Microbiology (2008), **154**, 81–93 DOI 10.1099/mic.0.2007/011569-0

The complete coenzyme B_{12} biosynthesis gene cluster of *Lactobacillus reuteri* CRL1098

Filipe Santos,[1] Jose L. Vera,[2] René van der Heijden,[3] Graciela Valdez,[2] Willem M. de Vos,[1] Fernando Sesma[2] and Jeroen Hugenholtz[1]

[1]Kluyver Centre for Genomics of Industrial Fermentation, TI Food and Nutrition, and NIZO Food Research, Kernhemseweg 2, PO Box 20, 6710 BA Ede, The Netherlands

[2]Centro de Referencia para Lactobacilos (CERELA-CONICET), Chacabuco 145 (4000), San Miguel de Tucumán, Tucumán, Argentina

[3]Center for Molecular and Biomolecular Informatics, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands

Correspondence
Jeroen Hugenholtz
jeroen.hugenholtz@nizo.nl

Received 12 July 2007
Revised 29 August 2007
Accepted 25 September 2007

The coenzyme B_{12} production pathway in *Lactobacillus reuteri* has been deduced using a combination of genetic, biochemical and bioinformatics approaches. The coenzyme B_{12} gene cluster of *Lb. reuteri* CRL1098 has the unique feature of clustering together the *cbi*, *cob* and *hem* genes. It consists of 29 ORFs encoding the complete enzymic machinery necessary for *de novo* biosynthesis. Transcriptional analysis showed it to be expressed as two tandem transcripts of approximately 22 and 4 kb, carrying *cobD, cbiABCDETFGHJ*, *cobA/hemD*, *cbiKLMNQOP*, *sirA*, *hemACBL*, and *cobUSC*, *hemD, cobT,* respectively. Both transcripts appear to be similarly regulated, and under the conditions assayed are induced in the late-exponential growth phase. Evidence for a regulatory mechanism of negative feedback inhibition by vitamin B_{12} itself was observed. Comparative genomics analysis of the coding sequences showed them to be most similar to those coding for the anaerobic coenzyme B_{12} pathways previously characterized in a few representatives of the genera *Listeria* and *Salmonella*. This contrasts with the trusted species phylogeny and suggests horizontal gene transfer of the B_{12} biosynthesis genes. G+C content and codon adaptation index analysis is suggestive that the postulated transfer of these genes was not a recent event. Additional comparative genomics and transcriptional analysis of the sequences acquired during this study suggests a functional link between coenzyme B_{12} biosynthesis and reuterin production, which might be implicated in *Lb. reuteri*'s success in colonizing the gastrointestinal tract. This information on gene organization, gene transcription and gene acquisition is relevant for the development of (fermented) foods and probiotics enriched in B_{12}.

INTRODUCTION

Lactobacillus reuteri is a Gram-positive, heterofermentative lactic acid bacterium, widespread throughout the gastrointestinal tract (GI tract) of humans and other animals (Walter *et al.*, 2003). Although this bacterium is currently marketed as a probiotic, human intervention studies showing relevant benefits remain to be reported (Saxelin *et al.*, 2005). Nonetheless, potential probiotic effects have been demonstrated; they include lowering blood cholesterol levels in mice (Taranto *et al.*, 2000), and stimulating anti-inflammatory activity in human cell lines (Ma *et al.*, 2004).

Lb. reuteri possesses the unique ability to produce and excrete reuterin (Talarico *et al.*, 1988). This broad-spectrum antimicrobial compound is a mixture of monomeric, hydrated monomeric and cyclic dimeric forms of 3-hydroxypropionaldehyde (3-HPA) (Talarico & Dobrogosz, 1989). The synthesis of reuterin is mediated by glycerol dehydratase (EC 4.2.1.30), a vitamin B_{12}-dependent enzyme, which is involved in catalysing the conversion of glycerol to 3-HPA (Daniel *et al.*, 1998).

We have reported the isolation of a compound from *Lb. reuteri* CRL1098 capable of fulfilling the auxotrophic vitamin B_{12} requirements of three different indicator strains (Taranto *et al.*, 2003). In the same study, *Lb. reuteri* genomic DNA was found to contain sequences homologous to genes involved in the anaerobic coenzyme B_{12}

Abbreviations: 3-HPA, 3-hydroxypropionaldehyde; RT-PCR, reverse transcriptase PCR; Q-RT-PCR, quantitative reverse transcriptase PCR; GI tract, gastrointestinal tract; GRAS, generally regarded as safe.

The GenBank accession number for the sequence reported in this paper is AY780645.

Two supplementary figures are available with the online version of this paper.

biosynthesis pathway, including *cysG/hemD* from *Selenomonas ruminantium* (Anderson *et al.*, 2001), *Listeria innocua* and *Listeria monocytogenes* (Glaser *et al.*, 2001), and *cbiK* and *cbiJ* from *Salmonella typhimurium* (Raux *et al.*, 1996).

Vitamin B_{12} consists of a tetrapyrrolic-derived corrin ring with a cobalt ion chelated at the core. Along with chlorophyll, coenzyme F_{430} and haem, amongst others, it constitutes one of the most structurally complex classes of cofactors. Various B_{12} derivatives with different upper axial ligands act as essential cofactors in many important enzymic reactions responsible for the catalysis of methyl transfers and carbon backbone rearrangements (Maggio-Hall & Escalante-Semerena, 1999). Coenzyme B_{12} biosynthesis is limited to a few representatives of bacteria and archaea (Martens *et al.*, 2002). It appears that B_{12}-dependent enzymes are absent from plants and fungi, but widespread in prokaryotes, protists and animals (Croft *et al.*, 2005; Rodionov *et al.*, 2003).

In humans, vitamin B_{12} deficiency leads to pernicious anaemia and neurological dysfunction, amongst other complications (Stabler, 1999). Three proteins are known to participate in the uptake and transport of vitamin B_{12}, namely haptocorrin, intrinsic factor and transcobalamin II. Absorption of vitamin B_{12} occurs by receptor-mediated endocytosis in the terminal ileum, where the specific receptor cubulin complexes with intrinsic factor bound to B_{12} (Banerjee, 2006). As a consequence, B_{12} produced by colonic bacteria is most likely inaccessible to the host. However, it has been suggested that B_{12} produced by a micro-organism capable of colonizing proximal to the ileum, such as *Lb. reuteri*, would potentially be host-accessible (Albert *et al.*, 1980).

Lb. reuteri was the first lactic acid bacterium reported to be able to produce B_{12} (Taranto *et al.*, 2003). Increasing our understanding of how this GRAS (generally regarded as safe) organism encodes, acquired and maintains a biosynthetic pathway of such complexity and magnitude is of great importance to the medical field and for the food and feed industries.

In this study, we extend the analysis of the presumed coenzyme B_{12} biosynthesis gene cluster of *Lb. reuteri* and describe the presence of a complete gene cluster encoding all the enzymic machinery necessary for the *de novo* synthesis of this important cofactor. Additional comparative genomics and transcriptional analysis of the new sequences acquired during this study suggests a functional link between coenzyme B_{12} biosynthesis and reuterin production, which might be implicated in *Lb. reuteri*'s success in colonizing the GI tract.

METHODS

Strains, media and culture conditions. *Lb. reuteri* CRL1098, isolated from sourdough, was obtained from the CERELA stock culture collection. It was cultivated at 37 °C in MRS medium and in vitamin B_{12} assay medium (Sigma) supplemented when mentioned with 1 mg l^{-1} of cyanocobalamin (Sigma-Aldrich). *Escherichia coli* strain XL-1 Blue MRA (P2) was obtained from Stratagene and cultivated at 37 °C under aerobic conditions in TY medium. *Salmonella enterica* serovar Typhimurium LT2 derivative strains TT25720 (*metE2119*::MudJ) and TT25722 [*metE2119*::MudJ, *cobS2621*::Frt(sw)] were kindly provided by Professor John R. Roth (Section of Microbiology, University of California at Davis, USA) and cultivated at 37 °C in TY medium or minimal E medium (Maloy *et al.*, 1996), supplemented with 100 nM cyanocobalamin when required.

Nucleotide sequence analysis. The sequence of the B_{12} biosynthesis gene cluster of *Lb. reuteri* was obtained by screening two genomic λ-phage libraries, and finalized by both inverted PCR and genomic primer walking. Total genomic DNA was isolated from *Lb. reuteri* according to standard molecular biology techniques (Sambrook & Russell, 2001).

A Southern blot analysis of a partial digestion of *Lb. reuteri*'s chromosomal DNA with the restriction enzymes *Bam*HI and *Bgl*II, using a *cysG/hemD* (Taranto *et al.*, 2003) probe amplified from the same strain, showed that the signals obtained corresponded to DNA fragments larger than 15 kb for both restriction enzymes (data not shown). Based on this knowledge, two *Lb. reuteri* genomic λ-phage libraries were constructed by the separate ligation of *Bgl*II- and *Bam*HI-digested *Lb. reuteri* genomic DNA into Lambda-DASH II/*Bam*HI vector and packaged with a Gigapack III Gold packaging extract (Stratagene) according to the manufacturer's recommendations. For the amplification of the Lambda-DASH II/*Bam*HI libraries a lysogenic P2 strain, *E. coli* XL-1 Blue MRA (P2), was used. Titre determination of bacteriophages, blotting of plaques on nylon membranes and λ-DNA isolation were all performed according to the manufacturer's recommendations. Probes purified through the JETPURE PCR Product Purification kit (GENOMED), were amplified, radioactively labelled with [α^{32}P]ATP (GE Healthcare Europe), and hybridized on membranes according to standard procedures (Sambrook & Russell, 2001). Membranes were exposed to BioMax MS or BioMax MR X-ray film (Kodak) for at least 5 h at −80 °C before developing. Sequencing of two ~15 kb non-overlapping inserts containing B_{12}-related DNA was carried out at Greenomics (Wageningen, The Netherlands).

For gap-closure between the two inserts, we resorted to inverted PCR. By standard procedures (Sambrook & Russell, 2001), *Hin*dIII-digested genomic DNA of *Lb. reuteri* was ligated to pNZ8048 (Sybesma *et al.*, 2003) digested with the same endonuclease. The ligation mixture was directly used as a template in a PCR using a primer designed on the vector and another based on the 5′ flanking region of the known sequence at the time. The resulting amplicon was isolated from an agarose gel and sequenced directly at Baseclear, The Netherlands.

Further sequencing efforts aimed at closing gaps and extending the flanking regions of the known sequence were done by genomic primer walking carried out at GATC Biotech, Germany.

The new sequence information obtained using the three different approaches described above was analysed and assembled resorting to in-house scripts, and online programs available from the Biology WorkBench of the San Diego Supercomputer Centre (http://workbench.sdsc.edu/). Standard RNA regulatory motif searches were performed in Rfam (Griffiths-Jones *et al.*, 2003) and using Riboswitch finder (Bengert & Dandekar, 2004). Predicted ORFs were manually annotated based on homology searches using the BLAST algorithm (Altschul *et al.*, 1997). All sequence information was deposited at GenBank under accession no. AY780645.

Complementation studies. A fragment containing *cobS* was amplified from *Lb. reuteri*'s genomic DNA using Herculase II DNA

polymerase (Stratagene), and primers LREf28196_28215 and LREr29724_29704 (Tables 1 and 2). Additionally, a fragment containing the native *cob* operon promoter (O'Toole *et al.*, 1993) was amplified from *Salmonella enterica* strain TT25720 using primers 5′-GACACCATTGTGGATGAGGTGGAG-3′ and 5′-GATGATCG-ATCATACCGGCTCCTGATGT-3′ (*Cla*I cleavage site underlined). 3′-A overhangs were added to both fragments by incubating the PCR reactions directly with 1 unit of *Taq* DNA polymerase for 3 min at 72 °C. The A-tailed fragments were then purified with the JETPURE PCR Product Purification kit (GENOMED) and digested with *Cla*I. The modified fragments were again purified by the same method and simultaneously cloned in pGEM-T Easy Vector (Promega Benelux), resulting in pNZ7749. The *Salmonella enterica* strain TT25722 was transformed with this vector as previously described (Sambrook & Russell, 2001), and its phenotype was characterized in minimal E medium (Maloy *et al.*, 1996).

Transcriptional analysis. The transcriptional organization of the vitamin B_{12} gene cluster of *Lb. reuteri* was determined by Northern blots, reverse transcriptase PCR (RT-PCR) and quantitative RT-PCR (Q-RT-PCR). Cells were cultured in batch fermentations of MRS medium, vitamin B_{12} assay medium (commercial rich broth lacking B_{12}) and vitamin B_{12} assay medium supplemented with vitamin B_{12} to a final concentration of 1 mg l^{-1}. RNA was isolated according to standard procedures (Sambrook & Russell, 2001) from samples collected at the mid-exponential, late-exponential and stationary phases. The integrity and concentration of the RNA were analysed with a 2100 Bioanalyser (Agilent Technologies). Northern blotting of RNA obtained from late-exponential-phase cells cultivated in MRS was performed as previously described (Kuipers *et al.*, 1993; Roest *et al.*, 2005). Probes were amplified from genomic DNA of *Lb. reuteri* by PCR using primer pairs designed to locate them throughout the cluster, namely on *cbiC*, *cbiP* and *cobT*. Subsequent hybridization with radiolabelled probes was carried out according to standard techniques (Sambrook & Russell, 2001). RT-PCR analysis of samples obtained from cells cultured in MRS was performed by systematically amplifying overlapping fragments throughout the full extent of the B_{12} biosynthesis cluster and flanking regions. All RNA samples were diluted to the same concentration and an extra DNase I (Invitrogen) treatment was implemented to eliminate possible remaining chromosomal DNA contamination. First-strand cDNA synthesis was carried out using Superscript III reverse transcriptase from Invitrogen according to the manufacturer's recommendations. Primers were manually designed and are listed in Table 1. To quantify the differential expression of the two operons within the B_{12} biosynthesis gene cluster between late- and mid-exponential phases and in the presence or absence of B_{12}, we performed Q-RT-PCR. Amplification was carried out in 96-well plates in an ABI Prism 7700 (Applied Biosystems) using the fluorescent agent SYBR Green for detection. Reactions were set up using the SYBR Green Master Mix from the same manufacturer, following its recommendations. Specificity and product detection were checked after amplification by determining the temperature-dependent melting curves. Primers were designed with the Primer Express software package (Applied Biosystems) to have a T_m between 59 and 61 °C and an amplicon size of 100 ± 20 bp (Table 2). Comparisons were made between the different growth phases and the different culture media.

Phylogenetic analysis. Each individual B_{12}-related amino acid sequence reported in this study was entered as a string to search for distantly related homologues using the PSI-BLAST algorithm (Altschul *et al.*, 1997). Sequence entries identified as homologues were retrieved in March 2007 from ERGO (http://ergo.integratedgenomics.com/ERGO/) (Overbeek *et al.*, 2003), and separately aligned using the MUSCLE algorithm (Edgar, 2004). From the sequence alignment of the proteins encoded by the coenzyme B_{12} biosynthesis cluster, a neighbour-joining tree was obtained using CLUSTAL W (Thompson *et al.*, 1994), analysed in LOFT (van der Heijden *et al.*, 2007), and visualized in MEGA3 (Kumar *et al.*, 2004). An identical exercise was carried out for the predicted product of the *rpsO* gene that is located downstream of the vitamin B_{12} gene cluster of *Lb. reuteri*, and for the 16S RNA gene. Finally, the topology of all trees was compared.

G+C content and codon adaptation index. G+C content and codon adaptation index (Sharp & Li, 1987) was calculated using the geecee, cusp and cai scripts, part of EMBOSS: European Molecular Biology Open Software Suite (Rice *et al.*, 2000). Comparisons were made between the coenzyme B_{12} biosynthesis gene cluster of *Lb. reuteri* presented here and the draft genome sequence of *Lb. reuteri* JCM1112 obtained by the DOE Joint Genome Institute and deposited at GenBank under accession no. CP000705. A similar exercise was performed for *Listeria innocua* Clip11262 (Glaser *et al.*, 2001) and *Salmonella enterica* Typhi Ty2 (Deng *et al.*, 2003), for which we compared the G+C content and codon usage of their vitamin B_{12} clusters to their published genomes.

RESULTS

Operon organization

A sequence of approximately 43.4 kb was assembled from the *Lb. reuteri* genome through the combined effort of the different molecular biology techniques, and was found to harbour a coenzyme B_{12} gene cluster encoding the complete enzymic machinery necessary for its biosynthesis. An overview of the organization of this gene cluster (Fig. 1) reveals that all predicted genes are in the same orientation, with only a few intergenic regions. Similar to what has been reported for *Salmonella typhimurium* (Roth *et al.*, 1993), we observed that approximately half of the genes (46 %) are overlapping and predicted to be translationally coupled.

The previously published sequence encoding the fusion protein homologous to CysG/HemD (Taranto *et al.*, 2003) is flanked by the large cluster of 17 *cbi* genes (Fig. 1). The *cbi* gene order is conserved amongst different B_{12} producers, notably representatives of *Listeria* and *Salmonella* (see Fig. 1). Quite unexpectedly, the *hem* genes are located directly downstream of the *cbi* genes. To our knowledge this genomic organization has not been described previously. These genes are predicted to encode uroporphyrinogen III synthesis from 5-aminolaevulinate, a derivative of glutamyl-tRNA. A cluster of five *cob* genes is located further downstream. This cluster is predicted to be involved in the attachment of the amino-propanol arm and assembly of the nucleotide loop, which connects the lower cobalt ligand to the corrin ring. Upstream of the B_{12} biosynthesis gene cluster are several genes predicted to be involved in the formation of polyhedral bodies, including *pduU* and *pduV* (Bobik *et al.*, 1999).

Detailed comparison of the predicted coding sequences of *Lb. reuteri* CRL1098 and the draft genome sequence of *Lb. reuteri* JCM1112, recently released by the DOE Joint Genome Institute, demonstrates that they are mostly identical (Table 3). The few exceptions are due to minor changes in the N-terminus (CbiA and CbiB), or in the C-

Table 1. Oligonucleotide primers used in RT-PCR reactions

Primer	Nucleotide sequence (5′–3′)	Strand*	Location
LREf5899_5921	GCACCGTCGCAACAATATCCCAC	+	IS
LREf6317_6340	CGTTTTCTTTGATTTTAGTAGGTG	+	*cobD*
LREr7366_7343	CTGCCACTCGATAGTATTGTCGGC	−	*cobD*
LREf7343_7366	GCCGACAATACTATCGAGTGGCAG	+	*cobD*
LREr8862_8843	CACGAATGAGGGTCACCAAG	−	*cbiB*
LREf8843_8862	CTTGGTGACCCTCATTCGTG	+	*cbiB*
LREr10437_10416	GGTGTGACGGCCATACTCATCA	−	*cbiC* and *cbiD*
LREf10416_10437	TGATGAGTATGGCCGTCACACC	+	*cbiC* and *cbiD*
LREr11947_11927	GTTCGCCCATGACTACTTGTC	−	*cbiE*
LREf11927_11947	GACAAGTAGTCATGGGCGAAC	+	*cbiE*
LREr13430_13411	CACCTAAGAACTTACCAACC	−	*cbiF*
LREf13411_13430	GGTTGGTAAGTTCTTAGGTG	+	*cbiF*
LREr14920_14900	GAGCAGCAGCTGCAATACTTG	−	*cbiH*
LREf14900_14920	CAAGTATTGCAGCTGCTGCTC	+	*cbiH*
LREr16411_16391	CTAGCCCAGCAATTGCACTAG	−	*cysG/hemD*
LREf16391_16411	CTAGTGCAATTGCTGGGCTAG	+	*cysG/hemD*
LREr17886_17864	GTAAAAGCACTATGCGCTGTTCC	−	*cbiK*
LREf17864_17886	GGAACAGCGCATAGTGCTTTTAC	+	*cbiK*
LREr19409_19388	CGACTAACTTTCATTGCTCGAC	−	*cbiM*
LREf19388_19409	GTCGAGCAATGAAAGTTAGTCG	+	*cbiM*
LREr20906_20887	CGCCGTAAATTTCGAAGTCC	−	*cbiO*
LREf20887_20906	GGACTTCGAAATTTACGGCG	+	*cbiO*
LREr22471_22451	GCTACTAAATCTGCGTTCGTG	−	*cbiP*
LREf22451_22471	CACGAACGCAGATTTAGTAGC	+	*cbiP*
LREr24150_24130	CCTTGCTAAAGCCCATATTGC	−	*hemA*
LREf24130_24150	GCAATATGGGCTTTAGCAAGG	+	*hemA*
LREr25669_25648	CCTTAGCCAATAACTGATCAGC	−	*hemC*
LREf25648_25669	GCTGATCAGTTATTGGCTAAGG	+	*hemC*
LREr26711_26683	GGTGGGTTTGTTTTGAGTAAATTAGATAC	−	Intergenic region
LREf26684_26718	GTGGGTTTGTTTTGAGTAAATTAGATACAACAAAG	+	Intergenic region
LREr27526_27499	CTGGGAGTCCACCACCGATTACTTTGCC	−	*hemL*
LREf27499_27518	GGCAAAGTAATCGGTGGTGG	+	*hemL*
LREr27987_27967	CTTGGTTGCCGCATTAAATGC	−	*hemL*
LREr28269_28250†	GTTCATCGACGTGCTGATAC	−	*cobU*
LREf28250_28269	GTATCAGCACGTCGATGAAC	+	*cobU*
LREr29724_29704‡	GGTATAGGTTAATGGAGCTGC	−	*cobC*
LREf29704_29724	GCAGCTCCATTAACCTATACC	+	*cobC*
LREr31148_31129	CTGCTATCGACATTGCTGGT	−	*cobT*
LREf31129_31148	ACCAGCAATGTCGATAGCAG	+	*cobT*
LREr31756_31737	GAAGTCCATCTCCTGCAATG	−	*cobT*
LREr32212_32193	CCTTGTGGCAACAGTCTTCT	−	Hypothetical
LREf32193_32212	AGAAGACTGTTGCCACAAGG	+	Hypothetical
LREf33354_33373	GGAATTCGCAACTCACGAAG	+	*rpsO*
LREr33545_33526	GCAGGTAAATCAGTCCGACG	−	*rpsO*
LREr33809_33790	TCGCGTACACCACCAAAAGG	−	Metallolactamase
LREf33790_33809	CCTTTTGGTGGTGTACGCGA	+	Metallolactamase
LREr35516_35496	CCACGACCACGATGATGTTCT	−	Metallolactamase
LREf35496_35516	AGAACATCATCGTGGTCGTGG	+	Metallolactamase
LREr36538_36518	CGCAATCAAAGCAGTTGAACG	−	Hypothetical

*Primers were designed to the coding strand (+) or non-coding strand (−).
†Primer also used in Q-RT-PCR experiment.
‡Primer also used in complementation studies.

Table 2. Oligonucleotide primers used in Q-RT-PCR reactions

Primer	Nucleotide sequence (5′–3′)	Strand*	Location
LREf65_84	CAATAACGCCAAGTGAAGCC	+	*pduU*
LREr211_192	CCACATGACGCAAAGCTGAT	−	*pduU*
LREf2543_2562	ATTCAATGTCGGCAGGGTCT	+	*pduV*
LREr2628_2609	GGCTGGCTTCTGTTCAATGT	−	*pduV*
LREf7315_7334	CGCCAATGTGATGATTACGC	+	*cobD*
LREr7431_7412	CAGCTCACGTCGTAACACTT	−	*cobD*
LREf8403_8422	GCAGAGTGTGGTGGCTTAAT	+	*cbiA*
LREr8505_8486	GCGGTGTCATCTCACTCATA	−	*cbiA*
LREf9418_9437	TCCAGCACGAATCACATGGT	+	*cbiB*
LREr9561_9542	CCTGCAACAACAGCTTCACT	−	*cbiB*
LREf22014_22033	GCTGATGCACCAGTAATCCT	+	*cbiP*
LREr22119_22100	TAATGCGTTGCTGGTCTTCG	−	*cbiP*
LREf26313_26332	CGTGATGCTGCTGATGGTTC	+	*hemB*
LREr26410_26391	GCTACTTCGCGCAATGCTTC	−	*hemB*
LREf27503_27522	AAGTAATCGGTGGTGGACTC	+	*hemL*
LREr27605_27586	GACAACGTTCCGGCATGATA	−	*hemL*
LREf28196_28215†	TCGAATTCAGCGTCACCAAG	+	*cobU*
LREf31622_31641	GCTCTCGGTCTTGATCCTTA	+	*cobT*
LREr31723_31704	GAGCATTGCCTTCACTCCAT	−	*cobT*
LREf33369_33388	CGAAGGAGACACTGGTTCTA	+	*rpsO*
LREr33511_33492	AGTTACGACGGTGACCAATC	−	*rpsO*

*Primer were designed to the coding strand (+) or non-coding strand (−).
†Primer also used in complementation studies.

terminus (CbiD and CobD), or due to the neutral replacement of residues with the same chemical properties (CbiC and CobU).

Complementation studies

To experimentally support our functional annotation of the newly sequenced coenzyme B_{12} biosynthesis gene cluster of *Lb. reuteri*, we performed complementation studies in *Salmonella enterica* mutants TT25720 (*metE2119*::MudJ) and TT25722 [*metE2119*::MudJ, *cobS2621*::Frt(sw)] (see Methods). When cultured in minimal medium lacking methionine, both strains are dependent on the B_{12}-dependent methionine synthase (MetH), since they lack MetE activity. However, due to the additional *cobS* mutation, strain TT25722 has auxotrophic requirements for B_{12}, while strain TT25720 can rely on its own native production of this cofactor.

We transformed the double mutant TT25722 with pNZ7749, harbouring a fragment containing *cobS* amplified from *Lb. reuteri* under control of the native *cob* operon promoter from *Salmonella enterica* (O'Toole *et al.*, 1993). Growth experiments were then performed on minimal E plates using strains TT25720 and TT25722 as a positive and negative control, respectively. Complementation of the double mutant with *cobS* from *Lb. reuteri* reconstituted its ability to grow in minimal medium lacking methionine without the exogenous supplementation of vitamin B_{12}, and therefore relying solely on its own native coenzyme B_{12} production (Fig. 2).

Transcription analysis

In order to determine the transcriptional organization of the B_{12} biosynthesis gene cluster we performed Northern blot analysis (see Supplementary Fig. S1, available with the online version of this paper). As a consequence of the relative rarity of the transcripts encoding B_{12} biosynthesis enzymes and their remarkably large size, we could predict that technical difficulties with the Northern blots would not allow conclusive determination of the exact size of the different operons within this gene cluster. Nonetheless, probes were designed to be complementary to sequences from the beginning and end of the predicted operons and their use in Northern hybridizations revealed the presence of two transcripts, one with a size over 20 kb and another of 4 kb (Fig. S1).

To further characterize the transcriptional organization of this gene cluster, a RT-PCR based strategy was implemented. It consisted of systematically amplifying overlapping RT-PCR fragments throughout the full extent of the cluster and flanking regions. To validate the specificity of the designed primer pairs, all reactions were tested in parallel using chromosomal DNA of *Lb. reuteri* as a positive control. The absence of any chromosomal DNA contamination was established by carrying out all reactions

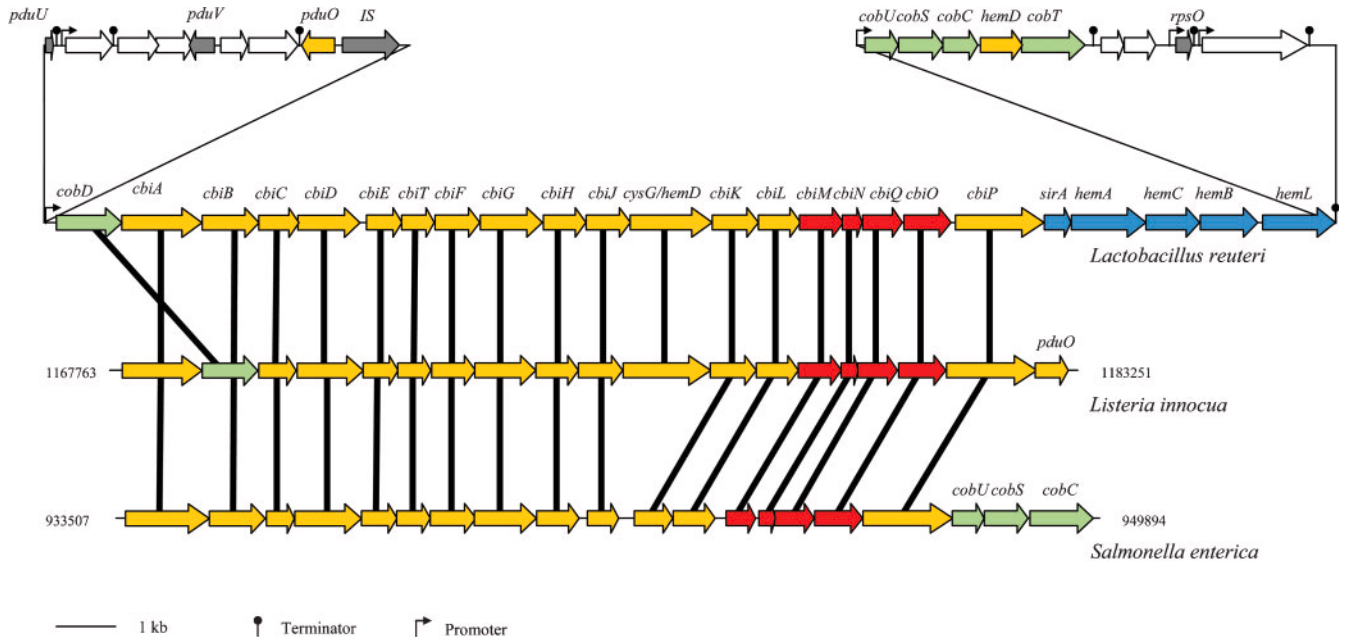

Fig. 1. Schematic representation of the vitamin B_{12} gene cluster of *Lb. reuteri*, and comparison of gene order with *Listeria innocua* and *Salmonella enterica*. The arrows represent genes that are involved in the synthesis of uroporphyrinogen III if depicted in blue; involved in the synthesis of adenosylcobinamide if depicted in orange; involved in the synthesis of the lower ligand if depicted in green; involved in cobalt transport if depicted in red; not related to B_{12} biosynthesis if depicted in grey; and not studied here if depicted in white. Functional annotation: *pduO*, ATP : Co(I)rrinoid adenosyltransferase (EC 2.5.1.17); *cobD*, threonine-phosphate decarboxylase (EC 4.1.1.81); *cbiA*, cobyrinic acid a,c-diamide synthase (EC 6.3.1.–); *cbiB*, adenosylcobinamide-phosphate synthase (EC 6.3.1.10); *cbiC*, precorrin-8X methylmutase (EC 5.4.1.2); *cbiD*, precorrin-5B C1-methyltransferase (EC 2.1.1.–); *cbiE*, precorrin-6Y C5,15-methyltransferase [decarboxylating] subunit CbiE (EC 2.1.1.132); *cbiT*, precorrin-6Y C5,15-methyltransferase [decarboxylating] subunit CbiT (EC 2.1.1.132); *cbiF*, precorrin-4 C^{11}-methyltransferase (EC 2.1.1.133); *cbiG*, precorrin-5A C^{20}-acyltransferase (EC 2.3.1.–); *cbiH*, precorrin-3B C^{17}-methyltransferase (EC 2.1.1.131); *cbiJ*, precorrin-6X reductase (EC 1.3.1.54); *cysG/hemD*, uroporphyrin-III C-methyltransferase (EC 2.1.1.107)/uroporphyrinogen-III synthase (EC 4.2.1.75); *cbiK*, sirohydrochlorin cobaltochelatase (EC 4.99.1.3); *cbiL*, precorrin-2 C^{20}-methyltransferase (EC 2.1.1.130); *cbiM*, cobalt transport protein; *cbiN*, cobalt transport protein; *cbiQ*, cobalt transport protein; *cbiO*, cobalt transport ATP-binding protein; *cbiP*, adenosylcobyric acid synthase (EC 6.3.5.10); *sirA*, precorrin-2 dehydrogenase (EC 1.3.1.76); *hemA*, glutamyl-tRNA reductase (EC 1.2.1.–); *hemC*, porphobilinogen deaminase (EC 2.5.1.61); *hemB*, δ-aminolaevulinic acid dehydratase (EC 4.2.1.24); *hemL*, glutamate-1-semialdehyde 2,1-aminomutase (EC 5.4.3.8); *cobU*, adenosylcobinamide kinase (EC 2.7.1.156)/adenosylcobinamide-phosphate guanylyltransferase (EC 2.7.7.62); *cobS*, adenosylcobinamide-GDP ribazoletransferase (EC 2.7.8.26); *cobC*, α-ribazole-5′-phosphate phosphatase (EC 3.1.3.73); *hemD*, uroporphyrinogen-III synthase (EC 4.2.1.75); *cobT*, nicotinate-nucleotide–dimethylbenzimidazole phosphoribosyltransferase (EC 2.4.2.21).

using RT-negative samples as a template, for a negative control. The results from the RT-PCR experiments (Table 4, Fig. 3) confirmed that the B_{12} biosynthesis gene cluster is expressed in two separate, but tandem, operons of approximately 22 and 4 kb. The large transcript includes the genes *cobD*, *cbiABCDETFGHJ*, *cobA/hemD*, *cbiKLMNQOP*, *sirA* and *hemACBL*. The 4 kb transcript derives from the *cobUSC*, *hemD* and *cobT* genes.

The intensities of the RT-PCR amplicons were compared between samples collected from the same MRS culture at different time points (see Fig. 3 for illustration). This suggested that for cells cultured in MRS the expression of the B_{12} gene cluster is strengthened during late-exponential phase in comparison to mid-exponential phase. The *rpsO* gene, located immediately downstream from the B_{12} gene cluster, served as a control for the transcriptional analysis. RT-PCR samples of this gene collected from the same culture at different time points showed that it is expressed constitutively throughout the growth curve, in contrast to the neighbouring B_{12} genes.

To quantify the differential expression first evidenced by the RT-PCR experiments, Q-RT-PCR was carried out on different loci throughout the entire cluster, using a locus on the *rpsO* gene as a reference. The results are in accordance with the previous RT-PCR-based observation, and confirm that for cells cultivated in MRS the cluster is indeed strongly induced during late-exponential growth (Fig. 4a). The operon carrying the *cbi* and *hem* genes is

Table 3. ORFs of the coenzyme B_{12} biosynthesis gene cluster of *Lb. reuteri* CRL1098: comparison on amino acid level to *Lb. reuteri* JCM1112, *Listeria monocytogenes* and *Salmonella typhimurium*

See legend of Fig. 1 for functional annotation of the genes.

Lb. reuteri CRL1098		*Lb. reuteri* JCM1112		*Listeria monocytogenes*		*Salmonella typhimurium*	
Name	**Length (aa)**	**Length (aa)**	**Identity (%)**	**Length (aa)**	**Identity (%)**	**Length (aa)**	**Identity (%)**
cobD	369	362	99	361	41	364	34
cbiA	454	454	99	452	48	459	44
cbiB	319	319	99	315	52	319	44
cbiC	227	227	99	210	56	210	55
cbiD	349	383	97	373	51	379	46
cbiE	200	200	100	198	51	201	38
cbiT	184	184	100	189	49	192	42
cbiF	253	253	100	249	72	257	65
cbiG	351	351	100	343	41	351	32
cbiH	241	241	100	241	61	241	58
cbiJ	252	252	100	250	39	263	29
cysG/hemD	464	464	100	493	38	457	46
cbiK	259	259	100	261	46	264	45
cbiL	237	237	100	236	46	237	34
cbiM	248	248	100	244	61	245	54
cbiN	110	103	100	98	58	93	53
cbiQ	225	225	100	225	37	225	34
cbiO	267	267	100	268	49	271	48
cbiP	501	501	100	511	56	506	52
sirA	152	152	100	159	37	311	37
hemA	421	421	100	435	40	418	25
hemC	305	305	100	309	48	318	39
hemB	323	323	100	324	64	324	53
hemL	412	431	100	429	58	426	52
cobU	186	196	94	185	40	181	39
cobS	253	253	100	248	36	247	32
cobC	196	271	100	191	32	202	28
hemD	236	236	100	493	26	246	23
cobT	356	350	100	–	–	356	41

upregulated 4.56 ± 0.92-fold during late-exponential growth when compared to mid-exponential phase. Similarly, the smaller operon carrying the *cob* and *hemD* genes is upregulated by a factor of 5.03 ± 0.32 between the late- and mid-exponential phases. The same approach was performed on two loci upstream of the B_{12} biosynthesis gene cluster predicted to encode PduU and PduV. We observed an average upregulation of 5.91 ± 3.25 for these transcripts, similar to that observed for the B_{12} gene cluster (Fig. 4b).

In order to confirm that the observed upregulation of the B_{12} biosynthetic genes during the late-exponential phase is not caused by the exhaustion of the vitamin B_{12} present in MRS, further experimentation was carried out. We analysed by Q-RT-PCR samples obtained from cells grown in B_{12} assay medium, which is B_{12} free, and in B_{12} assay medium supplemented with cyanocobalamin (Fig. 4b). In the absence of B_{12} we observed that during the late-exponential phase the *cbi* and *hem* operon was upregulated 6.89 ± 0.93, slightly more than what was observed for MRS.

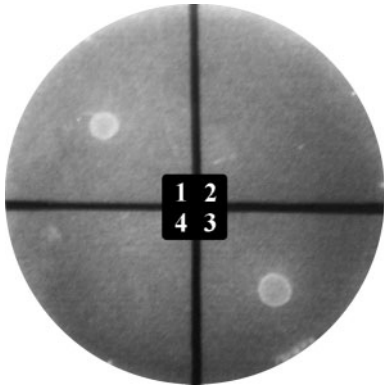

Fig. 2. Phenotypic characterization of complemented *Salmonella* mutant. Minimal E agarose plate with (1) *Salmonella enterica* TT25720 (*metE2119*::MudJ); (2) *Salmonella enterica* TT25722 [*metE2119*::MudJ, *cobS2621*::Frt(sw)]; (3) *Salmonella enterica* TT25722 pNZ7749 (harbouring *cobS* from *Lb. reuteri*); (4) empty.

Table 4. Summary of RT-PCR reactions

Primer pair		Product*		Size (bp)	Location
		Mid-exp.	Late exp.		
LREf5899_5921	LREr7366_7343	−	−	1467	IS, *cobD*
LREf6317_6340	LREr7366_7343	+	+ + +	1049	*cobD*
LREf7343_7366	LREr8862_8843	+	+ + +	1519	*cobD*, *cbiAB*
LREf8843_8862	LREr10437_10416	+	+ + +	1594	*cbiBCD*
LREf10416_10437	LREr11947_11927	+	+ + +	1531	*cbiDE*
LREf11927_11947	LREr13430_13411	+	+ + +	1503	*cbiETF*
LREf13411_13430	LREr14920_14900	+	+ + +	1509	*cbiFGH*
LREf14900_14920	LREr16411_16391	+	+ + +	1511	*cbiHJ*, *cysG/hemD*
LREf16391_16411	LREr17886_17864	+	+ + +	1495	*cysG/hemD*, *cbiK*
LREf17864_17886	LREr19409_19388	+	+ + +	1545	*cbiKLM*
LREf19388_19409	LREr20906_20887	+	+ + +	1518	*cbiMNQO*
LREf20887_20906	LREr22471_22451	+	+ + +	1584	*cbiOP*
LREf22451_22471	LREr24150_24130	+	+ + +	1699	*cbiP*, *sirA*, *hemA*
LREf24130_24150	LREr25669_25648	+	+ + +	1539	*hemAC*
LREf25648_25669	LREr26711_26683	+	+ + +	1063	*hemCB*
LREf26684_26718	LREr27526_27499	+	+ + +	842	*hemBL*
LREf27499_27518	LREr27987_27967	+	+ + +	488	*hemL*
LREf27499_27518	LREr28269_28250	−	−	770	*hemL*, *cobU*
LREf28250_28269	LREr29724_29704	+	+ + +	1474	*cobUSC*
LREf29704_29724	LREr31148_31129	+	+ + +	1444	*cobC*, *hemD*, *cobT*
LREf31129_31148	LREr31756_31737	+	+ + +	627	*cobT*
LREf31129_31148	LREr32212_32193	−	−	1083	*cobT*, hypothetical
LREf32193_32212	LREr33545_33526	−	−	1352	Hypothetical
LREf33354_33373	LREr33545_33526	+ +	+ +	191	*rpsO*
LREf33354_33373	LREr33809_33790	−	−	455	*rpsO*, metallo-β-lactamase
LREf33790_33809	LREr35516_35496	+ +	+ +	1726	Metallo-β-lactamase
LREf35496_35516	LREr36538_36518	−	−	1042	Metallo-β-lactamase, hypothetical

*Symbols refer to relative abundance of products on agarose gel. Increasing number of '+' corresponds to increasing intensity of band (see Fig. 3 for illustration); '−' corresponds to absence of expected band, indicating that the two genes are not part of the same transcript.

In the presence of an excess of exogenous B_{12} there was a 3.47 ± 0.90-fold change between the late- and mid-exponential growth phases for this same operon. Even though this upregulation is diminished in comparison to that in the absence of exogenous B_{12}, it is still quite considerable. Similar results were observed for the *cob* operon, upregulated 6.35 ± 0.73 and 3.31 ± 0.56 in the absence and presence of exogenous B_{12}, respectively. The upregulation in MRS of the transcript levels of the *pdu* loci was also observed in the absence of B_{12} from the medium. For these we observed a fold change of 3.76 ± 2.2 in the absence of exogenous B_{12} and 1.98 ± 0.68 when there was an excess of B_{12}.

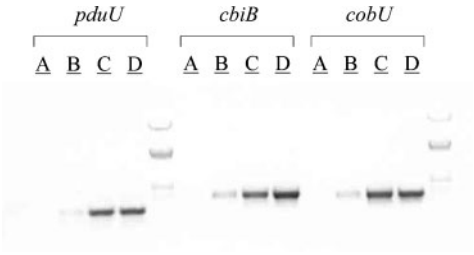

Fig. 3. RT-PCR amplicons from different loci collected from MRS cultures at different time points. Lanes A, RT-negative samples (negative control); lanes B, sample collected at mid-exponential phase; lanes C, sample collected at late-exponential phase; lanes D, *Lb. reuteri* chromosomal DNA (positive control).

To characterize in greater detail the specific impact of B_{12} supplementation for each growth phase, additional comparisons were made between the cultures lacking exogenous B_{12} and those with an excess of cyanocobalamin (Fig. 4c). During the exponential phase, expression of the B_{12} biosynthesis genes does not vary significantly with the absence of exogenous vitamin B_{12} (average fold change of 0.96 ± 0.08 for the *cbi* and *hem* operon, and 0.92 ± 0.08 for the *cob* operon). During the late-exponential phase, even though in both conditions the abundance of B_{12} biosynthesis transcripts is increased relative to the mid-exponential phase, in the absence of vitamin B_{12} supplementation,

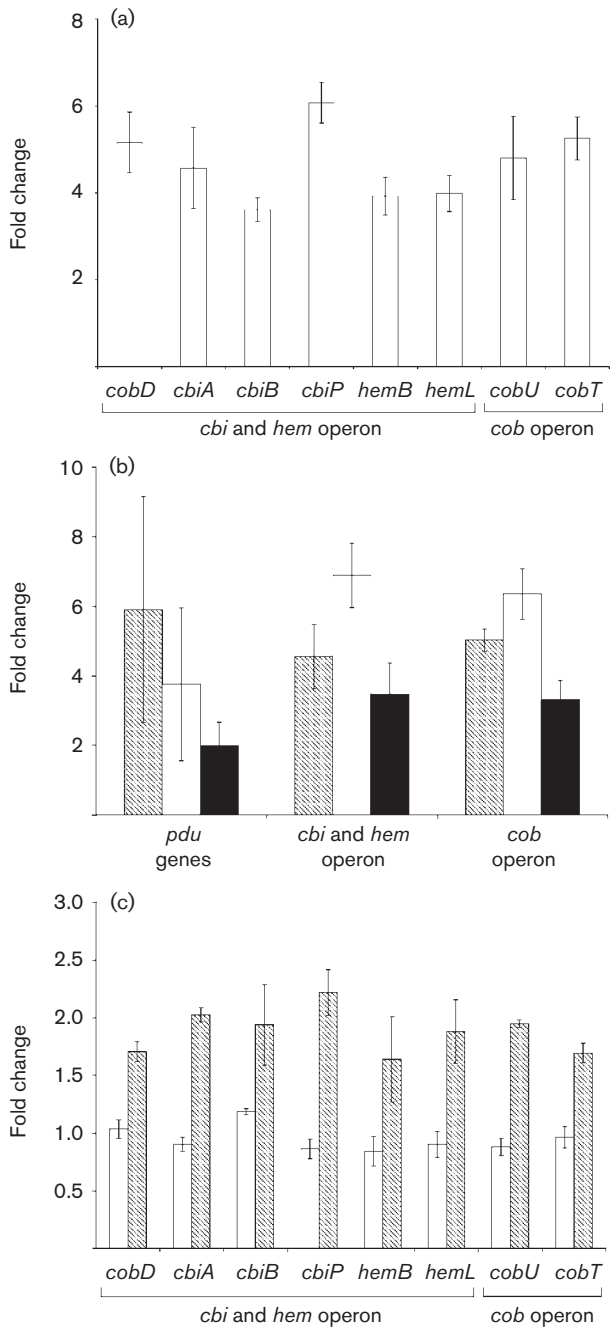

Fig. 4. Differential gene expression as determined by Q-RT-PCR. (a) Differential expression of the coenzyme B_{12} biosynthesis genes in late-exponential relative to mid-exponential phase in cells cultivated in MRS. (b) Average fold change between late-exponential and mid-exponential growth phases of *pdu* genes, the *cbi* and *hem* operon and the *cob* operon from *Lb. reuteri* cells cultured in MRS (hatched bars), B_{12}-free medium (white bars) and B_{12}-free medium supplemented with 1 mg cyanocobalamin l^{-1} (black bars). (c) Differential expression of B_{12} biosynthesis genes between cells cultivated in B_{12}-free medium and in B_{12}-free medium supplemented with 1 mg cyanocobalamin l^{-1} during mid-exponential (white bars) and late-exponential (hatched bars) growth phases.

their induction is stronger. For the late-exponential phase, when we compared the levels for cells cultured in the absence of exogenous B_{12} in relation to those cultivated in its presence, we determined an average fold-change of 1.90 ± 0.22 for the *cbi* and *hem* operon, and 1.82 ± 0.08 for the *cob* operon.

In silico analysis and comparative genomics

We determined the phylogeny of each predicted individual amino acid sequence encoded by the B_{12} gene cluster. We then compared them amongst each other, and with the deduced protein sequence of a control gene, *rpsO*, for which we performed the same exercise. The RpsO protein tree resembled the canonical phylogenetic topology deduced from 16S rRNA sequences (see Supplementary Fig. S2, available with the online version of this paper). In contrast, the predicted B_{12} proteins of *Lb. reuteri* were found to repeatedly cluster together with those of the genus *Listeria*, and closely neighboured by those of the genus *Salmonella* and other closely related γ-Proteobacteria (see Fig. 5 for illustration). This is suggestive of a common origin for the coenzyme B_{12} production pathway in these organisms. Variations from the mentioned tree topology were observed for *sirA*, *hemACBL* and *cobT*, and are addressed in the Discussion.

Both the G+C content and codon adaptation index (Sharp & Li, 1987) of the B_{12} cluster were compared with the draft genome sequence of *Lb. reuteri* JCM1112. The average G+C content of the coenzyme B_{12} biosynthesis gene cluster (36 mol%) does not differ significantly from the average of the draft genome sequence of *Lb. reuteri* available at the date of analysis (39 mol%). Concerning codon usage, again we did not observe any significant differences between the coenzyme B_{12} gene cluster of *Lb. reuteri* and other *Lb. reuteri* sequences. The average codon adaptation index for the genes of this cluster was calculated to be 0.69 ± 0.026, and we did not detect the usage of any rare codon. We also compared the G+C content and codon usage of the B_{12} biosynthesis clusters of *Listeria innocua* Clip11262 (Glaser *et al.*, 2001), 39 mol%, and *Salmonella enterica typhi* Ty2 (Deng *et al.*, 2003), 56 mol%, with their published genomes, 38 mol% and 52 mol% respectively.

DISCUSSION

The biosynthesis of coenzyme B_{12} from uroporphyrinogen III, the last shared metabolic precursor of the various tetrapyrrolic cofactors, requires about 25 enzymes, and has two different routes described: (i) the aerobic pathway studied in *Pseudomonas denitrificans* (Battersby, 1994); and (ii) the anaerobic pathway partially resolved in *Salmonella enterica*, *Bacillus megaterium* and *Propionibacterium shermanii* (Roessner & Scott, 2006). This biosynthetic pathway is commonly divided into three parts: (i) the synthesis of uroporphyrinogen III from either glutamyl-tRNA or

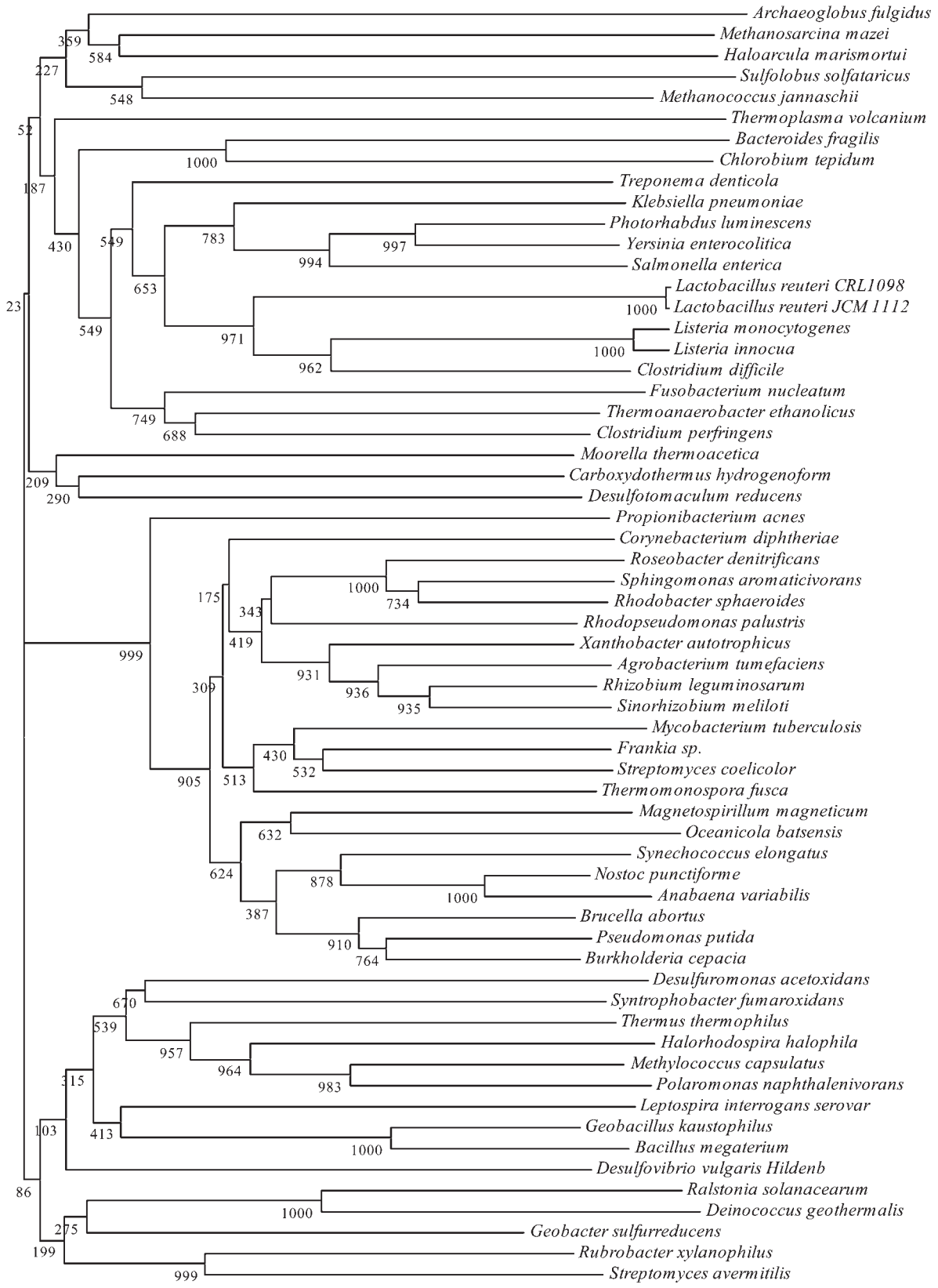

Fig. 5. Bootstrapped neighbour-joining phylogenetic tree of the CbiC protein.

glycine and succinyl-CoA; (ii) the corrin ring synthesis, which differs between the anaerobic pathway, starting with the insertion of cobalt into precorrin-2, and the aerobic pathway, where the cobalt chelation reaction occurs only after corrin ring synthesis; and (iii) the corrin ring adenosylation, attachment of the amino-propanol arm and assembly of the nucleotide loop bridging the lower ligand to the cobalt at the core of the corrin ring.

In *Lb. reuteri* we have found all the genes necessary to encode the complete anaerobic biosynthesis pathway of coenzyme B_{12}. Remarkably, and unlike the situation in other B_{12}-producing prokaryotes studied, genes for all three parts of the B_{12} biosynthetic pathway are clustered together in one continuous stretch of the chromosome. This presents a great advantage if considering metabolic engineering strategies aiming at transferring B_{12} production capability, as has been done before for other complex B vitamins (Sybesma *et al.*, 2004; Wegkamp *et al.*, 2004).

Based on the homology paradigm, our functional annotation of the newly sequenced coenzyme B_{12} biosynthesis gene cluster of *Lb. reuteri* was experimentally verified for *cobS* by the complementation of *Salmonella* mutant TT25722 (see Methods), lacking MetE and CobS activity. If cultured in minimal medium lacking methionine, this strain relies on the B_{12}-dependent methionine synthase (MetH), and has auxotrophic requirements for this cofactor. When we transformed TT25722 with pNZ7749, harbouring a fragment containing *cobS* amplified from *Lb. reuteri*, we reconstituted its ability to grow in minimal medium, depending on its own native coenzyme B_{12} production, and indirectly showed the functionality of *cobS* from *Lb. reuteri* (Fig. 2). Another example of functional evidence can be found in the recent report of the crystal structure of the PduO-type ATP : Co(I)rrinoid adenosyltransferase (St Maurice *et al.*, 2007) also sequenced within the course of this study.

Northern blotting and RT-PCR have shown that both the *cbi* genes, responsible for corrin ring synthesis, and the *hem* genes, responsible for the synthesis of uroporphyrinogen III, are transcribed together as part of a nearly 22 kb multicistronic operon. Although remarkably large, similar-sized transcripts have been detected in other lactic acid bacteria (van Kranenburg *et al.*, 2000). The *cob* genes are clustered in the same orientation, but expressed in a different operon of approximately 4 kb, situated just downstream of the previously mentioned *cbi* and *hem* transcript (Fig. 1).

The results from the Q-RT-PCR experiment corroborated the hypothesis emergent from the RT-PCR studies, that the B_{12} biosynthesis gene cluster is strongly induced during the late-exponential growth phase (Fig. 4a). Both operons are approximately five-fold upregulated in late-exponential when compared to mid-exponential growth, as determined by Q-RT-PCR for cells cultured in MRS broth. To ensure that the observed induction of the B_{12} biosynthesis genes in the late-exponential phase is not due to the depletion of B_{12} pools in MRS, we carried out additional experiments in B_{12}-free medium. We compared the induction of these genes between late- and mid-exponential phase, for cultures in the absence or presence of excess exogenous B_{12}. Although there was some variation in the levels of induction, it was clear that in all conditions assayed the B_{12} biosynthesis transcripts are more abundant in the late-exponential than in the mid-exponential phase (Fig. 4b).

The lower induction of the B_{12} biosynthesis genes during the late-exponential phase in the medium supplemented with B_{12} (Fig. 4c) suggests the presence of a regulatory feedback mechanism that inhibits the biosynthesis of this costly co-factor when it is available from the environment. Vitamin B_{12} metabolism has been shown to be often regulated by a conserved RNA structural element, known as riboswitch (Vitreschak *et al.*, 2003). We searched the coenzyme B_{12} biosynthesis gene cluster of *Lb. reuteri* for such conserved motifs using Rfam (Griffiths-Jones *et al.*, 2003) and Riboswitch finder (Bengert & Dandekar, 2004), but none could be found. The presence of a transposase immediately upstream of the first gene of the B_{12} cluster might have disturbed the riboswitch. The regulatory gene *pocR* (Bobik *et al.*, 1992), which is often between the B_{12} biosynthesis and *pdu* clusters, is not in this location in the chromosome of *Lb. reuteri*. In fact, this common regulator can be found at the far end of the adjacent *pdu* operon in the recently released genome of *Lb. reuteri* JCM1112. Its presence is in agreement with the experimental evidence gathered during this study suggesting co-regulation between the B_{12} cluster and the *pdu* genes located immediately upstream. PocR has been shown to be an activator of the coenzyme B_{12} biosynthesis cluster (Bobik *et al.*, 1992), and is likely to be involved in the observed negative feedback phenomena. Furthermore, PocR itself has been shown to be activated under carbon and redox control (Ailion *et al.*, 1993), which explains why we observed in all conditions assayed an induction of the B_{12} biosynthesis cluster during the late-exponential in comparison to the mid-exponential phase.

The topology of the phylogenetic tree obtained for the predicted product of the *rpsO* gene (data not shown) is similar to the canonical phylogenetic trees deduced from 16S rRNA sequences (see Fig. S2). In contrast, the phylogenetic comparison of all predicted amino acid sequences related to B_{12} biosynthesis showed that *Lb. reuteri* systematically clusters together with members of the genus *Listeria*, and closely neighbours the genus *Salmonella* and closely related *Enterobacteriaceae*. An illustration of a B_{12} biosynthesis protein phylogenetic tree is here depicted for CbiC, which was found to follow this topological pattern (Fig. 5). Exceptions to this topology include the products of *sirA* and *hemABCL*, for which *Lb. reuteri* clusters with *Listeria* and related genera of Gram-positive bacteria, while the *Enterobacteriaceae* now cluster with other γ-Proteobacteria, probably because their *hem* genes are properly adapted to aerobic conditions as well. In

addition the CobT protein is not encoded by the *Listeria* genomes, which may have suffered gene loss, while *Lb. reuteri* still clusters with *Salmonella* and closely related genera.

Lb. reuteri was the first lactic acid bacterium reported to produce coenzyme B_{12}, and the recently released genome sequences of a dozen lactic acid bacteria show no traces of genes related to B_{12} production. This observation, combined with the great differences in topology of the B_{12}-related trees and the canonical phylogenetic tree, suggests the acquisition of this capability by horizontal gene transfer. This promiscuity related to B_{12} metabolism between some genera of the Firmicutes and γ-Proteobacteria has been noted before when the phylogeny of the B_{12} regulatory motifs was being investigated (Vitreschak *et al.*, 2003).

The G+C content of *Lb. reuteri*'s B_{12} biosynthesis gene cluster does not clearly differ from the rest of its available genomic sequence, and the average codon adaptation index of this cluster is elevated, indicating that it is well suited to *Lb. reuteri*'s translational machinery. The same holds true for the B_{12} gene homologues of *Listeria* and *Salmonella*, indicating that the postulated horizontal gene transfer is not a recent event.

Associated with its survival strategy, *Lb. reuteri* is capable of producing and excreting reuterin, a broad-spectrum antimicrobial (Talarico *et al.*, 1988; Talarico & Dobrogosz, 1989). The production of this key component for its competitiveness is mediated by a B_{12}-dependent enzyme, glycerol dehydratase, responsible for catalysing the conversion of glycerol to 3-HPA, an intermediate of 1,3-propanediol in the glycerol catabolism pathway. The hypothesis that the acquisition of reuterin production and production of coenzyme B_{12} was a single event is supported by the following observations: (i) the genes involved in reuterin production are located just upstream of the B_{12} biosynthesis gene cluster; (ii) both sets of genes show similar phylogeny; and (iii) both sets of genes have similar expression patterns and seem to be part of the same regulon. This evolutionary event has presumably resulted in the speciation of *Lb. reuteri* from the other *Lactobacillus* species, and might have been important in its evolution to colonize the GI tract.

Lb. reuteri possesses the GRAS status and is an industrially relevant micro-organism. From a biotechnological point of view, the findings reported in this study can be applied for natural enrichment of (fermented) foods with B_{12}. Furthermore, they shed light on *Lb. reuteri* as a good candidate to investigate the possibility of *in situ* delivery of B_{12} in the GI tract.

ACKNOWLEDGEMENTS

We are grateful to Professor John R. Roth for kindly providing us with the *Salmonella enterica* mutants, and to Professor Roland Siezen for assistance in the initial annotation of the cluster.

REFERENCES

Ailion, M., Bobik, T. A. & Roth, J. R. (1993). Two global regulatory systems (Crp and Arc) control the cobalamin/propanediol regulon of *Salmonella typhimurium*. *J Bacteriol* **175**, 7200–7208.

Albert, M. J., Mathan, V. I. & Baker, S. J. (1980). Vitamin B_{12} synthesis by human small intestinal bacteria. *Nature* **283**, 781–782.

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

Anderson, P. J., Entsch, B. & McKay, D. B. (2001). A gene, *cobA* + *hemD*, from *Selenomonas ruminantium* encodes a bifunctional enzyme involved in the synthesis of vitamin B_{12}. *Gene* **281**, 63–70.

Banerjee, R. (2006). B_{12} trafficking in mammals: a for coenzyme escort service. *ACS Chem Biol* **1**, 149–159.

Battersby, A. R. (1994). How nature builds the pigments of life: the conquest of vitamin B_{12}. *Science* **264**, 1551–1557.

Bengert, P. & Dandekar, T. (2004). Riboswitch finder – a tool for identification of riboswitch RNAs. *Nucleic Acids Res* **32**, W154–W159.

Bobik, T. A., Ailion, M. & Roth, J. R. (1992). A single regulatory gene integrates control of vitamin B_{12} synthesis and propanediol degradation. *J Bacteriol* **174**, 2253–2266.

Bobik, T. A., Havemann, G. D., Busch, R. J., Williams, D. S. & Aldrich, H. C. (1999). The propanediol utilization (*pdu*) operon of *Salmonella enterica* serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B_{12}-dependent 1,2-propanediol degradation. *J Bacteriol* **181**, 5967–5975.

Croft, M. T., Lawrence, A. D., Raux-Deery, E., Warren, M. J. & Smith, A. G. (2005). Algae acquire vitamin B_{12} through a symbiotic relationship with bacteria. *Nature* **438**, 90–93.

Daniel, R., Bobik, T. A. & Gottschalk, G. (1998). Biochemistry of coenzyme B_{12}-dependent glycerol and diol dehydratases and organization of the encoding genes. *FEMS Microbiol Rev* **22**, 553–566.

Deng, W., Liou, S.-R., Plunkett, G., III, Mayhew, G. F., Rose, D. J., Burland, V., Kodoyianni, V., Schwartz, D. C. & Blattner, F. R. (2003). Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18. *J Bacteriol* **185**, 2330–2337.

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113.

Glaser, P., Frangeul, L., Buchrieser, C., Rusniok, C., Amend, A., Baquero, F., Berche, P., Bloecker, H., Brandt, P. & other authors (2001). Comparative genomics of *Listeria* species. *Science* **294**, 849–852.

Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. & Eddy, S. R. (2003). Rfam: an RNA family database. *Nucleic Acids Res* **31**, 439–441.

Kuipers, O. P., Beerthuyzen, M. M., Siezen, R. J. & De Vos, W. M. (1993). Characterization of the nisin gene cluster *nisABTCIPR* of *Lactococcus lactis*. Requirement of expression of the *nisA* and *nisI* genes for development of immunity. *Eur J Biochem* **216**, 281–291.

Kumar, S., Tamura, K. & Nei, M. (2004). MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.

Ma, D., Forsythe, P. & Bienenstock, J. (2004). Live *Lactobacillus reuteri* is essential for the inhibitory effect on tumor necrosis factor alpha-induced interleukin-8 expression. *Infect Immun* **72**, 5308–5314.

Maggio-Hall, L. A. & Escalante-Semerena, J. C. (1999). *In vitro* synthesis of the nucleotide loop of cobalamin by *Salmonella typhimurium* enzymes. *Proc Natl Acad Sci U S A* **96**, 11798–11803.

Maloy, S. R., Stewart, V. L. & Taylor, R. K. (1996). *Genetic Analysis of Pathogenic Bacteria: a Laboratory Manual*. Plainview, NY: Cold Spring Harbor Laboratory.

Martens, J. H., Barg, H., Warren, M. J. & Jahn, D. (2002). Microbial production of vitamin B_{12}. *Appl Microbiol Biotechnol* **58**, 275–285.

O'Toole, G. A., Rondon, M. R. & Escalante-Semerena, J. C. (1993). Analysis of mutants of *Salmonella typhimurium* defective in the synthesis of the nucleotide loop of cobalamin. *J Bacteriol* **175**, 3317–3326.

Overbeek, R., Larsen, N., Walunas, T., D'Souza, M., Pusch, G., Selkov, E., Jr, Liolios, K., Joukov, V., Kaznadzey, D. & other authors (2003). The ERGO genome analysis and discovery system. *Nucleic Acids Res* **31**, 164–171.

Raux, E., Lanois, A., Levillayer, F., Warren, M. J., Brody, E., Rambach, A. & Thermes, C. (1996). *Salmonella typhimurium* cobalamin (vitamin B_{12}) biosynthetic genes: functional studies in *S. typhimurium* and *Escherichia coli*. *J Bacteriol* **178**, 753–767.

Rice, P., Longden, I. & Bleasby, A. (2000). EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276–277.

Rodionov, D. A., Vitreschak, A. G., Mironov, A. A. & Gelfand, M. S. (2003). Comparative genomics of the vitamin B_{12} metabolism and regulation in prokaryotes. *J Biol Chem* **278**, 41148–41159.

Roessner, C. A. & Scott, A. I. (2006). Fine-tuning our knowledge of the anaerobic route to cobalamin (vitamin B_{12}). *J Bacteriol* **188**, 7331–7334.

Roest, K., Heilig, H. G., Smidt, H., de Vos, W. M., Stams, A. J. & Akkermans, A. D. (2005). Community analysis of a full-scale anaerobic bioreactor treating paper mill wastewater. *Syst Appl Microbiol* **28**, 175–185.

Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S. & Church, G. M. (1993). Characterization of the cobalamin (vitamin B_{12}) biosynthetic genes of *Salmonella typhimurium*. *J Bacteriol* **175**, 3303–3316.

Sambrook, J. & Russell, D. W. (2001). *Molecular Cloning: a Laboratory Manual*, 3rd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Saxelin, M., Tynkkynen, S., Mattila-Sandholm, T. & de Vos, W. M. (2005). Probiotic and other functional microbes: from markets to mechanisms. *Curr Opin Biotechnol* **16**, 204–211.

Sharp, P. M. & Li, W. H. (1987). The codon adaptation index – a measure of directional synonymous codon usage bias, and its potential applications. *Nucleic Acids Res* **15**, 1281–1295.

St Maurice, M., Mera, P. E., Taranto, M. P., Sesma, F., Escalante-Semerena, J. C. & Rayment, I. (2007). Structural characterization of the active site of the PduO-type ATP : Co(I)rrinoid adenosyltransferase from *Lactobacillus reuteri*. *J Biol Chem* **282**, 2596–2605.

Stabler, S. P. (1999). B_{12} and nutrition. In *Chemistry and Biochemistry of B_{12}*, pp. 343–365. Edited by R. Banerjee. New York: Wiley.

Sybesma, W., Starrenburg, M., Kleerebezem, M., Mierau, I., de Vos, W. M. & Hugenholtz, J. (2003). Increased production of folate by metabolic engineering of *Lactococcus lactis*. *Appl Environ Microbiol* **69**, 3069–3076.

Sybesma, W., Burgess, C., Starrenburg, M., van Sinderen, D. & Hugenholtz, J. (2004). Multivitamin production in *Lactococcus lactis* using metabolic engineering. *Metab Eng* **6**, 109–115.

Talarico, T. L. & Dobrogosz, W. J. (1989). Chemical characterization of an antimicrobial substance produced by *Lactobacillus reuteri*. *Antimicrob Agents Chemother* **33**, 674–679.

Talarico, T. L., Casas, I. A., Chung, T. C. & Dobrogosz, W. J. (1988). Production and isolation of reuterin, a growth inhibitor produced by *Lactobacillus reuteri*. *Antimicrob Agents Chemother* **32**, 1854–1858.

Taranto, M. P., Medici, M., Perdigon, G., Ruiz Holgado, A. P. & Valdez, G. F. (2000). Effect of *Lactobacillus reuteri* on the prevention of hypercholesterolemia in mice. *J Dairy Sci* **83**, 401–403.

Taranto, M. P., Vera, J. L., Hugenholtz, J., De Valdez, G. F. & Sesma, F. (2003). *Lactobacillus reuteri* CRL1098 produces cobalamin. *J Bacteriol* **185**, 5643–5647.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (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.

van der Heijden, R. T., Snel, B., van Noort, V. & Huynen, M. A. (2007). Orthology prediction at scalable resolution by phylogenetic tree analysis. *BMC Bioinformatics* **8**, 83.

van Kranenburg, R., Kleerebezem, M. & de Vos, W. M. (2000). Nucleotide sequence analysis of the lactococcal EPS plasmid pNZ4000. *Plasmid* **43**, 130–136.

Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. (2003). Regulation of the vitamin B_{12} metabolism and transport in bacteria by a conserved RNA structural element. *RNA* **9**, 1084–1097.

Walter, J., Heng, N. C., Hammes, W. P., Loach, D. M., Tannock, G. W. & Hertel, C. (2003). Identification of *Lactobacillus reuteri* genes specifically induced in the mouse gastrointestinal tract. *Appl Environ Microbiol* **69**, 2044–2051.

Wegkamp, A., Starrenburg, M., de Vos, W. M., Hugenholtz, J. & Sybesma, W. (2004). Transformation of folate-consuming *Lactobacillus gasseri* into a folate producer. *Appl Environ Microbiol* **70**, 3146–3148.

Edited by: P. W. O'Toole