International Journal of Food Microbiology 141 (2010) 222–228

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Characterization of pLAC1, a cryptic plasmid isolated from *Lactobacillus acidipiscis* and comparative analysis with its related plasmids

Ioanna-Areti Asteri [a,b,1], Konstantinos Papadimitriou [a,b,*,1], Effrossyni Boutou [b], Rania Anastasiou [a], Bruno Pot [c], Constantinos E. Vorgias [b], Effie Tsakalidou [a]

[a] Laboratory of Dairy Research, Department of Food Science and Technology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece
[b] Department of Biochemistry and Molecular Biology, Faculty of Biology, National and Kapodistrian University of Athens, Panepistimioupolis-Zographou, 157 84 Athens, Greece
[c] Applied Maths B.V.B.A., B-9830 Sint-Martens-Latem, Belgium

ARTICLE INFO

Article history:
Received 22 February 2010
Received in revised form 17 May 2010
Accepted 18 May 2010

Keywords:
Lactobacillus
Cryptic plasmid
Operon
Mob
Rep
Modular evolution

ABSTRACT

The pLAC1 plasmid of *Lactobacillus acidipiscis* ACA-DC 1533, a strain isolated from traditional Kopanisti cheese, was characterised. Nucleotide sequence analysis revealed a circular molecule of 3478 bp with a G + C content of 37.2%. *Ab initio* annotation indicated four putative open reading frames (*orf*s). *orf*1 and *orf*4 were found to encode a replication initiation protein (Rep) and a mobilization protein (Mob), respectively. The deduced products of *orf*2 and *orf*3 revealed no significant homology to other known proteins. However, *in silico* examination of the plasmid sequence supported the existence of a novel operon that includes *rep*, *orf*2 and *orf*3 in pLAC1 and that this operon is highly conserved also in plasmids pLB925A02, pSMA23, pLC88 and pC7. RT-PCR experiments allowed us to verify that these three genes are co-transcribed as a single polycistronic mRNA species. Furthermore, phylogenetic analysis of pLAC1 Rep and Mob proteins demonstrated that they may have derived from different plasmid origins, suggesting that pLAC1 is a product of a modular evolution process. Comparative analysis of full length nucleotide sequences of pLAC1 and related *Lactobacillus* plasmids showed that pLAC1 shares a very similar replication backbone with pLB925A02, pSMA23 and pLC88. In contrast, *mob* of pLAC1 was almost identical with the respective gene of plasmids pLAB1000, pLB4 and pPB1. These findings lead to the conclusion that pLAC1 acquired *mob* probably via an ancestral recombination event. Our overall work highlights the importance of characterizing plasmids deriving from non-starter 'wild' isolates in order to better appreciate plasmid divergence and evolution of lactic acid bacteria.

1. Introduction

Lactobacillus acidipiscis is a recently described species originally isolated from fermented fish (Tanasupawat et al., 2000). It has also been identified in soy sauce mash (Tanasupawat et al., 2002). A later heterotypic synonym of *L. acidipiscis*, *L. cypricasei* (Naser et al., 2006), was also described as part of the flora of Halloumi cheese (Lawson et al., 2001). In a previous study we reported the isolation of *L. acidipiscis* from traditional Kopanisti cheese, prepared from raw cow's milk (Asteri et al., 2009). According to our analysis, the isolate exhibited some interesting biochemical properties, especially regarding the production of volatile compounds important for the formation of cheese aroma (Asteri et al., 2009), as well as a distinct halotolerant phenotype (unpublished results). The study of the physiology of such unexplored 'wild' lactic acid bacteria (LAB) is of significant value for the food industry, since it may lead to the development of novel starters or adjunct cultures.

During our investigation of the overall characteristics of *L. acidipiscis*, we also focused on its plasmid content. It has clearly been established that in LAB plasmid-encoded genes are often involved in technologically important traits (e.g. lactose/galactose utilization, proteolysis, oligopeptide transport, bacteriophage resistance, citrate utilization, bacteriocin production and stress responses) (Schroeter and Klaenhammer, 2009). Even the discovery of cryptic plasmids has found application in the construction of shuttle vectors. Furthermore, there are currently 76 lactobacilli plasmid sequences deposited in NCBI that have been isolated from only 18 out of more than 120 *Lactobacillus* species (http://www.ncbi.nlm.nih.gov/sites/entrez). Therefore, more research may be necessary in order to obtain a more concise picture of *Lactobacillus* plasmid diversity. In this paper we report the isolation, cloning and sequencing of plasmid pLAC1 from *L. acidipiscis* ACA-DC 1533. We also present the comparative analysis of pLAC1 with its closely related plasmids, deriving from other *Lactobacillus* species.

* Corresponding author. Laboratory of Dairy Research, Department of Food Science and Technology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece. Tel.: +30 210 529 4661; fax: +30 210 529 4672.
E-mail address: kpapadimitriou@aua.gr (K. Papadimitriou).
[1] Authors contributed equally to this work.

0168-1605/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2010.05.018

2. Materials and methods

2.1. Bacterial strains and growth conditions

L. acidipiscis strain ACA-DC 1533 was grown in MRS medium (Oxoid Ltd., Basingstoke, Hampshire, UK) at 30 °C. *Escherichia coli* strain Mach1™ (Invitrogen, Carlsbad, CA) was used for routine transformations and was cultured in LB medium at 37 °C under aerobic conditions. When needed, ampicillin (Sigma, St. Louis, MO) was added to the medium at 100 μg/mL final concentration.

2.2. Plasmid isolation and cloning procedures

Plasmid DNA from *L. acidipiscis* was extracted using the NucleoSpin Plasmid kit (Macherey-Nagel GmbH and Co. KG, Düren, Germany) according to the manufacturer's instructions for Gram-positive bacteria with some modifications, which included the addition of lysozyme (10 mg/mL) and mutanolysin (100 U/mL) at the lysis step (both from Sigma). Plasmid extracts were then subjected to 0.8% (w/v) agarose gel electrophoresis. The target plasmid was excised from the gel and extracted using the QIAEX II Gel Extraction Kit (Qiagen Inc., Valencia, CA). Purified plasmid pLAC1 was digested with *Hind*III (New England BioLabs Inc., Beverly, MA) and the two acquired fragments of approximately 1.6 and 2.0 kb were ligated into pUC18. General cloning procedures were performed as described previously (Sambrook et al., 1989). Sequencing reactions were performed by Macrogen Inc. (Seoul, Korea). The complete nucleotide sequence of pLAC1 was determined through primer-walking across the gaps.

2.3. Annotation and comparative analysis of pLAC1

The complete sequence of pLAC1 was assembled using the CAP3 program (Huang and Madan, 1999). *Ab initio orf* finding was performed using several bioinformatics tools, including heuristic GeneMark™ (Besemer and Borodovsky, 1999), FGENESB (www.softberry.com) and MetaGeneAnnotator (Noguchi et al., 2006). Only *orf*s receiving high score and/or *orf*s common within the prediction results of all programs were further analysed. Ribosome binding sites (RBS) were predicted by MetaGeneAnnotator, while putative promoter and terminator sequences were analysed with BPROM and FindTerm, respectively (www.softberry.com). The deduced information from the aforementioned tools was imported in Kodon software (Applied Maths, Sint-Martens-Latem, Belgium). Kodon was used for the construction of the pLAC1 map and for DNA secondary structure predictions. The clustering and comparison options of full length plasmids offered by the program were employed for the comparative analysis among pLAC1 and its related plasmids. For this analysis, all plasmid sequences and annotations were synchronized by assigning the A nucleotide of the ATG start codon of the *rep* gene as the plasmids' start point. Sequence similarity searches were carried out using the BLAST suite (Altschul et al., 1997).

2.4. Reverse transcription (RT)-PCR experiments

Total RNA was extracted from log phase *L. acidipiscis* cells (OD_{610} ~0.7) with the RNeasy and RNAprotect kits (Qiagen) according to the manufacturer's instructions. RT-PCR experiments were performed using the SuperScript™ II Reverse Transcriptase (Invitrogen) and the DyNAzyme™ EXT DNA Polymerase (New England BioLabs Inc.). Two hundred nanograms of total RNA served as template. The primers used were designed to amplify a 100 bp fragment or the *rep* gene (REP_F: 5′-AGTATGCTTTGGCTGGTAC-3′ and REP_R: 5′-ACTAAGTCACCCTTATCAAC-3′), a 90 bp fragment of *orf*2 (ORF2_F: 5′-AATGGTTTTGATGTTGAT-3′ and ORF2_R: 5′-TACTTTTCCCTTTTCAAT-3′) and a 156 bp fragment of *orf*3 (ORF3_F: 5′-TACGATGAGCATGACTTAG-3′ and ORF3_R: 5′-GACTTCCGCTTTAACAAC-3′). Using the REP_F primer of the *rep* gene and the ORF3_R of *orf*3, we investigated whether all three genes are co-transcribed. In all cases, PCR reactions on appropriate negative controls for residual DNA contamination were performed.

2.5. Phylogenetic analysis

Multiple sequence alignments were built with MUSCLE (Edgar, 2004) and visualised with Jalview (Waterhouse et al., 2009). Only best match protein sequences according to BLAST analysis (*e*-value $<1.0e^{-10}$) were included in the alignments. Phylogenetic analysis was performed at Phylogeny.fr platform (Dereeper et al., 2008). This phylogeny pipeline incorporates MUSCLE for multiple sequence alignment, Gblocks for an automated curation of the alignment (Castresana, 2000), PhyML for tree building (Guindon and Gascuel, 2003) and TreeDyn for tree rendering (Chevenet et al., 2006). Supporting values for branches in the phylogenetic tree were calculated with the approximate likelihood-ratio test (aLRT) (Anisimova and Gascuel, 2006).

2.6. Southern hybridization for single-stranded DNA (ssDNA) detection

Southern hybridization was used for the detection of specific ssDNA, as described previously (Leenhouts et al., 1991). In brief, *L. acidipiscis* was grown to log phase (OD_{610} ~0.7) with or without rifampicin (Sigma) and total DNA was extracted. Rifampicin was applied to the cultures at final concentration of 100 μg/mL for 45 min. Total DNA from culture without rifampicin treatment, was incubated with S1 nuclease (2500 U/mL; Fermentas, St Leon-Rot, Germany) at 37 °C for 30 min. DNA was electrophorized in a 0.8% (w/v) agarose gel and was blot transferred to a positively-charged nylon membrane (Roche Diagnostics GmbH, Penzberg, Germany) without denaturation treatment. The hybridization probe was a 499 bp fragment of the *rep* gene synthesized using the primers 5′-GCTAAGAAGTTCATTGCGAGTT-3′ (forward) and 5′-ACCCTTATCAACGTCCTC-3′ (reverse). The probe was labeled using the PCR DIG probe synthesis mix (Roche). Hybridization and detection were carried out with the chromogenic DIG Nucleic Acid Detection Kit (Roche), according to the manufacturer's instructions.

2.7. Real-Time PCR for the relative copy number determination

The relative plasmid copy number (PCN) was determined by real-time PCR (Q-PCR) as described previously (Skulj et al., 2008). As a reference, a 61 bp fragment of the NifU protein coding gene was amplified with the forward primer 5′-AGCTCTGACAGCAACTCCATCC-3′ and the reverse primer 5′-CGGTATCCATCCCCATGATCT-3′. This is a single-copy gene in the chromosome of *L. acidipiscis* ACA-DC 1533, as verified by Southern hybridization (data not shown). A 61 bp fragment of the pLAC1 *rep* gene was amplified with the forward primer 5′-GAGCCAGTTGTGCATGTTGAAA-3′ and the reverse primer 5′-GCAGAGAACCAGTACCTTTGCG-3′. Q-PCRs were carried out on a MX3005P (Stratagene, La Jolla, CA) using Power SYBR Green PCR master mix (Applied Biosystems, Foster City, CA), according to the manufacturer's instructions. The analysis was performed in three independent experiments.

2.8. Nucleotide accession number

pLAC1 nucleotide sequence along with its annotated features was deposited in the EMBL database under accession number FN667595.

3. Results and discussion

3.1. Sequence analysis and annotation of pLAC1

Fig. 1A shows the plasmid content of *L. acidipiscis* strain ACA-DC 1533. Sequence analysis of pLAC1 revealed a circular DNA molecule of 3478 bp with a G + C content of 37.2%. The physical and genetic map of pLAC1 is depicted in Fig. 1B. *Ab initio* analysis indicated the existence of four

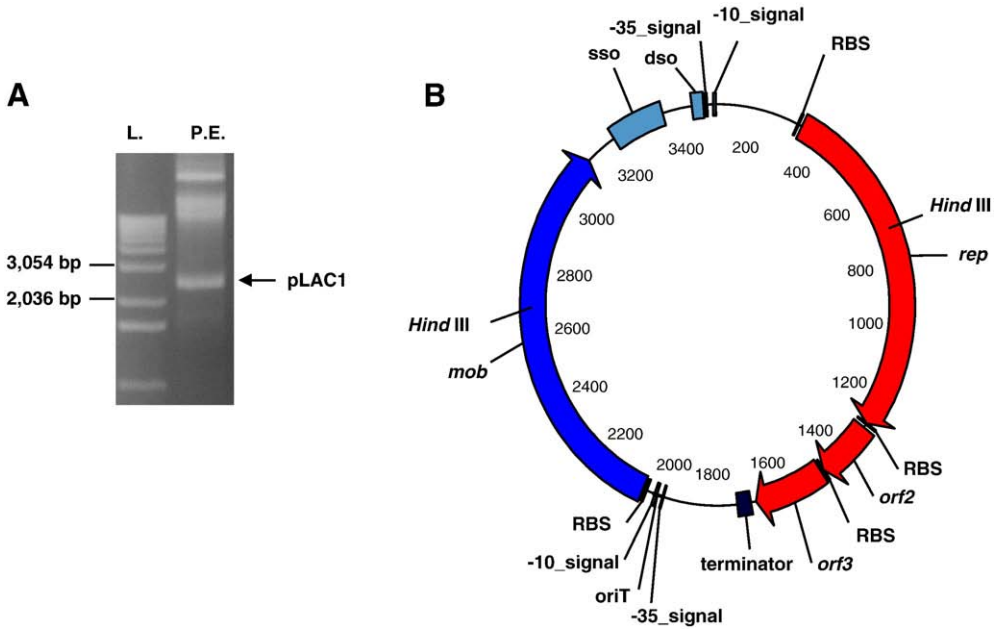

Fig. 1. Agarose gel electrophoresis of *L. acidipiscis* ACA-DC 1533 plasmid DNA (Lane P.E.). Lane L: 1 kb DNA ladder (A). Physical and genetic map of pLAC1 with its structural components (B).

putative *orf*s having the same orientation. *orf*1 and *orf*4 were found to encode a replication initiation protein (Rep) of 318 amino acids and a mobilization protein (Mob) of 362 amino acids, respectively, based on homology with previously characterized proteins found in other *Lactobacillus* plasmids. Putative promoter (-35 and -10) and RBS sequences were determined upstream the ATG start codons of both *rep* and *mob* genes (Fig. 1B).

*orf*2 and *orf*3 were located in a sequential fashion downstream the *rep* gene. Putative RBS were identified upstream of both *orf*s overlapping with the stop codon of the preceding *orf*, while no promoter regions could be detected. A potential terminator sequence was found downstream of *orf*3. Analysis with the FGENESB software of the *rep*, *orf*2 and *orf*3 region of pLAC1 predicted that they possess a typical operon structure. The deduced products of *orf*2 and *orf*3 revealed no significant homology to other known proteins after BLASTP searches. When TBLASTN queries were performed for both *orf* products, putative sequences were detected that could code for proteins homologous to Orf2 and/or Orf3 in the *Lactobacillus* plasmids pLB925A02 from *L. brevis* (Wada et al., 2009), pSMA23 from *L. casei* (Sudhamani et al., 2008), pLC88 from *L. casei* (GenBank Acc. No. U31333, direct submission) and pC7 isolated from *L. paraplantarum* (Park et al., 2004). In the original annotation of these plasmids, as deposited in GenBank, there was no reference of these two putative *orf*s. However, the existence of similar to *orf*2 and *orf*3 sequences has been reported for plasmids pSMA23 and pLC88 (Sudhamani et al., 2008). Previous analysis of these *orf*2 and *orf*3 sequences revealed frame-shift mutations and premature termination at nonsense codons for both *orf*s in the case of pSMA23, as compared to the respective sequences of pLC88 (Sudhamani et al., 2008). For that reason the authors argued that the putative proteins of *orf*2 and *orf*3 may not be functional for pSMA23. After re-annotating all afore-mentioned plasmids, we identified *orf*2 and/or *orf*3 in all cases and we were able to determine a very similar operon structure of *rep*, *orf*2 and *orf*3 for plasmids pLB925A02, pSMA23 and pLC88 with the exception of pC7, where *orf*3 was missing (Fig. 2A). In Fig. 2B, multiple sequence alignment shows that Orf2 protein was identical for pLAC1, pLB925A02 and pSMA23, while two conserved mutations were identified for pC7 (position 24: R to Q, position 45: V to I). Orf2 protein of pLC88 was found to be truncated at the first 12 amino acids of the N-terminal and the rest of the protein was almost identical to the other proteins with a conserved mutation (position 37: S to T). Orf3 protein was found identical in plasmids pLAC1, pLB925A02 and pSMA23, while the predicted product for pLC88 was truncated at the first 5 amino acids of the N-terminal end (Fig. 2C). Additionally, Orf3 of pLC88 had two non-conserved mutations (position 6: F to L and position 7: E to R) (Fig. 2C), the first of which is due to the alternative start codon TTG. In contrast to what has been previously suggested (Sudhamani et al., 2008), the afore-mentioned findings clearly indicate that Orf2 and Orf3 of pSMA23 follow the general consensus of these proteins and that Orf2 and Orf3 of pLC88 are both differentiated by being truncated at the N-terminal.

Our analysis strongly supported that *rep*, *orf*2 and/or *orf*3 exist as part of a functional operon in plasmids pLAC1, pLB925A02, pSMA23, pLC88 and pC7. To verify our predictions we performed RT-PCR experiments with pairs of primers specific for each of the three genes (Fig. 3). In all cases bands of the expected size were detected, proving that all three genes are expressed. In addition, using the distal primers REP_F and ORF3_R, we were able to amplify a 573 bp fragment that spans through these genes, validating that they are transcribed as a polycistronic mRNA species. It should be mentioned that no prediction can be currently drawn about Orf2 and Orf3 function, since they both lack similarity to other known proteins. The fact that they belong to an operon along with the Rep protein may lead to the hypothesis that they could somehow participate in the plasmid replication, a process whose control for pLAC1 related plasmids is still unknown (del Solar et al., 1998; Wada et al., 2009). Nevertheless, this hypothesis remains to be experimentally confirmed.

Further inspection of the pLAC1 nucleotide sequence revealed a number of important *cis*-acting elements. A putative double-strand origin (dso) consisting of 35 nt upstream the *rep* gene, that was highly similar to that of plasmids pLB925A02, pSMA23 and pLC88, was identified. The inverted repeat detected in this region with its stem-loop structure had a calculated free energy of −1.8 kcal/mol and is presented in Fig. 4A. In addition, a single-strand origin (sso), showing several inverted repeats ($\Delta G = -68.6$ kcal/mol) was predicted upstream the dso, as has been reported for plasmids pSMA23 and pLC88 (Fig. 4B). Finally, a putative oriT sequence (32 bp) was detected upstream of *mob* that was almost identical with that of the pSMA23 plasmid. The oriT consists of an inverted repeat ($\Delta G = -6.8$ kcal/mol), the loop of which carries the nick site (TGA′G) (Fig. 4C).

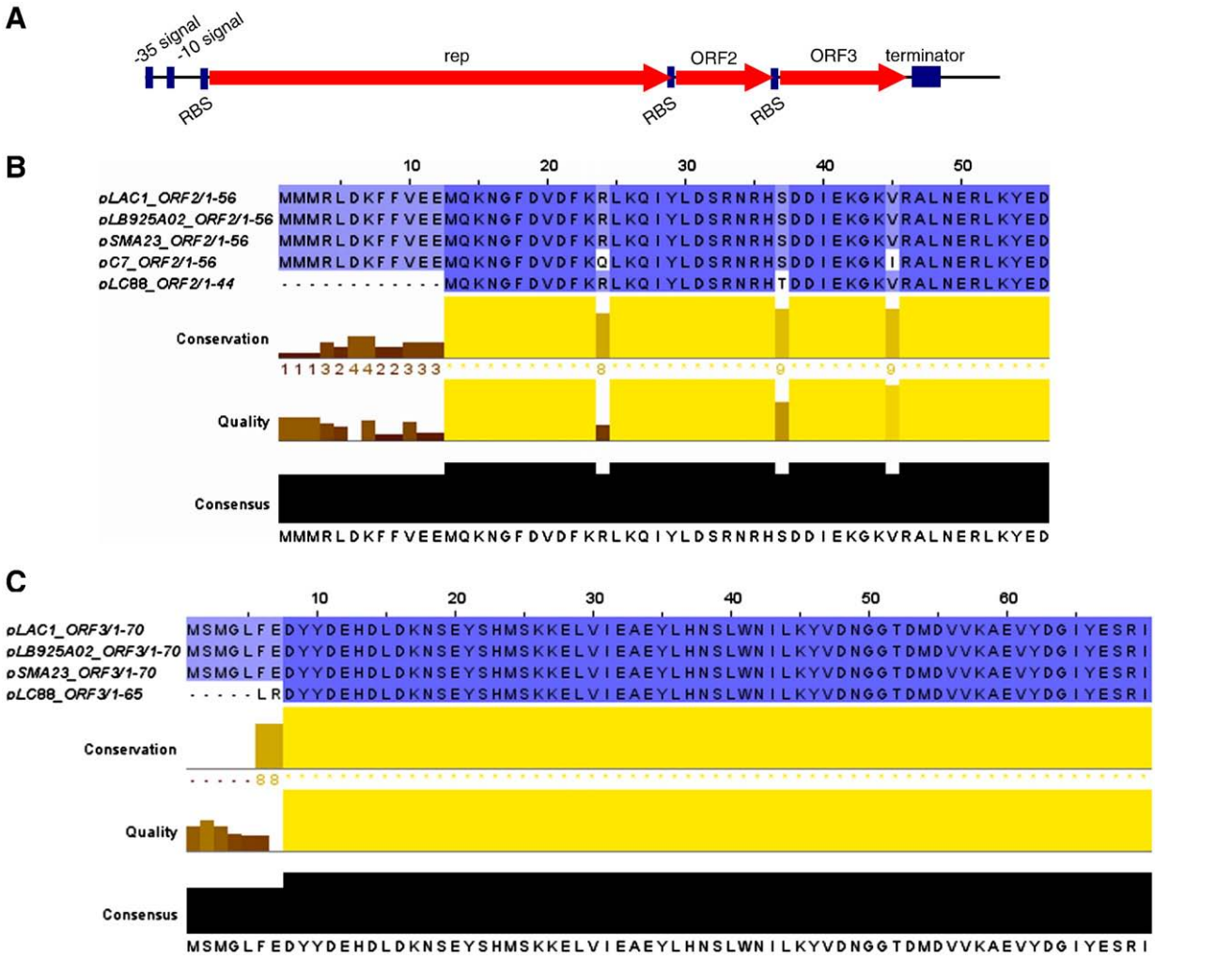

Fig. 2. Predicted operon structure of the *rep*, *orf*2 and *orf*3 genes (A). Multiple sequence alignment of *orf*2 (B) and *orf*3 (C) deduced products and their related predicted proteins as obtained by MUSCLE and visualized by Jalview.

3.2. Comparative analysis of pLAC1 and related plasmids

BLASTP queries for the pLAC1 Rep protein revealed 100% identity with the Rep protein of pSMA23. The phylogenetic analysis of the most similar Rep proteins clustered the pLAC1 Rep with the respective proteins of pSMA23, pLAB925A02 and pLC88 (Fig. 5A). All these plasmids have been previously characterised as belonging to the pC194 family of rolling-circle replicating (RCR) plasmids. Homology of pLAC1 Rep protein and dso sequence suggested that this plasmid also belongs to the pC194 family (Khan, 1997). Multiple sequence alignment of the pLAC1 Mob protein with its related proteins according to BLASTP homologies, revealed the presence of three conserved motifs (Motif I: GxxxHxxR; Motif II: NYD/EL; Motif III: HxDExxPHMHxGxxP) and thus classified Mob in clade A of the pMV158-superfamily of relaxases (data not shown) (Francia et al., 2004). The phylogenetic tree of Mob showed that unlike the Rep protein, Mob had maximum relatedness to those from plasmids pLAB1000 isolated from *L. hilgardii* (Josson et al., 1990) and pLB4, pPLA4 and pPB1, all isolated from *L. plantarum* (Bates and Gilbert, 1989; de las Rivas et al., 2004; Van Reenen et al., 2006) (Fig. 5B). Interestingly, pLB4 and pPB1 belong to the pE194 family of RCR plasmids, pPLA4 has been suggested to be a theta-replicating plasmid while only pLAB1000 belongs to pC194 family. Our findings therefore demonstrate two different patterns of similarity for pLAC1 Rep and Mob proteins, strongly indicating that these two plasmid regions may have originated from different sources as a result of a modular evolution process.

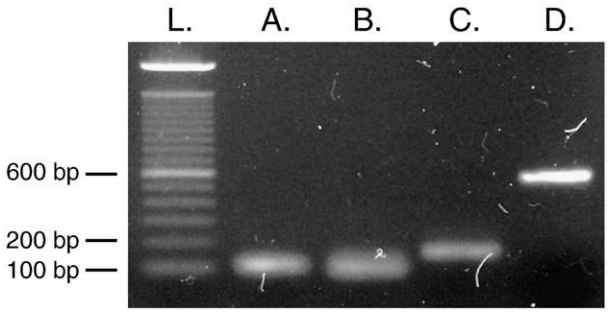

Fig. 3. RT-PCR experiments to validate the existence of the *rep*–*orf*2–*orf*3 operon. Lane L: 100 bp DNA ladder. Lanes A to D: linear internal fragments of *rep*, *orf*2, *orf*3 and *rep*–*orf*2–*orf*3 operon.

To further elucidate this observation, we investigated full length plasmid DNA sequence similarities. A comparative approach was followed so as to examine the relationship of pLAC1 with pSMA23, pLB925A02 and pLC88, whose Rep protein showed high similarity to pLAC1 Rep, as well as the pLAB1000, pLB4 and pPB1 whose Mob protein was homologous to pLAC1 Mob. We concentrated our analysis on plasmids with a similar gene content and molecule size (~3.5 kb, 2–4 genes) and excluded pPLA4 since it differed in both features (i.e. size of ~8 kb, 6 genes). Dot plot matrix was used to visualize highly similar regions in and between the plasmids (data not shown). This analysis identified distinct similarities as well as differences to the plasmids

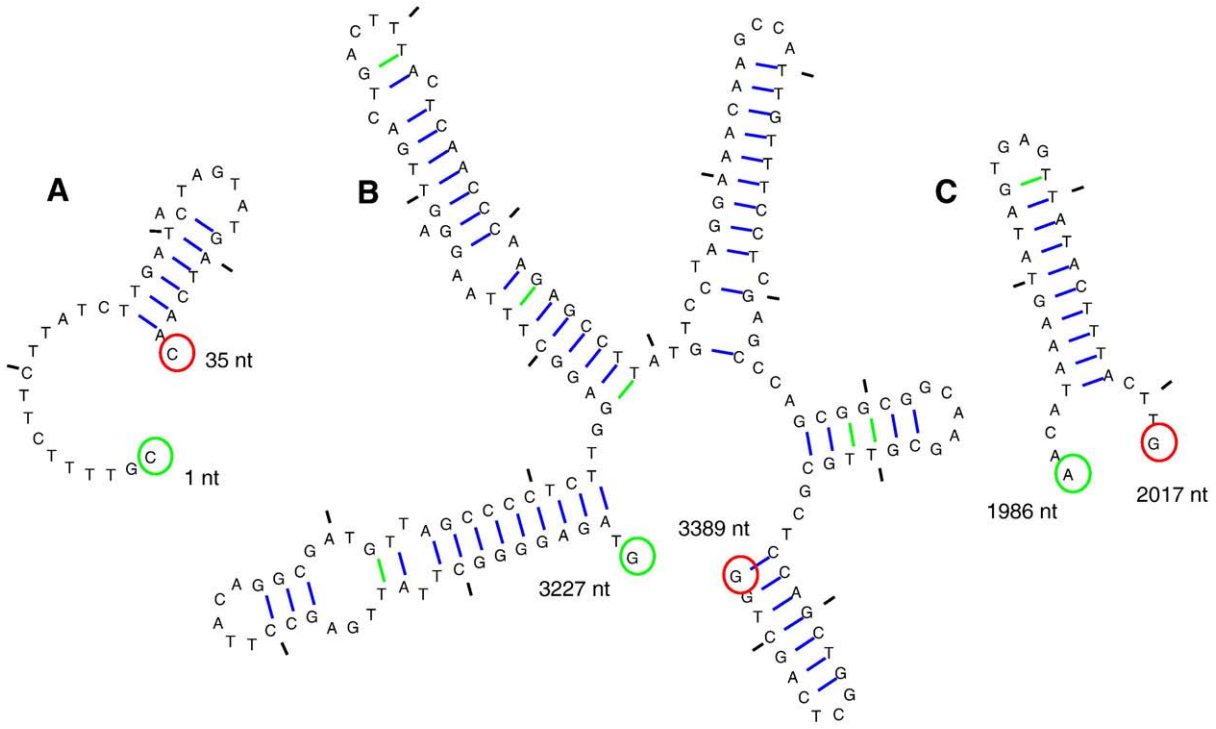

Fig. 4. Predicted secondary DNA structures of putative dso (A), sso (B) and oriT (C), as calculated by Kodon software. The inverted repeats and their stem-loops are displayed as arms.

architecture. Extensive sequence identity was restricted to pLAC1 Rep encoding region (including the dso) with pLB925A02, pSMA23 and pLC88 (in all cases ≥97%). On the contrary, pLAC1 Mob encoding region was proven almost identical solely with the pLAB1000, pLB4 and pPB1 corresponding region (in all cases ≥96% identity). It is worth noting, that while pLB925A02, pSMA23 and pLC88 were very similar overall, plasmids pLAB1000, pLB4 and pPB1 shared a stretch of homology only in the *mob* region. Fig. 6 shows the UPGMA (Unweighted Pair Group Method using Arithmetic averages) clustering of the plasmids, based upon the pairwise identity scores. In addition, the full length alignment of synchronised sequences displays stretches of homology for the coding and non-coding regions, using pLAC1 as the guiding sequence. As expected, the first cluster including pLAC1, pLB925A02, pSMA23 and pLC88 showed high sequence identity (≥97%) for the first 1450 nucleotides, corresponding to the *rep*, *orf*2 and *orf*3 operon. In the *mob* region only two well defined similar segments of approximately 150–200 bp were detected (≥80% identity). These segments may reflect a certain degree of conservation at the nucleotide level since all these proteins belong to the pMV158 superfamily of relaxases (Francia et al., 2004). A nearly identical region, variable in length, was also found at the dso and its downstream sequence. Interestingly, the second cluster exhibited identity (≥96%) strictly in the *mob* gene of pLAB1000, pLB4 and pPB1 with a full overlap of oriT only between pLAC1 and pLAB1000.

Our data demonstrate that pLAC1 shares a very similar replication backbone with plasmids pLB925A02, pSMA23 and pLC88 and that it has acquired the *mob* gene from a different source probably via an ancestral recombination event. In contrast, the replication backbone of pLAB1000, pLB4 and pPB1 do not show any similarity amongst them nor with pLAC1. However, the fact that the Mob protein of these plasmids shares high degree of homology, also extended at the gene level, supports that not only pLAC1 but also pLAB1000, pLB4 and pPB1 have separately acquired *mob* into their sequence and that this acquisition most likely derived from a common origin.

Our findings are in accordance with previous reports. In detail, pPB1 has been suggested to be a chimera consisting of two modules from pC1411 in the *rep* region and from pLB4 in the *mob* region (de las Rivas et al., 2004). In fact, this *mob* modular element has been previously proposed to be shared among plasmids pLAB1000 and pPB1, including the most divergent theta-replicating pPLA4 (Van Reenen et al., 2006). Furthermore, pSMA23 and pLC88 have been shown to be almost identical, although having two different origins of *mob* and *rep* (Sudhamani et al., 2008). In addition, pRS4 may have originated from pLC88 and pWCFS101 (Alegre et al., 2005).

3.3. Detection of ssDNA by southern hybridization

The generation of ssDNA intermediates, a distinct feature of RCR plasmids (del Solar et al., 1998), was confirmed by the positive signal observed in the S1 nuclease untreated DNA sample that disappeared after digestion with S1 nuclease (data not shown). When the bacterium was grown in rifampicin supplemented medium, the positive signal became much stronger. This clearly demonstrates that the inhibition of host RNA polymerase by this antibiotic resulted in the accumulation of ssDNA by preventing their conversion to dsDNA during plasmid replication, verifying the RCR mechanism of pLAC1.

3.4. Copy number determination of pLAC1

The relative PCN was determined using the equation: $PCN = (Ec)^{Ctc}/(Ep)^{Ctp}$, considering different amplification efficiencies (E) and Ct values for the two amplicons (chromosome-c and plasmid-p) (Skulj et al., 2008). Decimal serial dilutions of total DNA from *L. acidipiscis* were used to construct standard curves for both *rep* and *nifU* (data not shown). The generated standard curves were linear (R^2>0.99) in the range tested and their slopes were 3.451 and 3.412, respectively. The estimated PCN was found to be about 39 copies per cell. It should be mentioned that the segregational and physical stability of a pLAC1-like replicon, i.e. the pSMA23 plasmid has been previously investigated and two vectors based on the pSMA23 sequence have been constructed for the heterologous expression of proteins in Lactobacilli (Sudhamani et al., 2008). During this study *E. coli–Lactobacillus* shuttle vectors carrying the *rep* gene and the ori of pSMA23 (e.g. pL11) could efficiently replicate in other *Lactobacillus* species, like *L. casei* and *Lactobacillus gasseri*. The stability values for pL11 were lower than those reported for pWCFS101

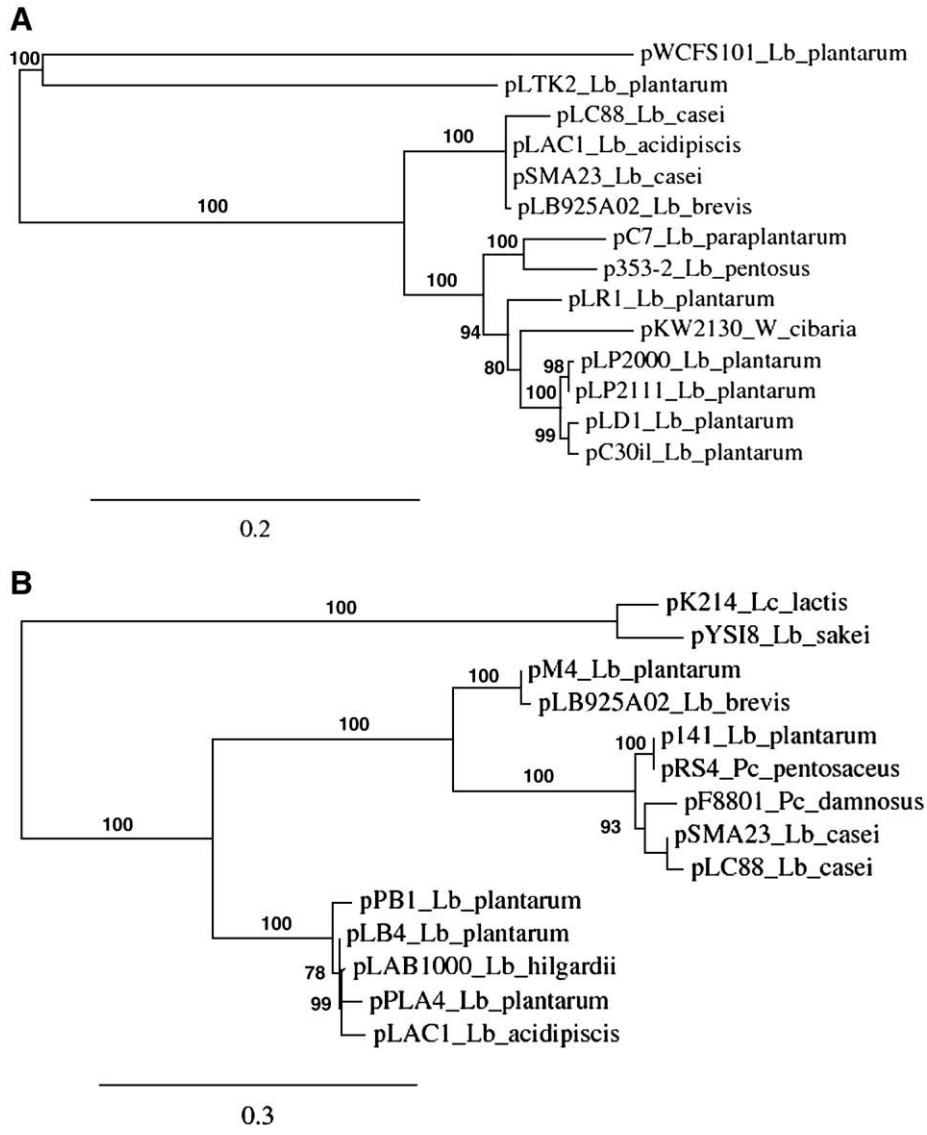

Fig. 5. Phylogenetic tree of the Rep (A) and the Mob (B) protein constructed with PhyML (maximum likelihood). Statistical validation for branch support (%) was conducted via a Chi^2-based parametric approximate likelihood-ratio test (aLRT) using the WAG substitution model.

derivatives but higher than those reported for the pLTK2-replicon and thus considered satisfactory for biotechnological purposes (Sudhamani et al., 2008). The authors also concluded that the *mob* gene is not essential for replication in *Lactobacillus*.

4. Conclusion

In this study we reported the characterization of the new RCR plasmid pLAC1 isolated from *L. acidipiscis* ACA-DC 1533. *In silico* evidence

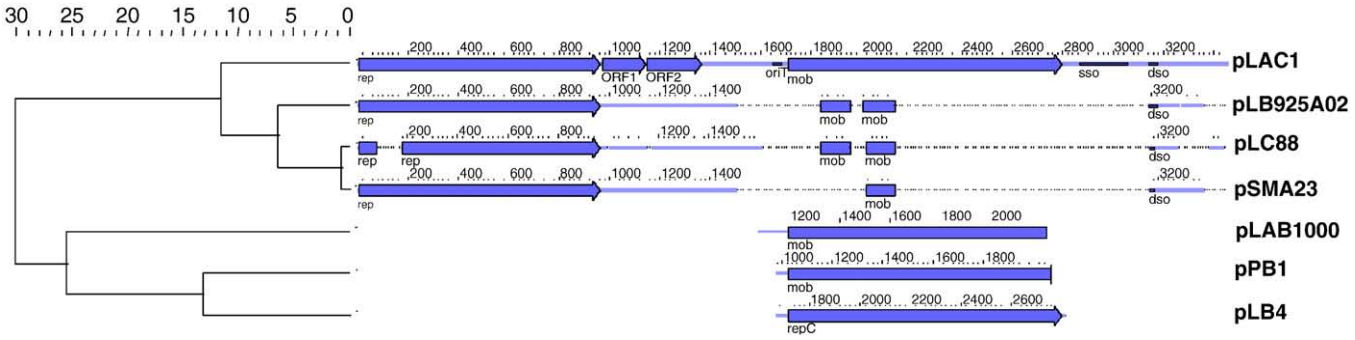

Fig. 6. Comparative analysis of pLAC1 and related plasmids. Full length alignment and clustering of synchronized plasmid sequences with annotation features. The clustering of the entries is based upon the pairwise identity scores and UPGMA was the clustering algorithm used. To facilitate comparison the plasmids were rearranged with respect to the GenBank entry and were all synchronised having the nucleotide A of the *rep* ATG codon as the start point. Corresponding genes are marked with the same colour. pLAC1 plasmid was used as guiding sequence and the multi-plasmid alignments were set up by pointing out this sequence as guiding reference and aligning the other sequences against this reference.

suggested the existence of a novel operon consisting of *rep*, *orf*2 and *orf*3 not only in pLAC1 but also in plasmids pSMA23, pLC88, pLB925A02 and pC7. We experimentally verified with RT-PCR that this operon truly exists and that these three genes are transcribed polycistronically. Our comparative analysis of pLAC1 along with its closely related *Lactobacillus* plasmids revealed evolutionary links among them and highlighted modular evolution as a mechanism for plasmid divergence in LAB. The proper identification of Mob as a 'modular unit' is important since it has been proposed that the mobilization regions may be very useful in an improved classification scheme of mobilizable plasmids, surpassing current criteria like Rep and dso homologies (Francia et al., 2004). Even though there is a clear tendency in the literature to identify putative donors or acceptors of modular elements, in our opinion such an attempt is seriously influenced by the availability of fully sequenced plasmids at the time of analysis and thus should be taken cautiously. For example, in contrast to what has been previously reported (Sudhamani et al., 2008), our results indicate that pSMA23 and pLC88 may not be modularly assembled, but they rather serve as donors. It is our strong feeling that in order to obtain a more accurate view of a plasmid's modular evolution, more plasmids need to be characterised.

Acknowledgements

Ioanna-Areti Asteri wishes to thank the State Scholarships Foundation of Greece (IKY-Idryma Kratikon Ypotrofion) for financial support.

References

Alegre, M.T., Rodriguez, M.C., Mesas, J.M., 2005. Nucleotide sequence, structural organization and host range of pRS4, a small cryptic *Pediococcus pentosaceus* plasmid that contains two cassettes commonly found in other lactic acid bacteria. FEMS Microbiol. Lett. 250, 151–156.

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

Anisimova, M., Gascuel, O., 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. Syst. Biol. 55, 539–552.

Asteri, I.-A., Robertson, N., Kagkli, D.-M., Andrewes, P., Nychas, G., Coolbear, T., et al., 2009. Technological and flavour potential of cultures isolated from traditional Greek cheeses — a pool of novel species and starters. Int. Dairy J. 19, 595–604.

Bates, E.E., Gilbert, H.J., 1989. Characterization of a cryptic plasmid from *Lactobacillus plantarum*. Gene 85, 253–258.

Besemer, J., Borodovsky, M., 1999. Heuristic approach to deriving models for gene finding. Nucleic Acids Res. 27, 3911–3920.

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

Chevenet, F., Brun, C., Banuls, A.L., Jacq, B., Christen, R., 2006. TreeDyn: towards dynamic graphics and annotations for analyses of trees. BMC Bioinform. 7, 439.

de las Rivas, B., Marcobal, A., Munoz, R., 2004. Complete nucleotide sequence and structural organization of pPB1, a small *Lactobacillus plantarum* cryptic plasmid that originated by modular exchange. Plasmid 52, 203–211.

del Solar, G., Giraldo, R., Ruiz-Echevarria, M.J., Espinosa, M., Diaz-Orejas, R., 1998. Replication and control of circular bacterial plasmids. Microbiol. Mol. Biol. Rev. 62, 434–464.

Dereeper, A., Guignon, V., Blanc, G., Audic, S., Buffet, S., Chevenet, F., et al., 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. 36, W465–W469.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Francia, M.V., Varsaki, A., Garcillan-Barcia, M.P., Latorre, A., Drainas, C., de la Cruz, F., 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. 28, 79–100.

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

Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res. 9, 868–877.

Josson, K., Soetaert, P., Michiels, F., Joos, H., Mahillon, J., 1990. *Lactobacillus hilgardii* plasmid pLAB1000 consists of two functional cassettes commonly found in other Gram-positive organisms. J. Bacteriol. 172, 3089–3099.

Khan, S.A., 1997. Rolling-circle replication of bacterial plasmids. Microbiol. Mol. Biol. Rev. 61, 442–455.

Lawson, P.A., Papademas, P., Wacher, C., Falsen, E., Robinson, R., Collins, M.D., 2001. *Lactobacillus cypricasei* sp. nov., isolated from Halloumi cheese. Int. J. Syst. Evol. Microbiol. 51, 45–49.

Leenhouts, K.J., Tolner, B., Bron, S., Kok, J., Venema, G., Seegers, J.F., 1991. Nucleotide sequence and characterization of the broad-host-range lactococcal plasmid pWVO1. Plasmid 26, 55–66.

Naser, S.M., Vancanneyt, M., Hoste, B., Snauwaert, C., Swings, J., 2006. *Lactobacillus cypricasei* Lawson et al. 2001 is a later heterotypic synonym of *Lactobacillus acidipiscis* Tanasupawat et al. 2000. Int. J. Syst. Evol. Microbiol. 56, 1681–1683.

Noguchi, H., Park, J., Takagi, T., 2006. MetaGene: prokaryotic gene finding from environmental genome shotgun sequences. Nucleic Acids Res. 34, 5623–5630.

Park, W.J., Lee, K.H., Lee, J.M., Lee, H.J., Kim, J.H., Lee, J.H., et al., 2004. Characterization of pC7 from *Lactobacillus paraplantarum* C7 derived from Kimchi and development of lactic acid bacteria—*Escherichia coli* shuttle vector. Plasmid 52, 84–88.

Sambrook, J., Fritisch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, New York.

Schroeter, J., Klaenhammer, T., 2009. Genomics of lactic acid bacteria. FEMS Microbiol. Lett. 292, 1–6.

Skulj, M., Okrslar, V., Jalen, S., Jevsevar, S., Slanc, P., Strukelj, B., et al., 2008. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. Microb. Cell Fact. 7, 6.

Sudhamani, M., Ismaiel, E., Geis, A., Batish, V., Heller, K.J., 2008. Characterisation of pSMA23, a 3.5 kbp plasmid of *Lactobacillus casei*, and application for heterologous expression in *Lactobacillus*. Plasmid 59, 11–19.

Tanasupawat, S., Shida, O., Okada, S., Komagata, K., 2000. *Lactobacillus acidipiscis* sp. nov. and *Weissella thailandensis* sp. nov., isolated from fermented fish in Thailand. Int. J. Syst. Evol. Microbiol. 50 (Pt 4), 1479–1485.

Tanasupawat, S., Thongsanit, J., Okada, S., Komagata, K., 2002. Lactic acid bacteria isolated from soy sauce mash in Thailand. J. Gen. Appl. Microbiol. 48, 201–209.

Van Reenen, C.A., Van Zyl, W.H., Dicks, L.M., 2006. Expression of the immunity protein of plantaricin 423, produced by *Lactobacillus plantarum* 423, and analysis of the plasmid encoding the bacteriocin. Appl. Environ. Microbiol. 72, 7644–7651.

Wada, T., Noda, M., Kashiwabara, F., Jeon, H.J., Shirakawa, A., Yabu, H., et al., 2009. Characterization of four plasmids harboured in a *Lactobacillus brevis* strain encoding a novel bacteriocin, brevicin 925A, and construction of a shuttle vector for lactic acid bacteria and *Escherichia coli*. Microbiology 155, 1726–1737.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., Barton, G.J., 2009. Jalview version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189–1191.