Assessing the adaptation to milk and the pathogenic potential of the dairy Streptococcus macedonicus ACA-DC 198 through comparative genomics

Konstantinos Papadimitriou, Rania Anastasiou, Eleni Mavrogonatou, Jochen Blom, Nikos C. Papandreou, Stavros J. Hamodrakas, Stéphanie Ferreira, Philip Supply, Bruno Pot, Effie Tsakalidou

1Department of Food Science and Human Nutrition, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece, 2Institute of Biosciences and Applications, National Centre for Scientific Research “Demokritos”, 153 10 Athens, Greece, 3Center for Biotechnology, Bielefeld University, Bielefeld, Germany, 4Department of Cell Biology and Biophysics, Faculty of Biology, University of Athens, Panepistimiopolis, 157 01 Athens, Greece, 5Genoscreen, Genomic Platform and R&D, Campus de l’Institut Pasteur, 1 rue du Professeur Calmette, 59000 Lille, France, 6Institut Pasteur de Lille, Center for Infection and Immunity of Lille (CIIL), F-59019 Lille, France, 7Inserm U1019, F-59019 Lille, France, 8CNRS UMR8204, F-59021 Lille, France, 9University of Lille de Nord France, F-59000 Lille, France
Laying the background: Fermented foods and Microorganisms

Raw material → lactic acid bacteria → Fermented foods
Laying the background: Fermented foods and Microorganisms

- Raw material
- Shelf life
- Nutritional Value
- Organoleptic characteristics

Fermented foods
Laying the background: Fermented foods and Microorganisms

Traditionally fermented dairies

Selection of novel starter cultures

Fermenting ecosystem

Experimental dairy production

Basic Research
- Taxonomy
- metabolism
- physiology
- genetics and genomics
- bioinformatics

Technological properties
- flavour bio-generation
- bio-texturant molecule development
- bio-preservative molecule production (antimicrobial peptides)
- probiotics
Laying the background: Fermented foods and Microorganisms

- Lactic Acid Bacteria: Gram-positive, low-GC, rods or cocci, non-sporulating, catalase-negative, aerotolerant but not aerophilic and they produce lactic acid as the major metabolic end-product of carbohydrate fermentation
  - Carnobacterium
  - Enterococcus
  - Lactococcus
  - Lactobacillus
  - Leuconostoc
  - Oenococcus
  - Pediococcus
  - Streptococcus
  - Weissella
Laying the background: Fermented foods and Microorganisms

- The *Streptococcus* genus (*sensu stricto*) mainly consists of commensals of humans and animals including both non-pathogenic and pathogenic bacteria.

- *Streptococcus thermophilus* is the only species in the genus used as a starter culture in food fermentations since all known lactic streptococci belong now to the *Lactococcus* genus.

- *Streptococcus macedonicus* which was originally isolated and characterized from traditional Greek Kasseri cheese (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).
• Streptococci that can be found growing in milk belong to the *Streptococcus bovis/Streptococcus equinus* complex (SBSEC)

Herrera et al. Anaerobe 2009
Sequencing the genome of *S. macedonicus* ACA-DC 198

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* its adaptation to the milk environment and its pathogenic potential.
Assessing the adaptation to milk and the pathogenic potential of the dairy
*S. macedonicus* ACA-DC 198 through comparative genomics

A. Complete genome sequencing and annotation of *S. macedonicus*

B. Comparative genomics of *S. macedonicus* against related streptococci

C. Assessing the adaptation of *S. macedonicus* to the milk environment

D. Assessing the pathogenic potential *S. macedonicus*
Complete genome sequencing and annotation of *S. macedonicus*

- **1st step:** shotgun pyrosequencing with 454 GS-FLX titanium (>100 contigs)
- **2nd step:** 3kb paired-end pyrosequencing with 454 GS-FLX titanium (7 scaffolds)
- **3rd step:** gap-closure and polishing with Illumina sequencing using the HiSeq 2000
- **4th step:** validation of the overall assembly (>200X coverage) with an *NheI* optical map

Chromosome: 2,130,034 bp
Plasmid: 12,728 bp (pSMA198)
Complete genome sequencing and annotation of *S. macedonicus*

- 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

Assessing the adaptation to milk and the pathogenic potential of the dairy Streptococcus macedonicus ACA-DC 198 through comparative genomics

A. Complete genome sequencing and annotation of S. macedonicus

B. Comparative genomics of S. macedonicus against related streptococci

C. Assessing the adaptation of S. macedonicus to the milk environment

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

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)
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 against related species within the *S. bovis/S. equinus* complex

![Genome comparison diagram](image-url)
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 against related species within the *S. bovis/S. equinus* complex

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

Core genome: 57.3%

1: Streptococcus_gallolyticus_subsp_gallolyticus_ATCC_43143_DNA
2: Streptococcus_infantarius_subsp_infantarius_CJ18
3: Streptococcus_macedonicus_ACA-DC_198_main_chromosome
4: Streptococcus_pasteurianus_ATCC_43144_NC_015600
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**

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

* not reported

1. *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. galloyticus* genome.

2. 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.
Assessing the adaptation of *S. macedonicus* to the milk environment

A. Complete genome sequencing and annotation of *S. macedonicus*

B. Comparative genomics of *S. macedonicus* against related streptococci

C. Assessing the adaptation of *S. macedonicus* to the milk environment

D. Assessing the pathogenic potential *S. macedonicus*
Assessing the adaptation of *S. macedonicus* to the milk environment

- *S. macedonicus* has deviated from the rumen environment

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

- *S. macedonicus* has an extra gene cluster responsible for lactose and galactose catabolism

![Diagram showing gene clusters of S. macedonicus ACA-DC 198, S. suis BM407, and L. lactis subsp. cremoris plasmid pLP712]
Assessing the adaptation of *S. macedonicus* to the milk environment

- *S. macedonicus* is particularly equipped with defenses against phages

<table>
<thead>
<tr>
<th>CRISPR label</th>
<th>Position</th>
<th>Nbr Spacers</th>
<th>Position</th>
<th>Nbr Spacers</th>
<th>Position</th>
<th>Nbr Spacers</th>
<th>Position</th>
<th>Nbr Spacers</th>
<th>Position</th>
<th>Nbr Spacers</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRISPR_1</td>
<td>1507890,</td>
<td>15</td>
<td>1484496,</td>
<td>29</td>
<td>1517213,</td>
<td>15</td>
<td>1412482,</td>
<td>50</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1508913</td>
<td></td>
<td>1486444</td>
<td></td>
<td>1518237</td>
<td></td>
<td>1415817</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRISPR_2</td>
<td>1515490,</td>
<td>12</td>
<td>1477224,</td>
<td>10</td>
<td>1515726,</td>
<td>12</td>
<td>1273106,</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1516317</td>
<td></td>
<td>1477919</td>
<td></td>
<td>1516570</td>
<td></td>
<td>1273801</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRISPR_3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRISPR_4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRISPR_5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRISPR_6</td>
<td>1395041,</td>
<td>37</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1397515</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Spacers 3, 5, 17, 46 in *S. macedonicus* CRISPR provide immunity against phages of the dairy *S. thermophilus* and *L. lactis*
Assessing the adaptation of *S. macedonicus* to the milk environment

- *S. macedonicus* is particularly equipped with defenses against phages

---

**S. macedonicus ACA-DC 198**

**S. pasteurianus ATCC 43144**

**S. infantarius CJ18**

---

**S. galloyticus UCN34**

**S. galloyticus ATCC 43143**

**S. galloyticus ATCC BAA-2069**
Assessing the adaptation of *S. macedonicus* to the milk environment

- The *S. macedonicus* plasmid pSMA198 belongs to the narrow host range pCl305 family of lactococcal plasmids
Assessing the adaptation of *S. macedonicus* to the milk environment

- The *S. macedonicus* plasmid pSMA198 belongs to the narrow host range pCI305 family of lactococcal plasmids

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

*S. macedonicus*
Assessing the adaptation to milk and the pathogenic potential of the dairy Streptococcus macedonicus ACA-DC 198 through comparative genomics

A. Complete genome sequencing and annotation of S. macedonicus

B. Comparative genomics of S. macedonicus against related streptococci

C. Assessing the adaptation of S. macedonicus to the milk environment

D. Assessing the pathogenic potential S. macedonicus
Assessing the pathogenic potential *S. macedonicus*

- *S. macedonicus* shows a diminished potential to bind to the extracellular matrix of the host.

<table>
<thead>
<tr>
<th>S. gallolyticus UCN 34</th>
<th>locus_tag</th>
<th>gene</th>
<th>function</th>
<th>S. gallolyticus ATCC BAA 2069</th>
<th>S. gallo. ATCC 43143</th>
<th>S. macedonicus ACA-DC 198</th>
<th>S. pasteurianus ATCC 43144</th>
<th>S. infantarius CJ18</th>
</tr>
</thead>
<tbody>
<tr>
<td>gallo_2179</td>
<td>-</td>
<td>accessory pilin (pil1)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gallo_2178</td>
<td>-</td>
<td>major pilin (pil1)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gallo_2177</td>
<td>-</td>
<td>sortase C (pil1)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gallo_1570</td>
<td>-</td>
<td>accessory pilin (pil2)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gallo_1569</td>
<td>-</td>
<td>major pilin (pil2)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gallo_1568</td>
<td>-</td>
<td>sortase C (pil2)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>gallo_2040</td>
<td>-</td>
<td>accessory pilin (pil3)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>gallo_2039</td>
<td>-</td>
<td>major pilin (pil3)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>gallo_2038</td>
<td>-</td>
<td>sortase C (pil3)</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Diagram:
- ECM
- CELL
- Cytoplasm
- Fibronectin
- Integrin
- Microfilaments of cytoskeleton
- Plasma membrane
- Proteoglycan molecule
- Proteoglycan complex
- Collagen fiber
Assessing the pathogenic potential *S. macedonicus*

* *S. macedonicus* has retained the hemolysin encoding genes

<table>
<thead>
<tr>
<th>Virulence factor</th>
<th>S. gallolyticus UCN34</th>
<th>S. gallolyticus ATCC 43143</th>
<th>S. gallolyticus ATCC BAA-2069</th>
<th>S. pasteurianus ATCC 43144</th>
<th>S. macedonicus ACA-DC 198</th>
<th>S. infantarius CJ18</th>
</tr>
</thead>
<tbody>
<tr>
<td>hemolysin TLY</td>
<td>GALLO_0630</td>
<td>SGGB_0605</td>
<td>SGGBAA2069_c05730</td>
<td>SGPB_0499</td>
<td>SMA_0591</td>
<td>Sinf_0511</td>
</tr>
<tr>
<td>hemolysin III</td>
<td>GALLO_1262</td>
<td>SGGB_1256</td>
<td>SGGBAA2069_c12530</td>
<td>SGPB_1172</td>
<td>SMA_1191</td>
<td>Sinf_1093</td>
</tr>
<tr>
<td>hemolysin A</td>
<td>GALLO_1799</td>
<td>SGGB_1786</td>
<td>SGGBAA2069_c17570</td>
<td>SGPB_1603</td>
<td>SMA_1706</td>
<td>Sinf_1530</td>
</tr>
</tbody>
</table>

![Image of bacterial culture](image-url)
Conclusions

1. *Streptococcus macedonicus* is evolving under genome decay processes suggesting adaptation to a rich in nutrients environment

2. Our analysis supports that the species shows traits of adaptation to the dairy environment

3. Even though *S. macedonicus* shows a diminished pathogenic potential compared to *S. gallolyticus*, several pathogenicity traits are still conserved

Functional analysis needed!!!
This work was performed in collaboration with:

- Prof. Stavros Hamodrakas
- Dr. Bruno Pot
- Dr. Philippe Supply
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!!!