Gene, 148 (1994) 125–129

GENE 08094

Cloning, sequence and in vitro transcription/translation analysis of a 3.2-kb *Eco*RI-*Hin*dIII fragment of *Leuconostoc oenos* bacteriophage L10

(Recombinant DNA; malolactic fermentation; wine; lactic acid bacteria)

Mark Sutherland[a], Hendrik J.J. van Vuuren[a] and Martha M. Howe[b]

[a]*Department of Microbiology and Institute for Biotechnology, University of Stellenbosch, Stellenbosch 7600, South Africa; and* [b]*Department of Microbiology and Immunology, University of Tennessee at Memphis, Memphis, TN 38163, USA. Tel. (1-901) 448-8215*

Received by J.A. Engler: 24 January 1994; Revised/Accepted: 7 March/9 March 1994; Received at publishers: 16 May 1994

SUMMARY

A 3.2-kb *Eco*RI-*Hin*dIII DNA fragment of *Leuconostoc oenos* bacteriophage L10 was cloned and sequenced. Computer-assisted analysis of the sequence identified eleven possible open reading frames (ORFs) that were all on the same strand. In vitro transcription/translation analysis of the full-length DNA fragment yielded five prominent proteins that were correlated with ORFs by their sizes and expression from deleted clones. Only those ORFs containing recognizable Shine-Dalgarno sequences coded for proteins. Neither the nucleotide sequence, nor deduced amino-acid sequences showed significant homology with other known sequences.

INTRODUCTION

Bacteriophages of *Lactococcus* and *Lactobacillus* spp. have been well studied due to their economic importance in the dairy industry (Sozzi et al., 1981; Jarvis, 1989; Prevots et al., 1990; Sèchaud et al., 1992). In contrast, the slow growth rates and low numbers of *Leuconostoc* spp. in mixed dairy starters have inhibited detection and analysis of their bacteriophages (Sozzi et al., 1978; Jarvis, 1989).

L. oenos is primarily responsible for the malolactic fermentation in wine (Davis et al., 1986). Stuck fermentations, which often occur, could be caused by bacteriophage infection (Davis et al., 1985; Henick-Kling et al., 1986; Nel et al., 1987). Bacteriophages were first observed microscopically in wine by Sozzi (Davis et al., 1985); however, bacteriophages able to form plaques on *L. oenos* have only recently been isolated from Australian red wines (Davis et al., 1985; Henick-Kling et al., 1986), South African red wines (Nel et al., 1987) and German cultures and wines (Arendt et al., 1991; Arendt and Hammes, 1992). Nel et al. (1987) isolated 20 *L. oenos* bacteriophages from wine and sugarcane and classified them into five genetic groups by restriction enzyme analysis. Using the same technique, Arendt and Hammes (1992) found four genetic groups among bacteriophages of *L. oenos* isolated from German wines.

The aim of this study was to characterize a 3.2-kb DNA fragment of *L. oenos* phage L10 (Nel et al., 1987) that is active against several commercially important *L. oenos* strains. This study represents the first reported nt sequence from an *L. oenos* phage.

Correspondence to: Dr. H.J.J. van Vuuren, Department of Microbiology, University of Stellenbosch, Stellenbosch 7600, South Africa. Tel. (27-21) 808-4866; Fax (27-21) 808-3611; e-mail: hjjvv@maties.sun.ac.za

Abbreviations: aa, amino acid(s); *B.*, *Bacillus*; bp, base pair(s); ds, double strand(ed); *E.*, *Escherichia*; GCG, Genetics Computer Group (Madison, WI, USA); IR, inverted repeat; kb, kilobase(s) or 1000 bp; *Lc.*, *Lactococcus*; *L.*, *Leuconostoc*; nt, nucleotide(s); ORF, open reading frame; PAGE, polyacrylamide-gel electrophoresis; SD, Shine-Dalgarno (sequence); SDS, sodium dodecyl sulfate.

SSDI 0378-1119(94)00305-C

EXPERIMENTAL AND DISCUSSION

(a) Cloning and nt sequence analysis

The 3.2-kb *Eco*RI-*Hin*dIII DNA fragment was purified from a digest of phage DNA isolated from mature phage particles, cloned into pUC18, and the nt sequence determined (Fig. 1). The G+C content of the DNA fragment is 38.4%, within the range of 38–42% reported for *L. oenos* strains (Dicks et al., 1990).

(b) ORF and protein analysis

Computer-assisted analysis of the nt sequence identified eleven possible ORFs, all located on the same strand (Fig. 2A). To determine which of the predicted ORFs encode protein we made deleted derivatives (pMS1–pMS7; Fig. 2B) of pMS10, containing the 3.2-kb fragment cloned in the plasmid expression vector pSP73 downstream from the SP6 promoter. In vitro transcription/translation of the intact plasmid (pMS10) using SP6 polymerase resulted in the production of several L10-specific proteins (Fig. 3). Similar analysis of the deletion plasmids allowed us to roughly locate the genes encoding proteins on the restriction map and correlate them with the predicted ORFs. The 26.6-kDa protein produced by plasmids pMS1 and 3 corresponds to ORF E (predicted size 21.8 kDa). The 23.7-kDa protein made only by pMS1, 2, 5 and 6 corresponds to ORF A (18.3 kDa). The 19.5, 12.5 and 10-kDa proteins made only by pMS1, 3, 5 and 7 arose from the *Sfu*I to *Ava*I region containing only ORF F (16.7 kDa). We presume that the largest protein corresponds to ORF F; the two smaller ones may result from initiation at internal Met within ORF F (13.8, 13.2, 10.3 kDa). The 16.8-kDa protein made only by pMS1 and 6 is derived from ORF I (14.2 kDa). The 10.6-kDa protein made by pMS3, 5 and 7 corresponds to ORF C and should terminate 42 bp into the vector, giving a predicted size of 11.3 kDa. Several protein bands were produced by only a single deletion plasmid and not by the intact plasmid. The bands unique to pMS2 and pMS6 are presumed to represent truncated products of ORF I and ORF E, respectively; the band unique to pMS5 is probably a fusion protein composed of the N terminus of ORF I and C terminus of ORF E.

The measured sizes of the L10-specific proteins were about 19% larger than those predicted from the ORFs. Differences between predicted and observed sizes for small proteins are common and may reflect effects of specific aa on protein folding and mobility (De Jong et al., 1978).

Prediction of protein hydropathy (Kyte and Doolittle, 1982) indicated that the proteins encoded by ORFs A, C and F are strongly hydrophilic, whereas those encoded by ORFs E and I are slightly hydrophilic. Homology searches of the GenBank database v72.0 and Swiss Protein database v22.0 using the GCG programs FASTA and TFASTA (Pearson and Lipman, 1988) conducted for the entire nt sequence and for the deduced protein sequences of ORFs A, C, E, F, I and K, respectively, did not reveal any significant similarity to other DNAs or proteins.

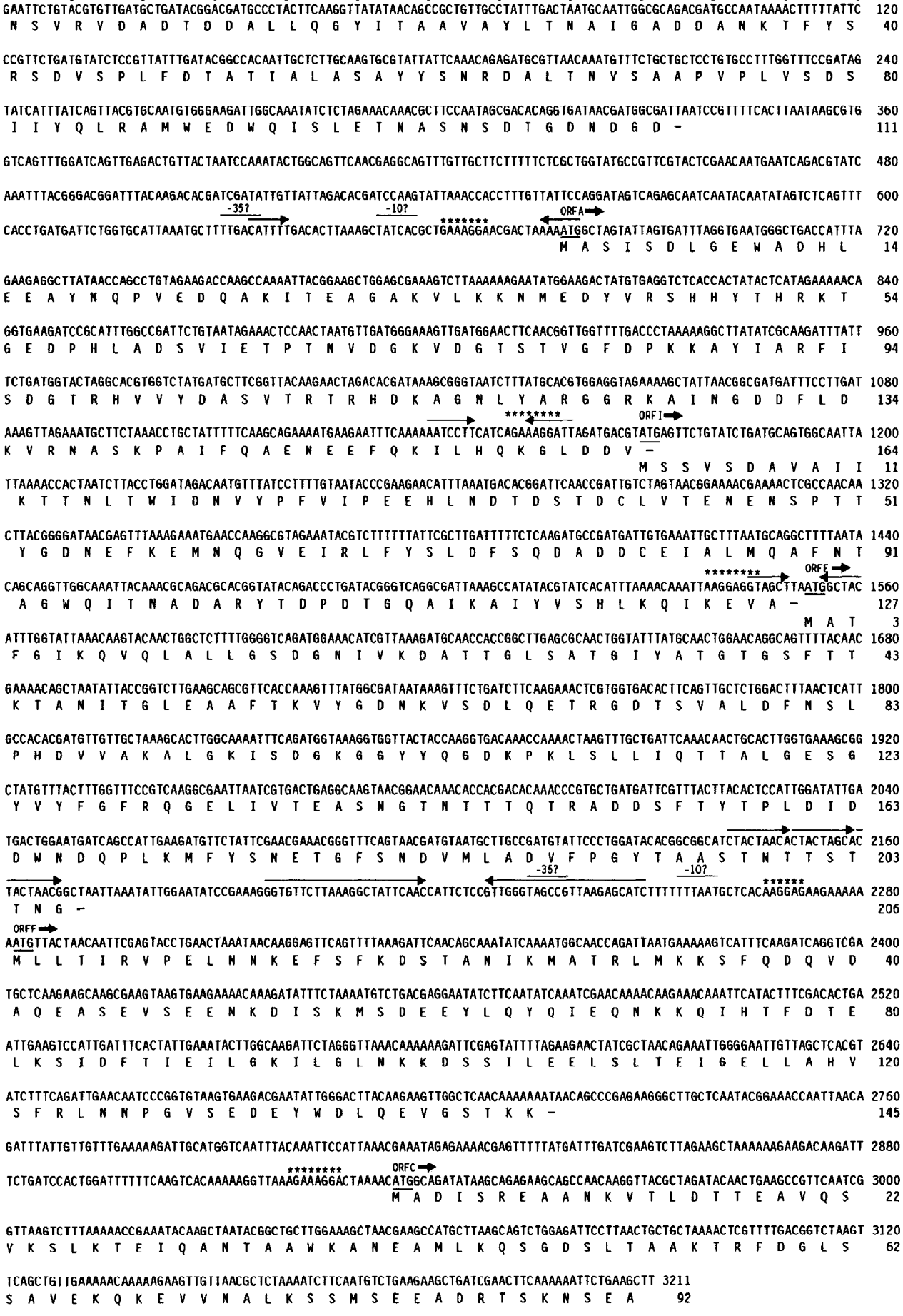
(c) Potential regulatory sequences and codon usage

Examination of the nt sequence upstream from the predicted ORFs for similarity to the consensus Gram$^+$ SD sequence for ribosome binding (5′-AGAAAGGAGGT: McLaughlin et al., 1981) revealed that only the five ORFs that produce detectable protein products (A, C, I, F and E) have recognizable SD sequences. As is typical of Gram$^+$ bacteria (McLaughlin et al., 1981), they have extensive complementarity (6 to 8 contiguous nt) to the 16S rRNA 3′ end and spacing of 5–9 nt to the ATG (Fig. 1). Examination of the 5′ flanking regions of the ORFs for similarity to the consensus Gram$^+$ promoter −10 (TATAAT) and −35 (TTGACA) hexamers (Graves and Rabinowitz, 1986) revealed only two potential promoter sequences, just upstream from ORF A and ORF F (Fig. 1). It is not clear, however, whether either candidate is responsible for transcription of the ORF, since each is unusually close (3 and 6 nt) to the SD sequence.

A search for direct and inverted repeats revealed several with potential regulatory function (Fig. 1), including short IR whose pairing could occlude access of the ribosome to the SD sequence or initiating ATG for ORFs A, I and E and a 9-nt sequence directly repeated three times near the end of ORF E that might serve as a protein binding site. A long IR followed by T residues, a structure similar to Rho-independent terminators (Brendel et al., 1986), located between ORFs E and F may terminate mRNA synthesis.

Comparison of codon usage of these five ORFs with that determined previously for *Lc. lactis* (Van de Guchte et al., 1992) and for eight recently sequenced genes from

Fig. 1. The nt sequence of the 3.2-kb *Eco*RI-*Hin*dIII DNA fragment (GenBank accession No. L13035) and deduced aa sequence for ORFs A, C, E, F, I and K. Both strands were completely sequenced, using synthetic primers and Sequenase v2.0 (US Biochemical, Cleveland, OH, USA) according to the manufacturer's recommendations. Asterisks indicate nt complementary to the 3′ end of *B. subtilis* and *Lc. lactis* 16S rRNA (Ludwig et al., 1985) that might serve as SD sequences for ribosome binding. Two sequences similar to the Gram$^+$ promoter hexamer −35 and −10 regions are indicated just upstream from ORF A and ORF F. Arrows indicate repeated sequences which may be involved in regulation due to their capacity for protein binding or RNA secondary structure.



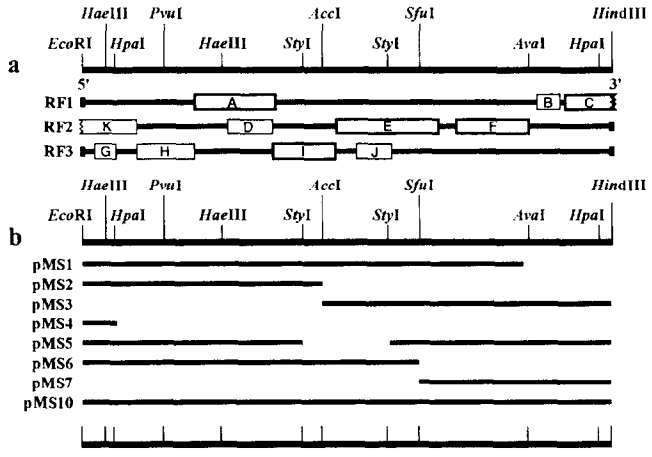


Fig. 2. Diagrammatic representation of the predicted ORFs (**a**) and deletion plasmids (**b**). (**a**) The ORFs were deduced in all six reading frames (RF1–RF6) by computer analysis of the nt sequence with Genepro v4.20 and ATG as the translation start codon. There were no significant ORFs in the opposite strand (RF4–RF6). The ORF designated K is lacking an initiating ATG; it probably represents an ORF interrupted by cloning at the *Eco*RI site. Heavy delineated boxes indicate the ORFs whose protein products were detected by in vitro transcription/translation analysis (Fig. 3). (**b**) Derivatives of the pMS10 plasmid containing deletions were generated for in vitro transcription/translation analysis by restriction endonuclease digestion, isolation of the DNA fragments from agarose gels by microspin filters (Macherey-Nagel, Düren, Germany) using the manufacturer's recommended procedures and ligation of the ends to recircularize the DNAs. The L10 DNA remaining in each deletion plasmid is indicated by the bar.

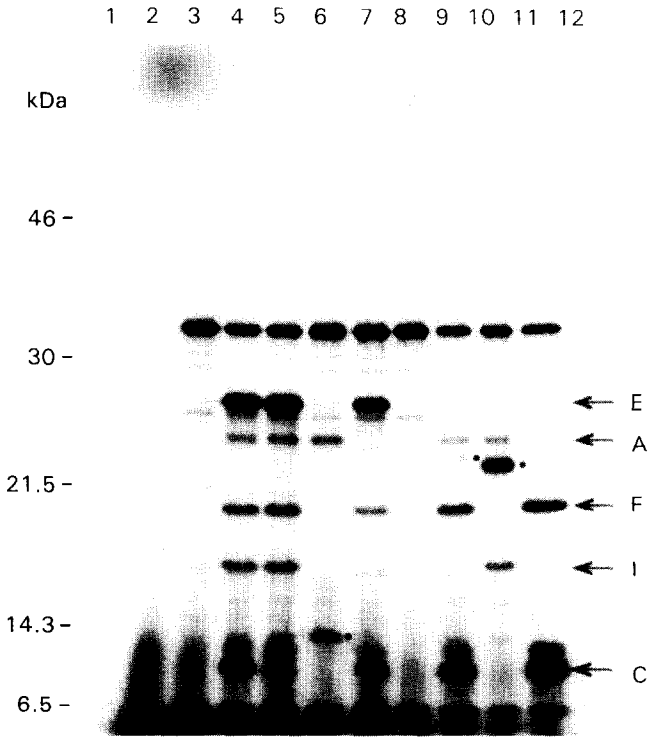


Fig. 3. In vitro transcription/translation products from various plasmid DNA templates. Lanes: 1 and 12, size markers; 2, no DNA control; 3, vector plasmid pSP73 (Promega Madison, WI, USA); 4, pMS10; 5–11, deletion derivatives pMS1–pMS7, respectively. Arrows designate the protein species arising from the indicated ORFs. Dots indicate unique species arising by truncation or fusion of ORFs in particular plasmids. In vitro transcription/translation analysis was performed using *E. coli* S30 extract (Amersham International, UK), [^{35}S]methionine (880 Ci/mmol, Amersham) and supercoiled plasmid DNA templates. The translation products were separated by 0.1% SDS-15% PAGE for 18 h. The gel was fixed (70% distilled H_2O/20% methanol/10% acetic acid) and dried before autoradiography.

Leuconostoc spp. (GenBank Nos.: M92281, M95954, M94060 and M64371) revealed considerable similarity in codon usage between these organisms as would be expected, since they are closely related and have similar G+C contents (Sneath et al., 1986).

(d) Genetic organization

The 4-nt overlap between the end of ORF A and the beginning of ORF I (5'-GTATGAGT) and the 1-nt overlap between the end of ORF I and the beginning of ORF E (5'-CTTAATGGC) suggest possible translational coupling of these ORFs during translation from a single polycistronic transcript (Normark et al., 1983) derived from an *AIE* operon. The presence of stop codons in all three reading frames as well as potential promoter and terminator sequences between ORFs E and F make it likely that F is encoded by a different transcript.

To test whether the regions between ORFs K and A and between ORFs F and C are really non-coding, we carried out analysis of those regions with the program TESTCODE (Fickett, 1982) to detect protein-coding regions by their nonrandom distribution of bases and with CODON PREFERENCE (Gribskov et al., 1984) to compare the codon usage with that from *Lc. lactis* and *Leuconostoc* sequences. The results for the K-A region containing ORF H were similar to that for ORF F (data not shown). In addition, ORF H and ORF K overlap by 11 nt, and ORF H and ORF A overlap by 8 nt, again suggestive of translational coupling. Therefore, we do not rule out the possibility that ORF H is a gene whose product is not detected due to the absence of a strong ribosome binding site. If this were the case, ORFs K, H, A, I and E might be contained within a single operon. Similar analysis for the region between ORFs F and C did not predict the presence of a coding region, leaving open the possibility that it may serve a structural or regulatory function. The lack of recognizable −35 and −10 hexamer sequences suggests that ORF C may be transcribed from a phage-specific promoter.

(e) Conclusions

A 3.2-kb DNA fragment from *L. oenos* phage L10 has been characterized by nt sequence and in vitro transcription/translation analysis, demonstrating the presence of at least five genes and several potential regulatory sequences.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the financial support of the FRD and the KWSI (H.J.J. vanV.), and a Van Vleet Professorship (M.M.H.). The oligodeoxyribonucleotides used in this work were provided by the University of Tennessee Molecular Resource Center. The authors thank D. Rawlings, M. Kahmeyer and R. Hill for assistance with the computer analysis.

REFERENCES

Arendt, E.K. and Hammes, W.P.: Isolation and characterization of *Leuconostoc oenos* phages from German wines. Appl. Microbiol. Biotechnol. 37 (1992) 643–646.

Arendt, E.K., Neve, H. and Hammes, W.P.: Characterization of phage isolates from a phage-carrying culture of *Leuconostoc oenos* 58N. Appl. Microbiol. Biotechnol. 24 (1991) 220–224.

Brendel, V., Hamm, G.H. and Trifonov, E.F.: Terminators of transcription with RNA polymerase for *Escherichia coli*: what they look like and how to find them. J. Biomol. Struct. Dyn. 3 (1986) 705–723.

Davis, C., Silveira, N.F.A. and Fleet, G.H.: Occurrence and properties of bacteriophages of *Leuconostoc oenos* in Australian wines. Appl. Environ. Microbiol. 50 (1985) 872–876.

Davis, C.R., Wibowo, D.J., Lee, T.H. and Fleet, G.H.: Growth and metabolism of lactic acid bacteria during and after malolactic fermentation of wines at different pH. Appl. Environ. Microbiol. 51 (1986) 539–545.

De Jong, W.W., Zweers, A. and Cohen, L.H.: Influence of single amino acid substitutions on electrophoretic mobility of sodium dodecyl sulfate-protein complexes. Biochem. Biophys. Res. Commun. 82 (1978) 532–539.

Dicks, L.M.T., Van Vuuren, H.J.J. and Dellaglio, F.: Taxonomy of *Leuconostoc* species, particularly *Leuconostoc oenos*, as revealed by numerical analysis of total soluble cell patterns, DNA base compositions, and DNA-DNA hybridizations. Int. J. Syst. Bacteriol. 40 (1990) 83–91.

Fickett, J.W.: Recognition of protein coding regions in DNA sequences. Nucleic Acids Res. 10 (1982) 5303–5318.

Graves, M.C. and Rabinowitz, J.C.: *In vivo* and *in vitro* transcription of the *Clostridium pasteurianum* ferridoxin gene. J. Biol. Chem. 256 (1986) 11409–11415.

Gribskov, M., Devereux, J. and Burgess, R.R.: The codon preference plot: graphic analysis of protein coding sequences and prediction of gene expression. Nucleic Acids Res. 12 (1984) 539–549.

Henick-Kling, T., Lee, T.H. and Nicholas, D.J.D.: Inhibition of bacterial growth and malolactic fermentation in wine by bacteriophages. J. Appl. Bacteriol. 61 (1986) 287–293.

Jarvis, A.W.: Bacteriophages of lactic acid bacteria. J. Dairy Sci. 72 (1989) 3406–3428.

Kyte, J. and Doolittle, R.F.: A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157 (1982) 105–132.

Ludwig, W., Seewaldt, E., Kilpper-Bälz, R., Schleifer, K.H., Magrum, L., Woese, C.R., Fox, G.E. and Stackebrandt, E.: The phylogenetic position of *Streptococcus* and *Enterococcus*. J. Gen. Microbiol. 131 (1985) 543–551.

McLaughlin, J.R., Murray, C.L. and Rabinowitz, J.C.: Unique features in the ribosome binding site sequence of the Gram-positive *Staphylococcus aureus* β-lactamase gene. J. Biol. Chem. 256 (1981) 11283–11291.

Nel, L., Wingfield, B.D., Van der Meer, L.J. and Van Vuuren, H.J.J.: Isolation and characterization of *Leuconostoc oenos* bacteriophages from wine and sugarcane. FEMS Microbiol. Lett. 44 (1987) 63–67.

Normark, S., Bergstrom, S., Edlund, T., Grundstrom, T., Jaurin, B., Lindberg, F.P. and Olsson, O.: Overlapping genes. Annu. Rev. Genet. 17 (1983) 499–525.

Pearson, W.R. and Lipman, D.J.: Improved tools for biological sequence comparison. Proc. Nat. Acad. Sci. USA 85 (1988) 2444–2448.

Prevots, F., Mata, M. and Ritzenthaler, P.: Taxonomic differentiation of 101 lactococcal bacteriophages and characterization of bacteriophages with unusually large genomes. Appl. Environ. Microbiol. 56 (1990) 2180–2185.

Sèchaud, L., Rousseau, M., Fayard, B., Callegari, M.L., Quènèe, P. and Accolas, J.-P.: Comparative study of 35 bacteriophages of *Lactobacillus helveticus*: morphology and host range. Appl. Environ. Microbiol. 58 (1992) 1011–1018.

Sneath, P.H.A., Mair, N.S., Sharpe, E.L. and Holt, J.G. (Eds.): Bergey's Manual of Systematic Bacteriology, Vol. 2. Williams and Wilkins, London, 1986.

Sozzi, T., Poulin, J.M., Maret, R. and Pousaz, R.: Isolation of a bacteriophage of *Leuconostoc mesenteroides* from dairy products. J. Appl. Bacteriol. 44 (1978) 159–161.

Sozzi, T., Watanabe, K., Stetter, K. and Smiley, M.: Bacteriophages of the Genus *Lactobacillus*. Intervirology 16 (1981) 129–135.

Van de Guchte, M., Kok, J. and Venema, G.: Gene expression in *Lactococcus lactis*. FEMS Microbiol. Rev. 88 (1992) 73–92.