TITLE **The *Lactobacillus casei* MaeKR two component system is required for L-malic acid utilization through a malic enzyme pathway.**

RUNNING TITLE **Two component MaeKR regulates malic enzyme pathway**

AUTHORS José María Landete[1‡], Luisa García-Haro[1‡†], Amalia Blasco[1], Paloma Manzanares[1], Carmen Berbegal[2], Vicente Monedero[1], and Manuel Zúñiga*[1]

[1] Departamento de Biotecnología de Alimentos, Instituto de Agroquímica y Tecnología de Alimentos (IATA), Consejo Superior de Investigaciones Científicas (CSIC) PO Box 73, 46100 Burjassot, Valencia, Spain

[2] Departamento de Microbiología y Ecología, Facultad de Biología, ENOLAB-Laboratorio de Microbiología Enológica, Universidad de Valencia, Burjassot, Valencia, Spain

† Present address: Instituto de Biomedicina de Valencia, Consejo Superior Investigaciones Cientificas CSIC, Jaime Roig 11, 46010 Valencia, Spain

* Corresponding author; Tel +34 963900022; Fax +34 963636301

E-mail: btcman@iata.csic.es

‡ These authors contributed equally to the work

ABSTRACT

Lactobacillus casei can metabolize L-malic acid via malolactic enzyme (MLF) or malic enzyme (ME). Whereas utilization of L-malic acid via MLF does not support growth, the ME pathway enables *L. casei* to grow on L-malic acid. In this work we have identified in the genomes of *L. casei* strains BL23 and ATCC 334 a cluster consisting of two diverging operons, *maePE* and *maeKR*, encoding a putative malate transporter (*maeP*), a ME (*maeE*) and a two-component (TC) system belonging to the citrate family (*maeK* and *maeR*). Homologous clusters were identified in *Enterococcus faecalis*, *Streptococcus agalactiae*, *Streptococcus pyogenes* and *Streptococcus uberis*. Our results show that ME is essential for L-malic acid utilization in *L. casei.* Furthermore, deletion of either the gene encoding the histidine kinase or the response regulator of the TC system resulted in the loss of the ability to grow on L-malic acid thus indicating that the cognate TC system regulates and is essential for the expression of ME. Transcriptional analyses showed that expression of *maeE* is induced in the presence of L-malic acid and repressed by glucose whereas the TC system expression was induced by L-malic acid and was not repressed by glucose. DNase I footprinting analysis showed that MaeR binds specifically to a set of direct repeats (5'-TTATT(A/T)AA-3') in the *mae* promoter region. The location of the repeats strongly suggests that MaeR activates the expression of the diverging operons *maePE* and *maeKR* where the first one is also subjected to carbon catabolite repression.

INTRODUCTION

The metabolism of L-malic acid by lactic acid bacteria (LAB) has brought about considerable interest because of its relevance in winemaking (24). The degradation of L-malate to L-lactate leads to a reduction of the acidity of wine and it provides microbiological stability by preventing the secondary growth of LAB after bottling. Most LAB decarboxylate L-malate to L-lactate by a NAD^+ and Mn^{2+}-dependent malolactic enzyme (MLE; Fig. 1); nevertheless, a few LAB species can also degrade L-malate to pyruvate by a ME (Fig. 1). This pathway was first detected in *Enterococcus faecalis* (20) and later in *Lactobacillus casei* (23,33) and *Streptococcus bovis* (14). In contrast to the utilization of L-malate through MLE, the utilization of the ME pathway enables these organisms to grow with L-malate as a carbon source (22). However, whereas MLE has been the focus of an extensive research effort, the physiological role and the regulation of ME remains largely unknown.

L. casei is a facultative heterofermentative lactic acid bacterium frequently used as a cheese starter culture and which is also employed as a probiotic. Extensive research has been carried out on the study of sugar catabolism (28,39-41), however the knowledge of the utilization of organic acids has received less attention. As previously indicated, physiological and biochemical studies identified two L-malate dissimilation pathways in *L. casei*. Furthermore, these studies showed that ME expression was induced in the presence of L-malate and very low concentrations of glucose (4,22). However, the regulatory system controlling ME expression was not identified. In other bacteria such as *Bacillus subtilis* (36) or *Escherichia coli* (11), utilization of some organic acids such as citrate, succinate or malate is regulated by two component (TC) systems. TC regulatory systems typically consist of a sensor kinase (HK) and a response regulator (RR) (34). Both proteins have a modular structure: HKs usually have two modules involved in the phosphorylation reaction, the kinase and H-box domains, and usually an N-terminal transmembrane sensory domain. On the other

hand, the RR has a receptor domain of the phosphoryl group and a C-terminal effector domain. The domains involved in the phosphorylation reaction of both HKs and RRs are homologous in all TC systems (16) whereas the sensor and effector domains are specific to individual TC systems and determine their specificity. HKs monitor environmental signals and in response to a stimulus, autophosphorylates at a histidine residue (H-box). The high-energy phosphate group is subsequently transferred to an aspartyl residue on the RR receptor domain. Phosphorylation of the RR in turn modulates the activity of the RR effector domain. In many cases, the RR receptor domains are DNA-binding domains so that RR acts as transcriptional repressors or activators.

Our research group has initiated two studies focused in the physiological role of TC systems and the metabolism of organic acids in *L. casei*. In this work we have identified in *L. casei* a gene cluster constituted by two putative operons. One of them encodes a TC system similar to other TC systems involved in the regulation of the utilization of organic acids in other bacteria. The other operon encodes a putative ME and a L-malate transporter. This work aimed to determine the possible regulatory role of the cognate TC system in the expression of ME. Results reported here show that this gene cluster accounts for L-malate utilization via ME pathway and that their expression is under control of the cognate TC system.

MATERIALS AND METHODS

Strains and growth conditions

The strains and plasmids used in this study are listed in Table 1. *L. casei* was routinely grown in MRS broth (Oxoid) at 37ºC under static conditions. *Escherichia coli* DH5α strains were grown in LB medium at 37ºC with aeration. Antibiotics used were 100 μg ml^{-1} ampicillin, 25 μg ml^{-1} kanamycin for *E. coli* and 5 μg ml^{-1} erythromycin for *L. casei*.

Growth assays, determination of L-malic acid degradation and gene expression analyses were carried out at 30ºC in malic enzyme induction medium (MEI (33)) modified as follows: tryptone replaced peptone and the medium was supplemented with cysteine, 0.5 g l^{-1} and Tween 80, 1 ml l^{-1}; and the pH was adjusted to 6.8. When required, filter-sterilised glucose or L-malate (adjusted to pH 6.8 with NaOH) were added at a final concentration of 5 g l^{-1}. Inoculation was performed with cells grown in MRS for 16 h and washed twice with 1 volume of sterile distilled water. Growth was monitored by measuring O.D. at 595 nm. At least three independent replicates of each growth curve were obtained. Results were expressed as averages and plus and minus standard deviations.

Analysis of organic acids

Samples of cultures grown in MEI medium were taken at different times during growth. The samples were centrifuged, the supernatant filtered through 0.22 μm pore size Millex-GV syringe driven filter units (Millipore) and stored at -80°C until use. Samples were analyzed using a HPLC equipment (Agilent series 1200) with isocratic pump (Agilent G1310A) following the procedure described by Frayne (9) with minor modifications. The mobile phase consisted of a solution of 0.75 ml of 85% H_3PO_4 per litre of deionised water, with a flow of 0.7 ml min^{-1}. An Agilent G1322A degasser was employed. Samples (5 μl) were injected automatically (Agilent G1367B). The

separation of the components was carried out using an Aminex HPX-87H precolumn (Biorad) coupled with two ion exclusion columns of 300 mm x 7.8 mm AMINEX HPX-87H (BioRad) thermostatically controlled at 65° C (Agilent G1316A). Compounds were detected by a Variable Wavelength Detector G1314B (Agilent) set to 210 nm and a Refractive Index detector (Agilent G1362A) in series. External calibration was performed.

Phylogenetic analyis

Bacterial genes encoding homologues of *maeK* and *maeR* were retrieved from whole genomes by using BLASTP and TBLASTN (2,3) and the genes carried by *L. casei* BL23 as query sequences. Representative species harbouring *maeK* and *maeR* homologues of all taxonomic divisions were selected. Some sequences were modified as follows. A possible frame-shift in the putative *maeK homolog* sequence of *Carboxydothermus hydrogenoformans* Z-2901 (GenBank accession number NC_007503) was corrected by deleting one A at position 1227250. A possible frame-shift in a *Bacillus anthracis* Ames ancestor *maeR* homolog (banth3; see Suppl. Fig. 1) was corrected by inserting an A after position 568015 in the genomic sequence (Acc. Nº NC_007530). Multiple alignments were obtained using ClustalW (37) and manually corrected where necessary. Positions of doubtful homology or that introduced phylogenetic noise were removed by using Gblocks software (6). The best fit models of amino acid substitution were selected using the program ProtTest (1). The Akaike Information Criterion (AIC) was adopted to select the best model that was LG (17) with a discrete gamma function with four categories plus invariant sites to account for substitution rate heterogeneity among sites for both protein sets. The selected model was implemented in PHYML 3.0 (12) to obtain maximum likelihood trees for the different alignments. Bootstrap support values were obtained from 500 pseudo replicates.

DNA techniques

Standard methods were used for cloning in *E. coli* (31). Restriction enzymes and T4 DNA ligase were purchased from New England Biolabs. Taq DNA polymerase for PCR screening was from Biotools (Biotools, B & M Labs, Madrid, Spain). Plasmids were isolated with the GFX Micro Plasmid Prep Kit (GE Healthcare). DNA from *L. casei* was isolated with the UltraClean Microbial DNA isolation kit (MoBio Laboratories, Solana Beach, CA). Southern hybridization analyses were carried out by transfer of DNA from agarose gels to Hybond-N membranes (GE Healthcare) (31). Probes were labelled with digoxigenin-dUTP by using PCR DIG labelling mix (Roche) in standard PCR reactions and detected by using antidigoxigenin-AP and CDP-Star (Roche). Hybridization, washing, and detection were performed as instructed by the supplier. *E. coli* strains were transformed by electroporation with a Gene Pulser apparatus (Bio-Rad), as recommended by the manufacturer, and *L. casei* strains were transformed as described previously (30).

RNA techniques

Total RNA was isolated from *L. casei* as described previously (42). Strains were grown in MEI (50 ml) supplemented with glucose and/or L-malate at 30°C. Unless otherwise indicated, samples were taken at mid exponential phase, beginning of the stationary phase and at late stationary phase. Sample preparation, denaturing agarose gel electrophoresis, and RNA transfer were performed by standard methods (31). Internal fragments of genes *maeE* and *maeK* were synthesized by PCR using primers MaeE1/MaeE2 and MaeKprob1/MaeKprob2 (Table 2) as probes and labelled with the PCR-digoxigenin labelling mix from Roche. Hybridization, washing and detection with the CDP-star (Roche) chemiluminiscent reagent were performed as recommended by the supplier.

Transcription initiation sites were determined with the 5'/3' RACE (rapid amplification of cDNA ends) kit (Roche), following the manufacturer's instructions. Reverse transcription reactions were performed with total RNA isolated from *L. casei* BL23 grown in MEI supplemented with 5 g/l L-

malic acid. To determine the transcriptional start site upstream from *maeP* reverse transcription was performed with primer MalicR-1 (Table2). The cDNA was dA-tailed, and then amplified by PCR using the primers oligo dT-anchor supplied in the kit and MalicR-2. The resulting PCR product was used in a second PCR with primers PCR-anchor (supplied with the kit) and MalicR-3. The amplified DNA fragment was purified and sequenced. For the determination of the transcriptional start site of *maeK*, the same strategy was followed by using primers MaeA, MaeG and RACEG2.

Real-time PCR.

RNA samples purified as described above were treated with the Ambion Turbo DNA-free™ kit (Applied Biosystems) using the routine DNase I treatment outlined by the supplier in order to remove contaminating DNA. The quality and concentration of the RNA samples was subsequently evaluated by using the Experion automated electrophoresis system (BioRad). Samples with 23S/16S ratios lower than 0.85 were discarded. First-strand cDNA was synthesized from 1 µg RNA using SuperScript VILO cDNA Synthesis kit (Invitrogen) as recommended by the manufacturer. Two retrotranscription reactions were performed for each RNA sample. Real-time PCR was performed using the Lightcycler 2.0 system (Roche) and the LC Fast Start DNA Master SYBR®Green I (Roche). Primers were designed by using the Primer-BLAST service (http://www.ncbi.nlm.nih.gov/tools/primer-blast) in order to generate amplicons ranging from 100 to 150 bp in size (Supplementary Table S1). Real-time PCR was performed for each cDNA sample in triplicate in 10 µl of the reaction mixture containing 1 µl of 10× master mix, 1.2 µl of $MgCl_2$ (25 mM), 0.5 µl of each primer (10 µM), and 1 µl of a 1/10 diluted sample from the cDNA synthesis reaction. Reaction mixtures without a template were run as controls. The cycling conditions were as follows: 95°C for 10 min, followed by 35 cycles of three steps consisting of denaturation at 95°C for 10 s, primer annealing at 55°C for 15 s, and the primer extension at 72°C for 20 s. For each set of primers, the cycle threshold values (crossing point [CP]) were determined by the automated

method implemented in the Lightcycler software 4.0 (Roche). In order to select appropriate reference genes, ten housekeeping genes (*fusA*, *ileS*, *lepA*, *leuS*, *mutL*, *pcrA*, *pyrG*, *recA, recG* and *rpoB*) were determined in all experimental conditions assayed and analysed using the geNorm approach (38). As a result of the analysis, genes *fusA*, *leuS*, *pyrG* and *recG* were selected as reference genes (see Supplementary Fig. 2). The relative expression based on the expression ratio between the target genes and reference genes was calculated using the software tool REST (relative expression software tool) (29). Linearity and amplification efficiency were determined for each primer pair. Every real-time PCR was performed at least six times.

RT-PCR

RNA and cDNA samples obtained above were used in PCR amplifications with oligonucleotides maeE1/maeE2 (Table 2) in 25 µl reactions containing 0.5 U of Paq5000 polymerase (Stratagene), 10 µM of each oligonucleotide, 200 µM dNTPs and 1 µl of a 1/10 dilution of each RNA or cDNA sample. The following PCR conditions were used: denaturation at 95ºC for 5 min followed by 25 cycles of denaturation at 94ºC for 15 s, annealing at 57ºC for 30 s and extension at 72ºC for 60 s and a final extension cycle at 72 ºC for five min. The amplification products were resolved by electrophoresis in 2% agarose gels.

Construction of strains

To construct a *maeE*-defective mutant, an internal fragment of the *maeE* gene was amplified by PCR using the oligonucleotide pair maeE1/maeE2 (Table 2) (18). The PCR product was digested with HindIII/SacII and ligated to the integrative vector pRV300 (17) digested with the same enzymes and transformed into *E. coli* DH5α. The resulting plasmid (pRVmaeE), was used to transform *L. casei* BL23 and single cross-over integrants were selected by resistance to

erythromycin and confirmed by Southern analysis using the *maeE* internal fragment used for inactivation as a probe. One of these integrants was selected and named BL321 (*maeE*::pRV300)

In order to obtain BL23 derivative strains harbouring complete deletions of either *maeK* or *maeR*, primer sets maeKup1/maeKup2 and maeKdown1/maeKdown2, and maeRup1/maeRup2 and maeRdown1/maeRdown2, (Table 2) were used to amplify the regions upstream and downstream of *maeK* and *maeR* (fragments maeKup, maeKdown, maeRup and maeRdown, respectively; Fig. 2 A) using Platinum Pfx DNA polymerase (Invitrogen). In order to generate single merged fragments (maeKdel and maeRdel), 100 ng of on one hand, maeKup and maeKdown and on the other, maeRup and maeRdown fragments were added to 50 µl of a PCR reaction mixture, without primers, after which 20 cycles (94°C, 15 s.; 60°C, 30 s; 72°C, 150 s) were performed. Amplified fragments were separated in a 0.8% agarose gel and the bands corresponding to the expected sizes of maeKdel and maeRdel fragments were excised from the gel and the DNA was purified as indicated above. Using 50 ng of each purified fragment, the fragments were amplified using their corresponding primers maeKup1/maeKdown2 and maeRup1/maeRdown2. After digestion with SpeI and XhoI, the fragments were cloned in pRV300 digested with the same enzymes resulting in plasmids pRVmaeK and pRVmaeR. *L. casei* was transformed with pRVmaeK or pRVmaeR, and for each plasmid, one erythromycin-resistant clone carrying the plasmid integrated by single cross-over was grown in MRS without erythromycin for approx. 200 generations. Cells were plated on MRS and replica-plated on MRS plus erythromycin. Antibiotic-sensitive clones were isolated and, among them, one was selected in which a second recombination event led to the deletion of either *maeK* (strain BL322) or *maeR* (strain BL315), as subsequently confirmed by sequencing of PCR amplified fragments spanning the deleted regions. The resulting sequences of the derivative strains are shown in Fig. 2 D.

In order to complement the MaeK deficiency, a PCR fragment spanning the *mae* promoter region and *maeK* was amplified with Platinum Pfx DNA polymerase by using primers MaeS-1 and MaeS-2. The amplified fragment was digested with BglII and ligated to the expression vector pT1NX digested with BglII and EcoRI (made blunt with the Klenow fragment), resulting in plasmid pT1maeK. In this construct *maeK* was expressed from its own promoter. The ligation mixture was used to transform *Lactococcus lactis* MG1363 by electroporation (13) and transformants were checked by restriction mapping and sequencing of the inserted fragment. Subsequently, plasmid pT1maeK was used to transform *L. casei* BL322.

Expression and purification of His-tagged MaeR.

The coding region of *maeR* was amplified by PCR using chromosomal DNA from *L. casei* BL23 as a template, primers MaeR-Nt and MaeR-Ct (Table 2), which added restriction sites to the 5' and 3' ends and Platinum Pfx DNA polymerase. The PCR fragment was cleaved with BamHI/SacI and cloned into plasmid pQE80 (Qiagen) digested with the same enzymes. The resulting plasmid, pQEmaeR, was used to transform *E. coli* M15(pREP4), and the correct sequence of the inserts was confirmed by DNA sequencing. One clone was selected for expression of MaeR. Bacterial cells were grown in 0.5 liter of LB medium supplemented with ampicillin and kanamycin at 37°C with agitation. When the culture reached an OD_{550} of 0.5, isopropyl-β-D-thiogalactopyranoside (1 mM) was added and incubation was continued for 1 h. Cells were harvested by centrifugation, washed with 10 ml of Tris-HCl 100 mM pH 7.4 and resuspended in 5 ml of Tris-HCl 100 mM (pH 7.4), 1 mg ml^{-1} lysozyme, 0.5 mM phenylmethylsulphonyl fluoride and 0.5 mM dithiothreitol. The cell suspension was incubated for 30 minutes at 37°C and sonicated. The cell debris was removed by centrifugation at 12,000 × *g* for 20 min at 4°C. The cleared extract was directly loaded onto Ni-nitrilotriacetic acid agarose (1 ml) column (Qiagen) equilibrated with buffer A (Tris-HCl 50 mM pH

7.4, Na_2SO_4 50 mM, 15 % glycerol). After the passage of the sample, the column was washed with 10 ml of buffer A, 30 ml of buffer B (buffer A supplemented with 30 mM imidazole) and finally MaeR was eluted with buffer C (buffer A supplemented with 300 mM imidazole). Fractions were analyzed by SDS-PAGE gels, pooled and dialyzed against 20 mM Bis-Tris HCl, pH 6, 25 mM Na_2SO_4, 3.75% glycerol. The dialyzed protein solution was applied to a RESOURCE S 1 ml column (GE Healthcare) equilibrated with 20 mM Bis-Tris HCl, pH 6, 25 mM Na_2SO_4. The column was washed extensively with the same buffer and proteins eluted with a linear NaCl gradient of 0 to 500 mM in the same buffer (total volume, 60 ml). The protein eluted in one peak and it was kept frozen at –80°C. Protein concentrations were determined with a Bio-Rad dye-binding assay.

Gel mobility shift and DNase I footprinting assays.

Primers FP1 and FP2 (Table 2) were 5'-labelled with T4 polynucleotide kinase (USB) and [γ-^{32}P]ATP and purified with MicroSpin G-25 columns (GE Healthcare). They were used in combination with the respective non-labelled oligonucleotide to amplify a 219 bp DNA fragment spanning the intergenic region between *maeP* and *maeK* (fragment *maepro3*; Fig. 2 B). In addition, labelled primer FP2 was used with primer FP4 and labelled primer FP1 with primer FP3 to generate fragments *maepro2* and *maepro1* containing two or one putative MaeR binding sites, respectively (Fig. 2 B). The amplified fragments were used in electrophoretic mobility shift assays with purified His-tagged MaeR. As negative controls the coding region of *maeR* and internal fragments of *maeK* and *maeE* were used. The binding assay was carried out in 10 µl of binding buffer (2.5 mM Tris-HCl [pH 7.5], 100 mM NaCl, 25 mM $MgCl_2$, 0.25 mM EDTA, 0.25 mM dithiothreitol, and 1.5% glycerol) with 10 ng of target DNA (10000 cpm approx.), 0.5 µg of salmon sperm DNA and different amounts of His-tagged MaeR. The binding mixtures were incubated for 1 hour at 37ºC and

separated on 6% non-denaturing polyacrylamide gels in 40 mM Tris-acetate pH 8.2, 1 mM EDTA buffer at 100 V for 1 h. The gel was dried and subjected to autoradiography.

For DNase I footprinting, binding reactions were performed as described for the gel retardation experiments in a total volume of 40 µl with 50000 cpm of each radiolabelled fragment. Complexes were allowed to form for 1 hour at 37ºC, after which 10 µl of 10 mM $MgCl_2$, 1 mM $CaCl_2$ and 0.025 U DNase I were added. Digestion was allowed to proceed for 2.5 min, after which the reactions were terminated by adding 140 µl of stop solution (192 mM Na-acetate, 32 mM EDTA, 0.14% SDS, yeast tRNA 64 µg ml^{-1}). Samples were extracted twice with phenol-chloroform, and the DNase I digestion products were precipitated with ethanol. The precipitates were resuspended in 6 µl of loading buffer and loaded on a 6% polyacrylamide-urea gel. A+G Maxam and Gilbert reactions were run on the same gel to locate sequence positions and protected regions (31).

RESULTS

In silico identification of gene clusters involved in L-malic acid metabolism in *L. casei*.

The inspection of the genomic sequences of *L. casei* ATCC 334 and *L. casei* BL23 led to the identification of two gene clusters possibly involved in L-malic acid metabolism. One cluster (Suppl. Fig. 3) is constituted by genes Lsei_0739/LCABL_08050 (encoding a putative transcriptional regulator of the LysR family), Lsei_0740/LCABL_08060 (encoding a putative malolactic enzyme) and Lsei_0741/LCABL_08070 (encoding a putative malate permease). The second cluster is constituted by two diverging putative operons (Fig. 2). The first one consists of genes Lsei_2866/LCABL_30690 and Lsei_2867/LCABL30700 encoding a putative malic enzyme and a transport protein (Pfam 03390), respectively. The second operon consists of genes Lsei_2868/LCABL_30710 and Lsei_2869/LCABL_30720 encoding a putative two-component

(TC) system closely related to TC systems involved in the regulation of the metabolism of dicarboxylic acids (8). On the basis of the similarities of the genes LCABL_30690 (77% identical, 86% conserved residues) and LCABL_30700 (77% identical, 88% conserved) to genes *maeE* and *maeP*, respectively, of *Streptococcus bovis* (14,15), we propose to rename them *maeE* and *maeP* and to rename genes LCABL_30710 and LCABL_30720, *maeK* and *maeR*, respectively.

Putative rho-independent terminators could be identified downstream of genes LCABL_030690 (*maeE*) and LCABL_30720 (*maeR*), suggesting that both couples of diverging genes constitute two operons (Fig. 2 A and C). Structurally identical gene clusters are present in *Enterococcus faecalis*, *Streptococcus agalactiae, Streptococcus pyogenes* and *Streptococcus uberis* (Suppl. Fig. 3). *Lactobacillus brevis* ATCC 367 also encodes a homologous TC system (Suppl. Fig. 3); however, these genes constitute a cluster with genes encoding a putative membrane protein and a D-lactate dehydrogenase. In fact, *L. brevis* does not code for a malic enzyme (26).

In order to gain insight into the evolutionary relationships of these TC systems, a phylogenetic analysis of selected MaeK and MaeR homologs was carried out. The analysis showed a remarkable congruency with the gene content: both MaeK and MaeR homologs of *L. casei*, *E. faecalis* and streptococci constitute strongly supported clusters whereas the *L. brevis* MaeK appears as distantly related and *L. brevis* MaeR is at a basal branch with low support (Suppl. Fig. 1). Furthermore, the phylogenetic reconstruction suggests a closer relationship to TC systems involved in the regulation of malate metabolism such as the *Bacillus subtilis yufLM* (36) and other TC systems of species belonging to family *Bacillaceae* which form clusters with genes homologous to *maeE* and *maeP* (Suppl. Fig. 1 and 3). In summary, the phylogenetic analysis indicate that the *mae* gene clusters present in LAB share a common origin although it cannot determine whether their distribution can be explained by horizontal gene transfer or lineage-specific gene losses and suggest a close relationship with *mae* clusters present in bacilli.

maeE enables *L. casei* BL23 to grow with L-malic acid as a carbon source and the TC system MaeKR is essential for this ability

The growth behaviour of *L. casei* BL23 in MEI medium supplemented with glucose and/or L-malate was investigated. As a control, *L. casei* BL23 was inoculated in MEI medium without glucose or L-malate. Under our experimental conditions, *L. casei* BL23 was able to grow in MEI medium without supplementation of glucose or L-malic acid to an O.D. of 0.28. When MEI medium was supplemented with L-malic acid, *L. casei* BL23 initially grew as in MEI medium but after a lag phase it resumed growth at a low rate (Fig. 3). In order to determine whether expression of gene *maeE* was responsible for the ability of *L. casei* BL23 to grow on L-malate, a mutant strain defective in *maeE* was constructed (BL321). The mutant strain grew in MEI medium supplemented with glucose or glucose and L-malic acid at growth rates similar to the wild type strain but it was not able to grow with L-malic acid alone (Fig. 3). Products of L-malic acid degradation were also analysed. During the initial growth phase the concentration of lactic acid exceeded that expected from L-malic acid consumption (Table 3) possibly due to fermentation of components of the MEI medium as expected from the observed basal growth in MEI medium. Subsequently, the consumption of L-malic acid correlated with the accumulation of lactic acid. A small amount of acetic acid was only detected in samples corresponding to late stationary cells (data not shown), indicating the activation of pyruvate dissimilative pathways alternative to lactate dehydrogenase (Fig. 1).

The structure of the *L. casei mae* cluster suggested that MaeKR might be involved in the regulation of the expression of *maePE*. This hypothesis was in agreement with previous results obtained by other researchers. On one hand, a number of studies had shown that expression of ME in *E. faecalis* and *L. casei* was induced by L-malic acid when glucose was at a concentration lower than 0.2%

(20,22,33). On the other hand, the expression in *Bacillus subtilis* of the malate transporter MaeN had been shown to be under control of the YufLM TC system (36). YufLM shares significant similarity with MaeKR (35% identical, 56% conserved residues for MaeK and 39% identical, 57% conserved residues for MaeR) and the phylogenetic analyses suggested an evolutionary relationship between these two systems. In order to ascertain the possible role of the MaeKR TC system, two BL23 derivative strains with deletions of *maeK* (BL322) or *maeR* (BL315) were constructed as indicated in Methods and their ability to grow with L-malic acid as a carbon source was assayed. Both derivative strains were unable to grow in MEI supplemented with 5 g/l L-malic acid whereas no significant differences were observed with the wild-type strain in MEI supplemented with glucose and L-malic acid (Fig. 3). In order to confirm that the loss of the ability to grow on L-malic acid was due to the inactivation of the MaeKR TC system, the mutant strain BL322 was transformed with plasmid pT1maeK. The expression of MaeK in the Δ*maeK* strain restored the ability to grow with L-malic acid (Fig. 3). Therefore, the MaeKR TC system is essential for growth with L-malic acid.

Transcriptional analysis of the *mae* gene cluster

In order to determine whether the MaeKR TC system is involved in the control of the expression of the *maePE* operon, Northern blot experiments were performed with RNA isolated from *L. casei* BL23 and the derivative strain BL315 (Δ*maeR*) grown in MEI containing glucose, L-malic acid or glucose plus L-malic acid. The results suggested that BL23 strain only expressed the *maePE* operon in the presence of L-malic acid and absence of glucose (Fig. 4 A) and only during active growth on L-malic acid since transcription was not detected in stationary phase. Transcription of *maePE* in the Δ*maeR* strain was not observed under any condition. The hybridization signals were at the level of the ribosomal RNAs, likely due to degradation or comigration of the RNA. These results were confirmed by RT-PCR analysis of *maeE*, which showed that the *maeE* cDNA was preferentially

amplified from BL23 strain actively growing in medium containing L-malic acid (Fig. 4 B). A faint signal was also observed in the cDNA sample obtained from cells of BL23 grown with glucose and L-malic acid (Fig. 4 B) possibly indicating a basal expression of *maeE* below the detection threshold of the Northern hybridization. Expression of the *maeKR* operon by using Northern blot analysis could only be detected after overexposure of the membrane (not shown) and changes in expression between different samples could not be reliably determined. Therefore, RT-qPCR assays were carried out in order to determine the expression pattern of genes *maeK* and *maeR*. RNA samples from BL23 cultures grown for 24 h in MEI medium supplemented with glucose, glucose and L-malic acid or L-malic acid were used. Results obtained are shown in Table 4. Taking as reference the transcript levels in MEI medium with glucose, a significant increase in expression of *maeK* and *maeR* was detected both in MEI medium with glucose and L-malic acid and in MEI with L-malic acid. Although both genes possibly are expressed as a single transcript, a difference in expression ratio was observed between *maeK* and *maeR* (Table 4). This was evidenced when the expression ratio in MEI with L-malic acid was determined using MEI with glucose and L-malic acid as the reference condition. Whereas no significant difference was observed for *maeR*, the expression of *maeK* was 1.8 fold lower in MEI with L-malic acid. Whether this result indicates a control system at post-transcriptional level such as mRNA processing, it requires further confirmation. In any case, the results obtained indicate that expression of *maeKR* is induced in the presence of L-malic acid and, as a difference with the *maePE* operon, glucose does not repress expression of *maeKR*.

The transcriptional initiation site of the *maePE* operon was determined by 5'-RACE and shown to be the T or A located at position 25 or 26 upstream the *maeP* translational start site (Fig. 4 C). The transcriptional initiation site could not be unequivocally identified due to an artefact of RACE when the start site is a T. On the other hand, the transcriptional initiation site of the *maeKR* operon could

not be reliably identified by this technique, possibly because of the low levels of transcript present in the samples. The inspection of the sequence allowed to identify a putative -10 box (5'-TATGCT-3') at -6 from the transcriptional initiation site of *maePE* (Fig. 4 C) although no -35 box could be identified. In concordance to the repressive effect of glucose on transcription, a sequence matching the consensus of the Gram-positive catabolite repression element (*cre*), the binding site of the CcpA regulator mediating carbon catabolite repression (27) as located upstream of the *maeP* -10 box. Finally, a second -10 box (5'-TTTACT-3') was located 39 bases upstream of the putative translational start site of *maeK* (Fig. 4 C).

MaeR binds specifically to the *mae* cluster promoter region

In order to verify the role of MaeR in the regulation of the expression of the *maePE* operon, gel retardation experiments were performed using fragment *maepro3* which spans the entire *mae* promoter region or nonspecific DNA fragments (MaeKup and MaeRdown; see Fig. 2 A) as controls of the specificity of the binding reaction. MaeR was able to bind to the *maepro3* fragment (Fig. 5, left panel) in the presence of a large excess of unspecific DNA competitor but did not bind to unspecific fragments (not shown). Treatment of MaeR with up to 25 mM acetyl-phosphate, which is an *in vitro* phosphate donor able to phosphorylate different RRs (25) did not change its binding activity (data not shown).

DNase I footprinting on DNA fragment *maepro3* was carried out to determine the sequences recognized by MaeR. Fig. 6 shows that MaeR protects two regions within *maepro3*. These regions contain three AT-rich direct repeats (5'-TTATT(A/T)AA-3') (Fig. 6), two of them separated by one base and the third one located 12 bases apart. Taking into account the contacts required by DNase I to cut the DNA at a certain position (35) and the presence of unprotected sites between the second and third repeats it can be inferred that protection is centred in the repeats. A remarkable difference

in protection could be appreciated between the forward (Fig. 6, right panel) and the reverse strand (left panel): whereas clear protected regions could be observed in the reverse strand a change in the pattern of sites preferentially digested by DNase I was observed. This result suggests that MaeR binds to the reverse strand leaving the forward strand relatively unprotected. Furthermore, the change in the DNase I digestion pattern indicates that MaeR binding results in an alteration of the structure of the forward strand.

In order to determine whether one or two MaeR binding sites would be enough for the formation of a MaeR:DNA stable complex, additional gel retardation experiments were performed using fragments *maepro2* and *maepro1*. Under our experimental conditions, MaeR did not form stable complexes either with fragment *maepro2* (Fig. 5, right panel) or *maepro1* (not shown) thus indicating that the three direct repeats are required for stable MaeR DNA binding.

DISCUSSION

Whereas MLF is widely distributed among LAB, ME is present in few species. The phylogenetic analysis reported here and the comparison of the respective *mae* gene clusters indicate that *mae* gene clusters present in LAB evolved from a common ancestor although data available do not allow to determine whether this cluster has been disseminated by horizontal gene transfer or the observed distribution is explained by multiple lineage specific gene losses. MLE decarboxylates L-malic acid to L-lactic acid which cannot be further utilized by most LAB. In contrast, biochemical characterization of *L. casei* ME showed that this enzyme yielded pyruvate and CO_2 from malate, utilized preferentially NAD over NADP and could also react with oxaloacetate (4). Pyruvate produced by L-malic acid can then be channelled to biosynthetic pathways via pyruvate phosphate dikinase or pyruvate carboxylase and PEP carboxykinase (Fig. 1) or through the lactate dehydrogenase, pyruvate dehydrogenase, pyruvate oxidase or pyruvate formate lyase pathways of dissimilation (Fig. 1). Therefore, *L. casei* possesses the required enzymatic complement to grow on

L-malic acid as a carbon source. Our results confirm this hypothesis and show that the ME encoded by *maeE* is essential for growth with L-malic acid. The analysis of the final products of L-malic acid utilization showed that most L-malic acid was degraded to lactate so that only a minor part was used for growth which would agree with the slow growth rate and low maximal O.D. reached by the cultures. Furthermore, *L. casei* codes for a putative fumarase (EC 4.2.1.2; LCABL_25800) which might convert malate into fumarate, however fumarate concentrations ranged between 0.8 and 4 μM and did not vary significantly during *L. casei* growth indicating that fumarase activity had not a significant role in L-malic acid metabolism under our experimental conditions.

Previous studies had shown that *L. casei* ME was inhibited by glycolytic intermediaries such as fructose-1,6-bisphosphate and 3-phosphoglycerate and that ME expression was induced in the presence of malic acid and low concentration of glucose (22). Studies conducted in *E. faecalis* also showed that ME synthesis was induced by L-malic acid and inhibited by glucose (21). This study also showed that ME and the L-malic acid transport system were coordinately expressed. Our results, in agreement with the aforementioned studies, showed that expression of gene *maeE* was repressed in the presence of glucose, which is congruent with the presence of a *cre* sequence in its promoter sequence. Therefore, expression of *maePE* is possibly regulated by the global carbon catabolite repression transcriptional regulator CcpA (27). Furthermore, we have shown that the cognate TC system encoded by *maeK* and *maeR* plays a role in the induction mechanism of *maePE* in the presence of L-malic acid. The TC system MaeKR is then a functional homolog of the *B. subtilis* system YufLM which was shown to be essential for growth on L-malic acid (36) and controls the expression of genes *maeN* and *yflS* encoding malate transporters (36) and *ywkA* (*maeA*) encoding a ME (7). Notwithstanding, there are some differences in the regulatory mechanisms operated by MaeKR and YufLM: deletion of the sensor encoding gene *yufL* resulted in constitutive expression of the transporter encoded by *maeN* (36) whereas deletion of *maeK* or *maeR* resulted in

inability to grow on L-malic acid or to express *maeE* and *maeP* in *L. casei*. There are also important differences in L-malic acid metabolism between *B. subtilis* and *L. casei* since the latter lacks of a functional tricarboxylic acid cycle (TCA). *L. casei* possesses MLE and ME whereas *B. subtilis* contains four putative ME encoding genes, *malS*, *mleA*, *ytsJ* and *maeA* (7). MaeE is essential for growth on L-malic acid; in contrast, absence of MaeA can be compensated by MalS or MleA in *B. subtilis* (7,19). In fact, *B. subtilis* can grow on L-malic acid, albeit slowly, even in the absence of the four ME since L-malic acid can be channelled into TCA via malate dehydrogenase (19).

The DNase I footprinting experiments showed that MaeR binds to a set of direct repeats in the *mae* promoter region. (5'-TTATT(A/T)AA-3'), two of them separated by one base and the third one twelve bases apart. In contrast, DNase I footprinting analyses of YufM binding sites identified two sets of direct repeats within the protected region of the *maeA* promoter, ATTAAAAAATTN11ATTAAAAAATT and TAAGTAN11TAAGTA (7) whereas no clear repeat sequences could be identified in the YufM-binding regions of the *maeN* and *yflS* promoter sequences (36). The inspection of the *mae* promoter regions in the homologous clusters of *E. faecalis*, *S. agalactiae*, *S. pyogenes* and *S. uberis* showed the presence of highly conserved repeats (Fig. 7) although with significant differences. *E. faecalis* conserves a similar array of repeats (two repeats separated by one base and a third one eleven bases apart). In contrast, in *Streptococcus* the order of the repeats is inverted and the gap separating the repeats is larger (19 bp for *S. agalactiae* and 28 bp for *S. pyogenes* and *S. uberis*). Furthermore, within this intervening region several degenerated MaeR binding sites are present (Fig. 7). These differences, together with the different spacing between the MaeR binding sites and the putative -10 boxes suggest that regulation of the *mae* cluster by MaeR may differ in streptococci from the regulation in *L. casei* and possibly, *E. faecalis*.

The distance (25 bp; Fig. 7) between the repeats and the -10 boxes suggest that MaeR act as a class II activator (5) in streptococci and for the *maePE* promoter of *E. faecalis* and *L. casei*. From the DNase I footprinting data it can be inferred that this is also the mechanism of activation of YufM (7). However, this seems not to be the case for the *maeKR* promoters of *E. faecalis* and *L. casei*. In these organisms, putative -10 boxes were identified but none of them is located at the expected distance from the most proximal *maeR* binding site (Fig. 7). Although we could not determine the transcriptional start site of the *maeKR* transcript, the transcriptional data indicate that expression of *maeKR* is induced by L-malic acid, thus strongly suggesting that MaeR also induces the expression of *maeKR*. If so, the mechanism of activation operated by MaeR must be different in the *maePE* promoter and the *maeKR* promoter, although further research will be required to ascertain this point.

In summary, our results have shown that expression of the *maeP* and *maeE* genes, required for growth on L-malic acid, is regulated by the cognate MaeKR TC system and repressed by glucose. We have shown that MaeR binds specifically to a set of three direct repeats in the *mae* promoter region which are conserved in promoter regions of homologous gene clusters in several bacteria. *In vitro* MaeR promoter binding activity did not require phosphorylation of the regulatory aspartate at the RR receptor domain. Further research is needed to determine how the environmental signal sensed by MaeK (most probably the presence of L-malic acid) is presumably transferred to MaeR.

ACKNOWLEDGMENTS

This work was financed by funds of the AGL2007-60975 and Consolider Fun-*C*-Food CSD2007-00063 from the Spanish Ministry of Science and Innovation.

We thank Isidoro Olmeda and José Vicente Gil for technical assistance.

Reference List

1. **Abascal, F., R. Zardoya, and D. Posada**. 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics. **21**:2104-2105.

2. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman**. 1990. Basic local alignment search tool. J. Mol. Biol. **215**:403-410.

3. **Altschul, S. F., T. L. Madden, A. A. Schaffer, J. Zhang, Z. Zhang, W. Miller, and D. J. Lipman**. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**:3389-3402.

4. **Battermann, G. and F. Radler**. 1991. A comparative study of malolactic enzyme and malic enzyme of different lactic acid bacteria. Can. J. Microbiol. **37**:211-217.

5. **Browning, D. F. and S. J. Busby**. 2004. The regulation of bacterial transcription initiation. Nat. Rev. Microbiol. **2**:57-65.

6. **Castresana, J.** 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. **17**:540-552.

7. **Doan, T., P. Servant, S. Tojo, H. Yamaguchi, G. Lerondel, K. Yoshida, Y. Fujita, and S. Aymerich**. 2003. The *Bacillus subtilis ywkA* gene encodes a malic enzyme and its transcription is activated by the YufL/YufM two-component system in response to malate. Microbiology **149**:2331-2343.

8. **Fabret, C., V. A. Feher, and J. A. Hoch**. 1999. Two-component signal transduction in *Bacillus subtilis*: how one organism sees its world. J. Bacteriol. **181**:1975-1983.

9. **Frayne, R. F.** 1986. Direct analysis of the major organic components in grape must and wine using high-performance liquid chromatography. Am. J. Enol. Vitic. **37**:281-287.

10. **Gasson, M. J.** 1983. Plasmid complements of *Streptococcus lactis* NCDO 712 and other lactic streptococci after protoplast-induced curing. J. Bacteriol. **154**:1-9.

11. **Golby, P., S. Davies, D. J. Kelly, J. R. Guest, and S. C. Andrews**. 1999. Identification and characterization of a two-component sensor-kinase and response-regulator system (DcuS-DcuR) controlling gene expression in response to C4-dicarboxylates in *Escherichia coli*. J. Bacteriol. **181**:1238-1248.

12. **Guindon, S. and O. Gascuel**. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. **52**:696-704.

13. **Holo, H. and I. F. Nes**. 1989. High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. Appl. Environ. Microbiol. **55**:3119-3123.

14. **Kawai, S., H. Suzuki, K. Yamamoto, M. Inui, H. Yukawa, and H. Kumagai**. 1996. Purification and characterization of a malic enzyme from the ruminal bacterium *Streptococcus*

bovis ATCC 15352 and cloning and sequencing of its gene. Appl. Environ. Microbiol. **62**:2692-2700.

15. **Kawai, S., H. Suzuki, K. Yamamoto, and H. Kumagai**. 1997. Characterization of the L-malate permease gene (*maeP*) of *Streptococcus bovis* ATCC 15352. J. Bacteriol. **179**:4056-4060.

16. **Koretke, K. K., A. N. Lupas, P. V. Warren, M. Rosenberg, and J. R. Brown**. 2000. Evolution of two-component signal transduction. Mol. Biol. Evol. **17**:1956-1970.

17. **Le, S. Q. and O. Gascuel**. 2008. An improved general amino acid replacement matrix. Mol. Biol. Evol. **25**:1307-1320.

18. **Leloup, L., S. D. Ehrlich, M. Zagorec, and F. Morel-Deville**. 1997. Single-crossover integration in the *Lactobacillus sake* chromosome and insertional inactivation of the *ptsI* and *lacL* genes. Appl. Environ. Microbiol. **63**:2117-2123.

19. **Lerondel, G., T. Doan, N. Zamboni, U. Sauer, and S. Aymerich**. 2006. YtsJ has the major physiological role of the four paralogous malic enzyme isoforms in *Bacillus subtilis*. J. Bacteriol. **188**:4727-4736.

20. **London, J. and E. Y. Meyer**. 1969. Malate utilization by a group D *Streptococcus*: physiological properties and purification of an inducible malic enzyme. J. Bacteriol. **98**:705-711.

21. **London, J. and E. Y. Meyer**. 1970. Malate utilization by a group D *Streptococcus*: regulation of malic enzyme synthesis by an inducible malate permease. J. Bacteriol. **102**:130-137.

22. **London, J., E. Y. Meyer, and S. Kulczyk**. 1971. Comparative biochemical and immunological study of malic enzyme from two species of lactic acid bacteria: evolutionary implications. J. Bacteriol. **106**:126-137.

23. **London, J., E. Y. Meyer, and S. R. Kulczyk**. 1971. Detection of relationships between *Streptococcus faecalis* and *Lactobacillus casei* by immunological studies with 2 Forms of malic enzyme. J. Bacteriol. **108**:196-201.

24. **Lonvaud-Funel, A.** 1999. Lactic acid bacteria in the quality improvement and depreciation of wine. Antonie van Leeuwenhoek **76**:317-331.

25. **Lukat, G. S., W. R. McCleary, A. M. Stock, and J. B. Stock**. 1992. Phosphorylation of bacterial response regulator proteins by low molecular weight phospho-donors. Proc. Natl. Acad. Sci. U. S. A **89**:718-722.

26. **Makarova, K., A. Slesarev, Y. Wolf, A. Sorokin, B. Mirkin, E. Koonin, A. Pavlov, N. Pavlova, V. Karamychev, N. Polouchine, V. Shakhova, I. Grigoriev, Y. Lou, D. Rohksar, S. Lucas, K. Huang, D. M. Goodstein, T. Hawkins, V. Plengvidhya, D. Welker, J. Hughes, Y. Goh, A. Benson, K. Baldwin, J. H. Lee, I. Díaz-Muñiz, B. Dosti, V. Smeianov, W. Wechter, R. Barabote, G. Lorca, E. Altermann, R. Barrangou, B. Ganesan, Y. Xie, H. Rawsthorne, D. Tamir, C. Parker, F. Breidt, J. Broadbent, R. Hutkins, D. O'Sullivan, J. Steele, G. Unlu, M. Saier, T. Klaenhammer, P. Richardson, S. Kozyavkin, B. Weimer, and D. Mills**. 2006. Comparative genomics of the lactic acid bacteria. Proc. Natl. Acad. Sci. USA **103**:15611-15616.

27. **Monedero, V., M. J. Gosalbes, and G. Pérez-Martínez**. 1997. Catabolite repression in *Lactobacillus casei* ATCC 393 is mediated by CcpA. J. Bacteriol. **179**:6657-6664.

28. **Monedero, V., A. Mazé, G. Boël, M. Zúñiga, S. Beaufils, A. Hartke, and J. Deutscher**. 2007. The phosphotransferase system of *Lactobacillus casei*: regulation of carbon metabolism and connection to cold shock response. J. Mol Microbiol. Biotechnol. **12**:20-32.

29. **Pfaffl, M. W., G. W. Horgan, and L. Dempfle**. 2002. Relative expression software tool (REST) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucleic Acids Res. **30**:e36.

30. **Posno, M., R. J. Leer, N. van Luijk, M. J. van Giezen, P. T. Heuvelmans, B. C. Lokman, and P. H. Pouwels**. 1991. Incompatibility of *Lactobacillus* vectors with replicons derived from small cryptic *Lactobacillus* plasmids and segregational instability of the introduced vectors. Appl. Environ. Microbiol. **57**:1822-1828.

31. **Sambrook, J., E. F. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

32. **Schotte, L., L. Steidler, J. Vandekerckhove, and E. Remaut**. 2000. Secretion of biologically active murine interleukin-10 by *Lactococcus lactis*. Enzyme Microb. Technol. **27**:761-765.

33. **Schütz, M. and F. Radler**. 1974. Das Vorkommen von Malatenzym und Malo-Lactat-Enzym bei verschiedenen Milchsäurebakterien. Arch. Microbiol. **96**:329-339.

34. **Stock, A. M., V. L. Robinson, and P. N. Goudreau**. 2000. Two-component signal transduction. Annu. Rev. Biochem. **69**:183-215.

35. **Suck, D., A. Lahm, and C. Oefner**. 1988. Structure refined to 2A of a nicked DNA octanucleotide complex with DNase I. Nature **332**:464-468.

36. **Tanaka, K., K. Kobayashi, and N. Ogasawara**. 2003. The *Bacillus subtilis* YufLM two-component system regulates the expression of the malate transporters MaeN (YufR) and YflS, and is essential for utilization of malate in minimal medium. Microbiology **149**:2317-2329.

37. **Thompson, J. D., D. G. Higgins, and T. J. Gibson**. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**:4673-4680.

38. **Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman**. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol **3**:RESEARCH0034.

39. **Yebra, M. J., V. Monedero, M. Zúñiga, J. Deutscher, and G. Pérez-Martínez**. 2006. Molecular analysis of the glucose-specific phosphoenolpyruvate : sugar phosphotransferase system from *Lactobacillus casei* and its links with the control of sugar metabolism. Microbiology **152**:95-104.

40. **Yebra, M. J., A. Veyrat, M. A. Santos, and G. Pérez-Martínez**. 2000. Genetics of L-sorbose transport and metabolism in *Lactobacillus casei*. J. Bacteriol. **182**:155-163.

41. **Yebra, M. J., M. Zúñiga, S. Beaufils, G. Pérez-Martínez, J. Deutscher, and V. Monedero**. 2007. Identification of a gene cluster enabling *Lactobacillus casei* BL23 to utilize myo-inositol. Appl. Environ. Microbiol. **73**:3850-3858.

42. **Zúñiga, M., M. C. Miralles, and G. Pérez-Martínez**. 2002. The product of *arcR*, the sixth gene of the *arc* operon of *Lactobacillus sakei*, is essential for expression of the arginine deiminase pathway. Appl. Environ. Microbiol. **68**:6051-6058.

FIGURE LEGENDS

Fig. 1. Schematic representation of the L-malate catabolic pathways and the putative gluconeogenic pathways in *L. casei* (steps shared by the gluconeogenic and glycolytic pathways have been omitted). ACDH, acetaldehyde dehydrogenase; ACK, acetate kinase; ADH, alcohol dehydrogenase; FBP, fructosebisphosphatase; LDH, lactate dehydrogenase; ME, malic enzyme; MLE, malolactic enzyme; PC, pyruvate carboxylase; PDH, pyruvate dehydrogenase; PFL, pyruvate formate-lyase; PEPCK, PEP carboxykinase; PGI, Glucose-6-P isomerase; PGM, phosphoglucomutase; POX, pyruvate oxidase; PPDK, pyruvate phosphate dikinase;

Fig. 2. (A) Schematic representation of the *mae* gene cluster. Fragments used for the deletion of genes *maeK* and *maeR* are shown as horizontal lines. (B) Positions of primers used to generate fragments *maepro1*, *maepro2* and *maepro3* for gel retardation and DNase I footprinting assays. Translational startsites of *maeP* and *maeK* are indicated. (C) Sequences of putative rho-independent terminators downstream of genes *maeE* and *maeR*. The coordinates correspond to the *L. casei* BL23 genomic sequence (Acc Nº Fm177140). (D) Sequence comparison of the Δ*maeK* and Δ*maeR* derivative strains with the parental strain BL23. Putative Shine-Dalgarno sequences (RBS), start and stop codons of *maeK* and *maeR* are indicated.

Fig. 3. Growth of *L. casei* BL23 and derivative strains *maeE*$^-$ (BL321), Δ*maeK* (BL322), Δ*maeR* (BL315) and Δ*maeK* harbouring pT1maeK in MEI medium supplemented with 5 g l^{-1} of glucose and 5 g l^{-1} of L-malic acid (MG) or 5 g l^{-1} L-malic acid (M). Values represent the means of three independent experiments; error bars represent standard deviations. Arrows indicate sampling points for transcriptional analyses.

Fig. 4. (A) Northern analysis of samples of *L. casei* BL23 (left) and Δ*maeR* (right) grown in MEI supplemented with: G, 5 g l^{-1} glucose; G+M, glucose and L-malic acid 5 g l^{-1}; M, L-malic acid, 5 g l^{-1} using a *maeE* specific probe. The upper pictures show the membranes stained with methylene blue, the lower picture the hybridization with the specific probe. Numbers indicate the $O.D._{595}$ when the samples were taken: mid exponential phase, onset of stationary phase and late stationary phase for cultures G and G+M, and onset of growth and late stationary phase for culture M (see Fig. 3). (B) RT-PCR analysis of *maeE* expression. The same RNA samples from panel A were used (C) DNA sequence of the promoter region of the *mae* gene cluster. Putative -10 boxes and a putative cre (catabolite responsive element) site are indicated; transcriptional start sites are indicated by vertical arrows; horizontal arrows indicate putative MaeR binding sites; RBS are indicated in bold characters; start codons are underlined. Horizontal lines under genes indicate probes used for transcriptional analyses.

Fig. 5. Left panel: binding of MaeR to the *maepro3* DNA fragment. A dash above a lane indicates no MaeR added. The triangle indicates the increase in the amounts of MaeR present in the binding reaction (0.025 mg ml^{-1}, 0.05 mg ml^{-1}, 0.1 mg ml^{-1} and 0.2 mg ml^{-1}). Right panel: binding of MaeR to the *maepro2* DNA fragment (0.2 mg ml^{-1}, 0.25 mg ml^{-1} and 0.3 mg ml^{-1} of MaeR). As a positive control, binding of MaeR (0.2 mg ml^{-1}) to fragment *maepro3* was run in the rightmost lane.

Fig. 6. Dnase I footprinting analysis of the MaeR-mae promoter region complex. Left and right panels correspond to each strand. A dash above a lane indicates no MaeR added. The triangles indicate the increase in the amountsof MaeR present in the binding reaction (0, 0.025 mg ml^{-1}, 0.05 mg ml^{-1}, 0.1 mg ml^{-1} and 0.2 mg ml^{-1}). AG: lane with the AG Maxam and

Gilbert reaction products of the same DNA fragment. Empty bars indicate the protected regions. Arrows indicate the position of AT-reach direct repeats. Black triangles indicate unprotected Dnase I sites between the set of direct repeats.

Fig. 7. Schematic representation of *mae* promoter regions in LAB. Translational startsites of *maeP* (5') and *maeK* (3') are underlined. MaeR binding sites are indicated by thick arrows and degenerated binding sites by thick dotted arrows. Putative -10 boxes are also indicated.

1 TABLE 1. Strains and plasmids used in this study

Strain	Characteristics or relevant genotype	Source or reference
Escherichia coli		
DH5α	F⁻ *endA1 hsdR17 gyrA96 thi-1 recA1 relA1 supE44* Δ*lacU169* (F80 *lacZ* DM15)	Stratagene
M15(pREP4)	Nal^s, Str^s, Rif^s, *Thi*$^-$, *Lac*$^-$, *Ara*$^+$, *Gal*$^+$, *Mtl*$^-$, F$^-$, *RecA*$^+$, *Uvr*$^+$, *Lon*$^+$	Qiagen
Lactobacillus casei		
BL23	wild type; genome sequenced at CNRS[a], INRA[b] and CSIC[c]	B. Chassy, U. Illinois
BL315	BL23 Δ*maeR*	This study
BL321	BL23 *maeE*::pRV300 Ery^r	This study
BL322	BL23 Δ*maeK*	This study
Lactococcus lactis		
MG1363	Plasmid-free derivative of NCDO712	(10)
pQE80	*cis*-repressed, IPTG inducible, N-terminal His_6-tagged recombinant protein overexpression vector; Amp^r	Qiagen
pQEmaeR	pQE80 with cloned *maeR* gene	This study
pRV300	Insertional vector for *Lactobacillus*, Amp^r, Ery^r	(18)

pRVmaeE	pRV300 containing a 0.6 kb internal fragment of *maeE*	This study
pRVmaeK	pRV300 containing a 2 kb with a *maeK* deleted fragment cloned at SpeI and XhoI sites	This study
pRVmaeR	pRV300 with a 2 kb fragment carrying a *maeR* deleted fragment cloned at SpeI and XhoI sites	This study
pT1NX	Expression vector for Gram-positive bacteria harbouring the constitutive P1 promoter, Ery^r	(32)
pT1maeK	pT1NX with cloned *maeK* expressed from its own promoter	This study

1

1

2 TABLE 2. Primers used in this study.

Primer	Sequence[a]
MaeE-F	TTTTAAGCTTACCAAGAAAAACACGGTTGC
MaeE-R	TTTTCCGCGGAATTTCCGGCTCTGGATTAG
MaeKup1	TTTTCTCGAGCGGTGCCAATGGACTGTTTCAG
MaeKup2	CACAAGATGGTTTCTCTACTGAACAGGCTGACACTAACATTA ACCGTGG
MaeKdown1	CCACGGTTAATGTTAGTGTCAGCCTGTTCAGTAGAGAAACCA TCTTGTG
MaeKdown2	TTTTGAGCTCTCAATTTCGCCTGTATCGTCTTTCATCG
MaeKprobe1	TTTTAAGCTTGCAAGTAGATGGCCAATTCG
MaeKprobe2	TTTTCCGCGGCATTAAGGTCACCTCACTG
MaeRup1	TTTTACTAGTGCGCTATTCGCCGGTACG
MaeRup2	GCCTTTTTAAAGATTTGAATGATGTTTATCAACGATTAGGATG TTCGTC
MaeRdown1	GACGAACATCCTAATCGTTGATAAACATCATTCAAATCTTTA AAAAGGC
MaeRdown2	TTTTCTCGAGGATAAAGTGCCCGCCAAGATG
MaeR-Nt	TTTTGGATCCATGACGAACATCCTAATCGTTG
MaeR-Ct	TTTTGAGCTCTTACGACAGCGCATGTGGTGCCG
MaeE1	TTTTAAGCTTACCAAGAAAAACACGGTTGC
MaeE2	TTTTCCGCGGAATTTCCGGCTCTGGATTAG
MalicR-1	GATAAAATATTGCCAACTTAGC

MalicR-2	CAGTGTCCCAACACCCATTG
MalicR-3	ACCGCCAAAGTTCTTCACAA
MaeS-1	TTGTAACGCCTTCATCAGGGGAAG
MaeS-2	TTTTAGATCTTCAACGATTAGGATGTTCGTATTTTG
MaeA	TTGAGCCTCGTGGTAGTCCTCATAATGTTAACAGCGTTTTGGTT
MaeG	TCAGCCTTGTTTGCTTCAAG
RaceG2	GCGGTTACACGGGGATCCTTTGC
FP1	AGCCCCATAGCCGCCATCC
FP2	CGTTGATTTGTTGGGTGTAACC
FP3	CCATAAACCCATTTATTTAAGCG
FP4	GGGTTTATGGTTTAATAAGTTAAATAA

1 [a]. Restriction sites used for cloning are underlined.

TABLE 3. Increment in L-malic acid and lactic acid during growth of *L. casei* BL23 in MEI medium supplemented with L-malic acid

Time (h)[a]	Δ[L-malic acid][b]	Δ[lactic acid][b]
4.5	-4.12 ± 0.11	5.01 ± 0.70
5.5	-1.09 ± 0.82	2.07 ± 0.17
23.0	-14.70 ± 1.99	14.88 ± 0.95
27.5	-3.29 ± 0.03	3.14 ± 0.28

[a] Time of sampling (see Fig. 3).

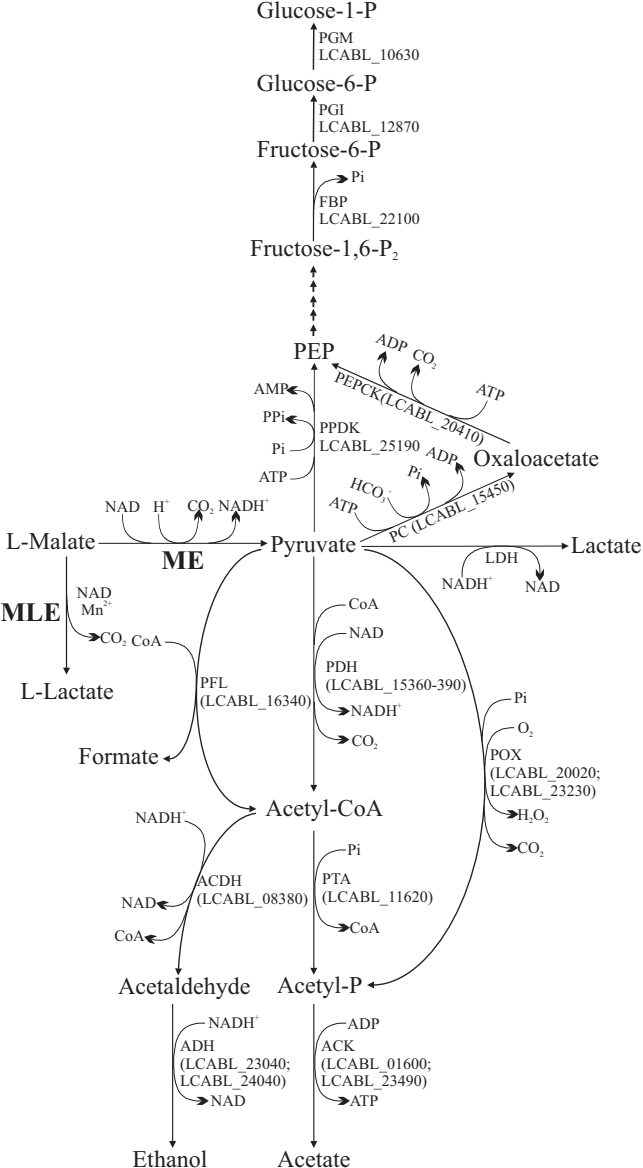
[b] For each sampling time t $\Delta[acid] = [acid]_t - [acid]_{t-1}$ (mM). Data express the average and standard deviation of two independent experiments.

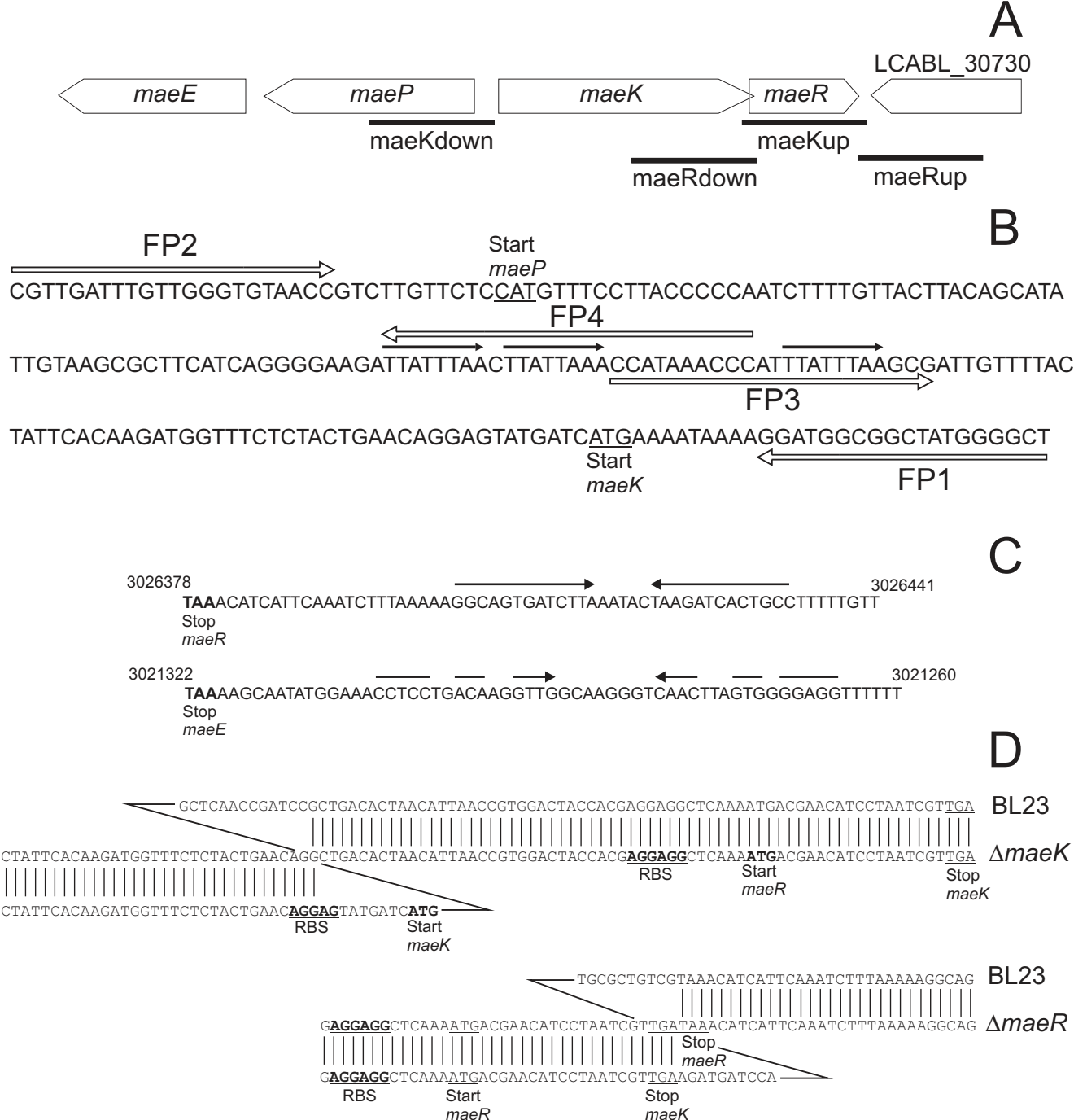
TABLE 4. RT-qPCR analysis of *maeR* and *maeK* transcription in *L. casei* BL23[a]

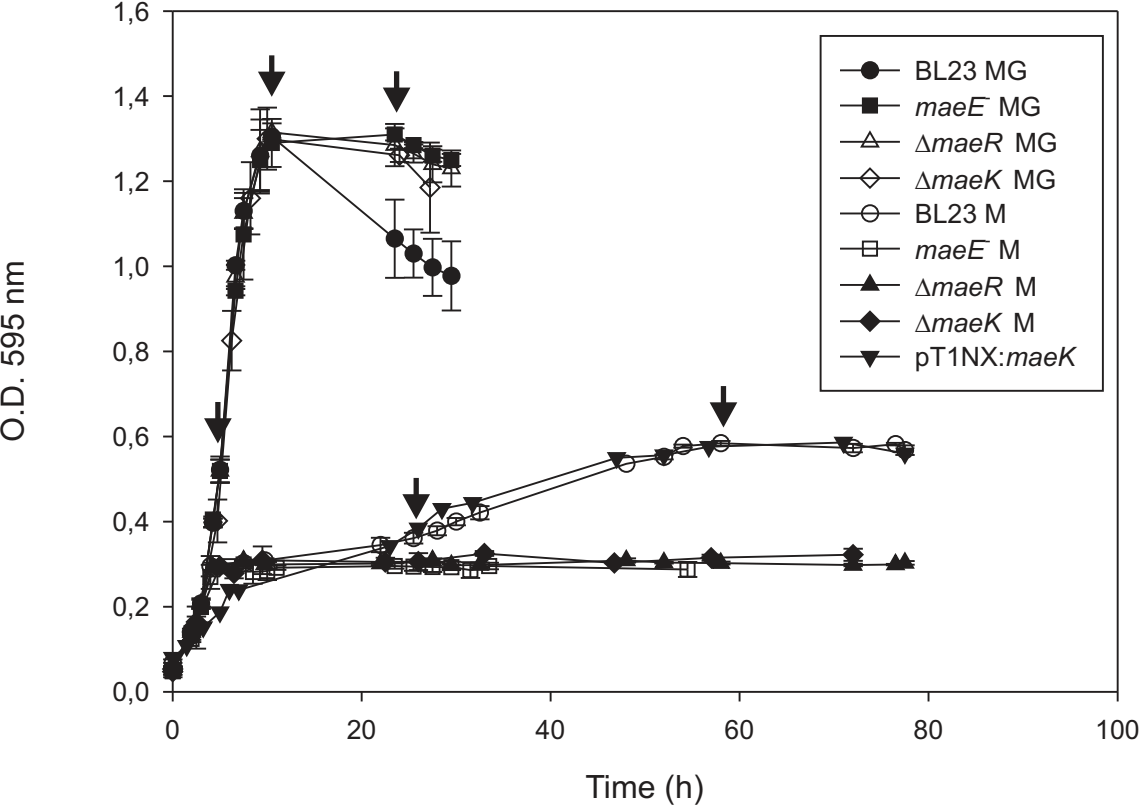
Reference	Sample							
	MEI+G+M[b]				MEI+M[b]			
	maeR		*maeK*		*maeR*		*maeK*	
	R.E.	S.E.	R.E.	S.E.	R.E.	S.E.	R.E.	S.E.
MEI+G[b]	7.56	4.27-12.81	16.58	12.24-21.87	7.47	5.56-10.53	9.03	6.36-12.13
MEI+G+M[b]					0.99	0.49-1.63	0.54	0.40-0.66

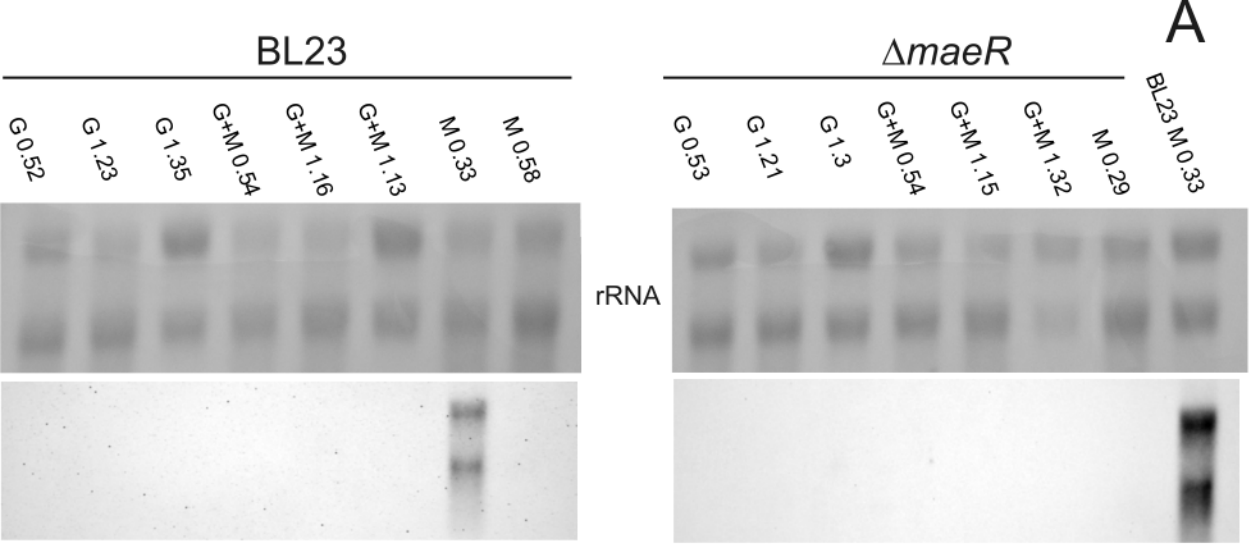
[a] R.E., relative expression; S.E., standard error.

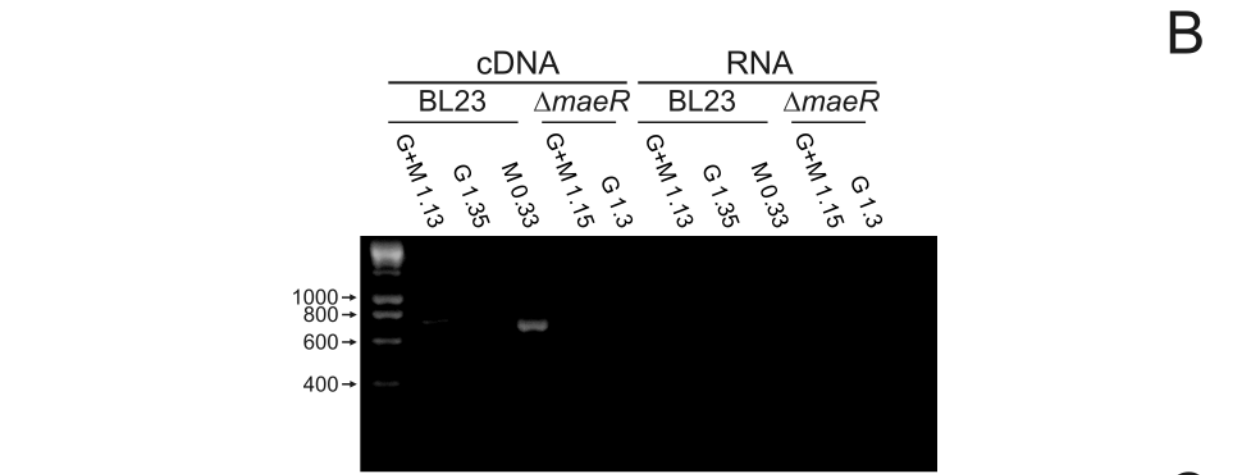
[b] RNA was extracted from *L. casei* BL23 cells grown in MEI medium plus glucose (G), glucose plus L-malic acid (G+M) or L-malic acid (M).

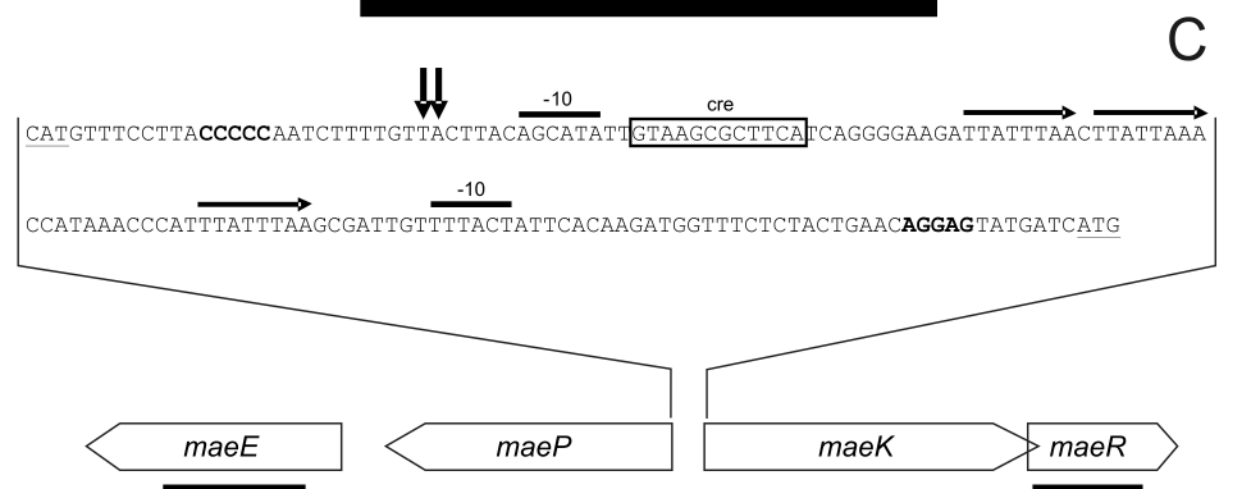

Glucose-1-P
PGM
LCABL_10630
Glucose-6-P
PGI
LCABL_12870
Fructose-6-P
Pi
FBP
LCABL_22100
Fructose-1,6-P2
PEP
ADP
CO2
PEPCK(LCABL_20410)
ATP
AMP
PPi
PPDK
LCABL_25190
Pi
ATP
Oxaloacetate
ADP
Pi
HCO3
ATP
PC (LCABL_15450)
NAD
H+
CO2
NADH+
L-Malate
ME
Pyruvate
Lactate
LDH
NADH+
NAD
MLE
NAD
Mn2+
CO2
CoA
L-Lactate
CoA
NAD
PDH
(LCABL_15360-390)
NADH+
CO2
PFL
(LCABL_16340)
Formate
Pi
O2
POX
(LCABL_20020;
LCABL_23230)
H2O2
CO2
Acetyl-CoA
NADH+
ACDH
(LCABL_08380)
NAD
CoA
Pi
PTA
(LCABL_11620)
CoA
Acetaldehyde
Acetyl-P
NADH+
ADH
(LCABL_23040;
LCABL_24040)
NAD
ADP
ACK
(LCABL_01600;
LCABL_23490)
ATP
Ethanol
Acetate

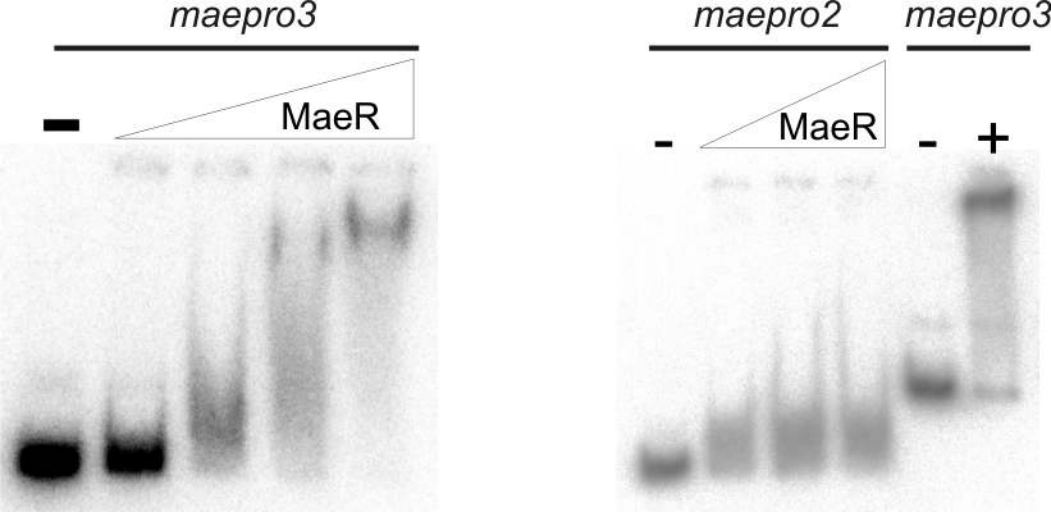

A
LCABL_30730
maeE
maeP
maeK
maeR
maeKdown
maeKup
maeRdown
maeRup
B
FP2
Start
maeP
CGTTGATTTGTTGGGTGTAACCGTCTTGTTCTCCATGTTTCCTTACCCCCAATCTTTTGTTACTTACAGCATA
FP4
TTGTAAGCGCTTCATCAGGGGAAGATTATTTAACTTATTAAACCATAAACCCATTTATTTAAGCGATTGTTTTAC
FP3
TATTCACAAGATGGTTTCTCTACTGAACAGGAGTATGATCATGAAAATAAAAGGATGGCGGCTATGGGGCT
Start
maeK
FP1
C
3026378
3026441
TAAACATCATTCAAATCTTTAAAAAGGCAGTGATCTTAAATACTAAGATCACTGCCTTTTTGTT
Stop
maeR
3021322
3021260
TAAAAGCAATATGGAAACCTCCTGACAAGGTTGGCAAGGGTCAACTTAGTGGGGAGGTTTTTT
Stop
maeE
D
GCTCAACCGATCCGCTGACACTAACATTAACCGTGGACTACCACGAGGAGGCTCAAAATGACGAACATCCTAATCGTTGA
BL23
CTATTCACAAGATGGTTTCTCTACTGAACAGGCTGACACTAACATTAACCGTGGACTACCACGAGGAGGCTCAAAATGACGAACATCCTAATCGTTGA
ΔmaeK
RBS
Start
maeR
Stop
maeK
CTATTCACAAGATGGTTTCTCTACTGAACAGGAGTATGATCATG
RBS
Start
maeK
TGCGCTGTCGTAAACATCATTCAAATCTTTAAAAAGGCAG
BL23
GAGGAGGCTCAAAATGACGAACATCCTAATCGTTGATAAACATCATTCAAATCTTTAAAAAGGCAG
ΔmaeR
Stop
maeR
GAGGAGGCTCAAAATGACGAACATCCTAATCGTTGAAGATGATCCA
RBS
Start
maeR
Stop
maeK

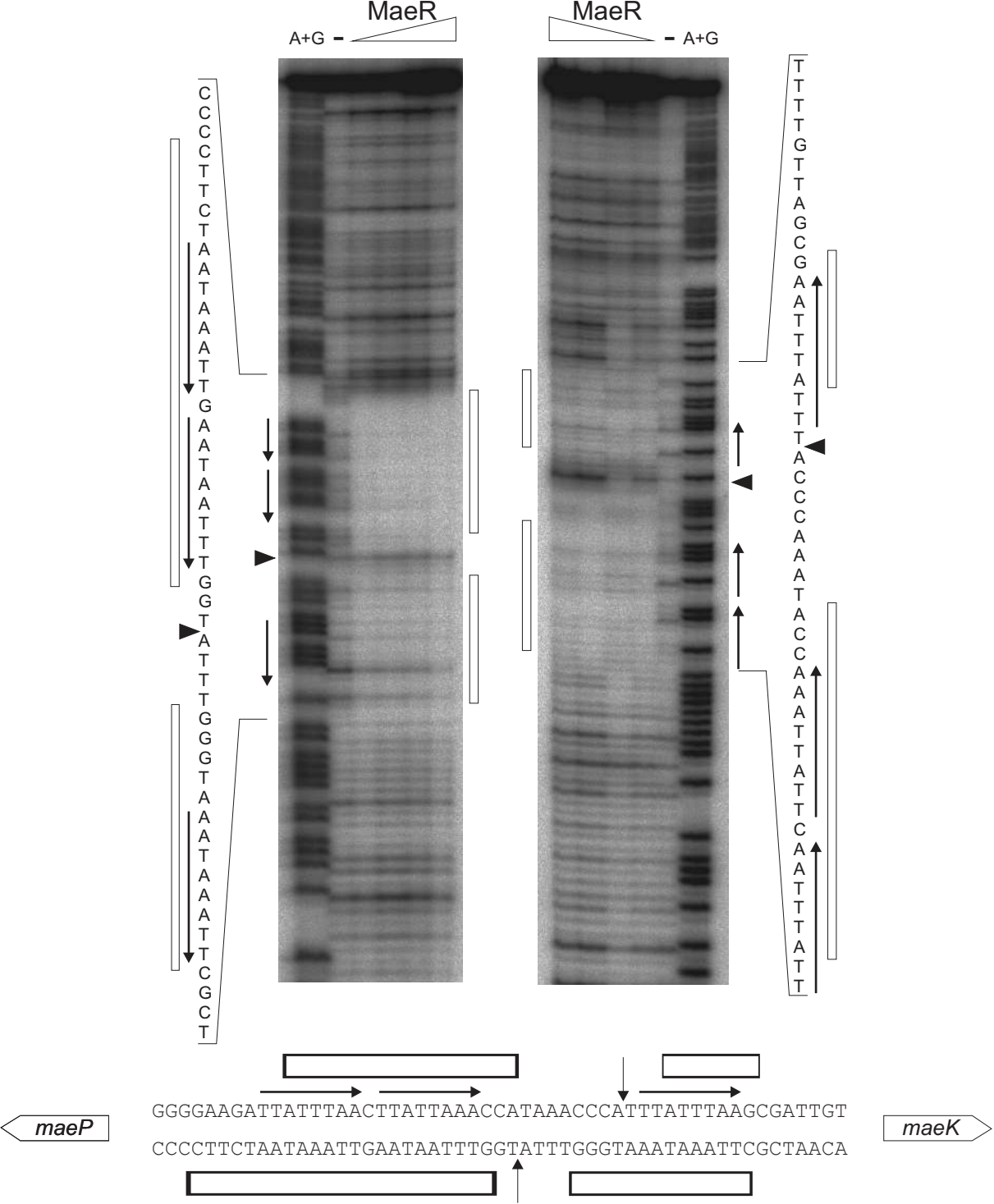

O.D. 595 nm
Time (h)
1,6
1,4
1,2
1,0
0,8
0,6
0,4
0,2
0,0
0
20
40
60
80
100
BL23 MG
maeE⁻ MG
ΔmaeR MG
ΔmaeK MG
BL23 M
maeE⁻ M
ΔmaeR M
ΔmaeK M
pT1NX:maeK

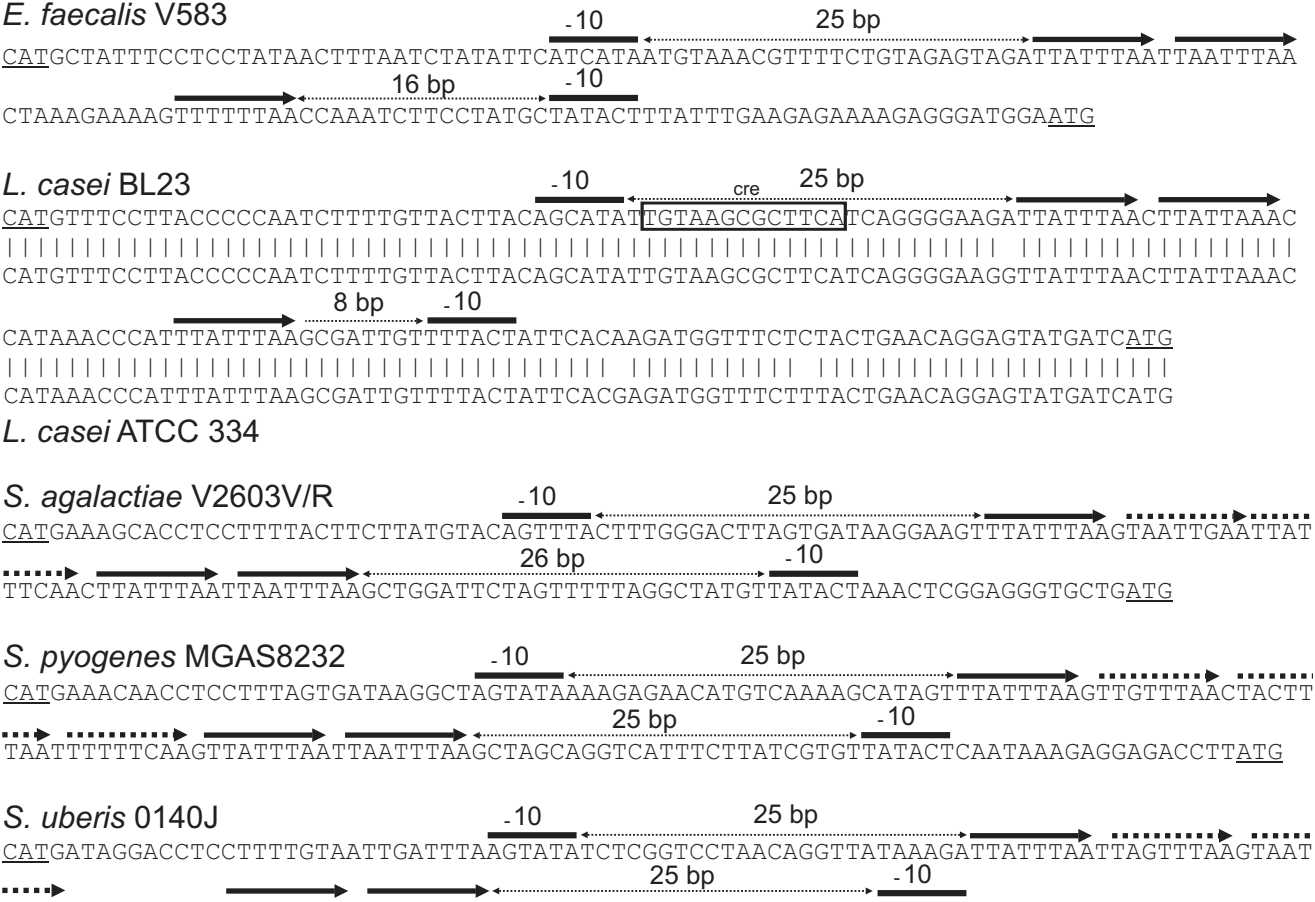

A
BL23
ΔmaeR
G 0.52
G 1.23
G 1.35
G+M 0.54
G+M 1.16
G+M 1.13
M 0.33
M 0.58
G 0.53
G 1.21
G 1.3
G+M 0.54
G+M 1.15
G+M 1.32
M 0.29
BL23 M 0.33
rRNA


B
cDNA
RNA
BL23
ΔmaeR
BL23
ΔmaeR
G+M 1.13
G 1.35
M 0.33
G+M 1.15
G 1.3
G+M 1.13
G 1.35
M 0.33
G+M 1.15
G 1.3
1000
800
600
400


C
-10
cre
CATGTTTCCTTACCCCCAATCTTTTGTTACTTACAGCATATTGTAAGCGCTTCATCAGGGGAAGATTATTTAACTTATTAAA
-10
CCATAAACCCATTTATTTAAGCGATTGTTTTACTATTCACAAGATGGTTTCTCTACTGAACAGGAGTATGATCATG
maeE
maeP
maeK
maeR


maepro3
MaeR
maepro2
maepro3
MaeR


MaeR
A+G
-
MaeR
-
A+G
CCCCTTCTAATAAATTGAATAATTTGGTATTTGGGTAAATAAATTCGCTAACA
TTTTGTTAGCGAATTTATTTACCCAAATACCCAAATTATTCAATTTATT
GGGGAAGATTATTTAACTTATTAAACCATAAACCCATTTATTTAAGCGATTGT
CCCCTTCTAATAAATTGAATAATTTGGTATTTGGGTAAATAAATTCGCTAACA
maeP
maeK


E. faecalis V583
-10
25 bp
CATGCTATTTCCTCCTATAACTTTAATCTATATTCATCATAATGTAAACGTTTTCTGTAGAGTAGATTATTTAATTAATTTAA
16 bp
-10
CTAAAGAAAAGTTTTTTAACCAAATCTTCCTATGCTATACTTTATTTGAAGAGAAAAGAGGGATGGAATG
L. casei BL23
-10
cre
25 bp
CATGTTTCCTTACCCCCAATCTTTTGTTACTTACAGCATATTGTAAGCGCTTCATCAGGGGAAGATTATTTAACTTATTAAAC
CATGTTTCCTTACCCCCAATCTTTTGTTACTTACAGCATATTGTAAGCGCTTCATCAGGGGAAGGTTATTTAACTTATTAAAC
8 bp
-10
CATAAACCCATTTATTTAAGCGATTGTTTTACTATTCACAAGATGGTTTCTCTACTGAACAGGAGTATGATCATG
CATAAACCCATTTATTTAAGCGATTGTTTTACTATTCACGAGATGGTTTCTTTACTGAACAGGAGTATGATCATG
L. casei ATCC 334
S. agalactiae V2603V/R
-10
25 bp
CATGAAAGCACCTCCTTTTACTTCTTATGTACAGTTTACTTTGGGACTTAGTGATAAGGAAGTTTATTTAAGTAATTGAATTAT
26 bp
-10
TTCAACTTATTTAATTAATTTAAGCTGGATTCTAGTTTTTAGGCTATGTTATACTAAACTCGGAGGGTGCTGATG
S. pyogenes MGAS8232
-10
25 bp
CATGAAACAACCTCCTTTAGTGATAAGGCTAGTATAAAAGAGAACATGTCAAAAGCATAGTTTATTTAAGTTGTTTAACTACTT
25 bp
-10
TAATTTTTTCAAGTTATTTAATTAATTTAAGCTAGCAGGTCATTTCTTATCGTGTTATACTCAATAAAGAGGAGACCTTATG
S. uberis 0140J
-10
25 bp
CATGATAGGACCTCCTTTTGTAATTGATTTAAGTATATCTCGGTCCTAACAGGTTATAAAGATTATTTAATTAGTTTAAGTAAT
25 bp
-10
TAAAGATATTCCATTTTTTTAATTAATTTAAGCTAGTTTCGAACCTATGCCTGTGATATACTCAAATAATAGGAGACACTATG