

Applications to GI microbiome data

1. GI microbime data preparation

We applied the newly defined LemonTree parameters recommendation on the GI microbiome data set from (Scher *et al.*, 2013). The representative sequences for each OTU and the OTU abundance table -with reads counts down to the genus classification- were downloaded from https://github.com/polyatail/scher_et_al_2013/tree/master/16S_Analysis. The raw sequencing data were summarized at the genus level, with a taxonomy from phylum to genus to well annotate the unclassified genera, and TMM normalized (Robinson and Oshlack, 2010). The abundance was further filtered by eliminating low abundance (average counts < 1) and low variance (SD < 0.5) microbes. The differential microbes were identified with LIMMA differential analysis, with the following criteria: i) summed abundance > 5; ii) expressed in at least 3 samples; iii) *p*-value < 0.05 and FC > 2. Finally, the expressions of the regulated and candidate regulator microbes were normalized to be zero centered with SD = 1. The taxonomical annotation can be found in Data S2, Sheet 1. The raw and TMM normalized counts are in Data S2, Sheet 2-3.

2. Clinical parameters analyses

The clinical parameters are in Data S2, Sheet 4.

2.1 Association of the clinical parameters with the left/right branches

Table S3-1. Fisher's exact test and t-test for the categorical and quantitative clinical parameters.

		left vs right (all MTX, Prednisone) <i>p</i> -value	left vs right (MTX only) <i>p</i> -value
Categorical (Fisher's exact test)	Disease	1	1
	Gender	1	1
	Race	0.17803	0.52381
	Country	0.272727	0.428571
	Rheumnodules	0.416667	1
	RF	0.416667	1
	CCP	1	0.444444
	Erosions	1	0.444444
	Pred5	0.204545	
	MTX15	0.204545	1
	NSAIDS	1	1
	Opioids	NA	NA
	Smoking	0.681818	1
	SE	0.575758	0.444444
PTPN22	1	1	

Quantitative (t-test)	Age	0.832069	0.29089
	VAS	0.902995	0.759054
	DAS28	0.367704	0.08511
	TJ	0.829908	0.892044
	SJ	0.729509	0.471898
	RFtitre	0.520543	0.47267
	CCPtitre	0.914051	0.713257
	ESR	0.152631	0.223164
	CRP	0.028144	0.041189
	RAduration	0.378784	0.322728

2.2 PCA analysis of the clinical parameters

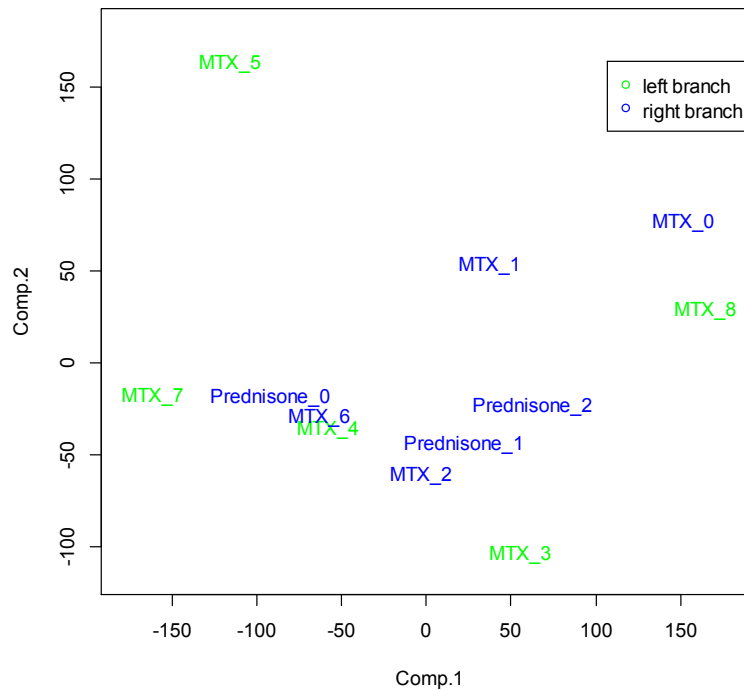


Fig S3-1. PCA plot of the clinical parameters (MTX, Prednisone).

3. LemonTree parameters

The regulator candidates are the differential microbes from the NORA-HLT comparison; the regulated candidates are all microbes but these candidates to avoid self-regulation; the samples are patients in the MTX and Prednisone treatment groups. LemonTree input parameters are:

- 1) $nc = 30$;
- 2) $nreg = reg = 14$ ("GENUS_71" filtered as low abundance in {MTX, Prednisone});
- 3) $perc = 70\%$.

Table S3-2. Candidate regulator microbes.

ID	Annotation (kingdom;phylum;class;order;family;genus)
GENUS_1	Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales;Prevotellaceae;Prevotella
GENUS_4	Bacteria;Firmicutes;Clostridia;Clostridiales;Lachnospiraceae;Roseburia
GENUS_7	Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales;Rikenellaceae;Alistipes
GENUS_8	Bacteria;Firmicutes;Clostridia;Clostridiales;Veillonellaceae;Phascolarctobacterium
GENUS_9	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;Faecalibacterium
GENUS_10	Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales;Prevotellaceae;unclassified
GENUS_15	Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales;Prevotellaceae;unclassified_Prevotellaceae
GENUS_16	Bacteria;Firmicutes;Clostridia;Clostridiales;Veillonellaceae;Dialister
GENUS_24	Bacteria;Firmicutes;Clostridia;Clostridiales;Veillonellaceae;unclassified
GENUS_32	Bacteria;Firmicutes;Erysipelotrichi;Erysipelotrichales;Erysipelotrichaceae;Catenibacterium
GENUS_43	Bacteria;Proteobacteria;unclassified;unclassified;unclassified;unclassified
GENUS_48	Bacteria;Proteobacteria;Gammaproteobacteria;Pasteurellales;Pasteurellaceae;Haemophilus
GENUS_53	Bacteria;Firmicutes;Erysipelotrichi;Erysipelotrichales;Erysipelotrichaceae;Coprobacillus
GENUS_56	Bacteria;Firmicutes;Clostridia;Clostridiales;Veillonellaceae;Veillonella

Table S3-3. Modules and member microbes generated from tight clustering.

ID	Annotation (kingdom;phylum;class;order;family;genus)
GENUS_33	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;unclassified_Ruminococcaceae
GENUS_5	Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales;Porphyromonadaceae;Parabacteroides
GENUS_17	Bacteria;Proteobacteria;Betaproteobacteria;Burkholderiales;Burkholderiales_incertae_sedis;unclassified
GENUS_65	Bacteria;Firmicutes;Clostridia;Clostridiales;Incertae_Sedis_XIII;Anaerovorax
GENUS_11	Bacteria;Firmicutes;unclassified;unclassified;unclassified;unclassified
GENUS_2	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;unclassified
GENUS_26	Bacteria;Proteobacteria;Gammaproteobacteria;Enterobacteriales;Enterobacteriaceae;unclassified
GENUS_18	Bacteria;Proteobacteria;Betaproteobacteria;Burkholderiales;Alcaligenaceae;Parasutterella
GENUS_31	Bacteria;Firmicutes;Bacilli;Lactobacillales;Streptococcaceae;Streptococcus
GENUS_3	Bacteria;Firmicutes;Clostridia;Clostridiales;Lachnospiraceae;unclassified
GENUS_58	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;Ruminococcus
GENUS_20	Bacteria;Firmicutes;Clostridia;Clostridiales;Lachnospiraceae;Coprococcus
GENUS_22	Bacteria;Firmicutes;Clostridia;Clostridiales;Incertae_Sedis_XIV;Blautia
GENUS_44	Bacteria;Actinobacteria;Actinobacteria;Coriobacteriales;Coriobacteriaceae;Collinseilla

GENUS_6	Bacteria;Firmicutes;Clostridia;Clostridiales;Lachnospiraceae;unclassified_Lachnospiraceae
GENUS_27	Bacteria;Proteobacteria;Betaproteobacteria;Burkholderiales;unclassified;unclassified
GENUS_63	Bacteria;Firmicutes;Clostridia;Clostridiales;Incertae_Sedis_XIII;unclassified
GENUS_35	Bacteria;Firmicutes;Clostridia;Clostridiales;Peptostreptococcaceae;Sporacetigenium
GENUS_60	Bacteria;Proteobacteria;Deltaproteobacteria;Desulfovibrionales;Desulfovibrionaceae;unclassified
GENUS_28	Bacteria;Firmicutes;Clostridia;Clostridiales;Veillonellaceae;Megasphaera
GENUS_13	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;Oscillibacter
GENUS_41	Bacteria;Firmicutes;Clostridia;Clostridiales;Peptostreptococcaceae;unclassified_Peptostreptococcaceae
GENUS_30	Bacteria;Firmicutes;Clostridia;Clostridiales;unclassified;unclassified
GENUS_12	Bacteria;unclassified;unclassified;unclassified;unclassified;unclassified
GENUS_69	Bacteria;Firmicutes;unclassified_Firmicutes;unclassified;unclassified;unclassified
GENUS_42	Bacteria;Bacteroidetes;Bacteroidia;Bacteroidales;unclassified;unclassified
GENUS_59	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;Butyricoccus
GENUS_40	Bacteria;Firmicutes;Clostridia;Clostridiales;unclassified_Clostridiales;unclassified
GENUS_14	Bacteria;Bacteroidetes;unclassified;unclassified;unclassified;unclassified
GENUS_19	Bacteria;Firmicutes;Clostridia;Clostridiales;Ruminococcaceae;Subdoligranulum
GENUS_36	Bacteria;Firmicutes;Clostridia;Clostridiales;Lachnospiraceae;Dorea

Here only one module (namely, mod_0) is generated.

Table S3-4. Assigned regulators (full list) to mod_0, with 10 instantiations.

Instantiation No.	Full list of assigned regulators	Corresponding scores
0	GENUS_4	14.975940
	GENUS_56	1.760507
	GENUS_9	0.568363
	GENUS_7	0.522100
1	GENUS_4	16.307376
	GENUS_7	12.704342
	GENUS_1	4.416815
	GENUS_56	2.793968
	GENUS_53	2.283879
	GENUS_9	2.259406
	GENUS_48	1.086739
	GENUS_43	1.059463
2	GENUS_4	17.259483
	GENUS_9	5.056355
	GENUS_7	4.725685
	GENUS_56	2.263539
	GENUS_43	2.118926
	GENUS_48	1.110911

	GENUS_53	1.044111
3	GENUS_4	23.537581
	GENUS_7	6.926109
	GENUS_9	5.635548
	GENUS_56	3.459837
	GENUS_1	2.208408
	GENUS_43	1.589195
4	GENUS_4	16.400015
	GENUS_7	2.871188
	GENUS_9	1.686254
	GENUS_48	1.642194
	GENUS_43	1.589195
	GENUS_56	1.172308
5	GENUS_4	23.201447
	GENUS_56	4.645876
	GENUS_7	2.978694
	GENUS_9	2.802565
	GENUS_48	1.079636
	GENUS_43	0.529732
6	GENUS_4	25.109613
	GENUS_7	6.433786
	GENUS_9	5.010833
	GENUS_56	2.912710
	GENUS_43	2.118926
	GENUS_53	0.619612
	GENUS_1	0.561231
	GENUS_8	0.513344
7	GENUS_4	14.114425
	GENUS_7	8.140855
	GENUS_9	3.383109
	GENUS_56	2.352057
	GENUS_48	2.212690
	GENUS_43	1.589195
	GENUS_53	1.043700
	GENUS_1	0.524715
8	GENUS_4	14.982218
	GENUS_56	4.673217
	GENUS_9	2.805546
	GENUS_7	1.761598
	GENUS_43	1.589195
9	GENUS_4	13.193747
	GENUS_56	1.760507
	GENUS_8	1.033015
	GENUS_9	0.551741

	GENUS_48	0.531283
	GENUS_7	0.522100

The choice for *perc* = 70% is made according to our conclusion from synthetic data that *perc* ≥ 30% can warrant good TPR (sensitivity), in this range 70% is the minimum value for a decent stability (≥ 80%). It thus ensures to obtain both stringency (100% stable microbe *Roseburia*) and flexibility (other two relatively stable microbes, *Veillonella* and *Faecalibacterium*).

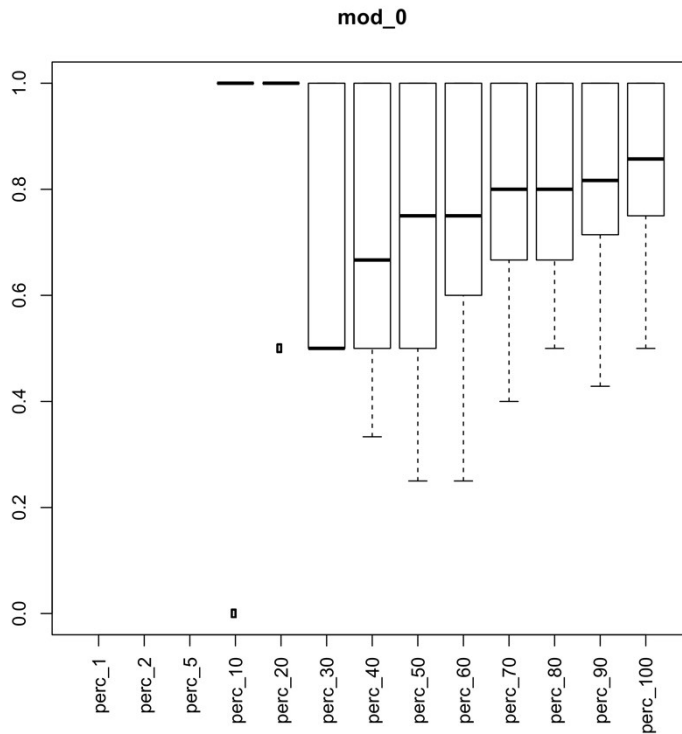


Fig S3-2. Regulators stability as LemonTree parameter *perc* varies.

Note: *perc* = 1 or 2 outputs no regulators due to the generally short length of the full list for each instantiation (4~8). LemonTree indeed generates empty files for the default top regulator setting (*perc* = 1%).

4. Community structure and Firmicutes/Bacteroides ratio analysis for the LemonTree results

4.1 Phylum community structure analysis

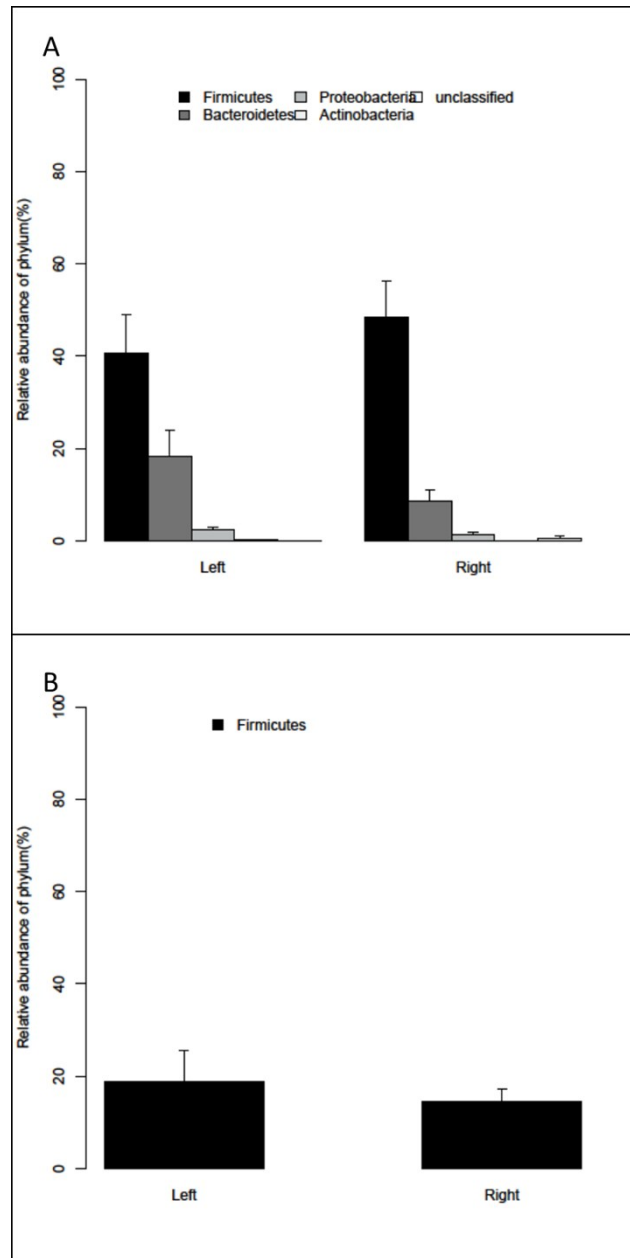


Fig S3-3. Phylum community structure in module (A) and assigned regulator (B) microbes in the left/right branches.

4.2 Firmicutes/Bacteriodes ratio analyses

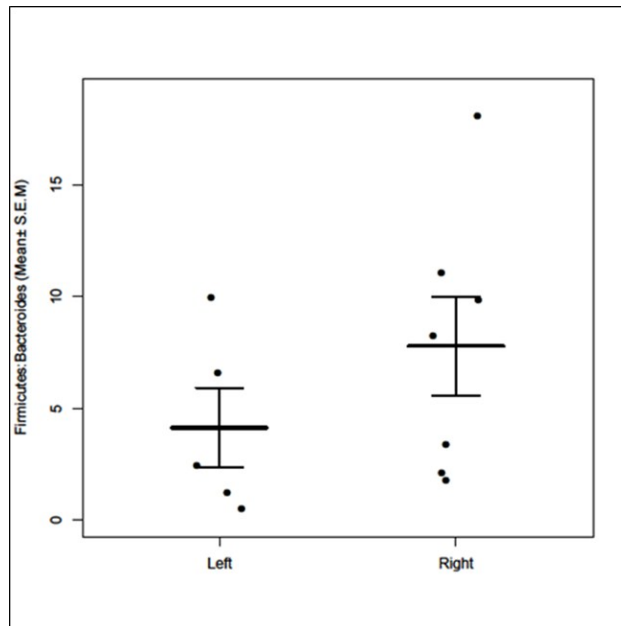


Fig S3-4. Firmicutes to Bacteroides ratios in left and right grouped samples of the module. No Bacteroides phylum in regulators.

Table S3-5. Phylum structure in module microbes.

Taxa	Left_MEAN	Right_MEAN	Left_SEM	Right_SEM
Firmicutes	40.59305	48.60127	8.527221	7.660571
Bacteroidetes	18.21999	8.615841	5.792842	2.322657
Proteobacteria	2.509126	1.303909	0.545723	0.563968
Actinobacteria	0.249582	0.034856	0.127894	0.028145
unclassified	0.078605	0.616817	0.044057	0.415848

Table S3-6. Phylum structure in regulator microbes.

Taxa	Left_MEAN	Right_MEAN	Left_SEM	Right_SEM
Firmicutes	14.81649	7.445971	6.626134	2.814312

Table S3-7. Firmicutes/Bacteroides ratio of samples in module mod_0.

SampleID	Firmicutes/Bacteroides	Branch
MTX_3	1.222222	Left
MTX_4	2.442402	Left
MTX_5	0.501488	Left
MTX_7	9.926471	Left
MTX_8	6.585657	Left
MTX_0	9.841432	Right
MTX_1	2.128904	Right

MTX_2	11.05805	Right
MTX_6	18.06897	Right
Prednisone_0	8.24821	Right
Prednisone_1	1.781481	Right
Prednisone_2	3.379108	Right

Reference

- Robinson,M.D. and Oshlack,A. (2010) A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.*, **11**, R25.
- Scher,J.U. *et al.* (2013) Expansion of intestinal *Prevotella copri* correlates with enhanced susceptibility to arthritis. *eLife*, **2**.