

**Colonization with the enteric protozoa *Blastocystis* is associated with increased diversity
of human gut bacterial microbiota**

Christophe Audebert^{1,2,+}, Gaël Even^{1,2,+}, Amandine Cian³, The *Blastocystis* Investigation Group[‡], Alexandre Loywick^{1,2}, Sophie Merlin^{1,2}, Eric Viscogliosi³, and Magali Chabé^{3,*}

¹GENES DIFFUSION, Douai, France

²PEGASE-Biosciences, Institut Pasteur de Lille, Lille, France

³Univ. Lille, CNRS, Inserm, CHU de Lille, Institut Pasteur de Lille, U1019 - UMR 8204 - CIIL
- Center for Infection and Immunity of Lille, Lille, France

*corresponding author: magali.chabe@univ-lille2.fr

+These authors contributed equally to this work.

‡See the list of The *Blastocystis* Investigation Group members at the end of the manuscript.

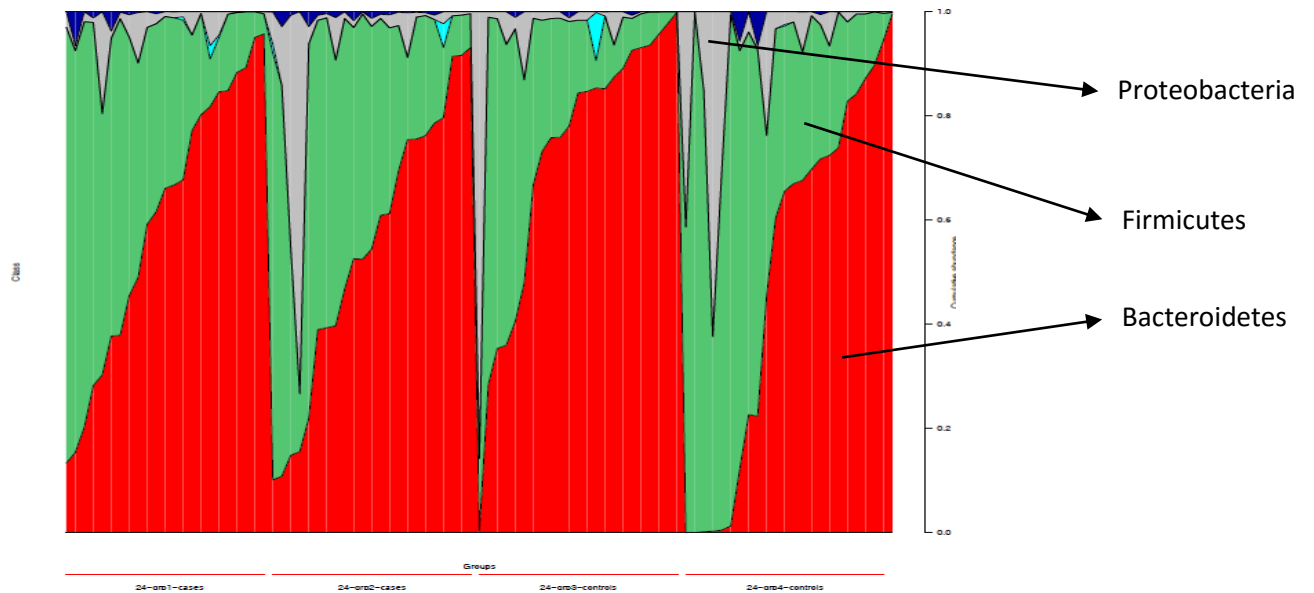
Supplementary information:

Supplementary Figure S1: Faecal bacterial communities present in the different samples (DNA samples from patients' groups 1, 2, 3 and 4) analyzed in this study. Plots of relative abundances of OTUs at four taxonomic levels: phylum (a), class (b), order (c) and family (d). The most abundant and interesting taxa are labeled with arrows.

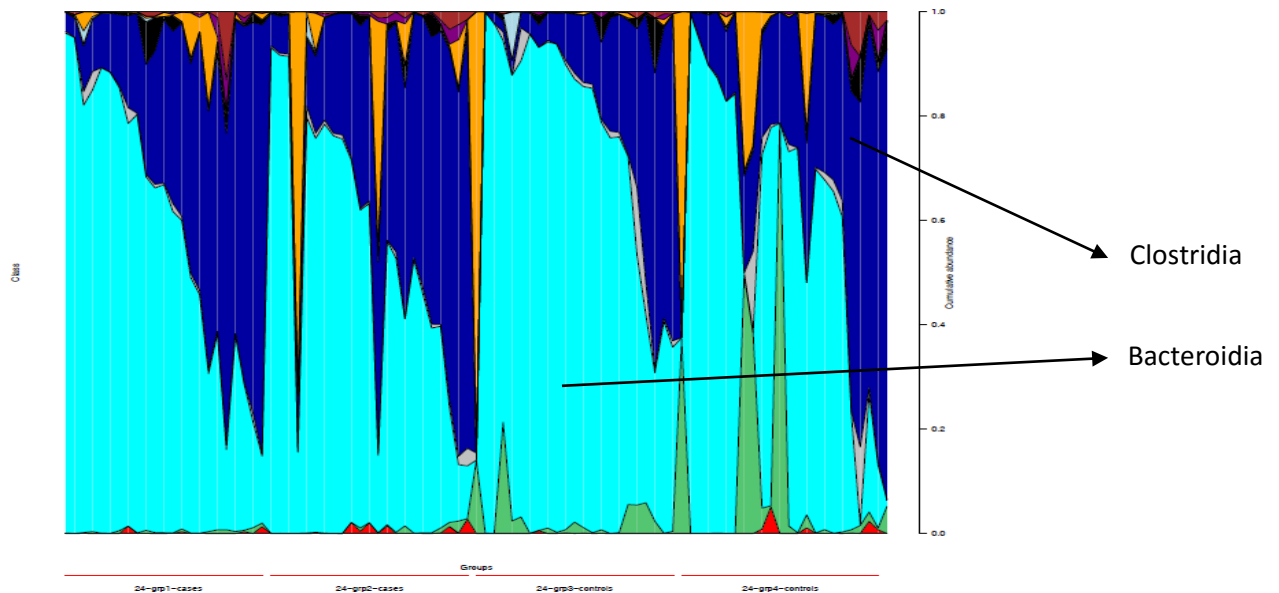
Supplementary Table S1: Reads metrics from output sequence data (PGM quality-approved, trimmed and filtered data) and mean number of preprocessed reads per index after the high-quality filtering approach. * Three outlier samples (indexes 18, 63 and 50 belonging to groups 1, 2 and 3 respectively), were discarded before the subsequent bioinformatics analyses.

Supplementary Figure S1. Faecal bacterial communities present in the different samples (DNA samples from patients' groups 1, 2, 3 and 4) analyzed in this study. Plots of relative abundances of OTUs at four taxonomic levels: phylum (a), class (b), order (c) and family (d). The most abundant and interesting taxa are labeled with arrows.

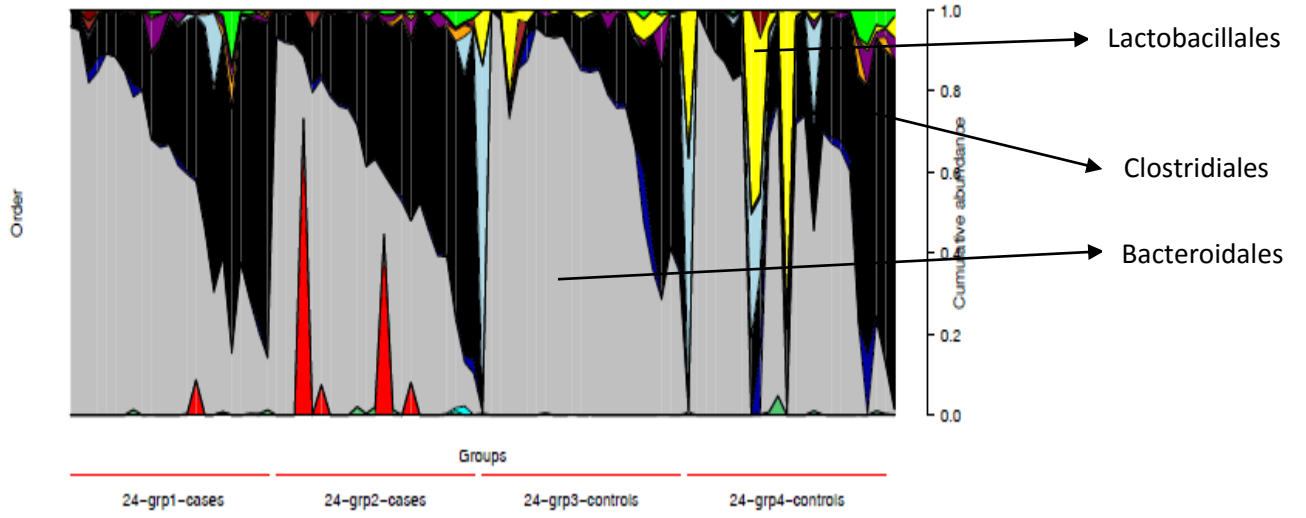
a)



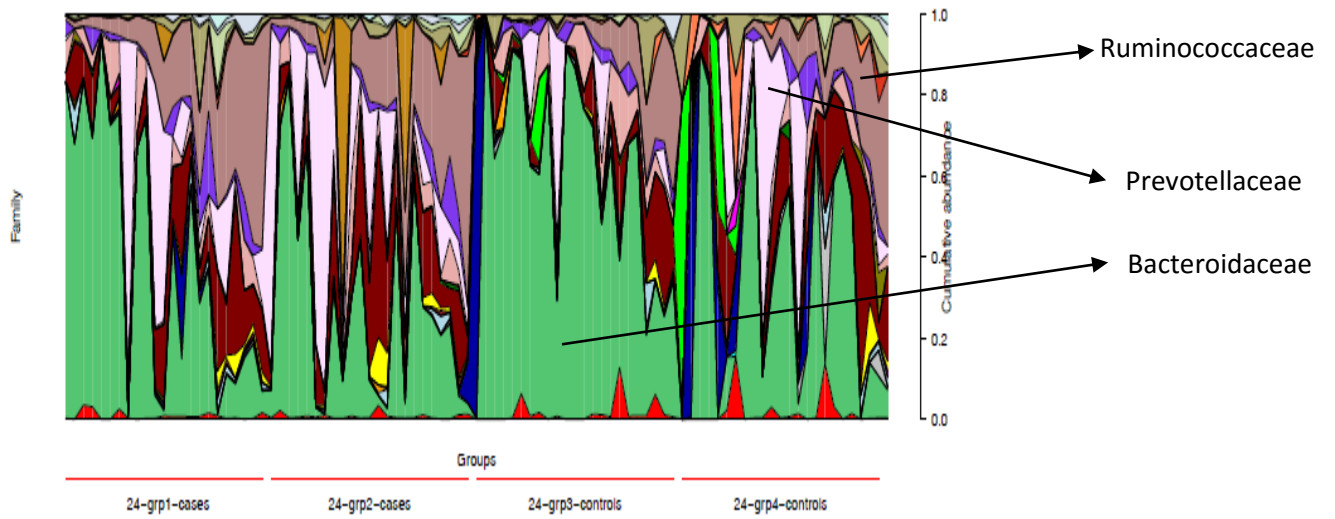
b)



c)



d)



Supplementary Table S1. Reads metrics from output sequence data (PGM quality-approved, trimmed and filtered data) and mean number of preprocessed reads per index after the high-quality filtering approach. * Three outlier samples (indexes 18, 63 and 50 belonging to groups 1, 2 and 3 respectively), were discarded before the subsequent bioinformatics analyses.

	Mean read length (bases)	Median read length (bases)	Minimum read number/index	Mean read number/index	Median read number / index	Mean preprocessed reads number/index
All reads ^a	272.67	273	14,617	42,603.26	41,956.00	29,485.03
Reads from <i>Blastocystis</i> - colonized patients	271.98	274	14,617	43,604.02	44,533.50	30,541.19
Reads from <i>Blastocystis</i> -free individuals	273.34	272	21 039	41,623.79	41,057.00	28,451.34