Evidence for horizontal gene transfer as origin of putrescine-production in *Oenococcus oeni* RM83

ÁNGELA MARCOBAL[a,b], BLANCA DE LAS RIVAS[a], M. VICTORIA MORENO-ARRIBAS[b], AND ROSARIO MUÑOZ[a*]

[a]Departamento de Microbiología and [b]Departamento de Caracterización de Alimentos, Instituto de Fermentaciones Industriales, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain.

Running Title

PUTRESCINE-PRODUCTION IN *OENOCOCCUS OENI*

* Corresponding author. Mailing address: Departamento de Microbiología, Instituto de Fermentaciones Industriales, CSIC, Juan de la Cierva3, 28006 Madrid, Spain. Phone: 34 91 562 2900. Fax: 34 91 564 4853. E-mail: rmunoz@ifi.csic.es

The nucleotide sequence of a 17.2-kb chromosomal DNA fragment containing the *odc* gene encoding the ornithine decarboxylase has been determined in the putrescine-producer *Oenococcus oeni* RM83. This DNA fragment contains 13 open reading frames, including genes coding for five transposases and two phage proteins. This description might represent the first evidence of a horizontal gene transfer event as the origin of a biogenic amine biosynthetic locus.

Wines are a highly selective media and support the growth of only a few species of lactic acid bacteria (LAB), mainly *Oenococcus oeni* and several lactobacilli. *O. oeni* is often responsible for wine malolactic fermentation and is frequently utilized as a starter culture to promote malolactic conversion.

In acidic media like wine, decarboxylation of amino acids to their corresponding amines is thought to provide energy through electrogenic transport as well as assist in maintaining an optimal internal pH (11). Some of these amines are considered "biogenic" and may cause intoxications when consumed. The biogenic amine putrescine, which can potentiate the action of histamine, is the most prevalent amine in wine, and is found in almost all analysed wines (12, 15).

Biogenic amines are primarily formed by decarboxylation of the corresponding amino acids by microorganisms through substrate-specific decarboxylases. The capability of biogenic amine production appears to be strain-dependent rather than species-specific. Previously, we reported the identification of the *odc* gene in the putrescine-producer *O. oeni* RM83 (formerly, *O. oeni* BIFI-83) for the first time (14). The *odc* gene encodes a

deduced 745-amino acid putative ornithine decarboxylase (ODC-EC 4.1.1.17) which catalyzes the conversion of ornithine to putrescine. The *odc* gene is seldom present in the *O. oeni* genome, as it has not been detected in a screen of 42 *O. oeni* strains tested to date (14). Moreover, *in silico* analysis of the draft *O. oeni* PSU-1 genome did not reveal the presence of any *odc* homologs (16).

Recently, Lucas et al. (2005) (13) described that the potential for producing histamine in *Lactobacillus hilgardii* 0006 is encoded on an unstable 80-kb plasmid; the authors further suggested that it is very likely that the histamine-producer *Tetragenococcus muriaticus* and *O. oeni* 9204, harbors the same plasmid. However, the localization of the *odc* gene in *O. oeni* RM83 remains unknown.

This study was undertaken to gain deeper insight into the origin of putrescine production in *O. oeni* RM83. Additionally, the *O. oeni* RM83 ODC has been expressed in *E. coli* and biochemically characterized.

Genetic location of the *odc* locus in *O. oeni* RM83. The putrescine-producer *O. oeni* RM83, formerly *O. oeni* BIFI-83, was isolated from lees of a Spanish red wine (14). Putrescine-production by *O. oeni* RM83 was maintained without ornithine pressure, suggesting that the *odc* locus was stable. To determine if *O. oeni* RM83 harbored any plasmids, total DNA was extracted and analysed by standard agarose gel electrophoresis (19). This assay revealed the absence of small plasmids in *O. oeni* RM83 (data not shown). Subsequently, native total DNA was embedded in agarose plugs and analysed by PFGE as described previously (1). Again, plasmids were not detected. Moreover, southern hybridization with a 1.4-kb DNA probe targeted to an internal *odc* fragment (14) yielded a

positive signal only in the chromosomal DNA (data not shown). Therefore, it is concluded that *odc* in *O. oeni* RM83 is located on the chromosome.

Characterization of the *odc* region in *O. oeni* RM83. Since i) *in silico* analysis of the *O. oeni* PSU-1 draft genome did not reveal the presence of an *odc* gene (16), ii) the presence of the *odc* gene appears to be infrequent in the *O. oeni* genome (13), and iii) the *odc* gene is chromosomically located in *O. oeni* RM83, we decided to identify the chromosomal DNA region involved in putrescine production in *O. oeni* RM83. The 17.2-kb sequence flanking the previously described 2.3 kb *odc* region was determined (14). This sequence was ascertained by creating a phage library of *O. oeni* RM83 genomic DNA and by several reverse PCR experiments.

To construct the *O. oeni* RM83 DNA library, chromosomal DNA was partially digested with Sau3AI restriction enzyme and ligated to the ZAP Express® vector (Stratagene, La Jolla, CA) digested with BamHI. The screening of the library using the 1.4-kb internal *odc* DNA fragment as a probe yielded five positive clones. Since the inserts of three of them were included in pAM4 and pAM8, only these plasmids were sequenced (Fig. 1). A total of 10,891 bp *O. oeni* RM83 DNA fragment was sequenced from the pAM4 and pAM8 plasmids. Two successive reverse PCR experiments, utilizing SnaBI and EcoRV, allowed us to sequence the 5´end of the fragment (Fig. 1). Similarly, a reverse PCR experiment using SpeI, allowed for the sequencing of the total 17,270-bp EcoRV-SpeI chromosomal DNA fragment surrounding the *O. oeni* RM83 *odc* gene (Fig. 1).

Sequence analysis of this DNA fragment showed the presence of 11 complete (albeit some interrupted) and 2 partial open reading frames (ORFs) in the *odc* region (Fig. 1, Table 1). Two interesting features were observed, the presence of two putative phage

proteins and the presence of five transposase-coding genes. The first incomplete ORF (*orf1*), is predicted to code for a protein showing the highest similarities (>30% identity) to *Streptococcus thermophilus* bacteriophage proteins. Notably, it has been reported that the genomes of currently characterized *S. thermophilus* phages exhibit homology to each other in a modular fashion (2). Furthermore, *orf2* is predicted to encode a protein similar to a DNA replication protein from an *Enterococcus faecalis* prophage. The *O. oeni* PSU-1 draft genome does not contain any intact temperate bacteriophage, or larger tracts of obvious bacteriophage origin, although several prophage integration sites have been found (16).

Contiguous to and divergently transcribed from *orf2*, we found a variant copy of the insertion sequence (IS) IS*1165* (99% nucleotide identity) (9). The existence of two additional copies of IS*1165*, at positions 3121-4675 and 12879-14433/c was observed. All three of the IS*1165* copies are identical and contain the canonical terminal inverted repeats. Although IS*1165* was originally described in *Leuconostoc mesenteroides* subsp. *cremoris*, copies of this IS element have been described in other LAB such as *Leuconostoc lactis*, *O. oeni*, *Pediococcus* sp., *Lactobacillus helveticus,* and *Lactobacillus casei* (10).

Upstream of *orf4*, there is a 876-bp region that might correspond to an insertion sequence-like element on the basis of sequence similarity (58.5% nucleotide identity to a *L. plantarum* transposase). Almost perfect 17-bp inverted repeats were found at positions 6124-6140 and 6984-7000. Another IS element, a copy of IS*Lpl4*, is found 722 nucleotides further upstream (4).

The *odc* gene is located downstream of IS*Lpl4*. The ODC protein is predicted to possess 745-amino acid residues including conserved residues involved in enzymatic activity as well as the consensus sequence containing the pyridoxal-5-phosphate binding domain (14). The next identified ORF is a putative *potE* gene. It encodes a 441-amino acid

residue, 47.6-kDa, protein showing 67.3% identity to the putrescine-ornithine antiporter (PotE) from *Haemophilus influenzae*. PotE can catalyse both the uptake and excretion of putrescine (8).

Functional expression of the *odc* gene from *O. oeni* RM83. To confirm that the *odc* gene from *O. oeni* RM83 encodes a functional ODC, we expressed this gene in *E. coli* HT414 (CGSC strain # 6856) as this strain is defective for ODC activity (20). Firstly, the gene was PCR amplified from *O. oeni* RM83 DNA by using Pfu DNA polymerase (Stratagene, La Jolla, CA) and oligonucleotides PIN-ODC-up (5´-GGAACTCTAGAGGGTATTAATAATGGATAGCGAAATAAATGATGATTC) and PIN-ODC-down (5´-CGCATTGCGTTCACGTCGTTGCTCAATTATCATCTTTTTTCTTCATCTTTTGAC). The purified PCR fragment was inserted into pIN-III(lpp^{p}-5)A3 (9) by using the strategy described by Geiser et al. (7).

Cell extracts for ODC enzymatic assay were obtained from *E. coli* HT414 cells harboring the control plasmid pIN-III(lpp^{p}-5)A3 or the recombinant plasmid pAM11 as previously described (18). The ODC assay was performed in 50 mM sodium phosphate buffer (pH 6.5) in the presence of 3.6 mM ornithine and 0.4 mM PLP. The reaction was incubated at 37 ºC for 1 hour. Subsequently, the putrescine formed in the reaction was derivatized, and detected by thin-layer chromatography (6) and by RP-HPLC (14). Extracts from the strains harboring pAM11 were able to decarboxylate the supplied ornithine to putrescine, whereas extracts prepared from control cells containing the vector plasmid alone did not (Fig. 2). Therefore, we have provided experimental evidence that the *odc* gene encodes a functional ODC.

Recombination as origin of the *odc* region in *O. oeni* RM83. Comparison of nucleotide positions 15271 to 17270 (2 kb) of the sequenced *O. oeni* RM83 DNA fragment, with positions 60644 to 62643 of the draft genome sequence of *O. oeni* PSU-1, accession NZ_AABJ03000005, reveals three distinct regions based on their nucleotide sequence identity (Fig. 3). The leftmost 649 nucleotides of both sequences showed a 47% nucleotide identity, the next 827 nucleotides are 61.5% identical between RM83 and PSU-1 and, the rightmost 524 nucleotide residues exhibiting perfect identity between both strains.

The proteins encoded by these 2-kb sequences are remarkable as well. From nucleotide position 61293, *O. oeni* PSU-1 encodes a 156 residue protein, Ooen02001059, annotated as a carbamoyl phosphate synthase and Ooen02001060, a 631 amino acid hypothetical protein (Fig. 3). This hypothetical protein contains a N-terminal putative signal peptide extending to amino acid 33, followed by five GW domains and a 130 amino acid residues C-terminal end containing a MucBP domain (Fig. 3).

It is noteworthy that *orf7* in *O. oeni* RM83 appears to be a chimeric protein originating from the fusion of a gene encoding a protein 50% identical to the *O. oeni* PSU-1 putative carbamoyl phosphate synthase (Ooen02002059) and a gene encoding a protein 55% identical on its MucBP domain to the PSU-1 hypothetical protein (Ooen02002060). Taking into account that GW and MucBP domains are found in a variety of bacterial proteins, it is possible that the unknown donor protein could have domains coded by DNA regions showing high nucleotide similarity with the corresponding *O. oeni* regions. This similarity could facilitate the crossover between this unknown donor DNA and *O. oeni* chromosomal DNA. Upon examination of regions of extreme identity, the crossover point appears to reside at nucleotide position 16747 of the *O. oeni* RM83 sequence described in

this work, and position 62120 of the *O. oeni* PSU-1 accession NZ_AABJ03000005. This recombination site is located in the junction of the MucBP domain and the first GW domain. Interestingly, in *O. oeni* PSU-1 the gene coding for the hypothetical protein Ooen02001060 is found less than 5 kb downstream of the gene *recP* coding for a transketolase. Recently, de las Rivas et al. (2004) (3) described in the *recP* locus a possible example of a recombinatorial event from an unknown source. The description of two recombinatorial events in the same DNA region indicate a region of great flexibility in the *O. oeni* chromosome as described recently in *L. plantarum* (17).

It is now understood that horizontal gene transfer provides an important mechanism for generating genotypic and phenotypic diversity in bacteria. This phenomenon has been studied extensively in relation to bacterial adaptability or fitness under certain growth conditions. Accordingly, it has been widely reported that adaptability traits can be encoded by mobile genetic elements. Genomic islands (GI) are clusters of chromosomal genes that have been described as horizontally acquired DNA regions (5). They often possess genes (or pseudogenes) coding for mobility-related elements such as phage genes, insertion sequence elements, transposases, and origins of replication. A typical GI carries genes encoding traits that may increase bacterial adaptability under certain growth conditions. All of these observations taken together, suggests that the 16.7-kb *O. oeni* RM83-specific DNA may be a fragment of a GI transferred by horizontal gene transfer.

Nucleotide sequence accession number. The DNA sequence of the *O. oeni* RM83 *odc* region has been deposited in the EMBL/GenBank database under accession number AJ746165.

This work was supported by grants AGL2005-00470 (CICYT), RM03-002 (INIA), and S-0505/AGR/000153 (CAM). We thank E. García and D. Llull for their help with the PFGE experiments. We thank the *E. coli* Genetic Stock Center (http://cgsc.biology.yale.edu) to generously provide *E. coli* HT414 strain. We also thank D. Sela for correcting the English version of the manuscript. The technical assistance of M.V. Santamaría and A. Gómez is greatly appreciated.

REFERENCES

1. **Arrecubieta, C., R. López, and E. García.** 1994. Molecular characterization of *cap3A*, a gene from the operon required for the synthesis of the capsule of *Streptococcus pneumoniae* type 3: sequencing of mutations responsible for the unencapsulated phenotype and localization of the capsular cluster on the pneumococcal chromosome. J. Bacteriol. 176:6375-6383.
2. **Brüssow, H., A. Probst, M. Frémont, and J. Sidoti.** 1994. Distinct *Streptococcus thermophilus* bacteriophages share an extremely conserved DNA fragment. Virology **200:**854-857.
3. **De las Rivas, B., A. Marcobal, and R. Muñoz.** 2004. Allelic diversity and population structure in *Oenococcus oeni* as determined from sequence analysis of housekeeping genes. Appl. Environ. Microbiol. **70:**7210-7219.
4. **De las Rivas, B., A. Marcobal, A. Gómez, and R. Muñoz.** 2005. Characterization of IS*Lpl4*, a functional insertion sequence in *Lactobacillus plantarum*. Gene **363:**202-210.
5. **Dobrindt, U., B. Hochnut, U. Hentschel, J. Hacker.** 2004. Genomic islands in pathogenic and environmental microorganisms. Nat. Rev. Microbiol. **2:**414-424.

6. **García-Moruno, E, A. V. Carrascosa, and R. Muñoz.** 2005. A rapid and inexpensive method for the determination of biogenic amines from bacterial cultures by thin-layer chromatography. J. Food Prot. **68:**625-629.
7. **Geiser, M., R. Cèbe, D. Drewello, and R. Schmitz**. 2001. Integration of PCR fragments at any specific site within cloning vectors without the use of restriction enzymes and DNA ligase. BioTechniques **31:**88-92.
8. **Igarashi, K., and K. and Kashiwagi, K.** 1999. Polyamine transport in bacteria and yeast. Biochem. J. **344:**633-642.
9. **Inouye, S., and M. Inouye.** 1985. Up-promoter mutations in the *lpp* gene of *Escherichia coli*. Nucleic Acids Res. **13:**3101-311.
10. **Johansen, E, and A. Kibenich.** 1992. Isolation and characterization of IS*1165*, an insertion sequence of *Leuconostoc mesenteroides* subsp. *cremoris* and other lactic acid bacteria. Plasmid **27:**200-206.
11. **Konings, W.N.** 2002. The cell membrane and the struggle for life of lactic acid bacteria. A. van Leeuw. **82:**3-27.
12. **Landete, J. M., S. Ferrer, L. Polo, and I. Pardo, I. 2005**. Biogenic amines in wine from three Spanish regions. J. Agric. Food Chem. **53:**1119-1124.
13. **Lucas, P. M., W. A. M. Wolken, O. Claisse, J. S. Lolkema, and A. Lonvaud-Funel.** 2005. Histamine-producing pathway encoded on an unstable plamsid in *Lactobacillus hilgardii* 0006. Appl. Environ. Microbiol. **71:**1417-1424.
14. **Marcobal, A., de las Rivas, B., Moreno-Arribas, M. V., and R. Muñoz.** 2004. Identification of the ornithine decarboxylase gene in the putrescine-producer *Oenococcus oeni* BIFI-83. FEMS Microbiol. Lett. **239:**213-220.

15. **Marcobal, A., P. J. Martín-Alvarez, M. C. Polo, R. Muñoz, and M.V. Moreno-Arribas**. 2006. Formation of biogenic amines during red wine manufacture. J. Food Prot. **70:**157-164.

16. **Mills, D. A., H. Rawsthorne, C. Parker, D. Tamir, and K. Makarova.** 2005. Genomic analysis of *Oenococcus oeni* PSU-1 and its relevance to winemaking. FEMS Microbiol. Rev. **29:**465-475.

17. **Molenaar, D., F. Bringel, F. H. Schuren, W. M. de Vos, R. J., Siezen, and M. Kleerebezem**. 2005. Exploring *Lactobacillus plantarum* genome diversity by using microarrays. J. Bacteriol. **187:** 6119-6127.

18. **Muñoz., R., R. López, M. de Frutos, and E. García.** 1999. First molecular characterization of a uridine diphosphate galacturonate 4-epimerase: an enzyme required for capsular biosíntesis in *Streptococcus pneumoniae* type 1. Mol. Microbiol. **31:**703-713.

19. **Sambrook, J., E. F. Fritsch, and T. Maniatis**. 1989. Molecular Cloning: a Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, HY.

20. **Tabor, H., C. W. Tabor, M. S. Cohn, and E. W. Hafner.** 1981. Streptomycin resistance (*rpsL*) produces an absolute requirement for polyamines for growth of an *Escherichia coli* strain unable to synthesize putrescine and spermidine [Δ(*speA-speB*) Δ*specC*]. J. Bacteriol. **147:**702-704.

Legends to Figures

FIG. 1. Genetic organization of the *O. oeni* RM83 17.2-kb DNA region containing the *odc* gene. Thick and thin arrows representing complete or interrupted ORFs, respectively. The localization of putative promoters (vertical bent arrow) and predicted transcriptional terminator regions (ball and stick) are indicated. Some of the plasmids used in this study are indicated, as are relevant restriction sites: *E*, EcoRV; *S*, Sau3AI; *Sp*, SpeI; *Sn*, SnaBI. Only some of the corresponding restriction sites present in this fragment are represented.

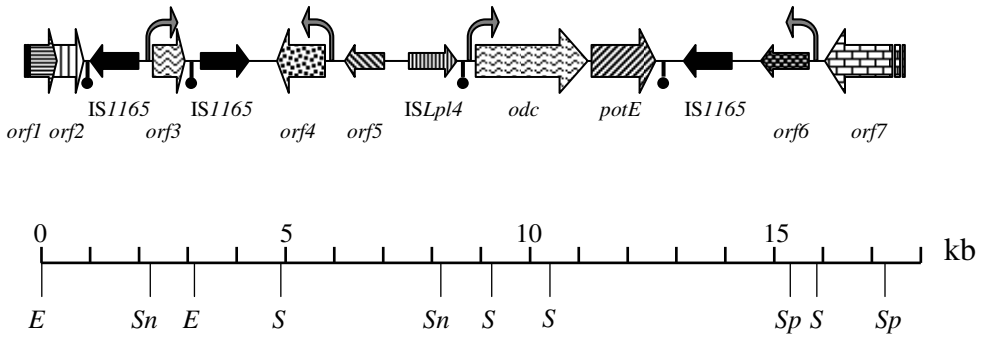
FIG. 2. Putrescine production by soluble cell extracts of IPTG-induced cultures of *E. coli* HT414 harboring pAM11. The putrescine produced during the enzymatic reaction was submitted to an automatic precolumn derivatization with *o*-ophthaldialdehyde prior to injection. Putrescine was determined by RP-HPLC as previously described (14). (A) Reaction from *E. coli* HT414 bearing the control pIN-III(lpp^{p}-5)A3 plasmid. (B) Reaction from *E. coli* HT414 bearing the recombinant pAM11 plasmid.

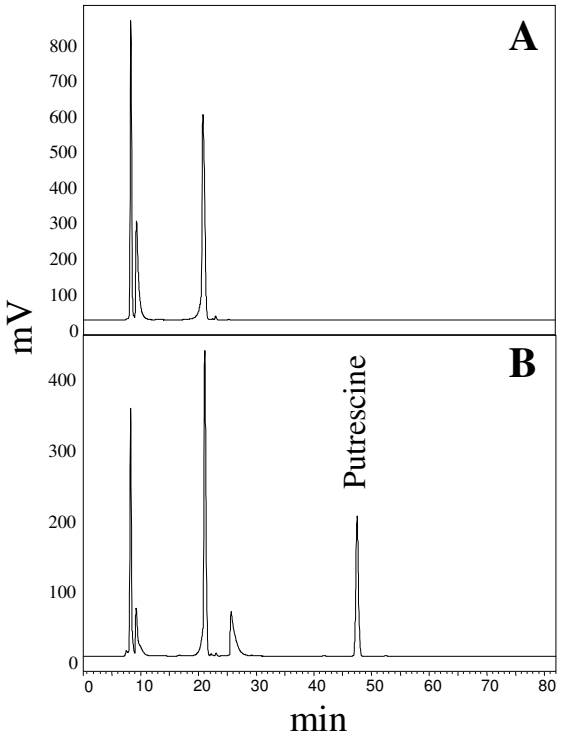
FIG. 3. Schematic overview of the sequence conservation between *O. oeni* PSU-1 and *O. oeni* RM83 chromosomal regions containing the proposed recombination site. Genes are represented as arrows. Thin arrow corresponds to interrupted gene. The genes present in these regions are indicated: *orf6* and *orf7* in *O. oeni* RM83, and *Ooen02001058* (encoding a putative transcriptional terminator), *Ooen02001059* (encoding the carbamoylphosphate synthase large subunit), and *Ooen02001060* (coding for a hypothetical protein ZP_00319317) in *O. oeni* PSU-1. The complete ORF coding for the hypothetical protein in *O. oeni* PSU-1 is also represented. Open squares and hexagons represent GW domains and

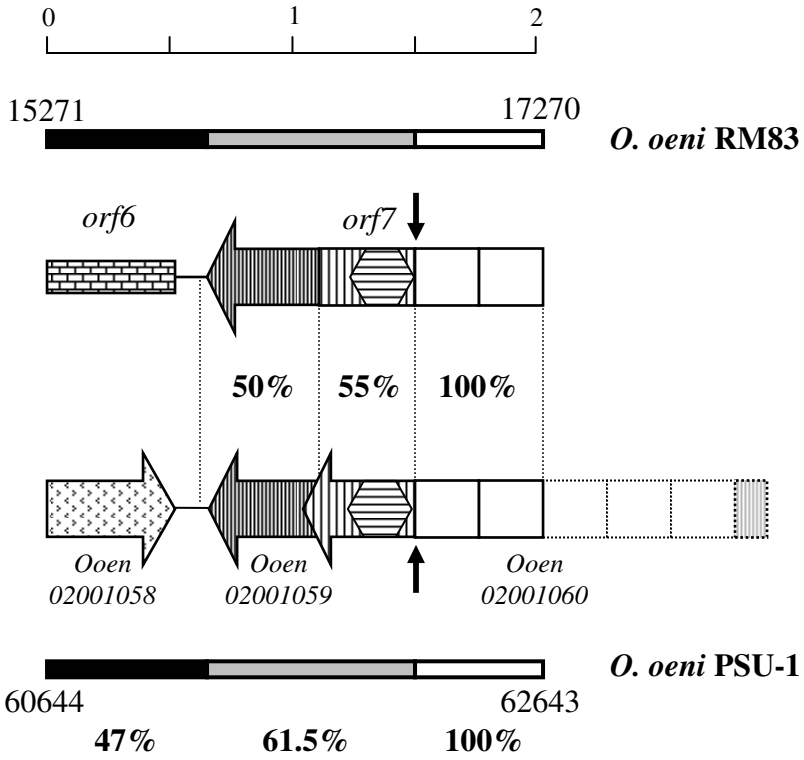
MucBP domains, respectively. ORF regions with identical shading correspond to regions having the same degree of sequence identity. The degree of amino acid identity between the protein fragments coded by these ORF is also shown. The color of the upper and lower bars indicates the degree of nucleotide identity between the 2 kb DNA regions: black, 47% identity; grey, 61.5%; and white, 100% identity. Two black arrows indicate the recombination site. The nucleotide positions corresponding to both sequences are also indicated. The *O. oeni* PSU-1 nucleotide sequence appears in the database under accession number NZ_AABJ03000005.

Table 1
odc region-encoded proteins: properties and similarities to proteins in the databases

Gene	Location in nucleotide sequence	G+C (%)	Predicted protein (aa/kDa)	Similar Polypeptide(s) (aa)	Proposed function	Database accession no.	Degree of identity (%)	Organism
orf1	< - 237	36.3	-	Str0776 (286)	Replication protein, phage-plasmid associated	Q5M098	32.5 (in 81 aa overlap)	*S. thermophilus* CNRZ 1066
			-	ORF12 (269)	Conserved protein in *S. thermophilus* phages	O34043	30.1 (in 63 aa overlap)	*S. thermophilus* temperate bacteriophage φ O1205
			-	ORF35 (271)	Hypothetical protein	Q9XJC8	30.5 (in 59 aa overlap)	*S. thermophilus* lytic bacteriophage DT1
orf2	230-784	32.8	184 / 20.5	EF1279	DNA replication protein	Q835U3	27.6	*Enterococcus faecalis* V583 putative prophage 02
IS*1165*	1175-2236/c	48.1	353 / 39.5	IS*1165* transposase	Transposase	Q48788	98.8 (in 335 aa overlap)	*Leuconostoc mesenteroides* subsp. *cremoris*
orf3	2502-3095/c	35.2	197 / 21.6	EfaeDRAFT_2583	Hypothetical protein	Q3Y3B9	55.7	*Enterococcus faecium* DO
				Spr0580	Hypothetical protein	Y580	55.6	*Streptococcus pneumoniae* R6
IS*1165*	3258-4319	48.1	353 / 39.5	IS*1165* transposase	Transposase	Q48788	98.8 (in 335 aa overlap)	*L. mesenteroides* subsp. *cremoris*
orf4	4767-5837/c	35.2	356 / 39.5	LJ1779	Major facilitator superfamily permease	Q74HG7	50.4	*Lactobacillus johnsonii* NCC533
orf5	6541-6927/c	46.0	128 / 13.6	lp_3570	Transposase	Q88S70	58.5	*Lactobacillus plantarum* WCFS1
IS*Lpl4*	7828-7953	43.2	41 / 4.7	IS*Lpl4* transposase	Transposase	CAI93853.1	97 (in 34 aa overlap)	*L. plantarum* CECT4645
odc	8866 - 11103	36.3	745 / 81	OdcI	Ornithine decarboxylase	P43099	67.1	*Lactobacillus* sp. strain 30a
potE	11184-12509	39.2	441 / 47.6	PotE	Putrescine-ornithine antiporter	P44768	67.3	*Haemophilus influenzae* Rd
IS*1165*	13235-14295/c	48.1	353 / 39.5	IS*1165* transposase	Transposase	Q48788	98.8 (in 335 aa overlap)	*L. mesenteroides* subsp. *cremoris*
orf6	15240-15800/c	30.4	186 / 22.0	BCE3325150	Probable membrane protein	Q630E9	36.8 (in 148 aa overlap)	*Bacillus cereus* ZK
orf7	15966 > /c	40.2	-	Ooen02001060	Hypothetical protein	ZP_00319317	78.2% (in 307 aa overlap)	*O. oeni* PSU-1


IS1165
IS1165
ISLpl4
odc
potE
IS1165
orf1
orf2
orf3
orf4
orf5
orf6
orf7
0
5
10
15
kb
E
Sn
E
S
Sn
S
S
Sp
S
Sp
pAM4
pAM8
pAM11


A
B
Putrescine
mV
min
800
700
600
500
400
300
200
100
0
400
300
200
100
0
0
10
20
30
40
50
60
70
80


0
1
2
15271
17270
O. oeni RM83
orf6
orf7
50%
55%
100%
Ooen 02001058
Ooen 02001059
Ooen 02001060
O. oeni PSU-1
60644
62643
47%
61.5%
100%