

Supplemental Materials

The immune profile of the fecal microbiome in graft versus host disease patients

Marcel A. de Leeuw & Manuel X. Duval, GeneCreek



Contents

Supplemental analyses	1
Linear model performance at aGvHD onset	1
GM changes through allo-HSCT	1
Most important differentially abundant species	2
GM modulation	3

List of Figures

S1	ROC analysis of the aGvHD onset linear model	1
S2	Strict anaerobe proportion change through allo-HSCT	1
S3	Gram staining balance change through allo-HSCT	2
S4	biological safety level at aGvHD onset	2
S5	Co-exclusion by and co-occurrence with QPS species	4
S6	Co-exclusion and co-occurrence between QPS species	5

List of Tables

S1	Top 50 differentially abundant species	3
----	--	---

Supplemental analyses

Linear model performance at aGvHD onset

The linear model from figure 6 as applied to the originating combined “onset” dataset, can be scored according to a receiver operating characteristic (ROC) model, since the response variable is binary, Fig. S1.

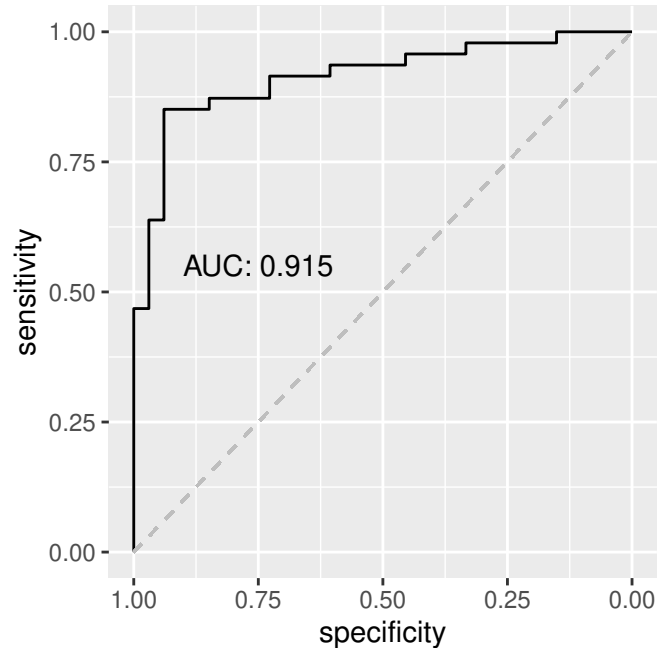


Figure S1: ROC analysis of the aGvHD onset linear model. Combination of datasets, 68 subjects, 80 samples.

GM changes through allo-HSCT

The proportion of strict anaerobes shows a paradoxical evolution across allo-HSCT, with non-aGvHD patients showing the lowest level, Fig S2. One could have expected to find the strongest depletion in aGvHD patients.

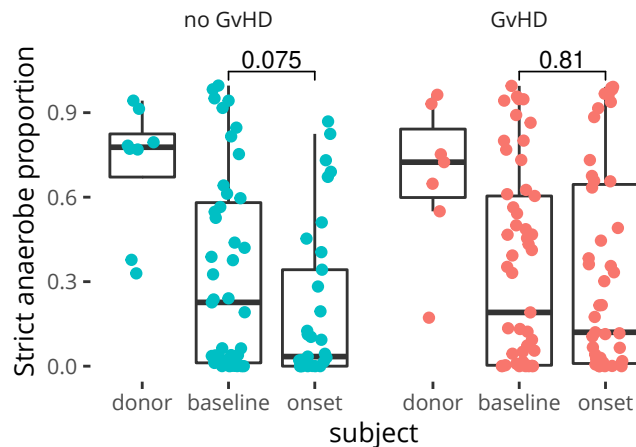


Figure S2: Strict anaerobe proportion change through allo-HSCT. Combination of datasets, 172 subjects, 191 samples. Numbers reflect Wilcoxon signed rank test p-values.

The gram staining balance also evolves through allo-HSCT with a boost of gram positive bacteria which is more tampered in controls at aGvHD onset, fig S3. Thus in controls, the antibiotics regime has eliminated more

anaerobes and gram positive bacteria.

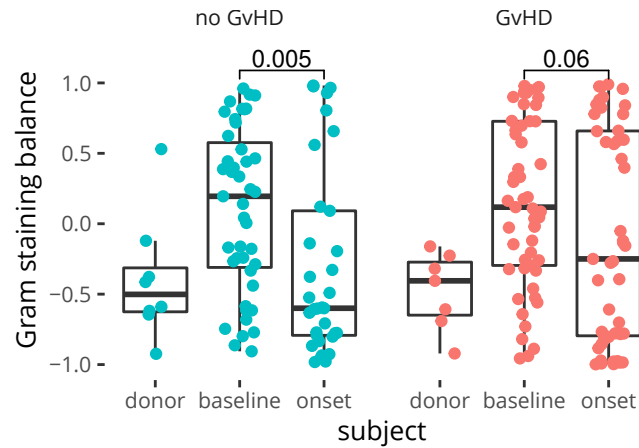


Figure S3: Gram staining balance change through allo-HSCT. Combination of datasets, 172 subjects, 191 samples. Numbers reflect Wilcoxon signed rank test p-values.

Another remarkable feature of the aGvHD control GM is the higher average biological safety level (BSL) which reaches significance at aGvHD onset, Fig. S4. This could be indicative of a different use of oral antibiotics.

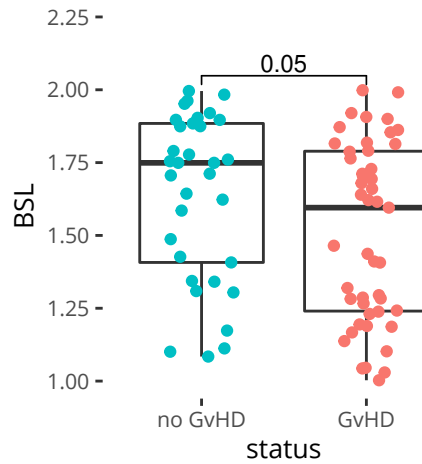


Figure S4: biological safety level at aGvHD onset. Combination of datasets, 68 subjects, 80 samples. Numbers reflect Wilcoxon signed rank test p-values.

Most important differentially abundant species

DeSeq analysis reveals over one hundred differentially abundant species when comparing aGvHD GMs to control GMs at onset, $p_{Adj} < 0.001$. Below, we provide the 50 most abundant of these species, Table S1.

Table S1: Top 50 differentially abundant species. Log2FC: base 2 logarithm of the fold change, bsl: biological safety level.

taxon	log2FC	association	bsl	gram	oxygen
Prevotella copri	17.6	GvHD	1	negative	anaerobe
Atopobium rimae	16.4	GvHD	1	positive	anaerobe
Bifidobacterium breve	15.5	GvHD	1	positive	anaerobe
Streptococcus constellatus	15.3	GvHD	2	positive	microaerophile
Gemella haemolysans	15.0	GvHD	1	positive	facultative
Alistipes putredinis	14.9	GvHD	2	negative	anaerobe
Sphingomonas aquatilis	14.8	GvHD	1	negative	aerobe
Solobacterium moorei	14.8	GvHD	2	positive	anaerobe
Bacteroides intestinalis	14.7	GvHD	1	negative	anaerobe
Clostridium perfringens	14.4	GvHD	2	positive	obligate anaerobe
Parvimonas micra	13.6	GvHD	1	positive	anaerobe
Gemella morbillorum	13.2	GvHD	2	positive	facultative
Alistipes inops	13.0	GvHD	1	negative	anaerobe
Gemmiger formicilis	12.7	GvHD	1	negative	obligate anaerobe
Lachnoanaerobaculum umeaense	12.5	GvHD	1	positive	anaerobe
Phascolarctobacterium faecium	12.4	GvHD	1	negative	anaerobe
Clostridium cellulovorans	12.3	GvHD	1	positive	anaerobe
Corynebacterium tuberculostearicum	11.7	GvHD	2	positive	aerobe
Trichococcus collinsii	11.4	GvHD	1	positive	facultative
Finegoldia magna	11.3	GvHD	1	positive	anaerobe
Anaerococcus vaginalis	11.2	GvHD	1	positive	anaerobe
Lactobacillus crispatus	10.7	GvHD	1	positive	anaerobe
Hungatella hathewayi	10.4	GvHD	1	positive	anaerobe
Turicibacter sanguinis	10.3	GvHD	1	positive	anaerobe
Butyricoccus pullicaecorum	10.0	GvHD	1	positive	anaerobe
Capnocytophaga sputigena	9.7	GvHD	2	negative	facultative
Lactobacillus fermentum	9.5	GvHD	1	positive	microaerophile
Enterococcus raffinosus	8.5	GvHD	2	positive	microaerophile
Faecalicoccus acidiformans	7.6	GvHD	1	positive	obligate anaerobe
Lactobacillus casei	5.3	GvHD	1	positive	facultative
Clostridium chauvoei	4.8	GvHD	2	positive	anaerobe
Stenotrophomonas maltophilia	4.8	GvHD	1	negative	aerobe
Bacteroides fragilis	4.6	GvHD	2	negative	obligate anaerobe
Ruminococcus bicirculans	-4.4	no GvHD	1	positive	anaerobe
Sutterella wadsworthensis	-5.7	no GvHD	2	negative	microaerophile
Blautia glucerasea	-5.8	no GvHD	1	positive	anaerobe
Blautia faecis	-5.9	no GvHD	1	positive	anaerobe
Parasutterella excrementihominis	-9.5	no GvHD	1	negative	anaerobe
Streptococcus cristatus	-9.9	no GvHD	2	positive	anaerobe
Megasphaera massiliensis	-11.5	no GvHD	2	negative	anaerobe
Staphylococcus haemolyticus	-12.0	no GvHD	2	positive	facultative
Clostridium sporosphaeroides	-12.6	no GvHD	1	positive	anaerobe
Shuttleworthia satelles	-12.7	no GvHD	2	positive	anaerobe
Clostridium butyricum	-13.0	no GvHD	1	positive	anaerobe
Blautia producta	-13.5	no GvHD	1	positive	anaerobe
Streptococcus anginosus	-13.6	no GvHD	2	positive	aerobe
Acidaminococcus intestini	-13.8	no GvHD	2	negative	anaerobe
Clostridium glycyrrhizinilyticum	-14.3	no GvHD	1	positive	anaerobe
Acidaminococcus fermentans	-14.7	no GvHD	2	negative	anaerobe
Anaerostipes caccae	-16.4	no GvHD	1	positive	anaerobe

GM modulation

To investigate if qualified presumption of safety (QPS), other than *Bifidobacterium longum* are of therapeutic interest, we computed pairwise χ^2 tests for all QPS with aGvHD case associated species reported above, Fig. 7, main manuscript. We used over 20,000 samples from several tens of studies as the basis for the test. Supplemental Fig. S5 provides the same analysis for control associated species and Fig. S6 provides pairwise χ^2 tests results between QPS species, the former to assess respect of control associated species and the latter to assess

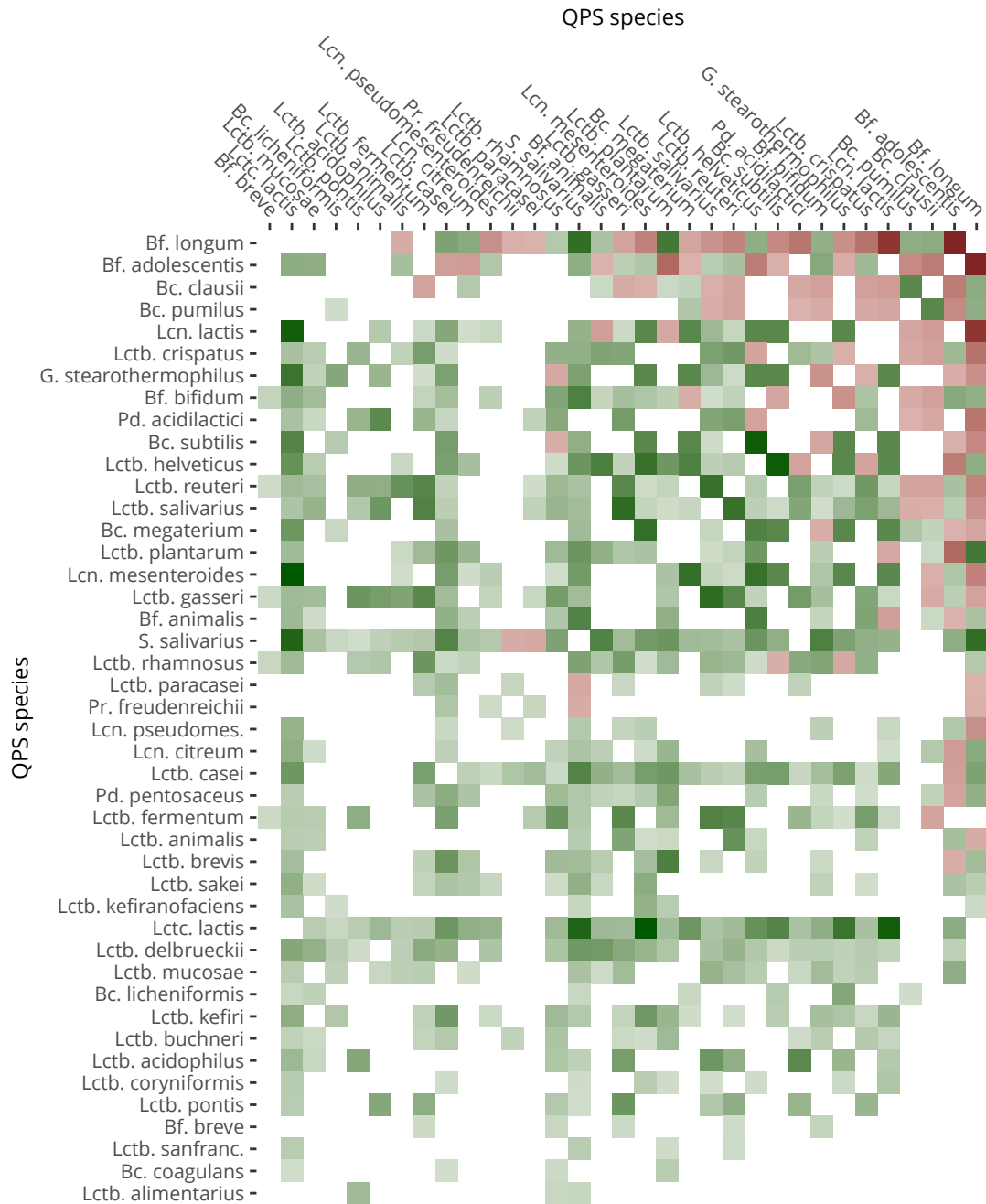


Figure S6: Co-exclusion and co-occurrence between QPS species. Putative inhibition is in shades of red, potential synergy in shades of green. White reflect neutrality or too little combined prevalence to make a call. Genera are abbreviated as follows: **Bcl.:** *Bacillus*, **Bf.:** *Bifidobacterium*, **Gb.:** *Geobacillus*, **Lcn.:** *Leuconostoc*, **Lctb.:** *Lactobacillus*, **Lctc.:** *Lactococcus*, **Pd.:** *Pediococcus*, **S.:** *Streptococcus*.