



Comparative genomics of *Streptococcus macedonicus* ACA-DC 198 against related species within the *Streptococcus bovis*/*Streptococcus equinus* complex

Papadimitriou K.¹, Anastasiou R.¹, Georgalaki M.¹, Ferreira S.³, Supply P.^{3,4}, Papandreou N.C.², Pot B.⁴ and Tsakalidou E.¹

¹Agricultural University of Athens, Department of Food Science and Technology, Laboratory of Dairy Research, Athens, Greece

²University of Athens, Faculty of Biology, Department of Cell Biology and Biophysics, Athens, Greece

³Genoscreen, Genomic Platform and R&D, Campus de l'Institut Pasteur, Lille, France,

⁴Institut Pasteur de Lille, Center for Infection and Immunity of Lille (CIIL), Lille, France



Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

•Introduction

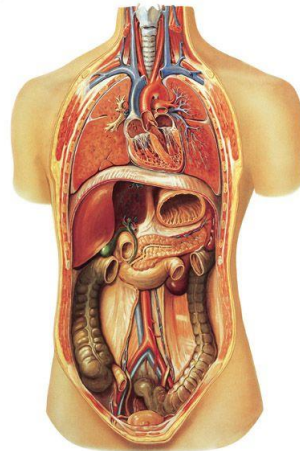
1. Lactic acid bacteria (LAB) are routinely used in food fermentations and as probiotics
2. Besides LAB that are of a benign nature there are several LAB that are notorious pathogens (e.g. Enterococci and Streptococci)
3. Streptococci are an extreme example since today only *S. thermophilus* is considered to be a non-pathogenic and food-grade species when compared to GAS, GBS and *S. pneumoniae* (Bolotin et al. Nature Biotechnol. 2004)
4. However, *S. thermophilus* is not the only streptococcal species found in fermented food especially of dairy origin...

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Introduction

5. Streptococci that can be found growing in milk belong to the *Streptococcus bovis*/*Streptococcus equinus* complex (SBSEC)

S. equinus
(formerly *S. bovis*)



colon cancer
infective
endocarditis
septicemia
meningitis

???

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

•Introduction

6. *Streptococcus macedonicus* which was originally isolated and characterized from traditional Greek Kasseri cheese belongs to the SBSEC (Tsakalidou et. al. Int J Syst Bacteriol. 1998) and seems to have milk and dairy products as its primary ecological niche (De Vuyst and Tsakalidou Int Dairy J. 2008).

7. The close phylogenetic relationship of *S. macedonicus* with pathogenic species of the SBSEC like *S. gallolyticus* and *S. pasteurianus* according to current taxonomy raises concern about the safety of the species

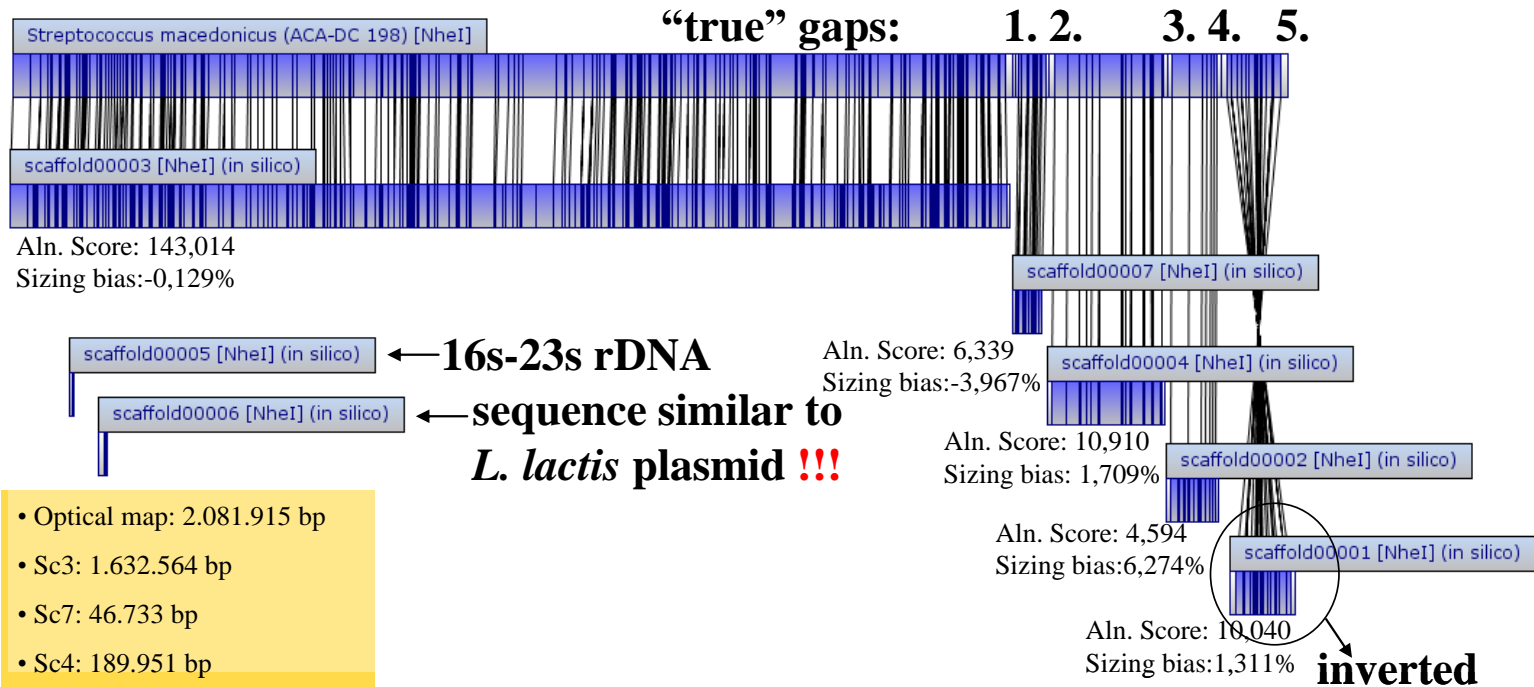
The aim of this work was to sequence the complete genome of the dairy isolate *S. macedonicus* ACA-DC 198 in order to assess *in silico* the pathogenic potential of the species and its potential adaptation to the milk environment

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Sequencing the genome of *S. macedonicus* ACA-DC 198

1st step: shotgun pyrosequencing with 454 GS-FLX titanium (>100 contigs)

2nd step: 3kb paired-end pyrosequencing with 454 GS-FLX titanium (7 scaffolds)



- Optical map: 2.081.915 bp
- Sc3: 1.632.564 bp
- Sc7: 46.733 bp
- Sc4: 189.951 bp
- Sc2: 84.486 bp
- Sc1: 107.539 bp
- Sc5: 5.020 bp
- Sc6: 12.099bp

Assembly: sc3/sc7/sc4/sc2/sc1(inv)/sc3

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

• Sequencing the genome of *S. macedonicus* ACA-DC 198

3rd step: gap-closure and polishing with Illumina sequencing using the HiSeq 2000 (1 chromosome and 1 plasmid)

4th step: validation of the overall assembly (>200X coverage) with an *NheI* optical map



Chromosome: 2,130,034 bp

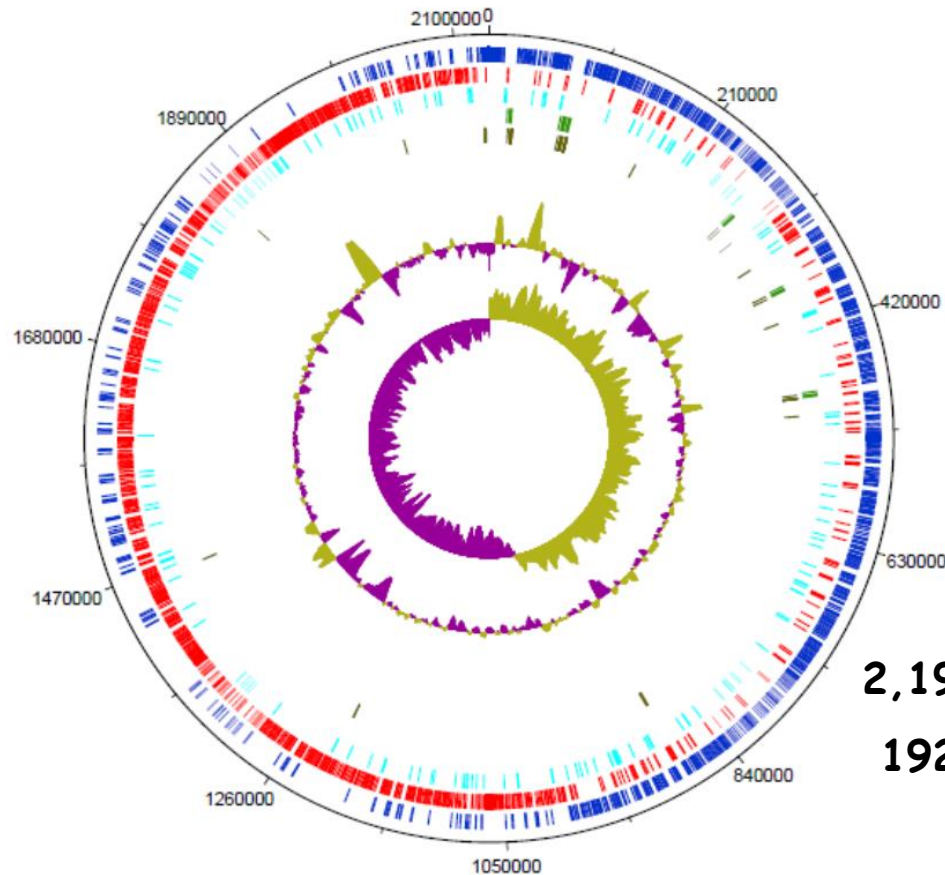
Plasmid: 12,728 bp (pSMA198)

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

•Annotating the genome of *S. macedonicus* ACA-DC 198

Annotation was performed with the BaSys and the RAST pipelines and Kodon software

Final corrections and quality assessment was performed using GenePRIMP



2,192 protein coding genes

192 potential pseudogenes

18 rRNA genes

70 tRNA genes

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

•Comparative genomics of *S. macedonicus* ACA-DC 198

1. In the SBSEC there are currently five additional complete genome sequences available:

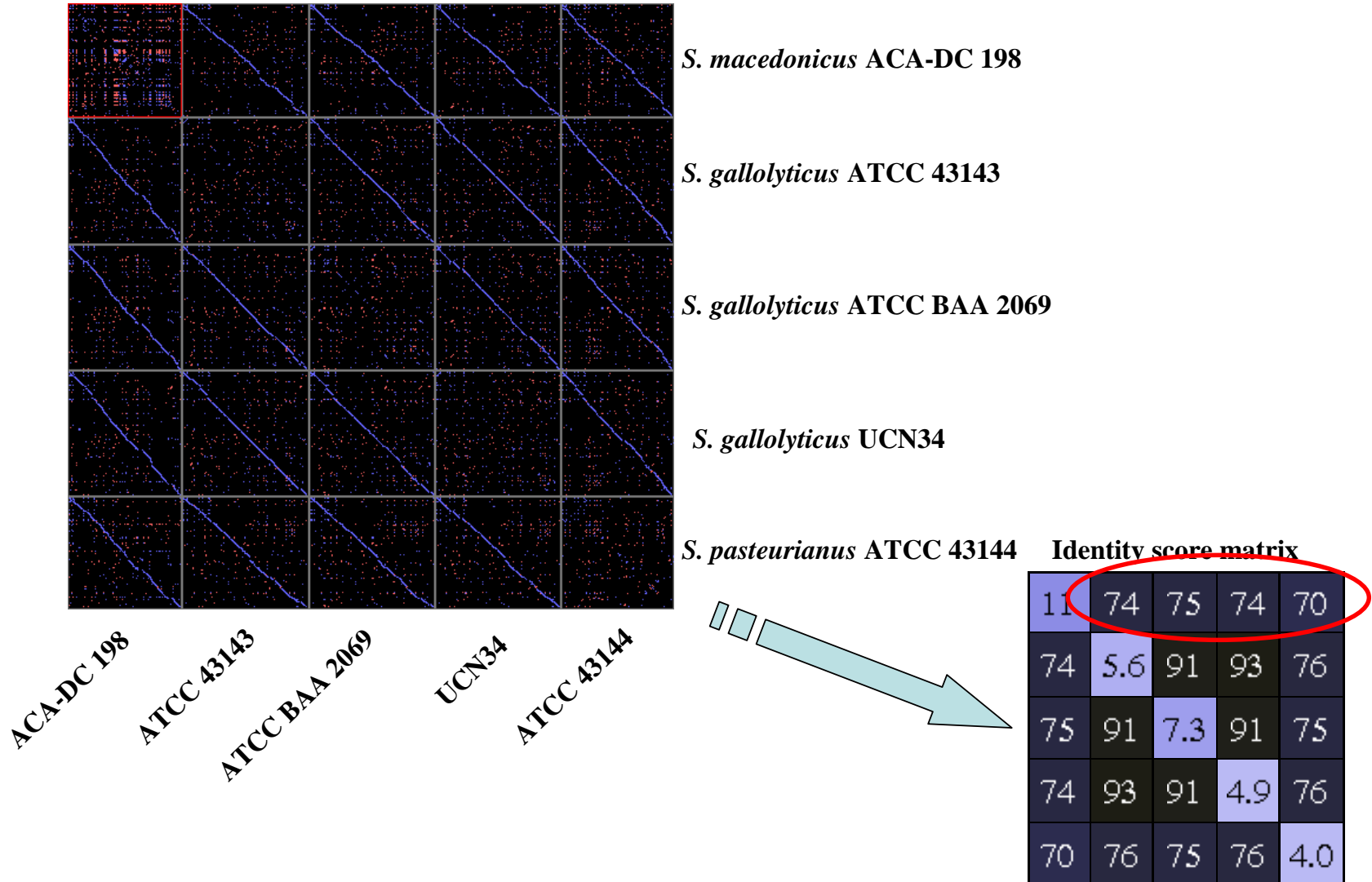
- S. gallolyticus* UCN34 (human blood)
- S. gallolyticus* ATCC BAA 2069 (human blood)
- S. gallolyticus* ATCC 43143 (human blood)
- S. pasteurianus* ATCC 43144 (human blood)
- S. infantarius* CJ18 (suusac fermented camel milk)

2. The inclusion of *S. macedonicus* and *S. pasteurianus* as subspecies of *S. gallolyticus* has been previously suggested (Schlegel et al. Int J Syst Evol Microbiol. 2003), but this taxonomic reappraisal has not been formally accepted due to low DNA-DNA hybridization relatedness values (<70%) (Whiley et al. Int J Syst Evol Microbiol. 2003)

3. The complete genome sequence of *S. macedonicus* offered new opportunities to investigate the properties of the species at the genomic scale

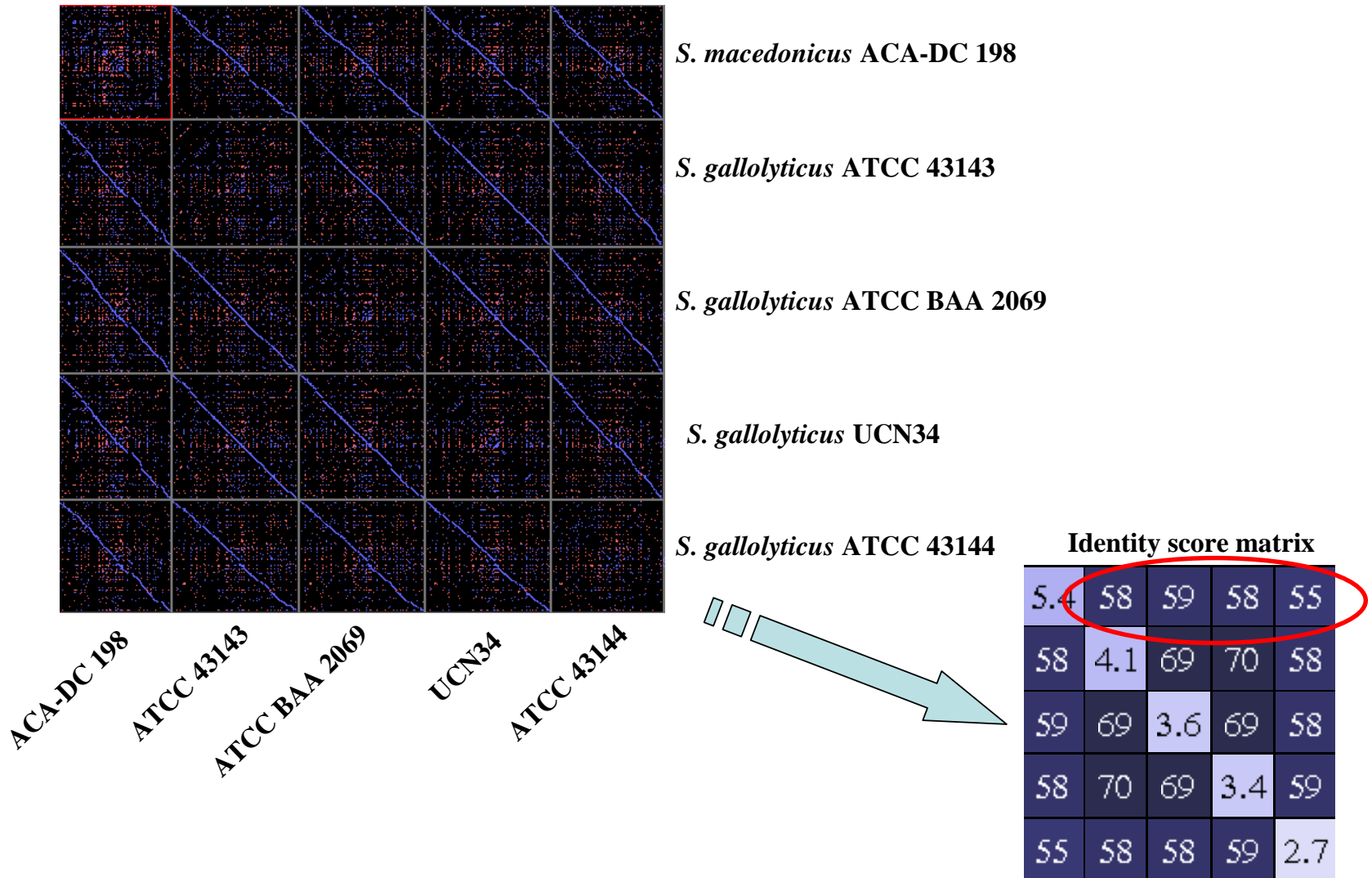
Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Comparative genomics of *S. macedonicus* ACA-DC 198



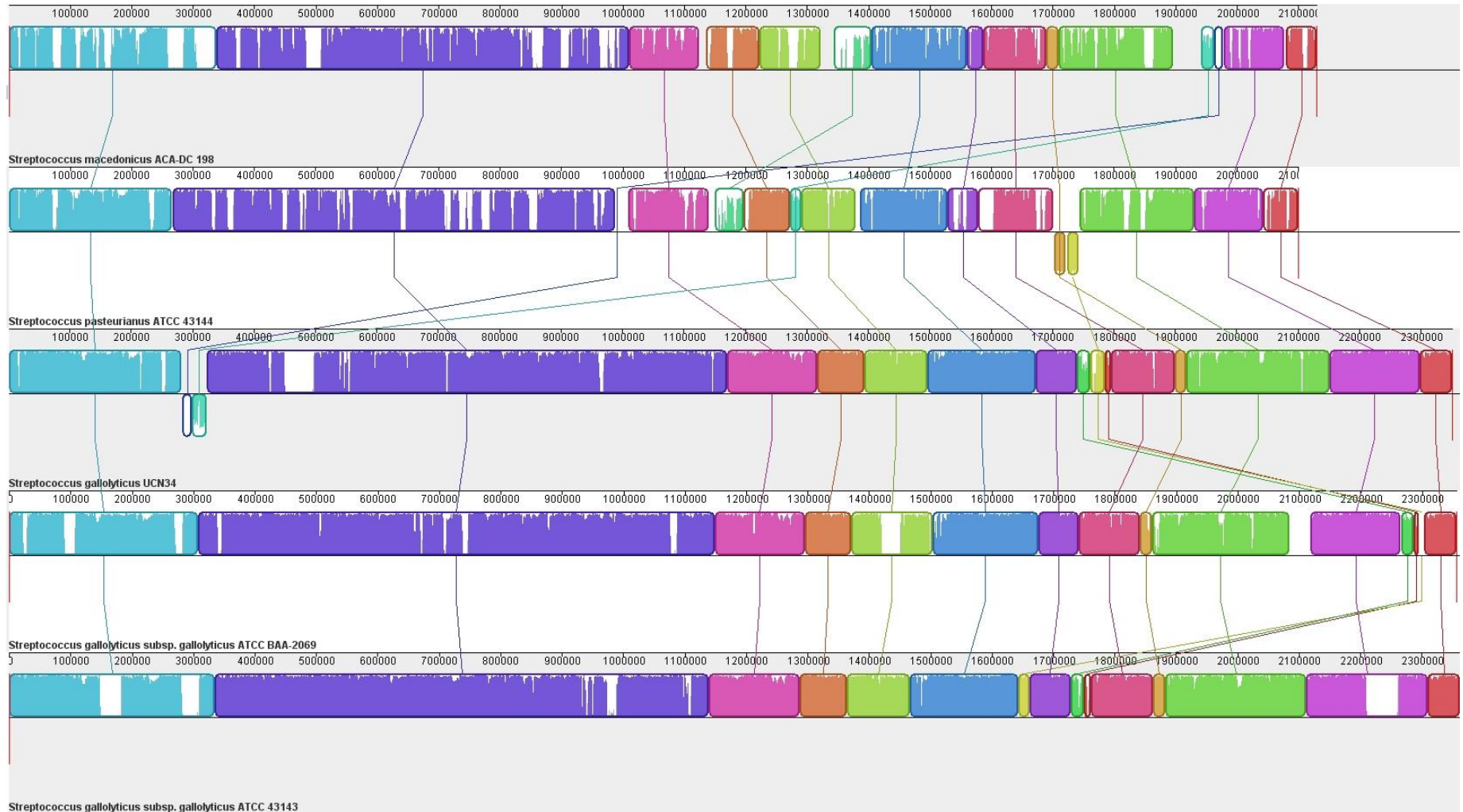
Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Comparative genomics of *S. macedonicus* ACA-DC 198



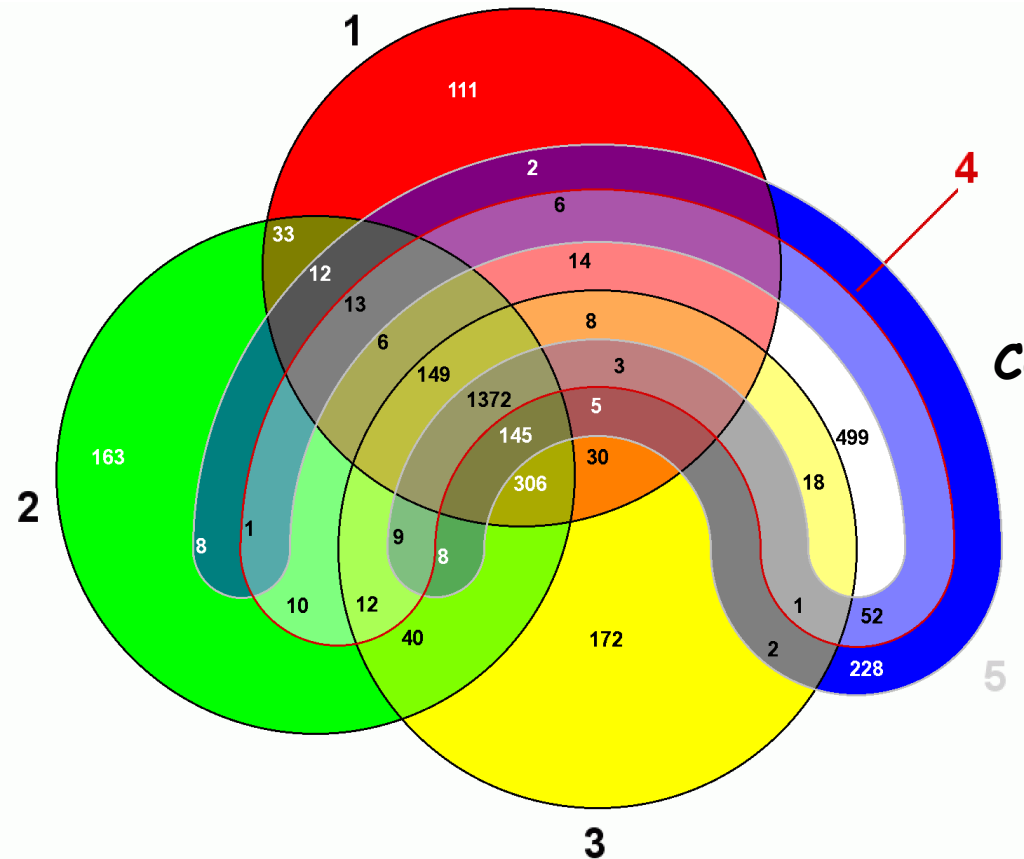
Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Comparative genomics of *S. macedonicus* ACA-DC 198



Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Comparative genomics of *S. macedonicus* ACA-DC 198

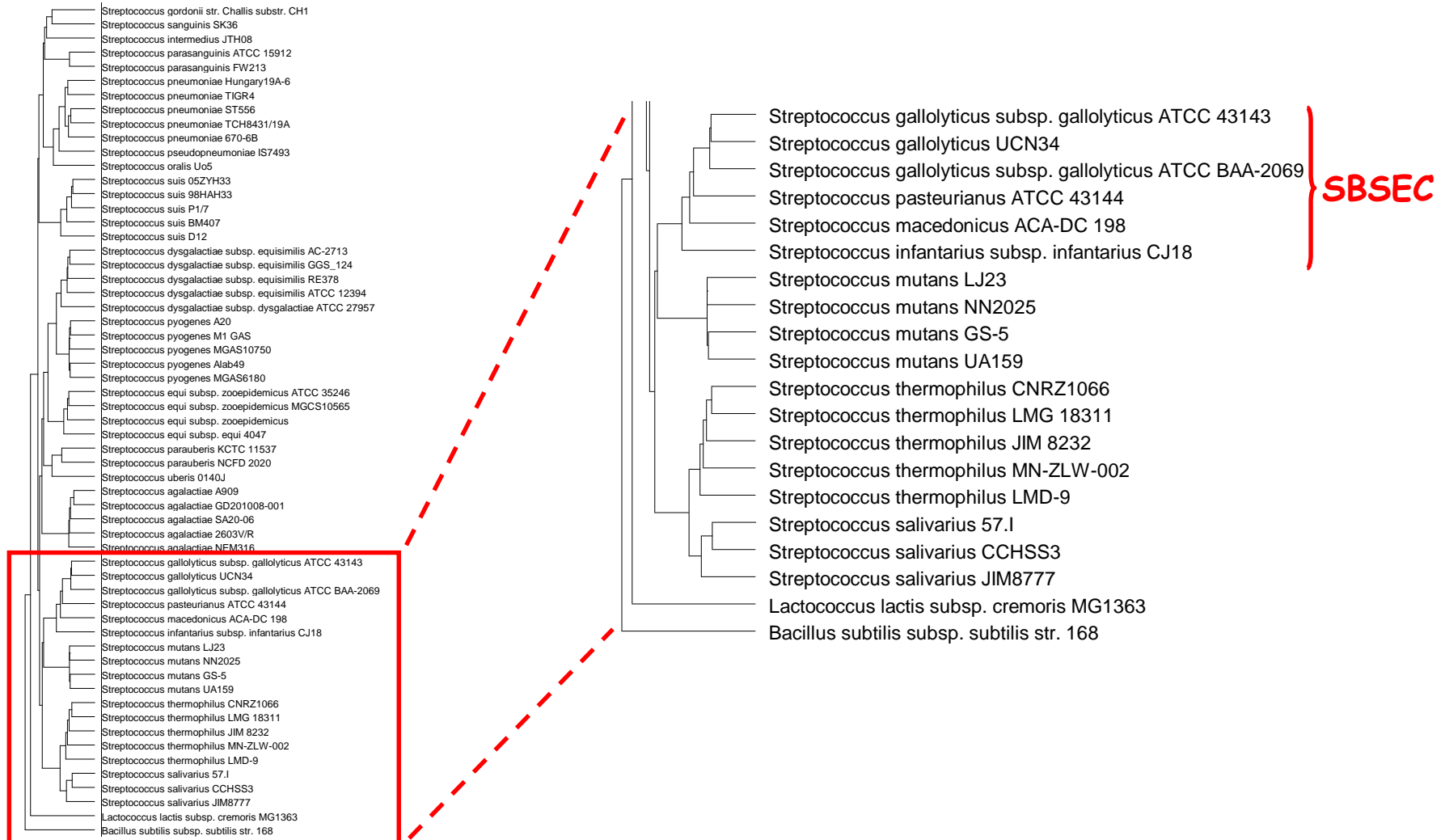


Core genome: **62.6%**

- 1: *Streptococcus_gallolyticus*_UCN34_NC_013798
- 2: *Streptococcus_gallolyticus*_subsp_gallolyticus_ATCC_43143_DNA
- 3: *Streptococcus_gallolyticus*_subsp_gallolyticus_ATCC_BAA-2069_NC_015215
- 4: *Streptococcus_macedonicus*_ACA-DC_198_main_chromosome
- 5: *Streptococcus_pasteurianus*_ATCC_43144_NC_015600

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Comparative genomics of *S. macedonicus* ACA-DC 198



Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Comparative genomics of *S. macedonicus* ACA-DC 198

1. Our findings support that most probably *S. macedonicus* is a separate species from *S. gallolyticus* as also suggested by the International Committee on Systematics of Prokaryotes
2. However, our findings are not conclusive since the results of the comparative genomics analysis are open to several interpretations and standards for this type of analysis are yet to be established

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Some additional characteristics of the genomes under investigation

Species	Genome size (Mb)	No. of protein coding genes	No. of potential pseudogenes/ (% percentage)
<i>S. gallolyticus</i> ATCC BAA 2069	2.35	2329	nr*/(nr)
<i>S. gallolyticus</i> ATCC 43143	2.36	2287	41(1.8)
<i>S. gallolyticus</i> UCN34	2.35	2251	28/(1.2)
<i>S. macedonicus</i> ACA-DC 198	2.13	2192	192/(8.7)
<i>S. pasteurianus</i> ATCC 43144	2.10	1869	157/(7.7)
<i>S. infantarius</i> CJ18	1.98	1964	nr/(4.6)

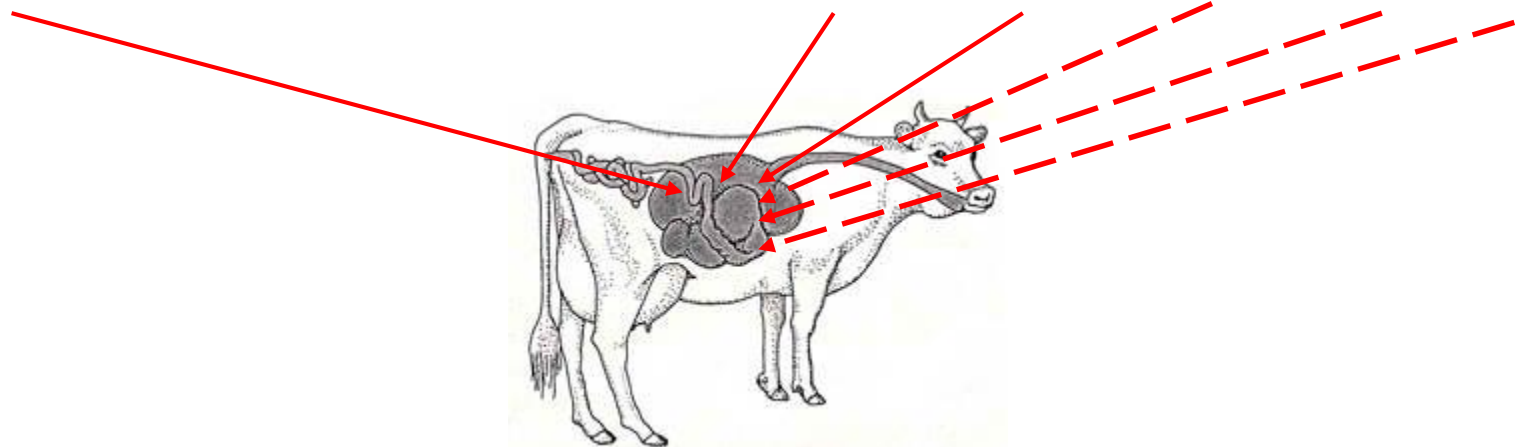
* not reported

- S. macedonicus*, *S. pasteurianus* and *S. infantarius* genomes are being shaped by selective pressures that favor extensive gene loss events and genome decay processes when compared to the *S. gallolyticus* genome
- This property (i.e. genome decay) has been linked to the adaptation of bacteria to rich in nutrients environments as in the case of *S. thermophilus* adaptation to the milk environment

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Niche-specific genes presence/absence

<i>S. gallolyticus</i> UCN 34 locus_tag	gene	function	<i>S. gallolyticus</i> ATCC BAA 2069	<i>S. gallolyticus</i> ATCC 43143	<i>S. macedonicus</i> ACA-DC 198	<i>S. pasteurianus</i> ATCC 43144	<i>S. infantarius</i> CJ18
gallo_0112	fruA	fructan hydrolase	✓	✓	-	-	-
gallo_0330	-	beta-1,4-endoglucanase (cellulase)	✓	✓	-	-	-
gallo_0757	-	α-amylase	✓	✓	-	-	-
gallo_0162	-	mannase	✓	✓	-	-	pseudo
gallo_0189	-	endo-beta-1,4-galactanase	✓	✓	pseudo	✓	pseudo
gallo_1577	-	pectate lyase	✓	✓	✓	pseudo	✓
gallo_1578	-	pectate lyase	✓	✓	✓	pseudo	✓
gallo_1632	amyE	α-amylase	✓	✓	✓	pseudo	✓
gallo_0933	tanA	tanins degradation	✓	✓	-	-	-
gallo_1609	similar to tanA	tanins degradation	✓	✓	-	-	-
gallo_2106	padC	gallic acid decarboxylation	✓	✓	✓	✓	-
gallo_0906		gallic acid decarboxylation	✓	✓	pseudo	✓	-
gallo_0818	bsh	bile salt hydrolase	✓	✓	pseudo	✓	✓



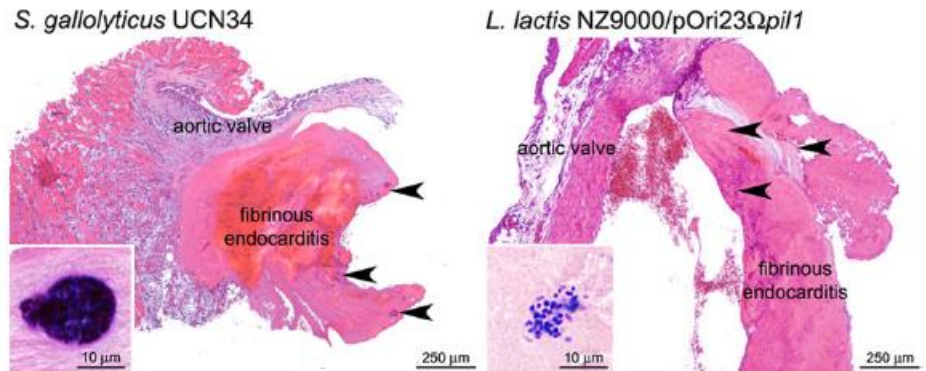
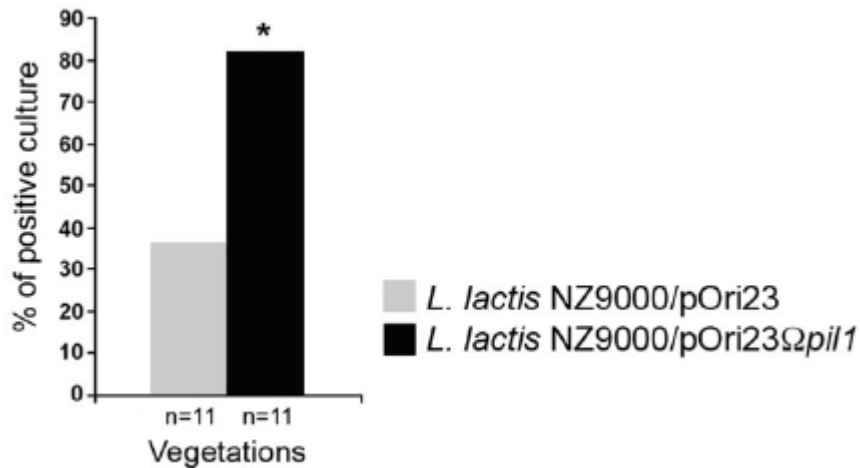
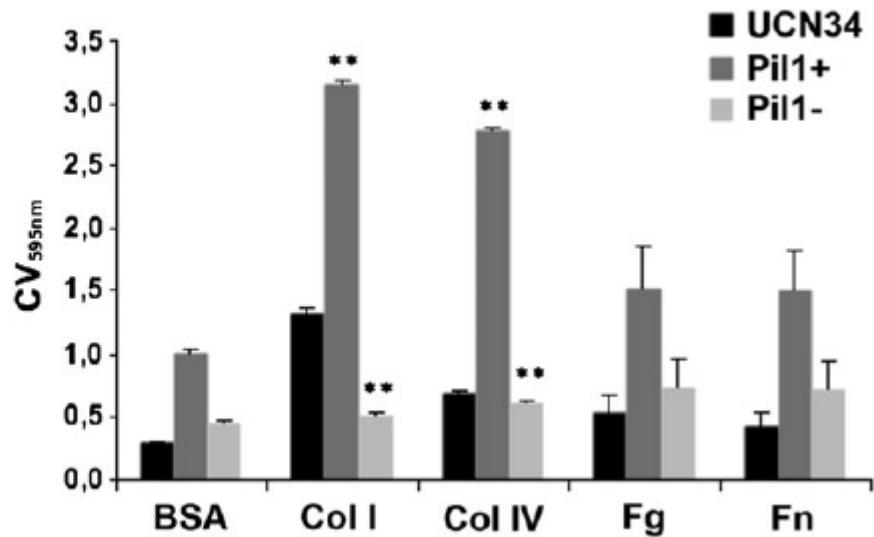
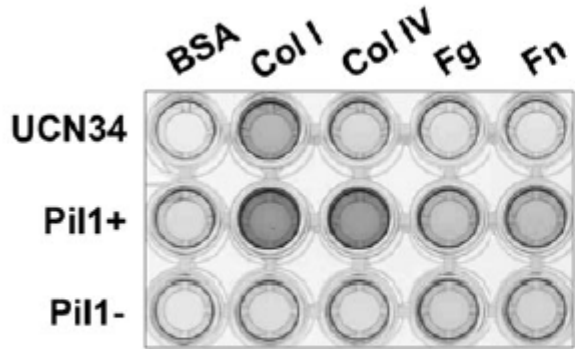
Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Pathogenicity genes presence/absence

1. Virulence factors have been rarely investigated in *S. bovis* despite its known pathogenicity
2. *In silico* analysis of *S. gallolyticus* revealed that it contains three pilus gene clusters (*pil1*, *pil2*, *pil3*), which may mediate its binding to the extracellular matrix (ECM)
3. Each pilus operon consists of two adhesins belonging to the MSCRAMM (microbial surface recognizing adhesive matrix molecules) family and a sortase C necessary for the polymerization of the adhesin filaments into a pilus
4. Danne et al. (J Infect Dis. 2011) recently demonstrated that *pil1* is most probably the major factor for adhesion of *S. gallolyticus* to ECM

Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

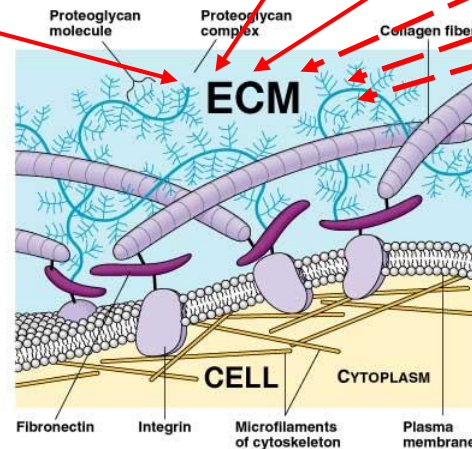
Pathogenicity genes presence/absence



Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Virulence genes presence/absence

<i>S. gallolyticus</i> UCN 34 locus_tag	gene	function	<i>S. gallolyticus</i> ATCC BAA 2069	<i>S. gallolyticus</i> ATCC 43143	<i>S. macedonicus</i> ACA-DC 198	<i>S. pasteurianus</i> ATCC 43144	<i>S. infantarius</i> CJ18
gallo_2179	-	accessory pilin (pil1)	✓	✓	-	-	-
gallo_2178	-	major pilin (pil1)	✓	✓	-	-	-
gallo_2177	-	sortase C (pil1)	✓	✓	-	-	-
gallo_1570	-	accessory pilin (pil2)	✓	✓	-	-	-
gallo_1569	-	major pilin (pil2)	✓	✓	-	-	-
gallo_1568	-	sortase C (pil2)	✓	✓	-	-	-
gallo_2040	-	accessory pilin (pil3)	✓	✓	✓	✓	✓
gallo_2039	-	major pilin (pil3)	✓	✓	✓	✓	✓
gallo_2038	-	sortase C (pil3)	✓	✓	✓	✓	✓



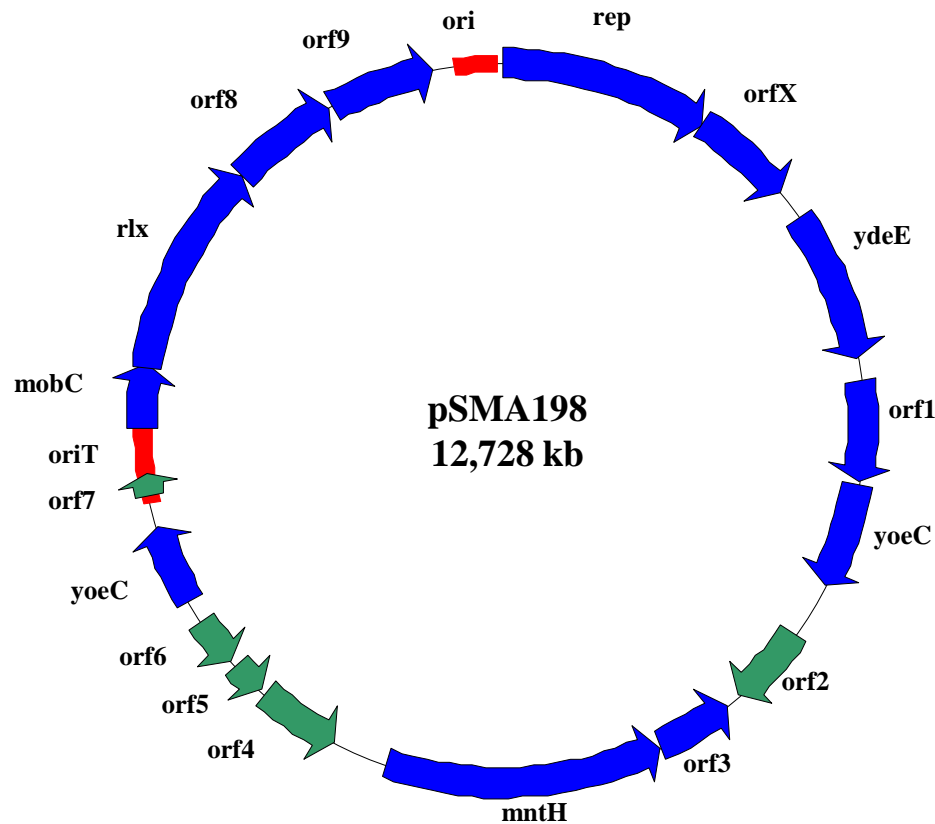
Comparative genomics of *S. macedonicus* ACA-DC 198 against related species within the *S. bovis*/*S. equinus* complex

Pathogenicity genes presence/absence

1. We have performed an elaborate search for known pathogenicity genes in the genomes of the species within the SBSEC against VFDB. Our findings also support a diminished pathogenicity potential for *S. macedonicus*, *S. pasteurianus* and *S. infantarius* as compared to *S. gallolyticus* (data not shown)
2. Interestingly, all SBSEC species seem to lack several major categories of streptococcal virulence factors (e.g. DNases, immunoreactive antigens, superantigens and toxins) supporting a reduced pathogenic potential for the complex as a whole

Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

- The presence of plasmid pSMA198 in *Streptococcus macedonicus* provided us with unexpected clues about the habituation of the species to the milk environment!!!



Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

1. pSMA198 belongs to the narrow host range pCI305/pWV02 family of lactococcal plasmids and it is the first such plasmid to be reported in *Streptococcus*

Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

Multiple sequence alignment of the pSMA198 Rep with RepB proteins of the pCI305 family

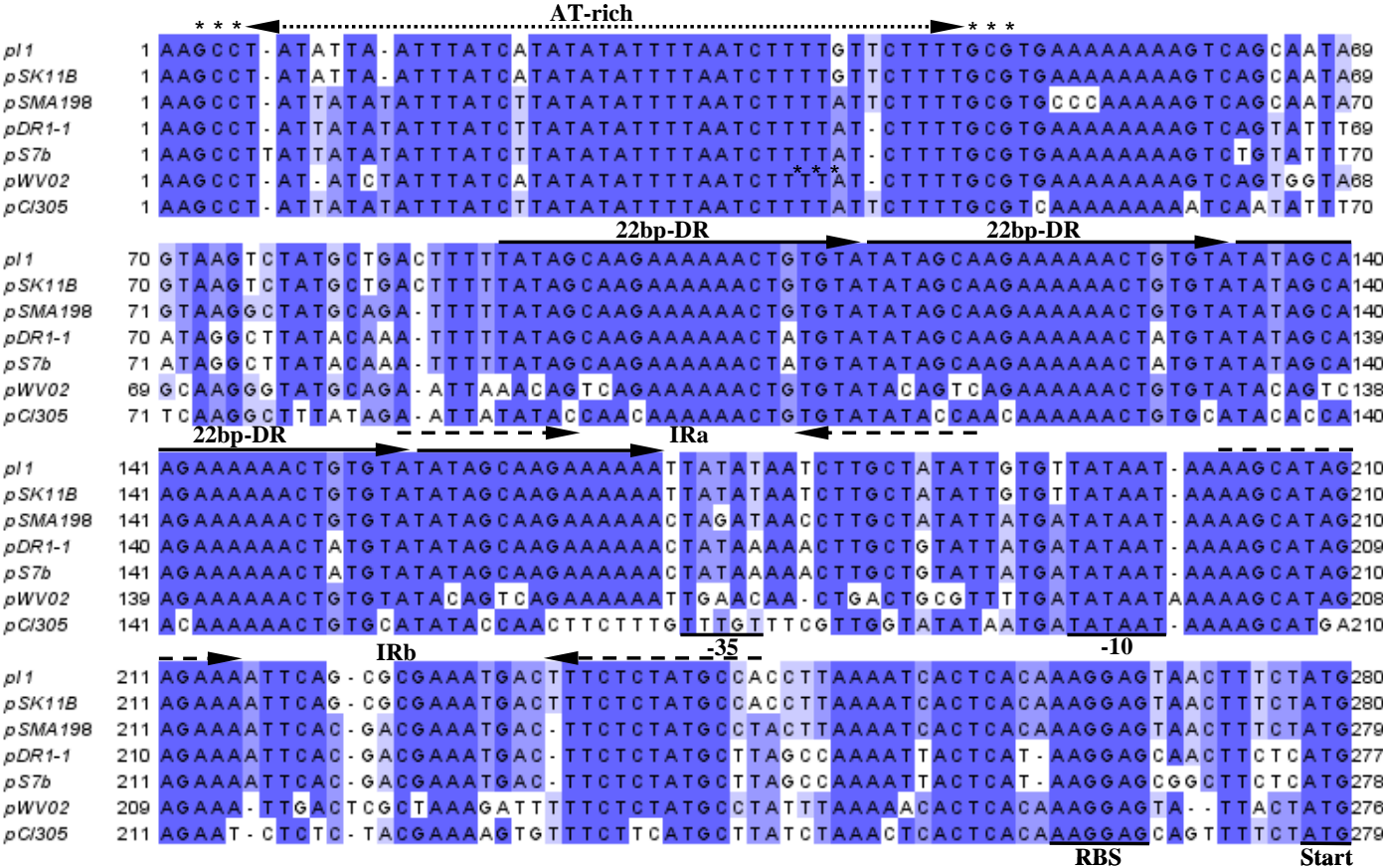
<i>pWW02</i>	1	MSI	AEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	AVYLSK	EE	LFTFFKVD	DDNDKHSR	FOA	AVKMQE	DAFFE	100		
<i>LLCRE1631_02531</i>	1	MSI	IPEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	FOA	AVEKMQ	QAAF	100		
<i>pIL7</i>	1	MP	I	SEKQNNQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	FOA	AVEKMQ	QAAF	100	
<i>pGL3</i>	1	MSS	I	TEKAKNQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRRF	KE	AVEKMQ	QAAF	100	
<i>pS7a</i>	1	MSS	I	TEKEENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRRF	KE	AVEKMQ	QAAF	100	
<i>pDBORO</i>	1	ML	I	TEKQENKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	FOA	AVEKMQ	QAAF	100	
<i>pK214</i>	1	MS	I	IPEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSK	AELFAF	FKVSDNDK	HRSR	FOA	AVEKMQ	QAAF	100	
<i>Ent_faecalis_TX1346</i>	1	MSS	I	PEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	KE	IEKMQ	QAAF	100	
<i>pCI305</i>	1	MSS	I	SKNEPNQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSK	EE	LFAF	FKVSDNDK	HRSR	FOA	AVENMQ	QAAF	100
<i>pSK11B</i>	1	MSS	I	PEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	KE	IEKMQ	QAAF	100	
<i>pLACR1</i>	1	MSS	I	PEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	KE	IEKMQ	QAAF	100	
<i>pSMA198</i>	1	MSS	I	PEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	KE	IEKMQ	QAAF	100	
<i>pVS40</i>	1	MS	I	IPEKQENQKQVLT	LNELSKRKVVEHNSL	ITS	IAKMDKTP	LKMFELAVSC	INTEEPPKDN	VYLSKR	DLFAF	FKVSDNDK	HRSR	FOA	AVEKMQ	QAAF	100	

<i>pWW02</i>	101	I	KEE	AGK	GFK	FRR	LP	I	P	T	V	E	W	T	D	Y	D	D	K	V	M	I	R	F	N	Q	I	M	P	Y	L	E	L	N	F	S	K	Y	A	L	S	E	I	M	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200				
<i>LLCRE1631_02531</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	N	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pIL7</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	N	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pGL3</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	D	D	V	E	S	Y	R	N	P	S	I	200
<i>pS7a</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	D	D	V	E	S	Y	R	N	P	S	I	200
<i>pDBORO</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pK214</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	N	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>Ent_faecalis_TX1346</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pCI305</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	D	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pSK11B</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pLACR1</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pSMA198</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	H	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200
<i>pVS40</i>	101	I	KEE	AGK	GFK	FKS	I	V	P	I	P	Y	E	W	T	D	Y	N	D	E	V	K	I	E	F	H	R	E	I	M	P	Y	L	I	N	L	K	N	F	T	O	H	A	L	S	D	I	A	E	L	N	S	K	Y	S	I	L	Y	R	W	L	S	M	N	Y	N	Q	E	H	Y	S	N	K	G	G	R	A	E	Q	V	E	S	Y	R	N	P	S	I	200

<i>pWW02</i>	201	S	V	K	E	L	R	T	I	D	T	V	N	E	Y	R	H	F	P	H	F	E	N	W	I	L	K	K	P	L	E	E	I	N	A	H	T	S	F	N	V	S	Y	E	K	I	K	G	R	S	I	D	S	I	V	F	H	I	E	K	K	R	A	D	D	N	S	Y	K	L	E	D	Q	T	V	I	E	G	K	A	K	E	E	T	K	D	L	Y	T	E	S	M	O	S	P	I	300
<i>LLCRE1631_02531</i>	201	S	I	K	E	L	R	I	M	T	D	T	V	N	S	Y	K	Q	F	T	R	F	N	D	V	L	K	T	P	L	D	E	I	N	A	H	T	S	F	N	V	T	Y	D	K	I	K	G	R	S	I	D	S	I	V	F	H	I	E	K	K	R	A	D	D	N	S	Y	K	L	D	D	R	A	Y	G	E	D	K	A	K	A	E	T	E	D	Q	L	V	L	O	A	M	S	P	I	300
<i>pIL7</i>	201	S	I	K	E	L	R	I	M	T	D	T	V	N	S	Y	K	Q	F	S	F	E	N	R	V	L	K	N	A	V	V	E	I	T	O	H	T	S	F	N	V	T	Y	D	K	I	K	G	R	S	I	D	S	I	V	F	H	I	T	K	K	R	A	D	D	N	S	Y	K	L	E	D	K	D	Y	Q	S	D	K	E	K	S	R	N	E	A	D	L	L	K	O	A	M	S	P	I	300
<i>pGL3</i>	201	S	V	K	E	L	R	I	M	T	D	T	V	N	E	Y	Q	R	F	T	N	F	T	K	K	I	L	D	I	P	L	K	E	I	N	D	N	T	F	K	Y	S	E	K	V	K	G	R	S	I	D	S	I	V	F	H	I	E	K	K	R	A	D	D	N	S	Y	K	L	E	D	K																									

Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

Multiple sequence alignment of the pSMA198 ori with the respective sequences of plasmids of the pCI305 family

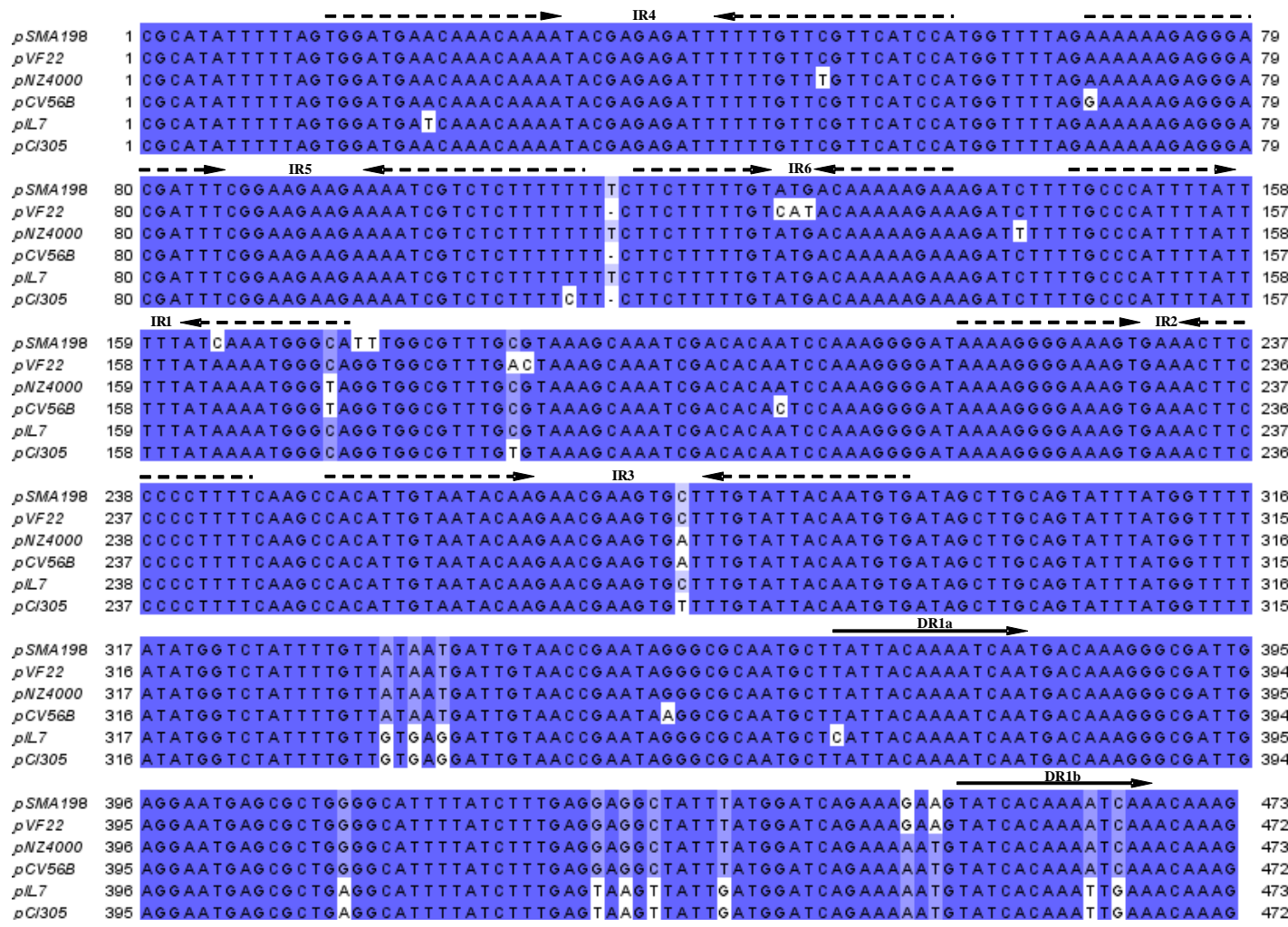


Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

1. pSMA198 belongs to the narrow host range pCI305 family of lactococcal plasmids and it is the first such plasmid to be reported in *Streptococcus*
2. *S. macedonicus* acquired pSMA198 from *L. lactis* and this acquisition took place most probably in the dairy environment

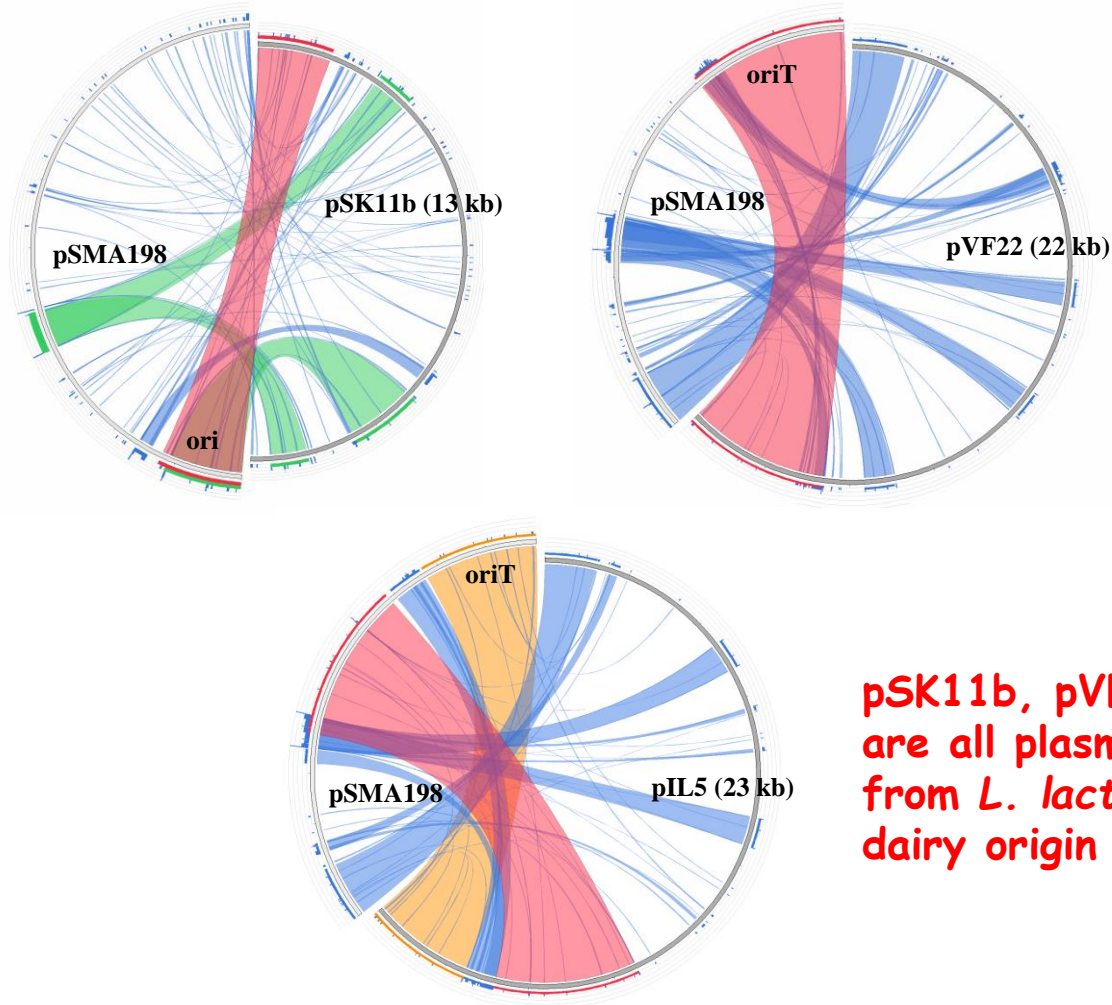
Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

Multiple sequence alignment of the pSMA198 *oriT* with the respective sequences of plasmids of the pCI305 family



Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

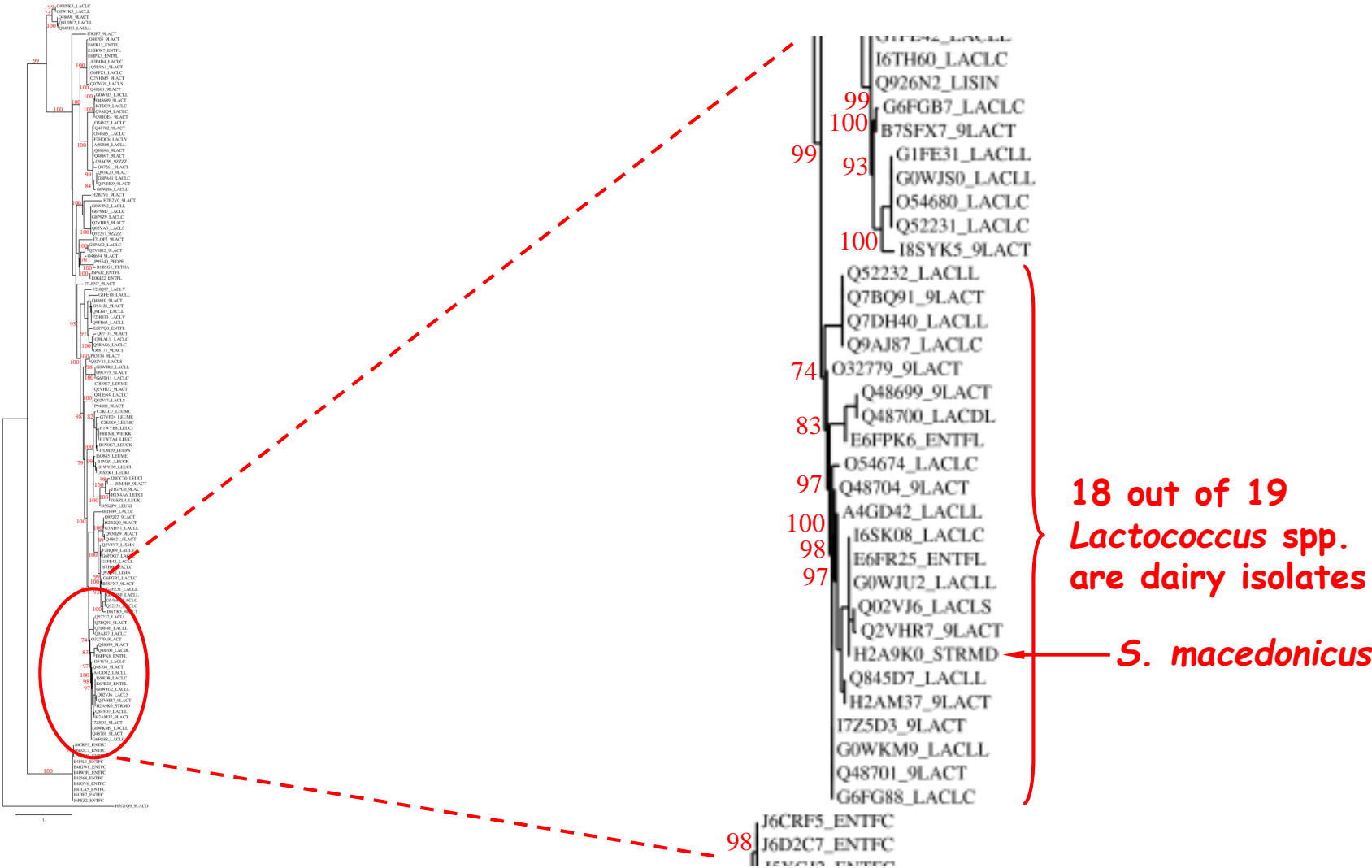
Comparative analysis of pSMA198 with its related plasmids



pSK11b, pVF22 and pIL5 are all plasmids isolated from *L. lactis* strains of dairy origin

Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

Phylogenetic analysis of the pSMA198 Rep with its related proteins



18 out of 19 *Lactococcus* spp. are dairy isolates

S. macedonicus

Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

1. pSMA198 belongs to the narrow host range pCI305 family of lactococcal plasmids and it is the first such plasmid to be reported in *Streptococcus*
2. *S. macedonicus* acquired pSMA198 from *L. lactis* and this acquisition took place most probably in the dairy environment
3. The acquisition of pSMA198 is most probably not a recent event

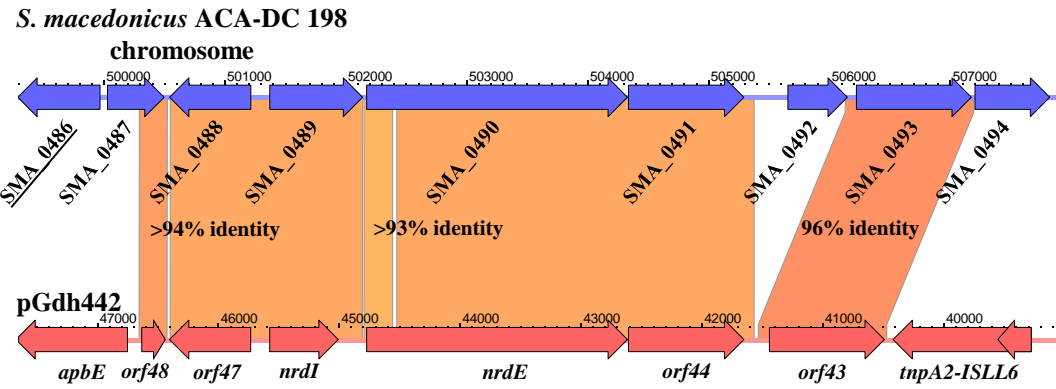
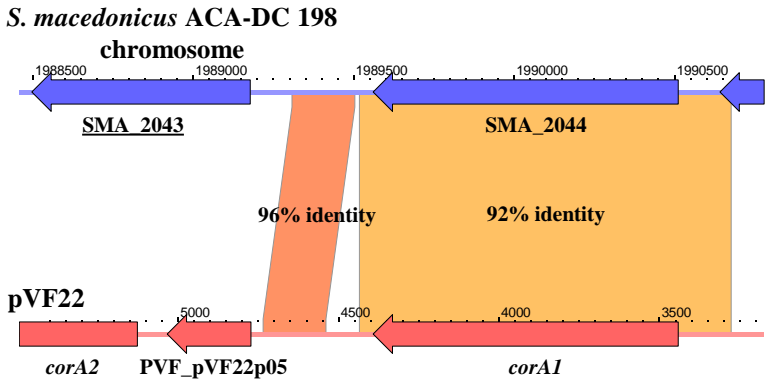
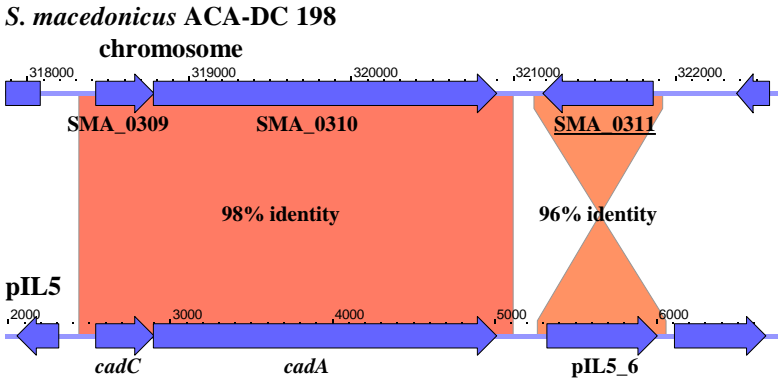
Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

Increased percentage of pseudogenes (5 out of 17) found in pSMA198 plasmid

locus_tag	gene	size nt	Best WU-Blastn hit (locus or locus_tag/ organism/ identity/ e-value)	Protein function
SMA_p0001	<i>rep</i>	1194	LACR_A06/ <i>Lactococcus lactis</i> subsp. <i>cremoris</i> SK11 plasmid 1/ 87%/ 1.8e-196	Initiator RepB protein
SMA_p0002	<i>orfX</i>	585	BN193_11490/ <i>Lactococcus raffinolactis</i> 4877/ 91%/ 2.3e-101	Replication associated protein
SMA_p0003	<i>ydeE</i>	858	ENT_30280 / <i>Enterococcus</i> sp. 7L76 / 99%/ 2.3e-184	AraC family transcriptional regulator
SMA_p0004	<i>orf1</i>	582	EfmE1039_1841/ <i>Enterococcus faecium</i> E1039/ 99%/ 4.7e-121	Integral membrane protein
SMA_p0005	<i>yoeC</i>	591	AF179848/ <i>Lactococcus lactis</i> subsp. <i>lactis</i> UC317 pCI305/ 88%/ 1.4e-97	Integrase/recombinase plasmid associated
SMA_p0006	<i>orf2</i>	459	CAC42047/ <i>Listeria innocua</i> Clip11262 pLI100/ 99%/ 3.7e-94	Putative pseudo ←
SMA_p0007	<i>orf3</i>	438	LACR_D31/ <i>Lactococcus lactis</i> subsp. <i>cremoris</i> SK11 plasmid 4/ 99%/ 4.9e-90	Universal stress protein family
SMA_p0008	<i>mntH</i>	1578	HMPREF0848_00725 / <i>Streptococcus</i> sp. C150/ 99%/ 0.0	Manganese transport protein MntH
SMA_p0009	<i>orf4</i>	480	llmg_pseudo_13/ <i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363 pseudogene/ 97%/ 3.3e-214	Putative pseudo ←
SMA_p0010	<i>orf5</i>	195	llmg_pseudo_13/ <i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363 pseudogene/ 97%/ 3.3e-214	Putative pseudo ←
SMA_p0011	<i>orf6</i>	276	llmg_pseudo_13/ <i>Lactococcus lactis</i> subsp. <i>cremoris</i> MG1363 pseudogene/ 97%/ 3.3e-214	Putative pseudo ←
SMA_p0012	<i>yoeC</i>	465	GMD1E_00300 / <i>Enterococcus</i> sp. GMD1E / 98%/ 2.2e-93	Integrase/recombinase plasmid associated
SMA_p0013	<i>orf7</i>	132	pIL7_28/ <i>Lactococcus lactis</i> subsp. <i>lactis</i> IL594 plasmid pIL7/ 84%/ 1.3e-13	Putative pseudo ←
SMA_p0014	<i>mobC</i>	366	HMPREF9519_01999/ <i>Enterococcus faecalis</i> TX1346/ 89%/ 5.9e-61	Mobilization protein
SMA_p0015	<i>rlx</i>	1233	CI5MOBPRO/ <i>Lactococcus lactis</i> subsp. <i>cremoris</i> UC503 pCI528/ 99%/ 3.4e-268	Mobilization protein
SMA_p0016	<i>orf8</i>	627	ENT_30400 / <i>Enterococcus</i> sp. 7L76/ 96%/ 6.3e-124	Conserved hypothetical protein
SMA_p0017	<i>orf9</i>	603	BN193_11500/ <i>Lactococcus raffinolactis</i> 4877/ 99%/ 3.5e-125	Fic family protein

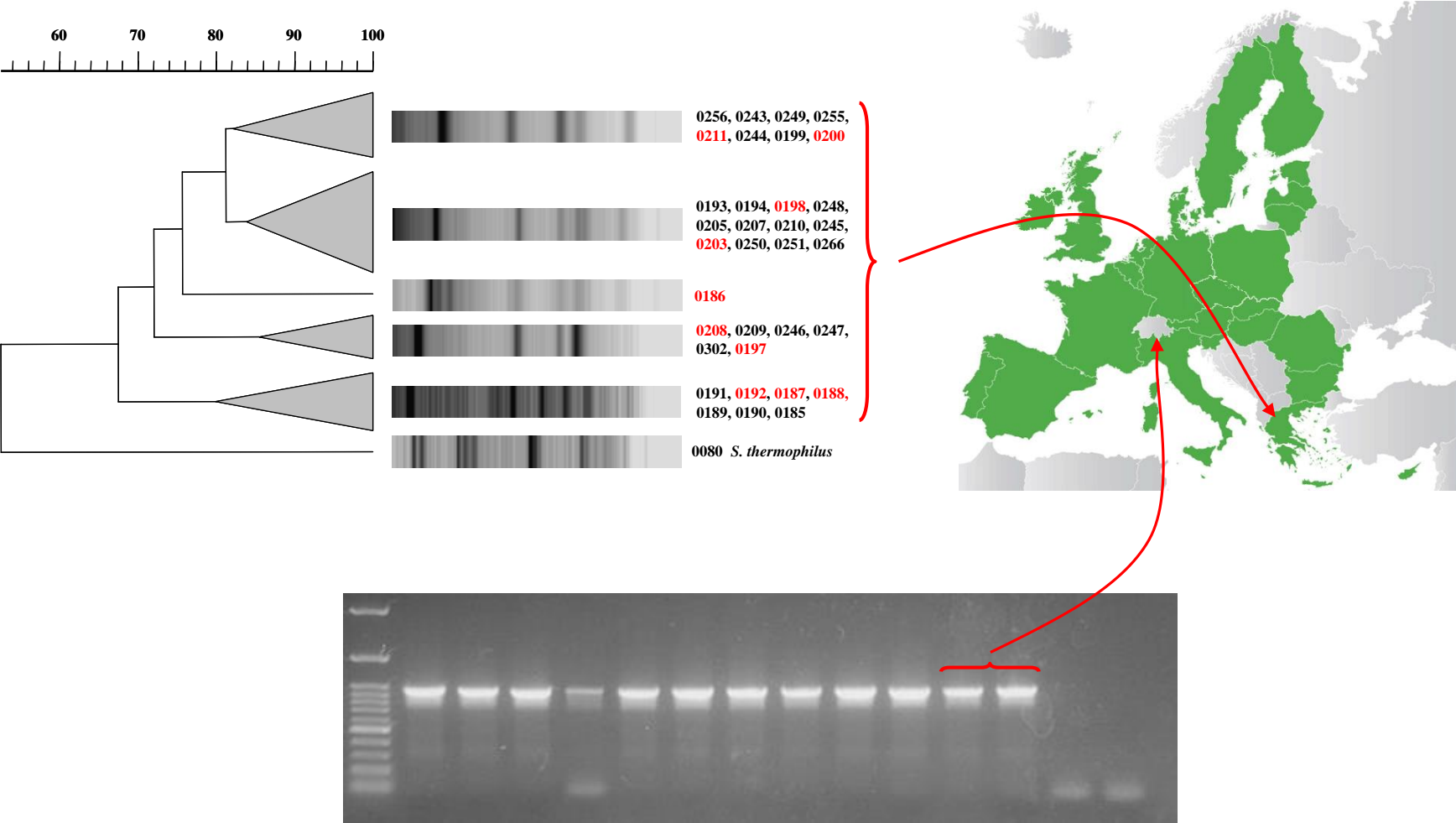
Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

S. macedonicus chromosome loci potentially originating from pSMA198



Analysis of the lactococcal plasmid pSMA198 points towards the habituation of *Streptococcus macedonicus* ACA-DC 198 to the dairy environment

The distribution of pSMA198 in different strains of *S. macedonicus*



Conclusions

1. *S. macedonicus* is most probably a separate species from *S. gallolyticus*
2. *In silico* analysis of *S. macedonicus* ACA-DC 198 suggests that:
 - The strain is at the process of adapting to a rich in nutrients environment
 - It shows a diminished capacity to live and survive in the GI tract of herbivores
 - It has a diminished pathogenic potential compared to *S. gallolyticus*
3. The acquisition of pSMA198 by *S. macedonicus* from *L. lactis* indicates that the species is habituated in the dairy environment

Acknowledgments

The present work was cofinanced by the European Social Fund and the National resources EPEAEK and YPEPTH through the Thales project



Thank you for your attention!!

