



# Whole genome sequencing of butyrate-producing bacteria

Jakub Mrazek, Francesco Strozzi, Katerina Fliegerova,  
Jiri Simunek, Hana Sechovcova, Diego Grilli, Jan Kopečný



## Goals and approaches

Bacterial whole-genome sequencing with an aim to understand fibrolytic enzyme diversities important for biomass degradation in anaerobic microbial systems. Key aspects of metabolism will be extracted in order to understand the basic biology of the species.

- RuminOmics **Task 6.1** Microbial genomics
- RuminOmics **Task 7.8** Genomic analyses of individual ruminal microorganisms

# Ruminal butyrate-producing bacteria

Bacteria – Firmicutes – Clostridia – Clostridiales - Lachnospiraceae

Bryant and Small (1956) genus *Butyrivibrio*

Moore *et al.* (1976) human isolate *B. crossotus*

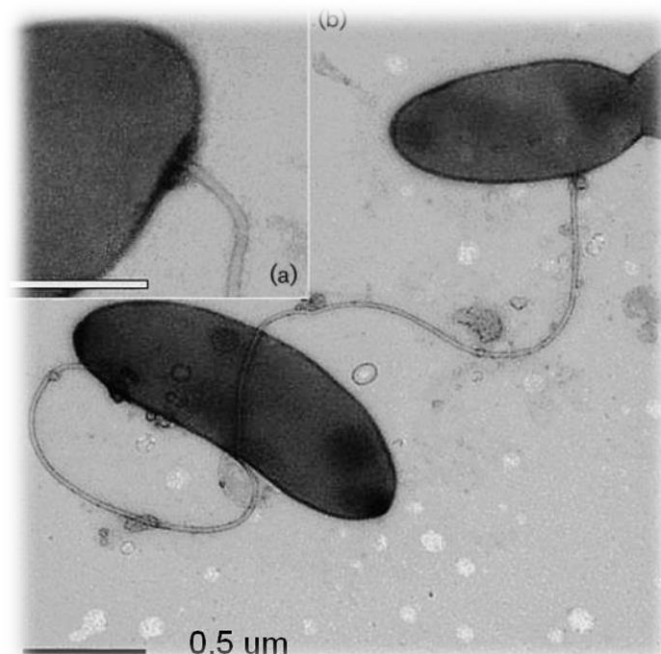
van Gylswyk *et al.* (1996) *Pseudobutyrvibrio ruminis*

Attwood *et al.* (1996) *B. proteoclasticus*

Kopečný *et al.* (2003) *B. hungatei* and *P. xylanivorans*

Kelly *et al.* (2010) *B. proteoclasticus* glycobiome

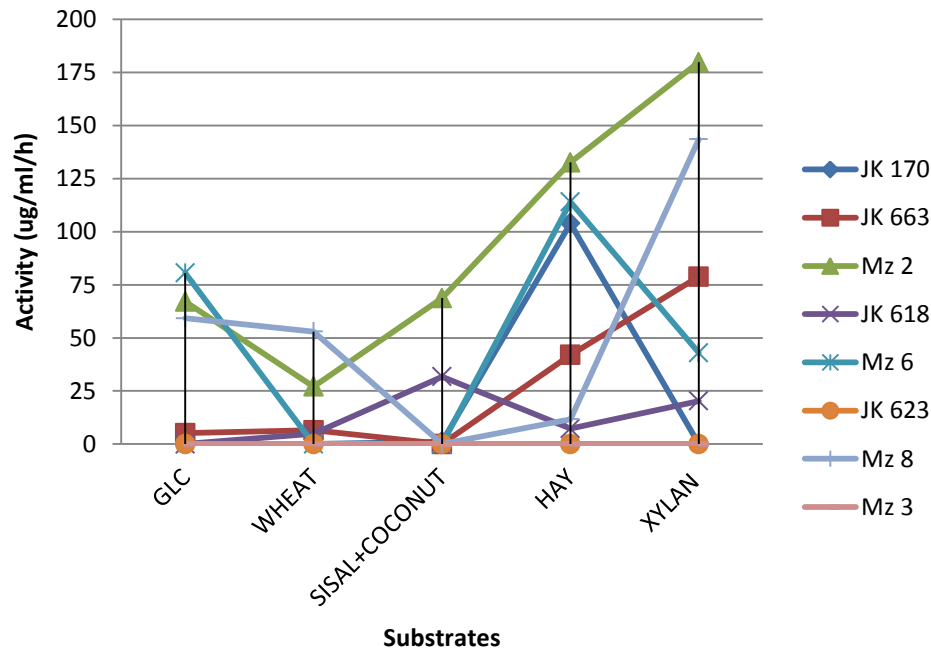
- anaerobic, Gram-positive, butyric-acid producing motile rods
- amylolytic; xylan and cellodextrin degradation
- fibre degradation, protein breakdown, biohydrogenation of lipids, degradation of plant structural carbohydrates



# Xylanolytic butyrate-producing bacteria

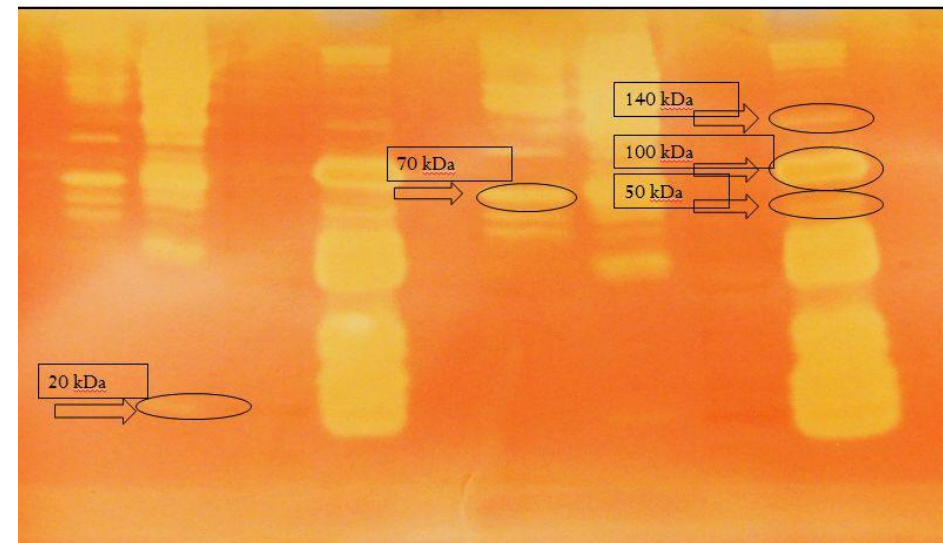
**Spectrophotometric** determination of enzyme activity (Lever 1977)

**Xylanolytic activity**  
5 days incubation



**Zymography - SDS-PAGE** with co-polymerized substrate (Flint, 1994)

UC 142		JK 170		UC 142		JK 170	
15 ul				20 ul			
X+G	X	X+G	X	X+G	X	X+G	X



# Butyrate-producing bacteria selection

Culture collection at IAPG (Prague): over 40 isolates from Czech Republic, Slovenia, United Kingdom, Australia...

## Bacterial strains

*Butyrivibrio proteoclasticus* P-18

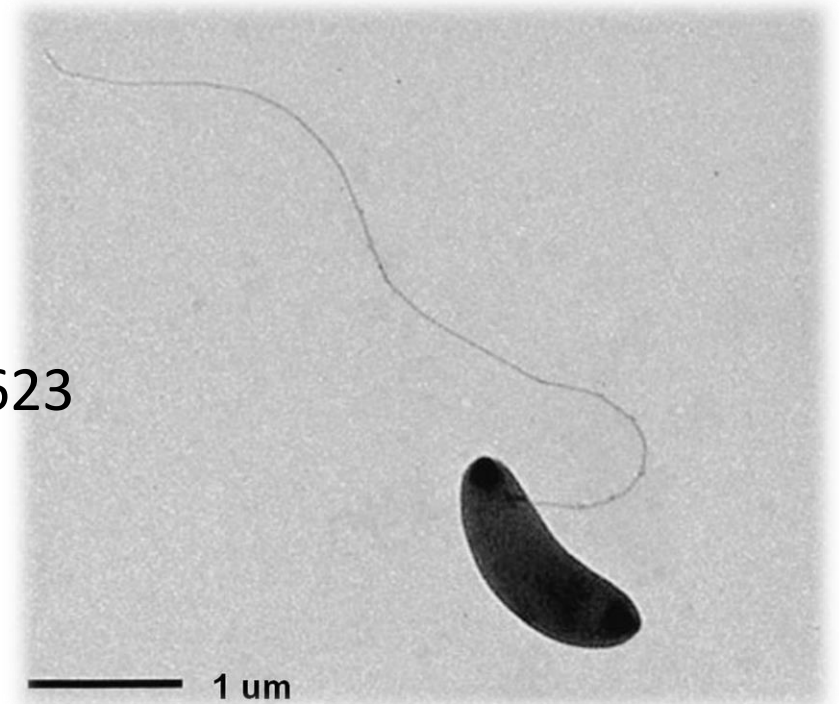
*Butyrivibrio fibrisolvens* 19171(D1)

*B. hungatei* JK615

*Pseudobutyrvibrio* sp. JK10-1

*Pseudobutyrvibrio xylanivorans* JK623

*Pseudobutyrvibrio ruminis* JK626



# DNA isolation

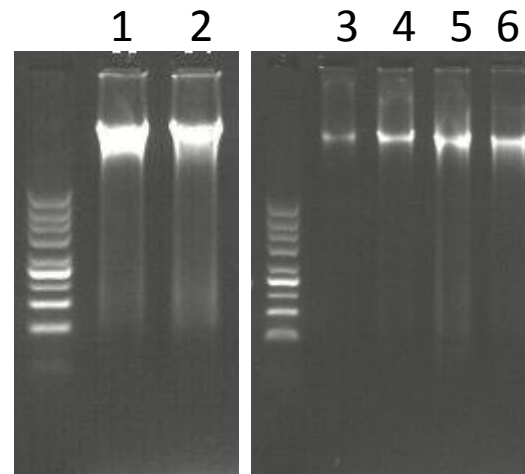
## Illumina vs PacBio sequencing technologies

**Illumina:** HiSeq equipment, “standard”, “easy”, short reads (2 x 250 bp)

**PacBio:** very long reads, high DNA quality requirements

### 1. Illumina sequencing (GATC)

Nucleospin Tissue kit (Macherey-Nagel, Germany)



- 1 – JK615
- 2 – JK623
- 3 – Mz5
- 4 – JK10/1
- 5 – D1
- 6 – JK626

# DNA isolation

## 2. PacBio sequencing

- requirements:
- (a) 7,5 µg gDNA
  - (b) A 260/280 between 1,75 and 1,9 (no RNA)
  - (c) no DNA fragmentation/degradation

Bacterial strains: *P. ruminis* Mz5 and *B. fibrisolvens* D1

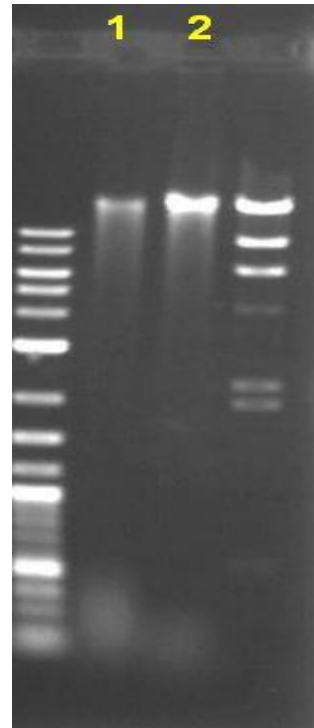
### Tested isolation methods

UltraClean Microbial DNA Isolation Kit (MoBio labs)

Nucleospin Tissue kit (Macherey-Nagel)

Blood and Tissue DNA Isolation Kit (Qiagen)

”Ruminomics“ method (FastPrep + QIAamp DNA Stool kit)



# Genome assembly and annotation

- Raw Illumina sequencing reads are trimmed to remove sequencing adapters using **Trimmomatic** software (Bolger *et al.* 2014)
- The trimmed reads are error-corrected and assembled using **SPAdes** assembler (Bankevich *et al.* 2012)
- The assembly quality is evaluated using **Quast** (Gurevich *et al.* 2013)
- The assembled scaffolds are finally annotated using **PROKKA**, to predict procaryotic genes (Seemann 2014)



# Illumina HiSeq sequencing

0001	Olsenella umbonata A2
0002	Eubacterium pyruvatorans I-6
0003	<i>Butyrivibrio proteoclasticus</i> P-18
0004	<i>Butyrivibrio fibrisolvens</i> D1=19171
0005	<i>Butyrivibrio hungatei</i> JK615
0006	<i>Pseudobutyrvibrio</i> sp. JK 10/1
0007	<i>Pseudobutyrvibrio xylanivorans</i> JK623
0008	<i>Pseudobutyrvibrio ruminis</i> JK626

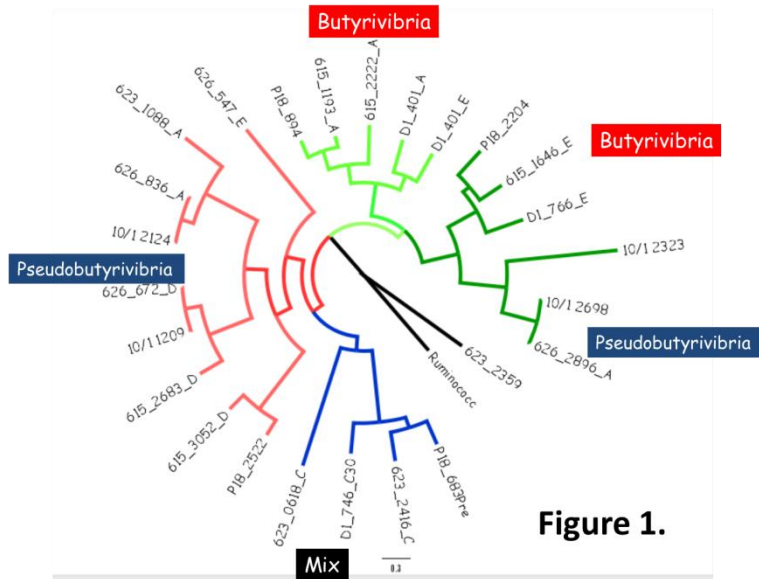
## Statistics

Illumina sequencing of all selected strains resulted in data including number of contigs, average and maximal contig size, number of documented ORF and a possible cumulative genome size. “Hypothetical proteins” represent about 30% of the ORF.



Table 1.	<i>Butyrivibrio hungatei</i> 615	<i>Pseudobutyrvibrio xylanivorans</i> 623	<i>Pseudobutyrvibrio</i> sp. 10/1	<i>Butyrivibrio fibrisolvens</i> D1	<i>Pseudobutyrvibrio ruminis</i> 626	<i>Butyrivibrio proteoclasticus</i> P18
Sequencing statistics						
Number of contigs	62	138	62	123	84	143
Max contig size (Kb)	279	420	151	126	586	275
Number of ORF	3117	2627	2740	4067	2922	3790
Total Size in Kb	3 445	2 866	2 986	4 829	3 153	4 221

## $\beta$ -1,4-Endoglucanase



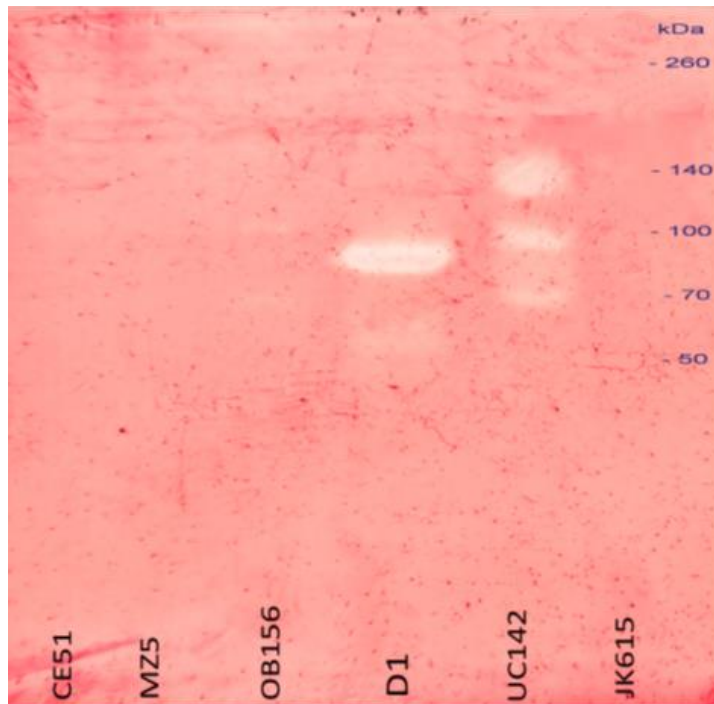
# Cellulolytic enzymes

Celulases in *Butyrivibrio* are separated in 6 groups (Figure 1). Most common are endoglucanases A, D, and E. Phylogenetic tree showed segregation by tested genera *Butyrivibrio* and *Pseudobutyrvibrio*.

Table 2.	Butyrivibrio	Pseudobutyrvibrio	Pseudobutyrvibrio	Butyrivibrio	Pseudobutyrvibrio	Butyrivibrio
Glycolytic enzymes	hungatei 615	xylanivorans 623	sp. 10/1	fibrisolvens D1	ruminis 626	proteoclasticus P18
Beta glucosidase	7	4	4	11	4	9
Endoglucanase A	2	1	3	1	3	1
Endoglucanase B				1		
Endoglucanase C307				1		1
Endoglucanase D	2	2	1		1	1
Endoglucanase E	1	1		2	1	1
Endoglucanase Z	1					
Exoglucanase/ xylanase	1					
Cellobiose phosphorylase	2	1	2	2	2	1
Sum of hemicellulases	16	9	10	18	11	14

# Cellulolytic enzymes

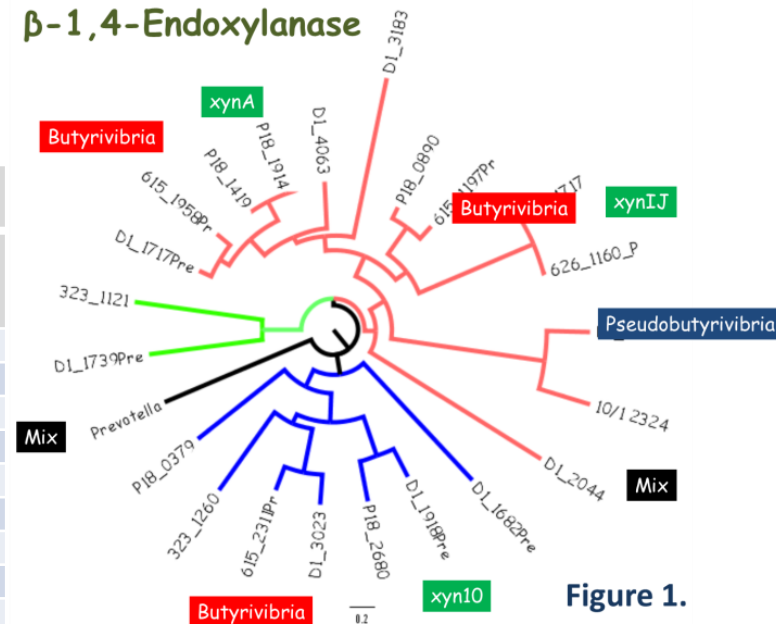
Zymogram with cellulase activities of extracellular and intra-cellular fractions of *Butyrivibrio* isolates.



Strains	Bands	MW of the most active EG
CE51	0	
D1	3	90
JK 615	0	
UC142	4	140, 95, 65
OB156	3	100
Mz5	0	100

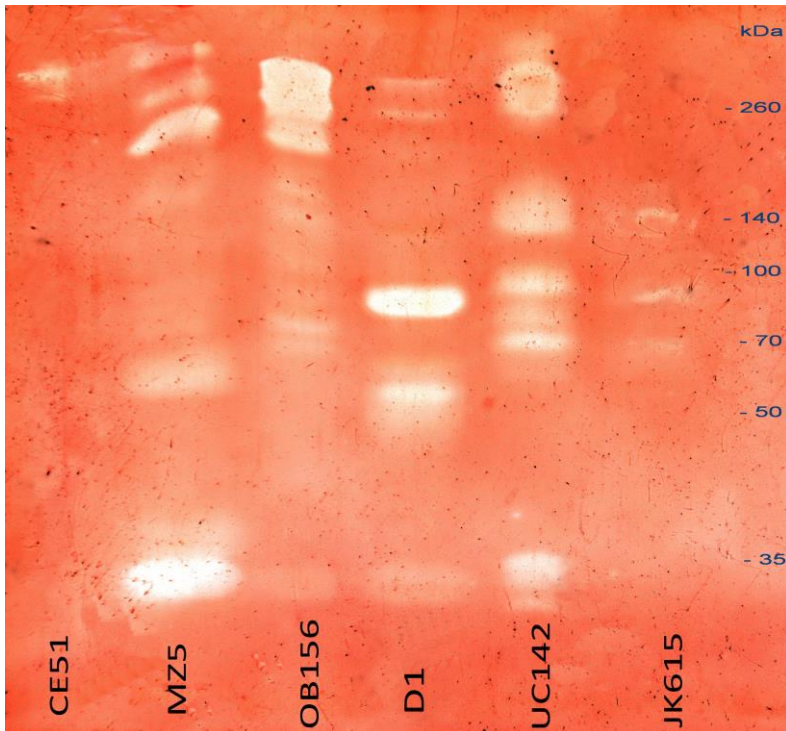
# Hemicellulolytic enzymes

Table 3.	Butyrivibrio	Pseudobuty	Pseudobuty	Butyrivibrio	Pseudobuty	Butyrivi
	hungatei	xylanivorans	sp. 10/1	fibrisolvens	ruminis	proteoclasticum
	615	623	sp. 10/1	D1	626	P18
<b>Glycolytic enzymes</b>						
<b><math>\alpha</math> glucosidase</b>				1		2
<b>Oligo1,6 glycosidase</b>					3	
<b>Exoglucanase/ xylanase</b>	1					
<b>Gal/glu bind protein</b>	1	5	6	1	2	2
<b>Oligo 1,6 glucosidase</b>			2			
<b>Glucurono xylanase</b>	1					
<b>Glucuronidase</b>	1					
<b><math>\beta</math> xylosidase</b>	8	3	2	7	3	1
<b><math>\alpha</math> xylosidase</b>	3	2		1		
<b>Xylosidase/arabinoxidase</b>	2	1	1	1		2
<b>Arabinoxylan arabinohydrolase</b>	3	1		1	1	1
<b>Oligoxylanase</b>		1	2		1	
<b>Endo 1.4 <math>\beta</math> xylanase A</b>	2	1			1	2
<b>Endo 1.4 <math>\beta</math> xylanase</b>						1
<b>Endo 1.4 <math>\beta</math> xylanase Y</b>	3	1		1		1
<b>Endo 1.4 <math>\beta</math> xylanase Z</b>	2	1		1		
<b>Bifunctional xylanase deacetylase</b>	1	1	1	1	1	1
<b>Endo 1,4 xylosidase</b>			2			1
<b>Sum of hemicellulases</b>	28	17	14	15	12	13



Butyrivibria are clustering with gene *xyn10* (cluster *xynB*; GH family 10) and gene *xynA* (cluster *xynA*; GH family 11) (Figure 1). Pseudobutyvibria usually produce *xynIJ* (cluster *xynA*; GH family 11).

# Hemicellulolytic enzymes



Zymogram with xylanase activities of extracellular and intra-cellular fractions of *Butyrivibrio* isolates

Strains	Bands	MW of the most active EX
CE51	2	295, 275
D1	6	90, 55
JK 615	3	140
UC142	9	90, 35
OB156	10	265, 255, 225
Mz5	10	245, 33

# B. fibrisolvens D1 PacBio sequencing

S. technology	No contigs	Max size (Kb)	Total size (Kb)	GC (%)	ORF	tRNA
Illumina	123	126	4829	~39	4076	32
PacBio	9	4656	5020	38.9	4313	?
PB + Illumina	6	4337	4908	39.9	4104	48

## Xylanases distribution

	Seq. technology	xylanases	signal peptides
<i>B. proteoclasticus</i> I-6	Illumina	9	3
<i>B. fibrisolvens</i> D1	Illumina	19	9
<i>B. fibrisolvens</i> D1	PacBio	6	?
<i>B. fibrisolvens</i> D1	PacBio + Illumina	14	?
<i>B. hungatei</i> JK615	Illumina	5	4
<i>Pseudobutyrvibrio</i> sp. JK10-1	Illumina	6	2
<i>P. xylanivorans</i> JK623	Illumina	6	1
<i>P. ruminis</i> JK626	Illumina	5	1



## Future plans

- Data mining from genome sequences
- Xylanase “xyn B” gene cloning and protein analyses
- Additional sequencing on Ion Torrent (400 bp reads):  
*Pseudobutyrvibrio xylanivorans* JK170, *Clostridium proteoclasticus* UC142
- Transcriptomics??



# Acknowledgement

Institute of Animal Physiology and Genetics, CAS, Czech Republic

Rowett Institute of Nutrition and Health University of Aberdeen (UNIABDN), United Kingdom

Parco Tecnologico Padano (PTP), Italy

