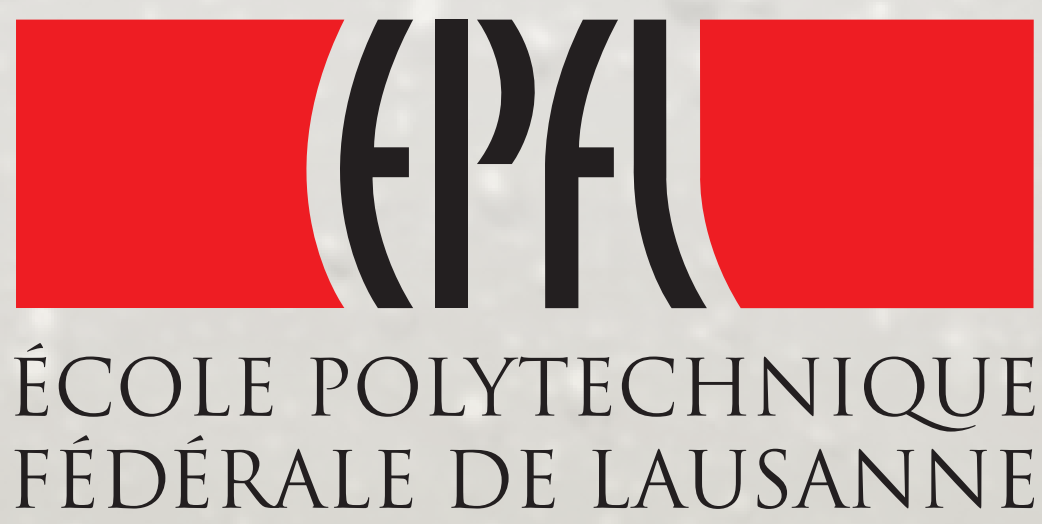# Phage-dependent variability of *Candidatus* 'Accumulibacter phosphatis' population in aerobic granular sludge

Aline Adler[a], E. Cabello[c], O. Cordero Sanchez[b], J. Maillard[a], J. Rougemont[c] and C. Holliger[a]

a Laboratory for Environmental Biotechnology, ENAC-IIE-LBE, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
b Microbial Population Biology Group, IBP - D-USYS, Eldgenössische Technische Hochscule Zürich (ETHZ), Switzerland
c Bioinformatics and Biostatistics Core Facility, SV - BBCF, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Objective

Study the main factors influencing the high variability of *Candidatus* 'Accumulibacter phosphatis' abundance in the metagenome of 46 granules from a lab-scale aerobic granular sludges (AGS) sequencing batch reactor performing N- and P-removal.

Context

>>> Aerobic granular sludges is a promising alternative to conventional activated sludge for wastewater treatment.

>>> The phosphate-accumulating organism (PAO) 'Accumulibacter' is often found with significant abundance in AGS as well as in other enhanced biological phosphorus removal (EPBR) systems and is generally associated with good EBPR performances.

>>> The bacteriophages EPV1 and EPV3 have previously been detected in an EBPR reactor dominated by 'Accumulibacter' (1) and were suspected to be responsible for decrease in 'Accumulibacter' proportions and phosphorus-removal performance (2).

>>> Bacteriophage infections may allow co-existence of losing-competitors increasing the genus diversity. It may also exert selection pressure on the host, modifying the diversity of the host.
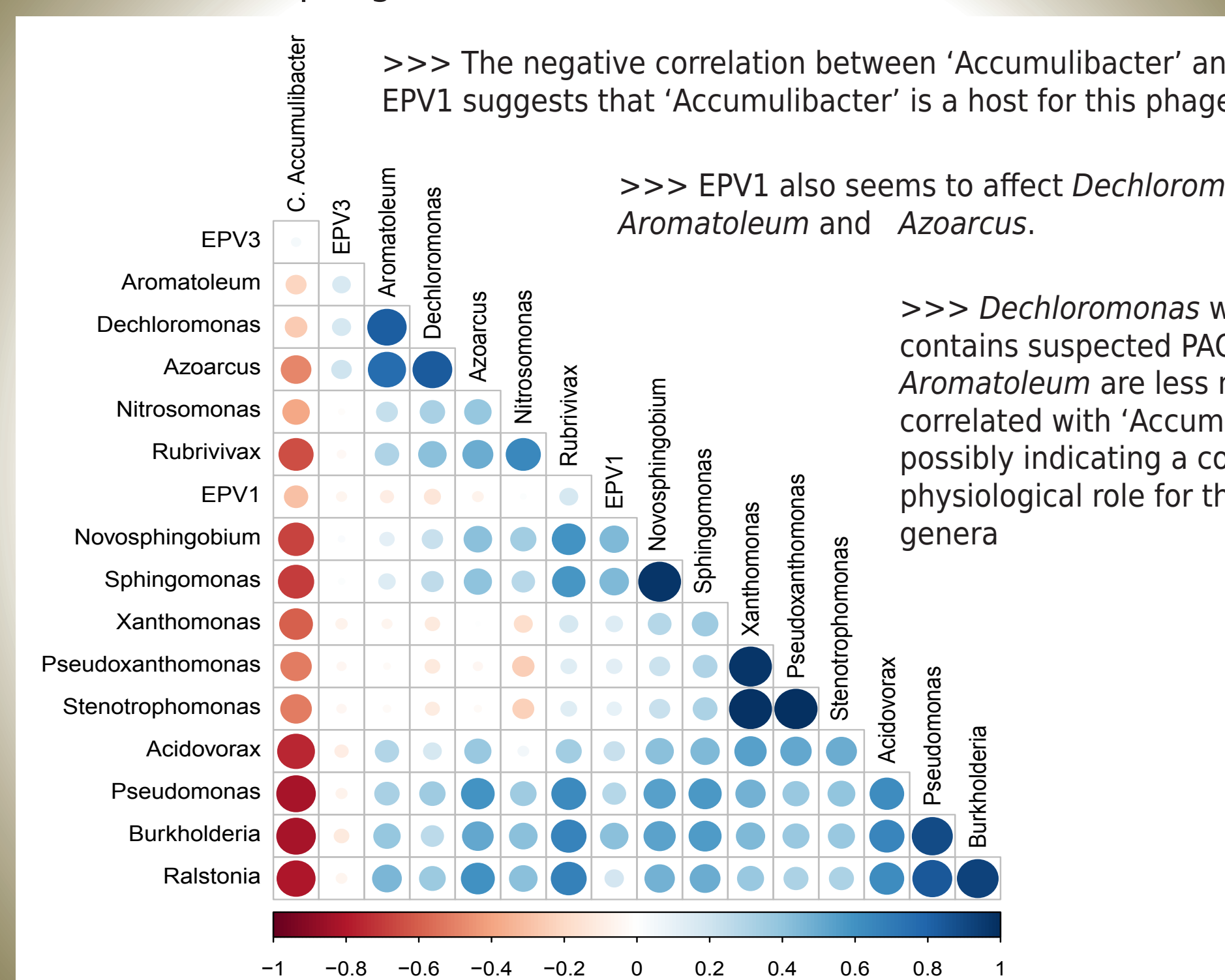
Methodology

>>> The DNA extracted from 46 granules of our routine lab-scale AGS reactor was sequenced with Illumina Mi-seq technology in paired-end mode.

>>> After merging overlapping pairs, sequences were aligned against the NCBI bacterial genomes database allowing for multiple alignments.

>>>The coverage of each strain was estimated by the normalized sum of sequences lengths mapping to its reference genome. The estimated coverage of a genus was obtained by summing the coverages of strains belonging to the genus.

>>> Reads hitting 100 different regions were considered non-specific and removed for this analysis.

>>> The relative proportion of each bacterial genus was estimated based on its relative estimated coverage.

>>> Non-mapping sequences were aligned against the genomes of EPV1 and EPV3 phages.

Results:~$

Genus proportions

>>> The proportions of the dominant genus 'Accumulibacter' show a high variability amongst the different granules (Fig.A).

>>> Only a part of this variability can be linked with the phenotype of the granules (Fig.B and Fig.C).

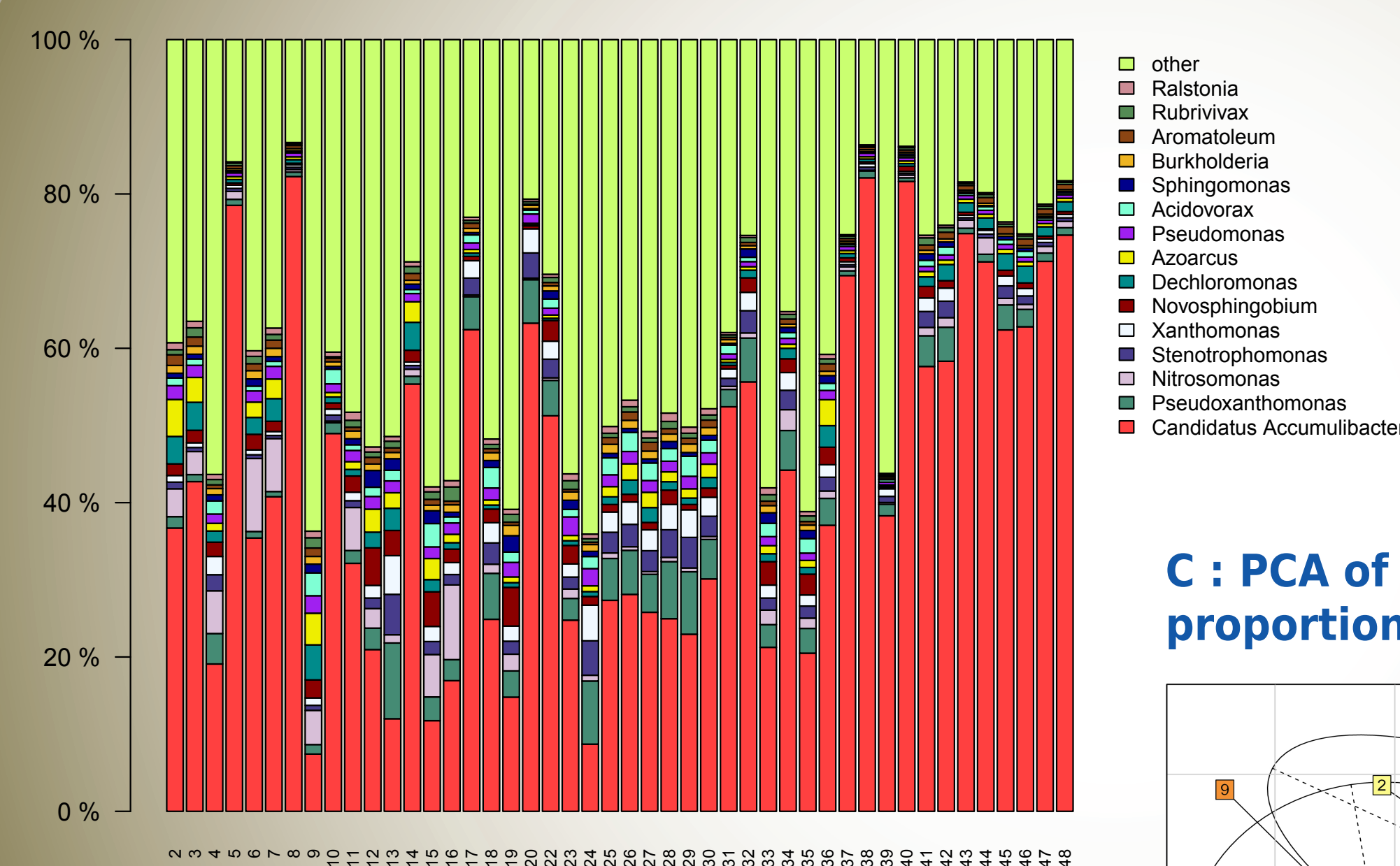### A : Genus estimated proportions in the granules

B : Granule's phenotypes

>>> Six different phenotypes of granules where selected for sequencing :

17-24		"Grainy"
9-16		"Amber"
2-8		"Transparent"
33-40		"white"
25-32		"Cauliflowery"
41-48		"Pink and slimy"

C : PCA of the granules genus proportions

Principal Componant Analysis of the estimated genus proportions in the 46 granules clustered according to their phenotype, 2 axis are shown.

Correlations

'Accumulibacter' proportion is negatively correlated with most other populations, including EPV1 phages.

>>> The negative correlation between 'Accumulibacter' and EPV1 suggests that 'Accumulibacter' is a host for this phage.

>>> EPV1 also seems to affect *Dechloromonas*, *Aromatoleum* and *Azoarcus*.

>>> *Dechloromonas* which contains suspected PAO strains and *Aromatoleum* are less negatively correlated with 'Accumulibacter', possibly indicating a concerted physiological role for these bacterial genera
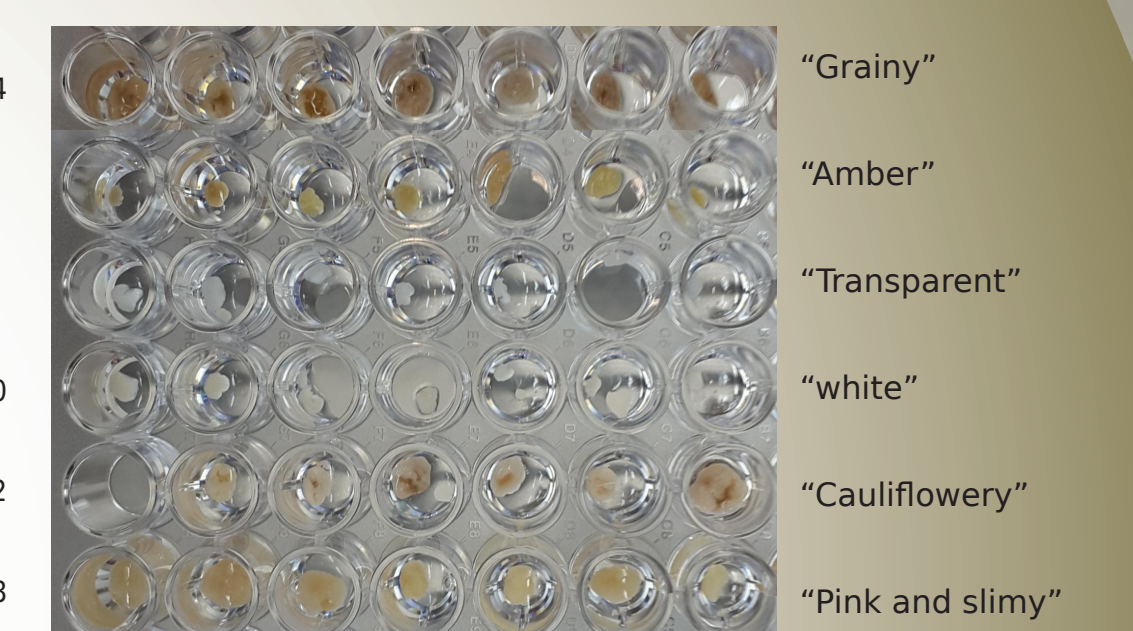
Correlation matrix between the 15 most abundant genera based on estimated proportions and the number of EPV1 and EPV3 related sequences.

Conclusions

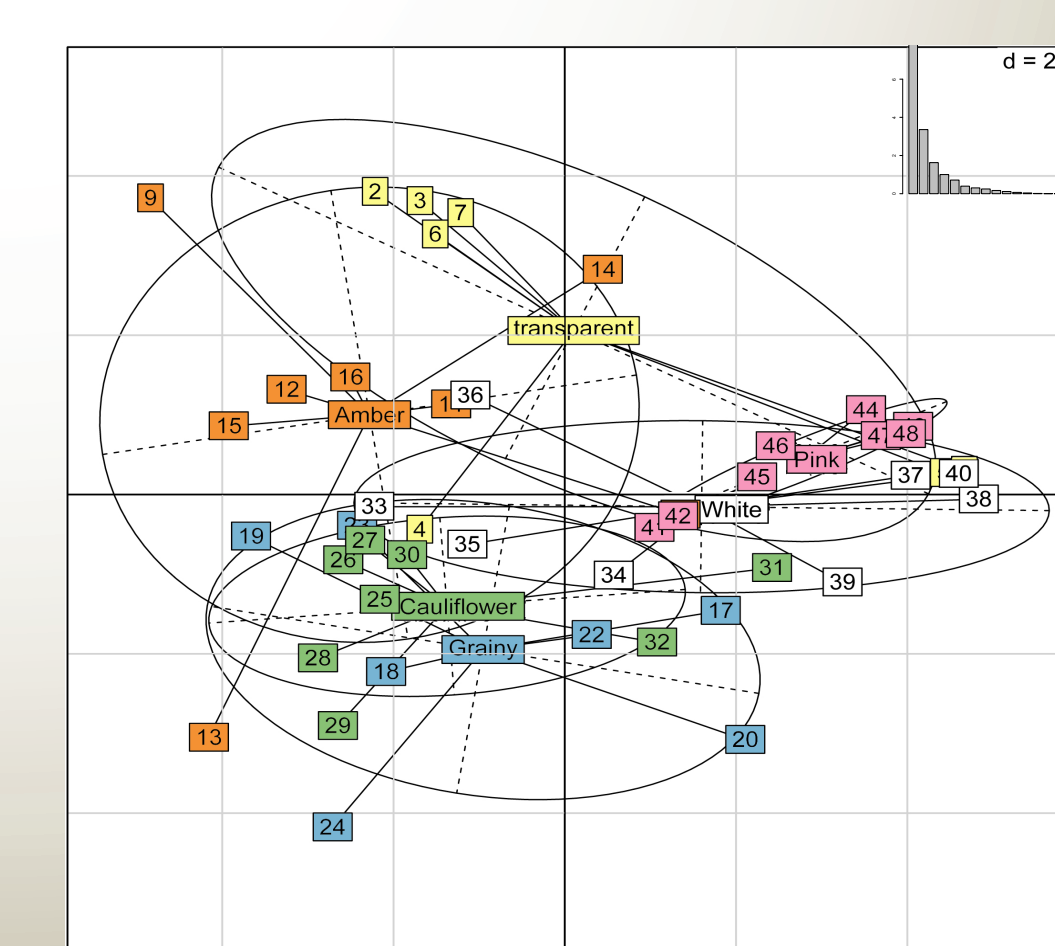>>> The negative correlation between the proportions of 'Accumulibacter' and the number of EPV1 related sequences suggest a phage attack was occurring in our lab scale AGS-reactor at sampling time, affecting the bacterial community composition of the granules.

>>> These results suggest that phage attacks can influence community compositions of the biomass and should be therefore taken into account while studying the variation of the bacterial communities composition of AGS.

Perspectives

>>> Continue to develop adapted bioinformatic tools to study the correlation between the different populations in the granules of our AGS-labscale reactor.

>>> Acquire a better understanding of the functional role of these populations inside the granules.

Long-term objectives

>>> Study the functional gene pool of the AGS to increase the understanding of the global metabolism in the granules and of biofilm formation (QS, EPS production, toxine/anti-toxine systems).

>>> Study the variation of the granule's microbial communities and metagenomes depending on the influent composition, from simple soluble to more complex particulate organic substrates.

>>> Link these variations with reactors performances.

>>>

(1) Skennerton et al., 2011, Plos One
(2) Barr et al., 2010, Fems Microbiology Ecology, 631-642
(3) Langmead B, Salzberg S. Fast gapped-read alignment with Bowtie 2. Nature Methods. 2012, 9:357-359
(4) Edgar, Robert C. (2004), MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Research 32(5), 1792-97
(5) Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics, btu170
(6) R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org