, one per subgroup (NW, ANT1 and ANT2).

2.2 - Sources of the metabolic database

We downloaded the **EMBL GEMs database** : A collection of GEnome-scale Models for bacterial species at ***.xml*** format [7]. It currently contains 5587 models which represent metabolic diversity across bacterial life at the strain level, listing all metabolites and reactions described within them, with cross-references to some well known databases such as KEGG. It was constructed with **CarveMe** tool, designed for reconstruction of species and community level metabolic models, that uses a top-down approach to build single-species and community models in a fast and scalable manner. This data constitute our bacterial data.

To get human metabolic data, we collected **Recon3D** database at ***.xml*** format, a computational resource that includes metabolite data and enables integrated analyses of metabolic functions in humans, with 13,543 metabolic reactions involving 4,140 unique metabolites [8].

2.3 - We parsed these metabolic databases to make our own

To have a better computing time, we decided to build our own database from **EMBL GEMs** and **Recon3D**, cited above, in the format of python dictionaries, callable from python scripts. It allows us to have an easier and quicker access to the metabolic data.

For this, we parsed all ***.xml*** files and selected data of interest. As an example, we made a dictionary using the name bacteria as key (`Genre_species`), and for each bacteria, a list of extracellular metabolites that are associated with it. It will allow us to compare the extracellular metabolites between multiple bacteria.

2.4 - Construction of co-occurrence graphs

We used **FlashWeave tool**, to construct three co-occurrence graphs, one per subgroup. FlashWeave predicts ecological interactions between microbes from large-scale compositional abundance data (i.e. OTU tables constructed from sequencing data) through statistical co-occurrence or co-abundance. It reports direct associations, with adjustment for bystander effects and other confounders [9]. We installed this plugin on **Julia**, a dynamic language for technical computing, and used it by command-line.

All bacteria stored in the database have their own genus and species and specific metabolic data. However, some bacteria in the abundance tables in the article only have the genus and not the species, which is not sufficient to match the database. For this reason, we had to keep only the bacteria referenced by species, which represent only 35% of the bacteria present in the abundance tables (123/341).

To make the graphs, we set sensitive parameter to true, to enable fine-graines associations, heterogeneous parameter to false, as data are not multi-habitat or multi-protocol source, and transposed parameter to true, as rows were variables and columns were samples. Three co-occurrence graphs were generated at the ***.gml*** format.

3 - Results

3.1 - Using the tool

We wanted our work to be automatised with python scripts, so that it could be useful for other datasets. The goal of these scripts, is to annotate co-occurrence graphs with metabolic data. The user can give one or multiple co-occurrence graphs at the ***.gml*** format.

- The first script `make_not_oriented_graph.py` generates non-oriented graphs : where each edge has been annotated with the **common extracellular metabolites of the bacteria of the two nodes.** Bigger is the number of these metabolites, larger is the edge. It includes metabolites from both reversible and irreversible reactions.

- The second script `make_oriented_graph.py` generates oriented graphs using exclusively data from known irreversible reactions, thus with a given direction. We selected the irreversible reactions where one or several metabolites are either entering the cell as reactants, or exiting the cell as products.

In the parameter file, the user have access to several options.

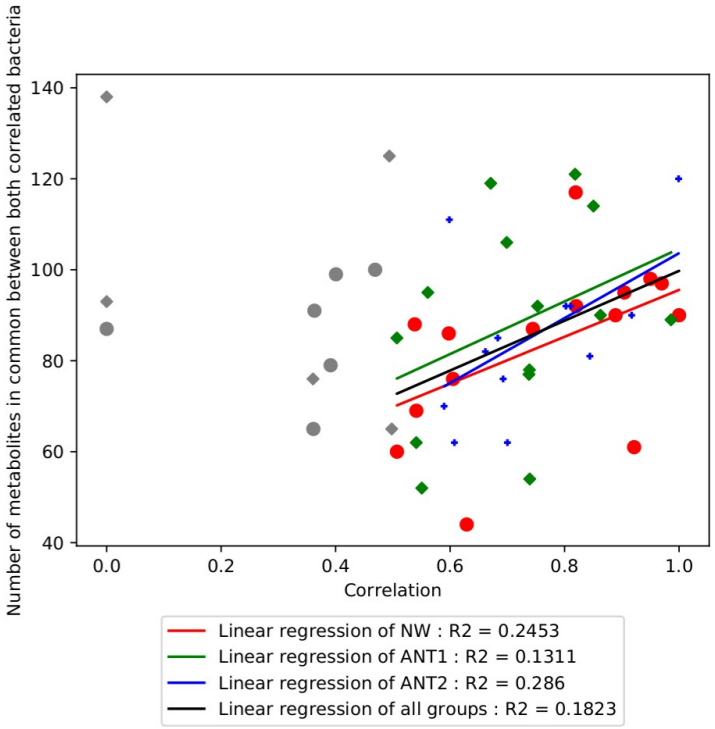
- **Threshold of correlation of the edges**. We can filter the edges by their correlation score
- **Add human as a node**. It is possible to add an extra node corresponding to Homo Sapiens metabolic data, to make it interact with bacteria of the graphs
- **Select bacteria of interest**. This options enables to highlight the indicated bacteria in the graph. Each bacteria of the database has its own number that the user must give to select them. To do this, an index file is included with the script.
- **Select metabolites of interest**. To focus on a biological pathway or specifical metabolites, the user can give their KEGG ID (Exemple : C00780 for serotonin). The edges annotated with the metabolite(s) given will be selected and a subgraph will be generated.
- Generate annotated graph at the format of **.gml** files

3.2 - Link between correlation score and common extracellular metabolites ?

Above a absolute correlation score of 0.5, we observe a slight trend of correlation between the correlation score and the number of shared extracellular metabolites between two bacteria. This is the trend we expected: it seems that the higher the correlation score, positive or negative, the more likely the bacteria are to communicate via extracellular metabolites. That's why we decided to use this threshold to filter our annotated graphs (**Figure 1**).

Figure 1 : Linear regressions between absolute correlation score and number of shared extracellular metabolites

We performed linear regresions to assess whether there is a potential link between the absolute correlation score and the number of shared extracellular metabolites between two bacteria. We did this once per subgroup, and then on the whole pooled data.
It allowed us to select a threshold of 0.5 even if the r scores are not significative.



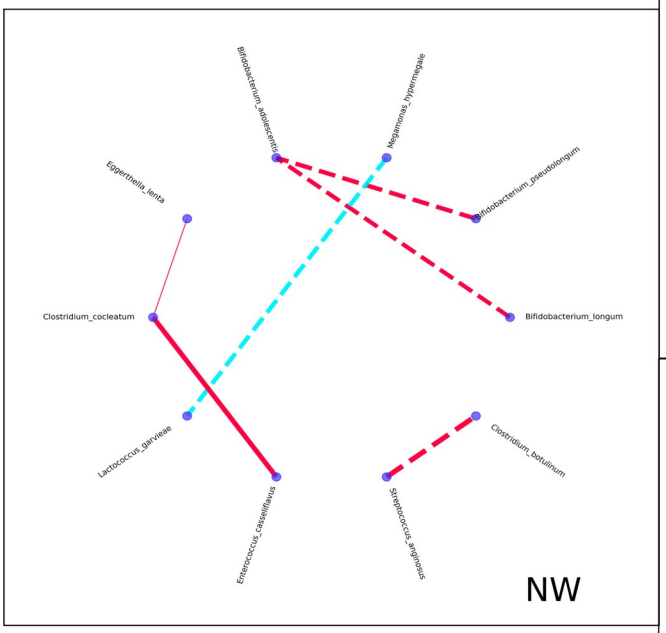
3.3 - Exemple of metabolite application : tryptophan

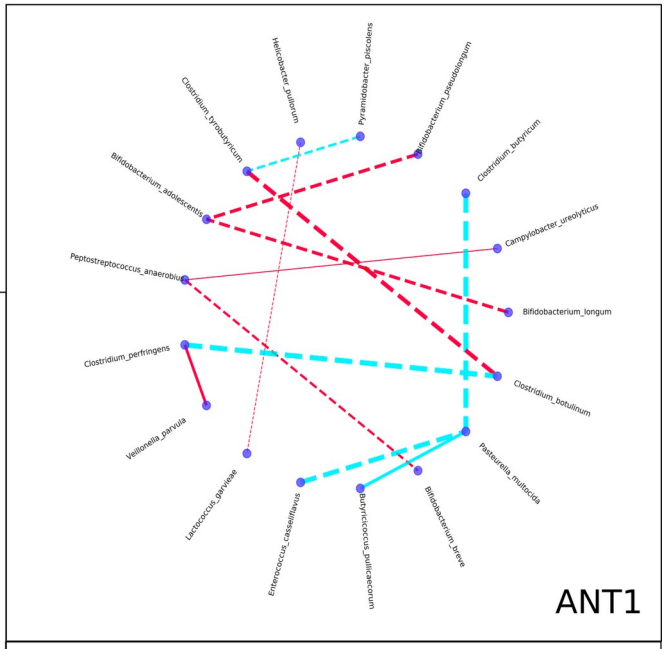
Multiple levels of evidence links disturbances in the serotonergic system and several psychiatric disorder such as depression, anxiety, and borderline personality disorder. For example, the metabolism of tryptophan, a precursor of serotonin, is potentially regulated by the gut microbiota thereby enabling it to influence brain function. Tryptophan is an essential amino acid derived from the diet, and tryptophan that is absorbed from the gut into the bloodstream passes the blood-brain barrier to contribute to serotonin synthesis *in situ*. The availability of tryptophan is strongly affected by the gut microbiota [6]. The serotonergic neurotransmission may thereby be influenced by the availability of tryptophan for serotonin production.

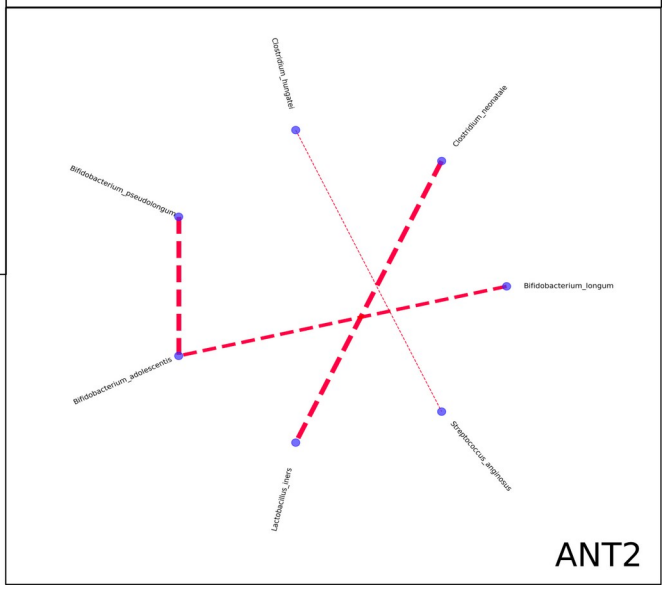
These statements insterested us into making our own investigation about tryptophan gut metabolism. We used our script to select specifically the edges that were annotated with tryptophan, and kept the directly and indirectly connected nodes with it (**Figure 2**).

Figure 2 : Tryptophan metabolite in annotated correlation graphs between NW, ANT1 and ANT2 patients

Larger is the width of the edges, bigger is the number of common extracellular metabolites between two nodes. We targeted tryptophan annotated edges here, and kept also the direct neighbors of these interactions.







Legend :

- Metabolite is present on the edge
- Negative correlation
- Positive correlation

There is an increase in the number of interactions with tryptophan in ANT1 patients, which decreases significantly in ANT2 patients. Could this be a sign of an increase in the flow of tryptophan within the microbiota of anorexic patients? It is conceivable that these important exchanges between bacteria may inhibit its release into the plasma and prevent it from playing its role as a precursor of serotonin. This is only one hypothesis, which needs to be studied more precisely, but it represents an example of the application of this tool.

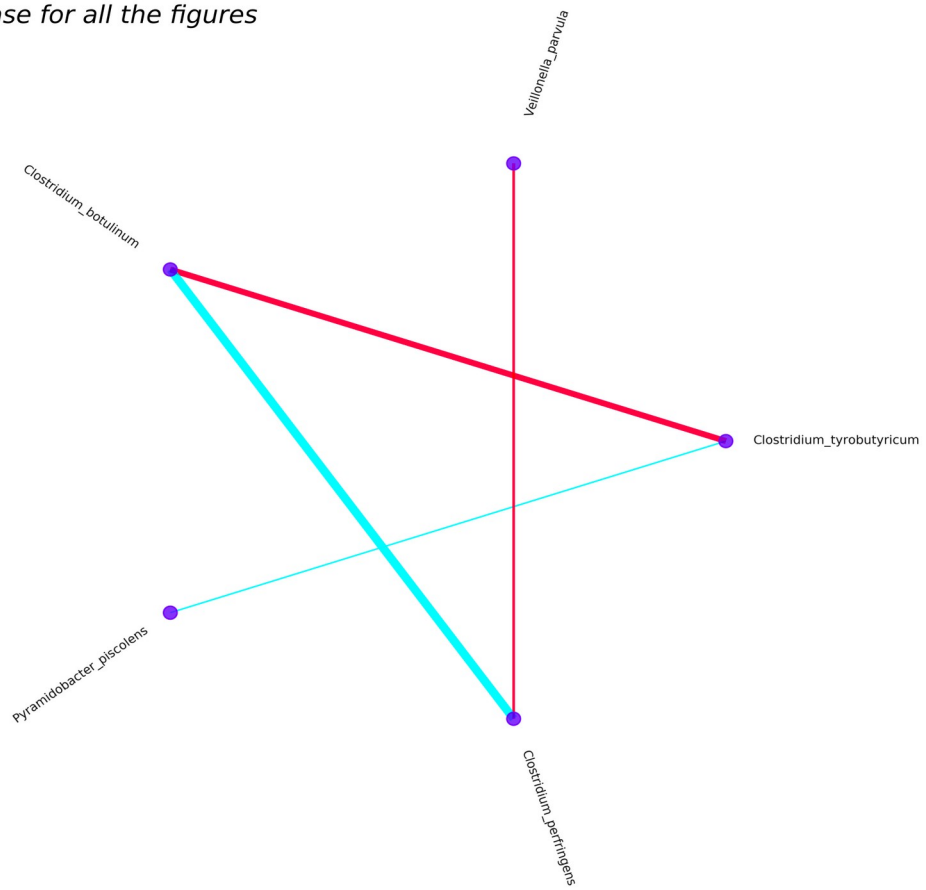
The zoomed figures are available in the appendices (Figures 1, 2 and 3).

3.4- Subnetworks with interest metabolites

We offer the possibility for the user to select metabolites and to look at the connected networks that are related to this metabolite (meaning that only one interaction involving this metabolite will retrieve the whole network associated to this interaction). We performed such analysis with KEGG identifier of SCFA, BCFA, Serotonin pathaways (Tryptophan and Serotonin) and other hormones such as Dopamin or Endorphins. However due to the fact that one interaction is composed of many possible metabolites exchange, all the metabolites retrieved the exact same networks, which were therefore only dependent on the original network. However it is worthy to note that not all connected networks were retrieved, meaning that maybe it could be used to restrict a little bit the sub-networks of interest. It is important for further use to find a way to not simply include metabolites in common but, maybe through flux variability analysis and a project such as the one conducted by Matthieu, Matthias and Léo, to look at metabolites that are more plausible to really be involved in an interaction between two bacteria.

Figure 3 : Example of extracted subnetwork from annotated graph

This figure represent one of the subnetwork obtained when extracting SCFA in ANT1 and was the same when extracting Tryptophan and Serotonin, whereas they are not specifically related in metabolomic terms, and it is the case for all the figures



3.5 Attempt to qualitatively narrow down metabolic data

We tried to search for other pathways involved differientially in our subgroups, thus potentially linked to Anorexia Nervosa, so we sorted the metabolites between those unique to a group and the others and we tried to focus our attention on the metabolites unique to a group. However the numerous pathways inhibit the clear view of one specific and important pathway. Moreover this analysis shows that our database is insufficient for certain analysis as D-Glucose 6-phosphate is only found in ANT1 while it is a center-piece of the glucose metabolism. However this might be because the secretion of G6P in the medium is an abnormal behavior of bacteria. You can have a view on this work in the appendices (Figure 5).

4 - Discussion

4.1 - Loss of data

Unfortunately, despite there are 5587 models described in embl database, it does not represent the total diversity of the bacteria present in the dataset. Because of this, we had to remove numerous nodes in the correlation graphs, because of the unavailability of their metabolic data. Furthermore, we noticed that some bacteria names have changed between the construction of embl database (2017) and the paper's date publication (2016). We are aware that it can provoke a big bias in the final results, as it can "break" some networks because of missing nodes, that would potentially lead to the loss of major interactions essential for a good interpretation. A perspective would be to merge more databases to complete the missing species. Also, we encountered a big loss of the listed bacterias in the abundance files as we had to keep only those ones with the genus and the species specified.

4.2 Limits directly linked to data

Another limit linked to our data is the fact that the ANT2 patients followed a particularly rich diet to come back to a normal weight, and so the composition of the stool sample and microbiota is highly linked to this diet. It would have been better to also have stool samples at another time of those patients, not just upon discharge. It would also be really interesting to know if those patients had a relapse, maybe so we could potentially identify a signature profile of patient that have a higher risk for such relapses. Finally we did not know the AN type that the patient suffered, but we know that a restrictive type or a binge-purging type does not have the same microbiota or physiological response [5], and we only had the abundance tables and no metabolic data.

Of course there are also the usual limitations of time and space : the stool sample was collected at only one moment though it is highly dependent on the diet and vary from one individual to another, from one time of the day to another and also a stool sample does not reflect what happens in all the parts of the gut and rarely include mucosally coherent microbes. Rare microbes are not detected and the 16S approach is sometimes to aspecific, where differentiating one strain from another could lead to a better understanding and a more accurate model. However, stool samples remains widely used and are still full of informations [1].

4.3 - Homo sapiens as a big bacteria

Human metabolic data collected from the Recon3D database do not describe the type of cell source of metabolites, the environmental context (ECM), and do not represent the variability of metabolism between individuals that can certainly add a large variation in interactions, so we must remain careful when using this data. It is interesting to include it in the graph, to show the vision of the global ecosystem between us and our microbiota. However, the results are difficult to analyze because of the amounts of extracellular metabolites shared between each bacterium and the homo sapiens node. As an illustration, an example of a graph including Homo Sapiens nodes can be viewed in the appendices (Figure 4).

4.4 - Big amounts of data

Due to the large number of extracellular metabolites shared between two nodes, we cannot easily and graphically annotate them directly on the graphs and must work with the txt files to get a better overview. We simply adjust the width of the edges according to this number. In addition, we cannot predict whether these metabolites are really objects of conversation between two nodes. There are many ways to study these graphs. The necessary point is to select a specific pathway, a family of metabolites or a group of bacteria not to be lost in the face of this amount of data.

4.5 - Quantitative dimension

We did not quantitatively study extracellular metabolites as a function of their stoichiometry in each reaction. It would be interesting to see if a metabolite is particularly relaxed in the environment, for example H+, which would increase pH and cause increased acidity and stomach aches in AN patients. We know that another group composed of Mathias, Matthieu and Leo worked on this point, so it would be worth merging our work.

4.6 - From correlation to causality

It is an interesting concept to add a causal layer on correlations, but it means that we have to trust the correlation data, because they represent a kind of support for causality. The matrix was mostly empty, so we could have used another tool such as SparCC that is known to manage well such matrix but we prefered to base our analysis on FlashWeave which is able to better recognize indirect interactions from direct interactions, and it seemed to us that it was easier to use, meaning we could spend more time on the development of our tool. However it is worth saying that negative correlations are more prone to false negative or positive, so we should keep that in mind when using our networks. It would be a good markdown to compare different tools as a support for metabolic annotation.

4.7 - Oriented graphs

We did not focus as expected on oriented graphs but on not oriented graphs. Oriented graphs can only be established on the basis of irreversible reactions, which considerably reduces metabolic data. For example, tryptophan was not present in any of the irreversible reactions and therefore cannot be studied in oriented graphs. We also chose not to focus on oriented graphs as we did not observe any strongly directed exchange. Indeed for all of the interactions we observed, there were reactions going one way and others the other way. We were not able to define if for instance having more metabolites exchanged could be a sign of dependance in any article, so we were not able to conclude anything on those graphs.

5 - Conclusion

Co-occurrence graphs alone are not sufficient to interpret the correlations observed between bacterial species. To solve this problem, we have created a script that automatically annotates the observed correlations with metabolic data. Caution should be exercised in interpreting such graphs, as they depend directly on the results of previous correlations, and the presence of annotated metabolites does not prove that the concerned bacterias are exchanging this metabolite. If a specific potential biological pathway is identified, further biological studies are required to investigate this point. This tool offers new perspectives for analyzing communication between bacteria. We hoped to find here bacteria that could explain some aspect of Anorexia Nervosa through the pathways in which they are involved. This could help to improve recovery protocols and the use of personalized treatments such as prebiotics and probiotics or to recognize specific signature. However we were not able to achieve this goal, but we did establish a link between correlation graphs and metabolic data. The possibility to select bacteria and/or metabolites of interest is, we hope, a feature of our tool that will help to narrow the enormous amount data that usually compose such networks. MicroPhona could be used in ecology-oriented studies, where the exchange in terms of metabolites would no longer be with human cells but with the environment. The MicroPhona tool will soon be available on GitLab. Contact us to get the link if you are interested.

6 - Bibliography

[1] C. Allaband *et al.*, « Microbiome 101: Studying, Analyzing, and Interpreting Gut Microbiome Data for Clinicians », *Clin. Gastroenterol. Hepatol.*, vol. 17, n° 2, p. 218-230, 2019.

[2] « Defining Dysbiosis for a Cluster of Chronic Diseases | Scientific Reports ». [Consulté le: 15-déc-2019].

[3] S. Weiss *et al.*, « Correlation detection strategies in microbial data sets vary widely in sensitivity and precision », *ISME J*, vol. 10, n° 7, p. 1669-1681, 2016.

[4] V. Mendez-Figueroa *et al.*, « Can Gut Microbiota and Lifestyle Help Us in the Handling of Anorexia Nervosa Patients? », *Microorganisms*, vol. 7, n° 2, févr. 2019.

[5] I. Mack *et al.*, « Weight gain in anorexia nervosa does not ameliorate the faecal microbiota, branched chain fatty acid profiles and gastrointestinal complaints », *Sci Rep*, vol. 6, n° 1, p. 1-16, mai 2016.

[6] M. Sjögren, S. Frostad, et K. K. Barfod, « Dysbiosis of the Microbiota in Anorexia Nervosa: Pathophysiological Implications », *Anorexia and Bulimia Nervosa*, mars 2019.

[7] D. Machado, S. Andrejev, M. Tramontano, et K. R. Patil, « Fast automated reconstruction of genome-scale metabolic models for microbial species and communities », *Nucleic Acids Res.*, vol. 46, n° 15, p. 7542-7553, 06 2018.

[8] « Recon3D enables a three-dimensional view of gene variation in human metabolism | Nature Biotechnology ». [En ligne]. Disponible sur: https://www.nature.com/articles/nbt.4072. [Consulté le: 15-déc-2019].

[9] J. Tackmann, J. F. Matias Rodrigues, et C. von Mering, « Rapid Inference of Direct Interactions in Large-Scale Ecological Networks from Heterogeneous Microbial Sequencing Data », *Cell Syst*, vol. 9, n° 3, p. 286-296.e8, sept. 2019.

7 - Acknowledgements

We would like to thank our Professor, M.Eveillard who guided us through this project and gave us interesting ideas.

8 - APPENDICES

Figure 1 : Zoomed figure : NW - Tryptophan metabolite in annotated co-occurence graph

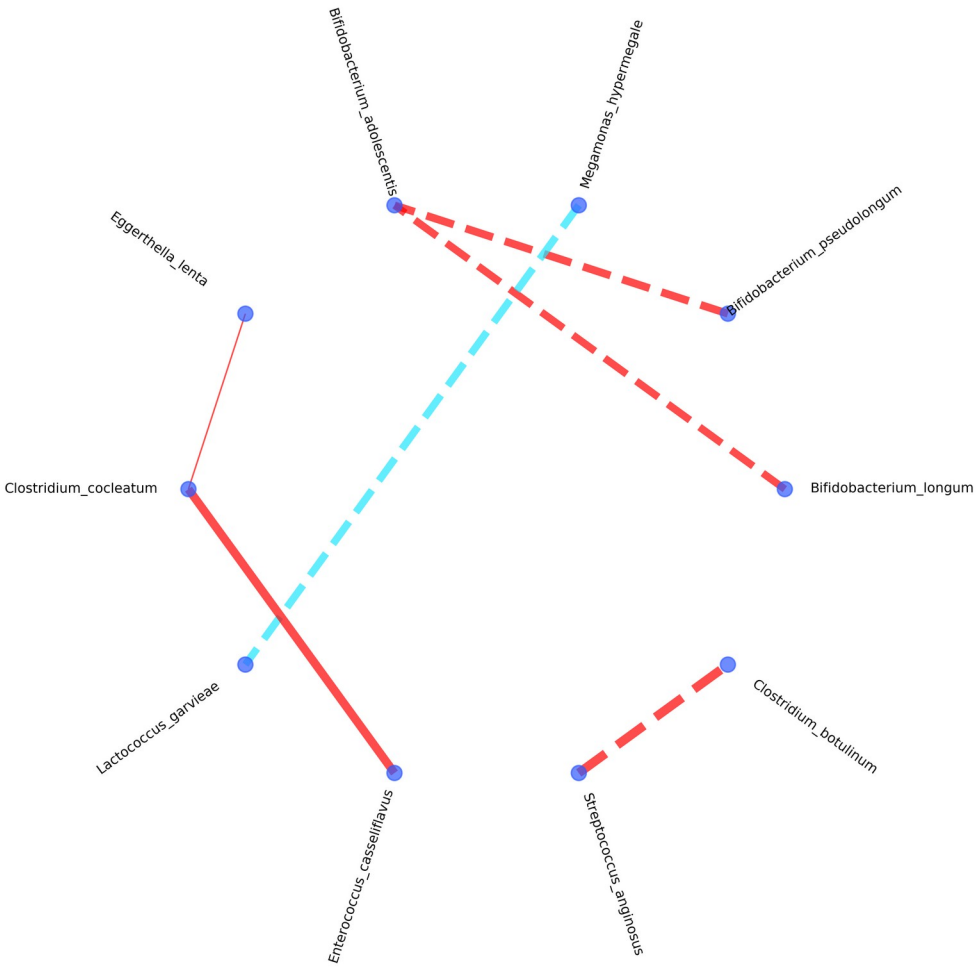


Figure 2 : Zoomed figure : ANT1 - Tryptophan metabolite in annotated co-occurence graph

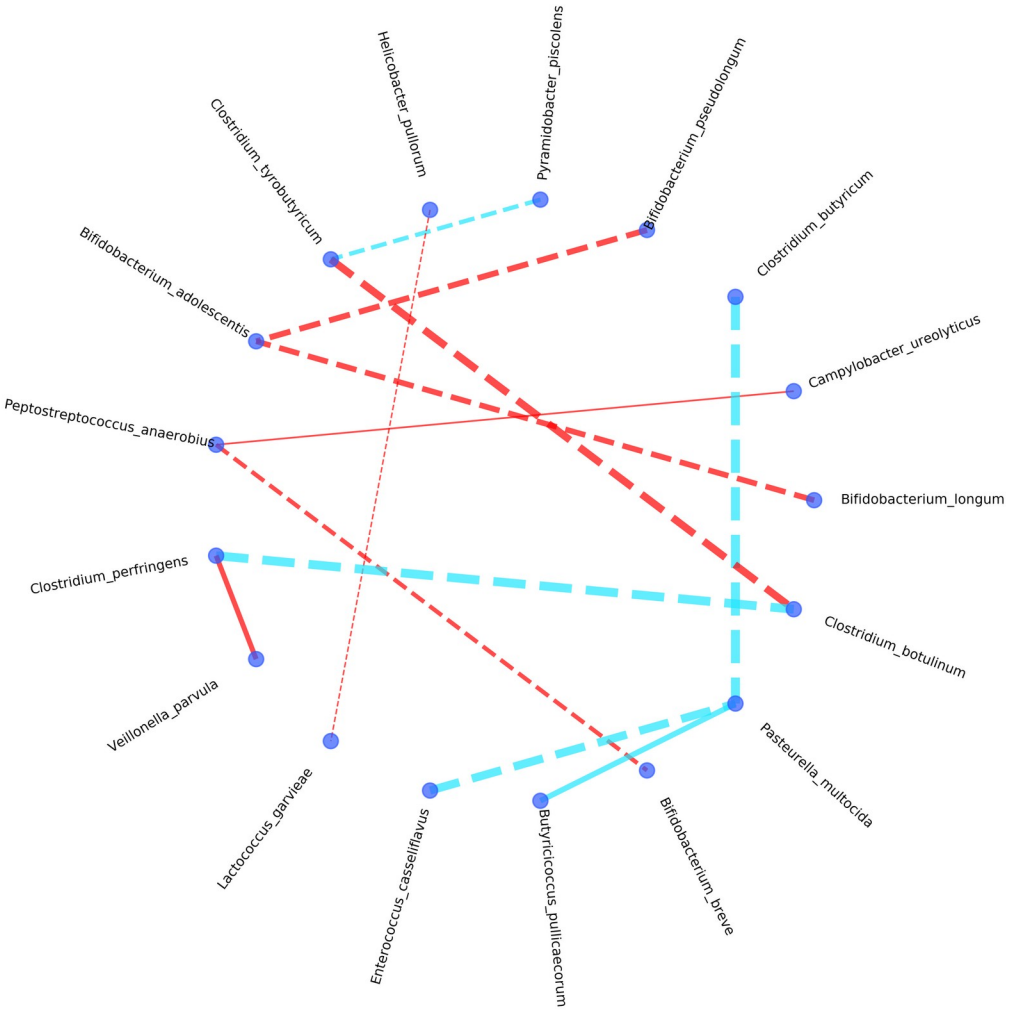


Figure 3 : Zoomed figure : ANT2 - Tryptophan metabolite in annotated co-occurence graph

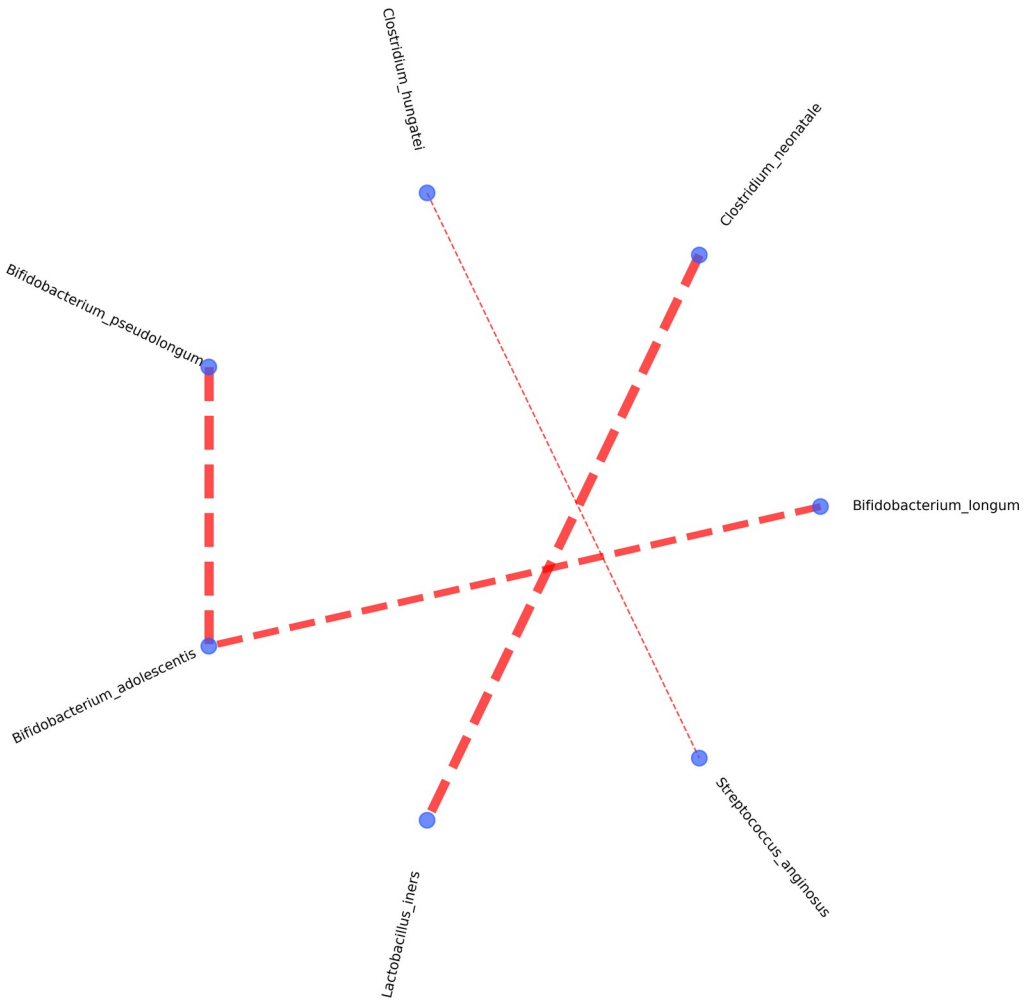


Figure 4 : Homo sapiens as a node - NW example - Threshold of 0.5 for absolute correlations

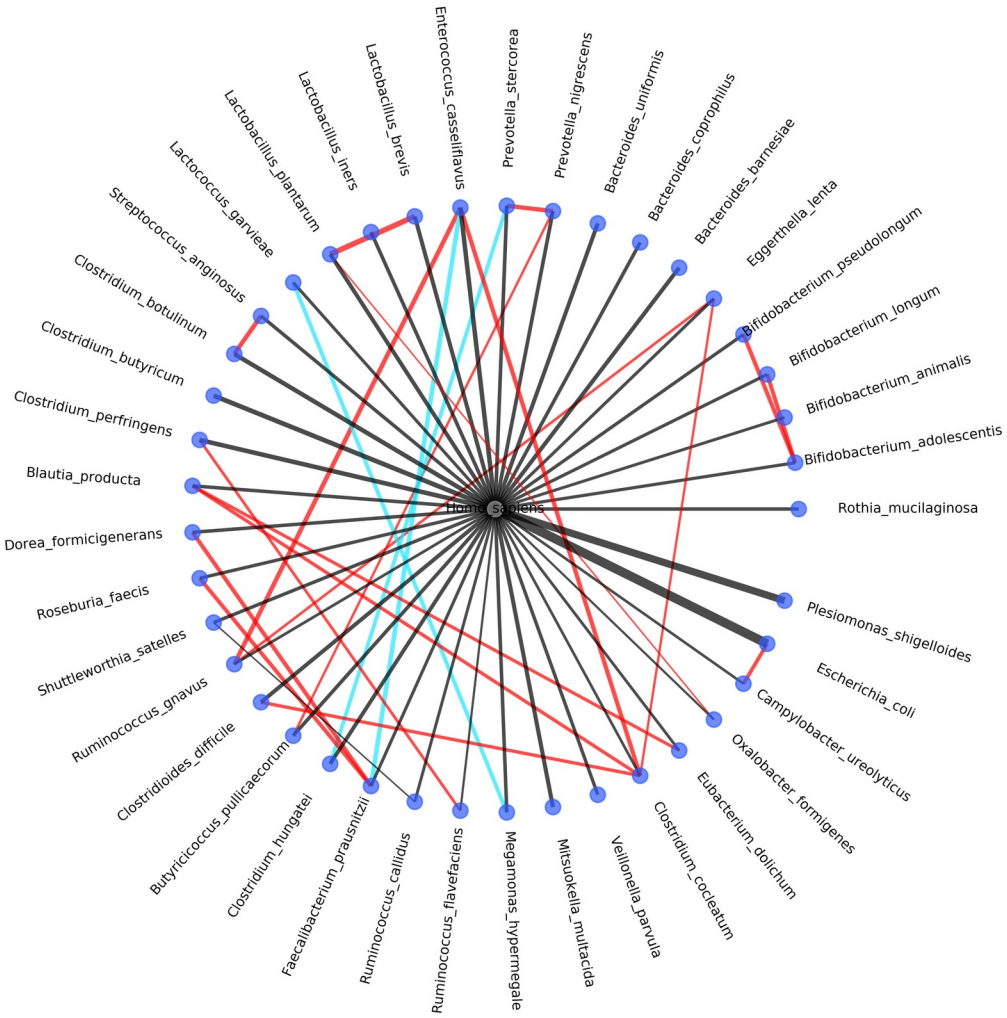


Figure 5 : Lists of selected metabolites and the biological pathways they are involved from the three subgroups :

Metabolites uniquely in groups and the pathways they are involved in :

NW :

- '3-hydroxycinnamic acid'
  - Ubiquinone and other terpenoid-quinone biosynthesis
  - Tyrosine metabolism
  - Phenylpropanoid biosynthesis
  - Isoquinoline alkaloid biosynthesis
- 'Xylan (12 backbone units, 3 glcur side chain)'
  - Carbohydrate digestion and absorption
- 'N-Acetylmuramate'
  - D-Glutamine and D-glutamate metabolism
  - Amino sugar and nucleotide sugar metabolism
  - Peptidoon glycan biosynthesis
- 'KDO(2)-lipid (A)'
  - Lipopolysaccharide biosynthesis
- 'Alpha-L-Arabinan (3 subunits)'
  - Amino sugar and nucleotide sugar metabolism
- 'Benzyl alcohol'
  - Toluene degradation
  - Metabolic pathways
  - Microbial metabolism in diverse environments
  - Degradation of aromatic compounds
- 'Trithionate'
  - Sulfur metabolism
- 'Gamma-butyrobetaine'
  - Lysine degradation
  - ABC transporters
- 'Xylan (8 backbone units, 2 glcur side chain)'
  - same as other xylan
- '2(alpha-D-Mannosyl)-D-glycerate
  - Fructose and mannose metabolism

ANT1 :

- '6-Phospho-D-gluconate'
  - Pentose phosphate pathway
  - Biosynthesis of secondary metabolites
  - Microbial metabolism in diverse environments
  - Biosynthesis of antibiotics
  - Carbon metabolism
- 'Salmochelin-SX'
  - part in reaction : apo-salmochelin esterase
- 'D-Glucose 6-phosphate'
  - Starch and sucrose metabolism

    - Biosynthesis of antibiotics
    - Phosphotransferase system (PTS)
    - Insulin secretion
    - Insulin resistance
    - Carbohydrate digestion and absorption
- 'Indole'
    - Tryptophan metabolism
    - Phenylalanine, tyrosine and tryptophan biosynthesis
    - Benzoxazinoid biosynthesis
- 'Carbonic acid'
    - Proximal tubule bicarbonate reclamation
    - Collecting duct acid secretion
    - Salivary secretion
    - Gastric acid secretion
    - Pancreatic secretion
    - Bile secretion
- 'D-Mannose 6-phosphate'
    - Fructose and mannose metabolism
    - Amino sugar and nucleotide sugar metabolism
    - Metabolic pathways
    - Biosynthesis of secondary metabolites
    - Biosynthesis of antibiotics
    - Phosphotransferase system (PTS)
    - Lysosome
- '2-Dehydro-3-deoxy-D-gluconate'
    - Pentose phosphate pathway
    - Pentose and glucuronate interconversions
    - Metabolic pathways
    - Carbon metabolism
- '1,3-Propanediol'
    - Glycerolipid metabolism
- 'Cellulose (n=6 repeating units)'
    - Starch and sucrose metabolism
- 'L Sorbose C6H12O6'
    - Fructose and mannose metabolism
- 'Salmochelin-S2-Fe-III'
    - involved in reaction of enzyme : apo-salmochelin esterase
- 'Galactomannan(n=6 repeat units mannose, alpha-1,4 man)'
- 'D-Fructose 6-phosphate'
    - Galactose metabolism
    - Starch and sucrose metabolism
    - Methane metabolism
    - Glucagon signaling pathway
    - Insulin resistance
- 'Trans 4 Hydroxycinnamate C9H7O3'
    - Ubiquinone and other terpenoid-quinone biosynthesis
    - Tyrosine metabolism

        - Phenylpropanoid biosynthesis
        - Isoquinoline alkaloid biosynthesis
    - 'D-Glucosamine 6-phosphate'
        - Alanine, aspartate and glutamate metabolism
        - Amino sugar and nucleotide sugar metabolism
        - Metabolic pathways
        - Biosynthesis of antibiotics
        - Phosphotransferase system (PTS)
        - Insulin resistance
    - 'L-Lyxose'
        - Pentose and glucuronate interconversions

ANT2 :

- 'Isethionate C2H5O4S'
    - Taurine and hypotaurine metabolism
- 'Sulfoacetate C2H2O5S'
    - Taurine and hypotaurine metabolism
- 'Butanesulfonate C4H9O3S'
    - next best thing : Methanesulfonate
        - Sulfur metabolism
- '4-Amino-4-deoxy-L-arabinose modified core oligosaccharide lipid A (G. metallireducens, variant 2)'
    - Amino sugar and nucleotide sugar metabolism

ANT1 and ANT2 (not NW) :

- ‘L methionine R oxide C5H11NO3S'
    - Cysteine and methionine metabolism
- '3&apos(=GTP);-GMP'
- '3&apos(=ATP);-AMP'
- 'D-Mannose 1-phosphate'
    - Fructose and mannose metabolism
    - Amino sugar and nucleotide sugar metabolism
    - Biosynthesis of enediyne antibiotics
- 'Phosphoethanolamine KDO(2)-lipid (A)’
- 'Glycerophosphoglycerol'
- 'Mannotetraose'
- 'Rifampin'
    - Bile secretion
- 'Novobiocin'
    - Novobiocin biosynthesis
    - Biosynthesis of antibiotics
- '3'-CMP'
    - Pyrimidine metabolism
- '(S)-Propane-1,2-diol'
    - Propanoate metabolism
- 'Methanesulfonate'
    - Sulfur metabolism