Molecular characterization and structural instability of the industrially important composite metabolic plasmid pLP712

Udo Wegmann, Karin Overweg, Sophie Jeanson,† Mike Gasson‡ and Claire Shearman§

Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, UK

The widely used plasmid-free Lactococcus lactis strain MG1363 was derived from the industrial dairy starter strain NCDO712. This strain carries a 55.39 kb plasmid encoding genes for lactose catabolism and a serine proteinase involved in casein degradation. We report the DNA sequencing and annotation of pLP712, which revealed additional metabolic genes, including peptidase F, α-lactate dehydrogenase and α-keto acid dehydrogenase (E3 complex). Comparison of pLP712 with other large lactococcal lactose and/or proteinase plasmids from L. lactis subsp. cremoris SK11 (pSK11L, pSK11P) and the plant strain L. lactis NCDO1867 (pGdh442) revealed their close relationship. The plasmid appears to have evolved through a series of genetic events as a composite of pGdh442, pSK11L and pSK11P. We describe in detail a scenario by which the metabolic genes relevant to the growth of its host in a milk environment have been unified on one replicon, reflecting the evolution of L. lactis as it changed its biological niche from plants to dairy environments. The extensive structural instability of pLP712 allows easy isolation of derivative plasmids lacking genes for casein degradation and/or lactose catabolism. Plasmid pLP712 is transferable by transduction and conjugation, and both of these processes result in significant molecular rearrangements. We report the detailed molecular analysis of insertion sequence element-mediated genetic rearrangements within pLP712 and several different mechanisms, including homologous recombination and adjacent deletion. Analysis of the integration of the lactose operon into the chromosome highlights the fluidity of the MG1363 integration hotspot and the potential for frequent movement of genes between plasmids and chromosomes in Lactococcus.

INTRODUCTION

Lactococcus lactis, a fermentative lactic acid bacterium, is one of the most widely used bacterial species in industrial food fermentations, and has more recently become the focus of diverse and novel biotechnology applications (Hugenholtz, 2008; Morello et al., 2008). Originally plant-associated, dairy L. lactis strains evolved through a process of reductive evolution (Kelly et al., 2010). Losing a lot of the key functions needed to metabolize plant-derived carbohydrates along the way, they rely on lactose, the sole sugar found in milk, as their energy source. Furthermore, milk constitutes a new protein-rich biological niche lacking free amino acids, allowing L. lactis strains to become auxotrophic for several amino acids provided they maintain their extensive proteolytic system to compensate for their limited synthetic capabilities. For example, in the case of L. lactis subsp. cremoris MG1363, a plasmid-free derivative of the dairy starter strain NCDO712, six amino acids (glutamate, leucine, isoleucine, valine, histidine and methionine) are essential for growth (Jensen & Hammer, 1993).

It is therefore somewhat counterintuitive that the genes encoding essential functions for the survival of dairy lactococci in milk, such as lactose metabolism and casein degradation, are plasmid-encoded, but this probably illustrates the fact that these genes were acquired when the bacteria entered the dairy environment (Fallada et al., 2011). Over time other more plant niche-relevant genes,
such as the glutamate dehydrogenase gene (Tanous et al., 2005), heavy metal-resistance and metal transporter genes (Fallico et al., 2011), have been lost from these plasmids.

Industrial strains of *L. lactis* used for dairy fermentations often contain a complex complement of different plasmids (Gasson, 1983; Mills et al., 2006). Spontaneous rearrangements in the different plasmids made the assignment of phenotypes and the molecular analysis of gene transfer processes very difficult. The plasmid-borne nature of the lactose metabolism and/or proteinase determinants on large plasmids in certain *Lactococcus* strains is well established (Gasson, 1983; Maeda & Gasson, 1986; McKay et al., 1976).

These large plasmids are subject to spontaneous loss or frequent deletions (Bachmann et al., 2012; McKay & Baldwin, 1974), with obvious consequences for the industrial processes that they are involved in. All these genetic events involving plasmids led to a wide diversity of phenotypes among lactococci strains. As technological and sensorial attributes are strain-dependent, it is therefore essential to increase knowledge about the diversity of strains, as it may contribute to particular technological properties and performance (Taïbi et al., 2010). Lactose utilization and casein breakdown are widely shared among the lactococcal plasmids, but more particular properties relevant to starter performance such as exopolysaccharide biosynthesis (van Kraanenburg et al., 1999), phage resistance, bacteriocin production (Mills et al., 2006) and antibiotic resistance (Flórez et al., 2008) have also been described as plasmid-borne. Because of their industrial importance, the problems of plasmid instability and phenotype loss in lactococci have been extensively studied. Initial work with the commercial dairy culture strain NCDO712 was severely impaired by extensive structural instability within its genome. Lactose and proteinase genes were linked to a 55 kb plasmid, pLP712 (Gasson, 1983). The lactose catalobic operon has been subcloned and sequenced (van Rooijen & de Vos, 1990; van Rooijen et al., 1991) but otherwise no sequence data have been published. The DNA rearrangements of pLP712 occur spontaneously, as a result of plasmid-curing techniques [protoplast regeneration, acriflavine and heat treatments (Gasson et al., 1987)] or during gene transfer processes such as transduction and conjugation (Gasson et al., 1992). Whilst restriction analysis has revealed some aspects of the molecular rearrangements involved, their detailed analysis has been limited by the lack of DNA sequence information.

Here we report the complete pLP712 DNA sequence, and analyse in detail the different genetic events and molecular processes responsible for the creation of pLP712, its continued instability and the integration of its *lac* operon into the chromosome of *L. lactis*.

**METHODS**

**Bacterial strains, plasmids and media.** The strains used in this study came from the Institute of Food Research collection (see Table S1 available with the online version of this paper). *Escherichia coli* JM109 was grown at 37 °C in Luria–Bertani medium; when appropriate, 200 μg ampicillin ml⁻¹ was added. *L. lactis* strains were grown at 30 °C in M17 medium supplemented with either 0.5% (w/v) glucose or lactose. McKay’s indicator plates, containing 0.5% (w/v) glucose or lactose and bromocresol purple indicator, were used to analyse lactic acid production (Kondo & McKay, 1982).

**Molecular techniques and DNA isolations.** Total lactococcal plasmid DNA from *L. lactis* was prepared according to the protocol of Anderson & McKay (1983), and total lactococcal DNA was isolated with a Qiagen genomic kit and 100/G columns (Qiagen).

**Transformation of *L. lactis*.** *L. lactis* electrocompetent cells were prepared and transformed essentially as described by Gerber & Solioz (2007).

**pLP712 library construction, DNA sequencing and PCR amplification.** A random library of pLP712 DNA fragments was constructed in *E. coli*. Plasmid DNA was hydro-sheared, end-repaired and tailed with an extra adenosine. A-tailed DNA was subsequently sequenced using standard pUC/M13 forward and reverse primers. Sequencing reactions were carried out using the BigDye Terminator Ready Reaction Cycle Sequencing kit from PE Applied Biosystems, and samples analysed on an ABI 3100 DNA sequencing system. Remaining gaps were closed by primer walking on clones or PCR products spanning the respective gaps. PCRs were carried out using Phusion (Finnzymes) according to the manufacturer’s recommendations. PCR products were purified using the Wizard PCR Prep kit (Promega).

**Sequence assembly and bioinformatic analysis.** Sequences were assembled using the STADEN software package in combination with the Phred/Phrap software (http://www.phrap.org). Putative ORFs were identified with the ARTEMIS program (Carver et al., 2005). Similarity searches against the UniProt database were carried out using BLASTP (Altschul et al., 1997). Predicted coding sequences (CDSs) were manually reviewed and alterations made on the basis of the presence of potential ribosome-binding sites, sequence alignments and available data in the literature. Dot plots against other plasmid sequences were carried out using DOTTER (Sonnhhammer & Durbin, 1995).

**Analysis of pLP712 deletion derivative plasmids.** Spontaneous deletions occur in pLP712 during gene transfer. Transductants and transconjugants carrying a variety of derivative plasmids were analysed by restriction analysis to map the deletions (Gasson et al., 1987, 1992) and additional unpublished data were employed (data not shown). Based on this information, primers were designed flanking the deletion regions of particular pLP712 derivatives and used to amplify the respective regions using total DNA extracted from plasmid-carrying strains. Fragments of interest were amplified by standard PCR, using Phusion High Fidelity DNA polymerase (New England Biolabs) and purified with a QIAquick PCR Purification kit (Qiagen) or SureClean (Bioline). Nucleotide sequencing was performed as described above. Exact positions of the deletion junctions were deduced by sequence comparison with the pLP712 sequence.

**Bacteriophage isolation for transduction.** Bacteriophage lysates from lysogenic strains were prepared as follows. A 200 ml volume of GM17 broth was inoculated with 20 ml overnight culture and incubated at 30 °C to OD₆₀₀ 0.6. Cells were harvested by centrifugation (3800 g, 5 min) and the pellet was resuspended in 100 ml quarter-strength Ringer’s solution (full strength 130 mM NaCl, 6 mM KCl, 2 mM NaHCO₃, 2 mM CaCl₂). Induction was carried out by a 5 s exposure to UV. Irradiated and control cells were harvested by
centrifugation and then resuspended in 10 ml quarter-strength Ringer’s solution. The suspension was added to 40 ml GM17 broth and subsequently incubated at 30 °C for 3 h. The optical density of lysogenic cultures fell after 1.5–2 h, and crude lysates were centrifuged at 1500 g for 10 min and the supernatant was passed through a 0.45 μm pore-size membrane filter. Filtered lysate was stored at 4 °C.

**Transduction.** Bacteriophage infection of recipient cells was achieved by mixing equal volumes of lysate and recipient (100 μl). Calcium borogluconate was added to a final concentration of 100 mM. The mixture was then incubated at room temperature for 20 min. Transductants were selected for their ability to utilize lactose as a carbohydrate source and produce acid on McKay’s indicator agar (McKay et al., 1973). In the case of selection for erythromycin resistance, it was necessary to include 30 min incubation in non-selective broth prior to plating on media containing antibiotic. Transduction frequency is expressed as transductants per p.f.u., with plaque counts derived by plating dilutions of bacteriophage lysates on bacteriophage-cured strains of *L. lactis*.

## RESULTS AND DISCUSSION

### General sequence features of the lactococcal plasmid pLP712

Plasmid pLP712 is a circular molecule of 55 395 bp with an average GC content of 37.39 mol%, which is higher than the GC content of the host chromosome *L. lactis* (35.8 mol%). The sequence was annotated with base-pair 1 situated 95 bp downstream from lacR rather than at the ‘single’ SalI site within the proteinase gene as used in the restriction maps described in the literature. Note that there are two SalI sites, at positions 8310 and 8344, both within the proteinase gene but sufficiently close to be identified as a single SalI site on the basis of previous restriction mapping analysis. As a result, the pLP712 sequence and figures described in this paper are reversed when compared with the plasmid restriction maps reported in the literature (Gasson, 1983). Plasmid pLP712 comprises metabolic cassettes punctuated by regions that encode proteins homologous to known transposases, including complete insertion sequence (IS) elements with the potential to promote transposition and rearrangement (Fig. 1). It carries a total of 61 CDSs, 17 of which were classified as pseudogenes. The remaining 44 CDSs, many of which can be found on other lactococcal plasmids, fall mainly into four categories: lactose metabolism, transposition, proteolysis and plasmid replication (see Table S2).

### Metabolic functions encoded by pLP712

In the main, pLP712 combines two essential functions necessary for survival in a dairy environment, lactose metabolism and casein proteolysis, and contains several other genes with potential effects on the bacterium’s ability to grow in milk. The D-lactate dehydrogenase (D-Ldh) encoded by the dld gene on pLP712 is highly similar to that of *E. coli* (Fig. S1), and putative D-LdhS are only found in *L. lactis* plasmids pSK11P and pGdh442. Siezen et al. (2005) suggested that this is a clear case of horizontal transfer between Gram-negative and Gram-positive bacteria [similar to the ycdB gene transfer between *Lactococcus* and *Salmonella* (Bolotin et al., 2004)]. D-LdhS from Gram-positive (320–350 aa) and Gram-negative bacteria (*E. coli*, 571 aa) have different sizes and metabolic roles. Gram-positive D-LdhS are NADH-dependent and produce D-lactate during anaerobic sugar fermentation. Gram-negative D-LdhS are FAD-dependent but NADH-independent, and utilize D-lactate under aerobic conditions. If the pLP712 D-Ldh causes the reaction from pyruvate to D-lactate it could boost the cell’s ability to produce lactate without the NADH requirement, unlinking it from glycolysis. Siezen et al. (2005) have speculated that the lactococcal plasmid-encoded D-Ldh could be involved in D-lactate utilization under aerobic conditions, shifting the balance of fermentation end products towards formate and acetate, thereby reducing the fall in pH and providing more ATP during acetate formation. However, a pGdh442-carrying strain could not grow on D-lactate as sole carbon source in M17, and no expression of the dld gene was detected by Northern blotting or real-time quantitative RT-PCR (Tanous et al., 2007). Hence, there is no proof that the D-Ldh homologues found on pGdh442 and pSK11L are functional (Tanous et al., 2007; Siezen et al., 2005). The pLP712 dld gene product has homology (Table S2) with an α-ketoacid dehydrogenase (E3 lipoamide dehydrogenase; Lpd), which is part of a family of multicomplex enzymes with many roles in metabolism that includes pyruvate dehydrogenase, α-ketoglutarate dehydrogenase, acetoin dehydrogenase, branched-chain α-ketoacid dehydrogenases (and the glycine cleavage system), depending on the bacterial species (Hein & Steinbüchel, 1994; Smith et al., 2002; Krüger et al., 1994). These Lpd enzymes are redox-active disulphide flavoproteins with an NADH-binding domain within a larger FAD domain. The complex of three enzymes catalyses the oxidative decarboxylation of the α-keto acid, with NAD reduction and the formation of a CoA adduct of the substrate (e.g. acetyl-CoA in the case of pyruvate dehydrogenase). The number of Lpd genes varies between bacteria, and some Lpd proteins are specific and some are interchangeable between different enzyme complexes (Burns et al., 1989; Hein & Steinbüchel, 1994). Different Lpd proteins are active under aerobic and anaerobic conditions.

*L. lactis* MG1363 has chromosomal genes encoding the components of the pyruvate dehydrogenase complex (including E3). The effect of a potential Lpd (pLP712 dld) on lactococcal metabolism will depend on the conditions under which the gene is expressed, whether the Lpd is specific or interchangeable, and what redox conditions (e.g. NAD/NADH levels) activate or repress its activity. It is possible that the pLP712 dld Lpd could be involved in branched-chain amino acid catabolism as part of the breakdown of casein.

### Replication of pLP712

The replication region of pLP712 resembles that of pGdh442 (accession no. NC_009435), featuring a single...
replication origin, the RepA replication initiation protein and the partitioning proteins ParA and ParB, thus confirming the prediction of Tanous et al. (2007) that this is the reason for the plasmid incompatibility observed between these two plasmids. Located upstream of this region is a short CDS transcribed in the opposite direction, which has been annotated as repC on lactococcal plasmids pND324 (U44843.1) and pH003 (AF247159.1) (Christensson et al., 2001), and is also present but not annotated on pIL105 (AF116286.2) (Anba et al., 1995). We suspect that the repC gene product is not involved in the process of pLP712 replication but rather is a remnant of a genetic exchange with another plasmid, as it is usually encoded in an operon with a RepB protein.

**Evolution and genetic stability of pLP712**

Genomic instability and transposition in *L. lactis* are prominent features that are compatible with the recent acquisition of new genetic traits. The pLP712 plasmid appears to have evolved as a collection of metabolic genes that are relevant to growth in a milk environment, and were probably acquired by *L. lactis* as it moved from a plant to a dairy environment. This can still be seen in the modal codon usage or mode of genes encoded on pLP712. Rather than the genome-wide average, the mode describes the codon usage that matches most of the genes of the genome (Davis & Olsen, 2010). In multireplicon genomes one can calculate the specific mode for each replicon and compare...
them with each other. Using the software provided by Davis & Olsen (2011) we calculated the distance between the MG1363 chromosome and pLP712 modes as 0.2634, a significant distance in comparison with the mean value of 0.1642 (SD 0.0319) obtained for 100 equivalently sized randomized gene sets from the total gene pool (chromosome and plasmid genes). In fact the ‘match_to_genome_frequencies’ algorithm from the same group links the pLP712 mode to Streptococcus agalactiae, a lactose-metabolizing commensal bacterium colonizing the intestinal tract of a significant proportion of the human population. These findings could explain why the key functions for survival of L. lactis in a dairy environment are plasmid-encoded and hint at where they might have been acquired from.

In the main, pLP712 combines two essential functions necessary for survival in a dairy environment, lactose metabolism and casein proteolysis. In common with many other lactose and/or proteinase plasmids its metabolic cassettes are bounded by transposons. As a result, pLP712 is very fluid and is capable of undergoing substantial rearrangements and deletions, as observed by the instability of the lactose metabolism and casein proteolysis phenotypes in dairy starter strains. We carried out a detailed DNA sequence comparison with other lactococcal plasmids, which revealed that pLP712 can be considered a composite of three other lactococcal plasmids or relatives thereof. They could all be related to common ancestor plasmids and have undergone rearrangements and deletions, and exchanged DNA regions with other plasmids and the chromosome. More than 80% of pLP712 DNA can be attributed to pGdh442 (NC_009435) (Tanous et al., 2000) (Siezen et al., 2005), which carries the genes necessary for the proteolysis of caseins, and pSK11L (DQ149244) (Siezen et al., 2005), which carries the lactose metabolism genes. In the following description, these plasmids and their relatives are referred to by the name of the published representative. Although some genetic events have left clues behind by which they can be identified, the high degree of plasmid plasticity and the complexity of the rearrangement events prevent us from ascertaining the exact nature and sequence of genetic events that led to the formation of pLP712. Nevertheless, we are able to describe a highly likely scenario in Fig. 1.

Based on the fact that the replication genes of pLP712 are identical (99%) to the respective genes of pGdh442, we conclude that a pGdh442-like plasmid was the starting point for the creation of pLP712. First, an inversion (inversion I) occurred through homologous recombination between IS946 elements D and E, resulting in ORFs 36–38 being transcribed in the same direction as the replication genes (repA–parB). Because the sequence homology between pLP712 and pSK11P stretches from the right inverted repeat of the IS946 element E to the right inverted repeat of the IS946 element A (disregarding insertion II), we postulate that inversion I was followed by homologous recombination between two divergently orientated IS946 elements (A and E) located on a pGdh442 derivative and pSK11P, resulting in the integration of the prtP-carrying fragment into pGdh442 (insertion I). We speculate that as a result of this event the opp–pepO operon normally located in this area of pGdh442 was lost from the pLP712 ancestor, but due to its crucial function, selective pressure favoured progeny that had the operon integrated into the hotspot region of the L. lactis genome (Wegmann et al., 2007). This idea is supported by the observation that a 12.6 kb fragment with a pGdh442-like gene organization, showing 98% sequence identity stretching from pGdh442_18 to pGdh442_p23, appears to be integrated into the MG1363 hotspot region (Wegmann et al., 2007). The flanking chromosome regions of this fragment do not indicate homologous recombination as the mechanism behind this event, but this area of the genome shows signs of heightened plasticity, which could have destroyed the evidence for a homologous recombination event. Alternatively, this fragment could have been integrated into the chromosome through illegitimate recombination. Following the integration of the prtP-carrying fragment, an inversion involving the IS946 element E and the pseudo IS946 element F must have occurred, flipping the dld-containing region of the plasmid as a result (inversion II). The integration of the lac operon-containing fragment (insertion II) from a pSK11L-type plasmid concludes the macro events that led to the formation of pLP712. Again, there is no evidence for homologous recombination and we can only offer illegitimate recombination as an explanation.

Smaller rearrangements have taken place within the plasmid structures since these major events, e.g. a minor rearrangement of the pepF region in pSK11L that accounts for the small break in DNA homology between pLP712 and pSK11L in the pepF region (seen on Fig. 1), occurred after the generation of pLP712. BLAST results confirmed the high level (~82%) of sequence homology in the pepF1 regions of the lactose–proteinase plasmids pLP712, pLP763 and pSK11L and their homology with the pepF2 regions of the L. lactis strain MG1363, NCD0763 and SK11 chromosomes (data not shown). Nardi et al. (1997) were the first to identify pepF homologues: the almost identical pepF1p and pepF1c on plasmid pLP763 from strain NCD0763 and on the chromosome of L. lactis strain IL1403, as well as pepF2 on the L. lactis strain NCD0763 chromosome. In contrast to pepF1p and pepF1c, pepF2 is only 80% identical to the corresponding chromosomal copy of L. lactis IL1403. Based on the different identities, Nardi et al. (1997) concluded that pepF1p and pepF1c diverged later than the two L. lactis subspecies, indicating a recent horizontal gene transfer. They suggested that the pepF region moves between the chromosome and plasmids through IS-directed mobilization, and our results support this idea.

Analysis of pLP712 deletion plasmid derivatives

The pLP712 plasmid deletion sequence is segmented by 10 IS elements (two IS904, five IS946, one pseudo IS946, one
IS981 and one IS1077) and two resolvase genes. The fluidity of this modular structure is highlighted by its similarity to, and interrelationships with, pGdh442, pSK11L and pSK11P. We hypothesize that the plasmid regions encoding transposases are involved in deletion generation, as the different pLP712 plasmid derivatives, which arose spontaneously during protoplast curing and transduction experiments, are varied in size and structure and cannot be accounted for solely on the basis of RecA-dependent homologous recombination. The derivatives have been classified into four groups, depending on restriction analysis results and their lactose catabolism and proteinase phenotype: group I, lac+/prt+; group II, a and b lac+/prt−; group III, lac−/prt+; and group IV lac−/prt− (see Table S1, Fig. 2). These groups/phenotypes are independent of the type of experiment that they originated from.

Restriction mapping of plasmid deletion derivatives (Gasson, 1983; Gasson et al., 1987) (additional unpublished data used) combined with the pLP712 DNA sequence reported here was used to identify approximate deletion end points. Using primers flanking the deletion regions of respective pLP712 derivatives, these regions were amplified and the resulting products were sequenced. Using the sequence data obtained, we were able to define precise deletion junctions and establish the precise molecular processes involved in their generation. Homologous recombination between the IS946 C and pseudo F gave rise to 1393 (transductant, group I). The same genetic mechanism involving the two IS904 elements located on pLP712 is behind the creation of 1440 (protoplast cure, group III). Alternatively, the transposon can act as one fixed end and delete adjacent DNA. This has occurred from IS946 elements B and D. With respect to the IS946 D element, the adjacent deletion results in strain 1441 (group III). Adjacent deletion starting from the IS946 B element results in derivative plasmids in two different phenotypic groups: strain 1432 (protoplast cure, group II) and strains 1370, 1422 and 1426 (all protoplast cure, group IV) (Figs 1 and 2).

**Fig. 2.** pLP712 deletions and rearrangements. Examples of groups I, II, III and IV deletions are represented in a linear manner. The top line is the original arrangement in pLP712 and the bottom line is the arrangement following deletion. Arrows represent ORFs: blue, transposases; dark blue, tnp946, pale blue, tnp904; green, pin resolvase; purple, lactose operon; orange, hypothetical protein or pseudogene. Inverted repeats are shown in black.
Plasmid transduction by the bacteriophage φTP712

The _L. lactis_ bacteriophage φTP712 genome is approximately 42 kb, and is predicted to package its DNA via a headful mechanism (data not shown). Bacteriophage φTP712 facilitates the transduction of chromosomal as well as plasmid DNA from _L. lactis_ NCDO712 and its derivatives. Transduction of relatively small plasmids such as pCK1 and pIL253 occurs at a frequency of $2.1 \times 10^{-3}$–$2 \times 10^{-4}$ transductants per p.f.u. However, transduction of pLP712 with selection for lactose utilization takes place at a relatively low frequency of approximately $2 \times 10^{-7}$ transductants per p.f.u.. However, transduction of pLP712 with selection for lactose utilization takes place at a relatively low frequency of approximately $2 \times 10^{-7}$ transductants per p.f.u. (Table 1). If these transductants are subsequently used as donors in a second round of transduction experiments the frequency is increased by four orders of magnitude to $10^{-3}$ transductants per p.f.u. We have shown that plasmids isolated from lactose-positive transductants are smaller than the original pLP712, with a size similar to that of the φTP712 genome. We have also demonstrated that the pLP712 plasmid suffers spontaneous deletions through homologous recombination and adjacent deletion, involving IS elements in both cases. We also believe that these processes account for the primary generation of plasmid molecules that are appropriately sized for packaging and subsequent transduction. Restriction enzyme analysis of the transduced plasmids showed that they are all approximately 40 kb in size. Following the selection of these shorter molecules by first-round transduction, they are available for highly efficient packaging in subsequent transduction rounds, and their smaller size explains the elevated transduction frequency. There are two possible mechanisms for efficient transduction: either plasmids are pre-matched in size through deletion to the capacity of the phage head or a specific region is needed for the packaging process to take place.

The possibility that bacteriophage DNA has been integrated into the pLP712 plasmid during transduction seems unlikely on the basis of our detailed restriction and deletion junction mapping. BLAST comparison of φTP712 and pLP712 sequences shows the presence of many short stretches of homology (data not shown). It is more likely that there is a requirement for a specific region of pLP712 essential for bacteriophage packaging. This has been reported for high-frequency pX3 plasmid transduction by _Lactobacillus delbrueckii_ bacteriophage LL-H, where there is no evident homology or recombination between the phage and the plasmid (Ravin _et al._, 2006).

**Characterization of the chromosomally stabilized lactose operon**

_L. lactis_ strain FI5267 was derived by the transduction of a plasmid-free strain of _L. lactis_ with the genes determining lactose metabolism and contains a chromosomally stabilized lactose operon. The strain was further adapted for use as host strain for a food-grade self-selecting host/vector system based on _lacF_ complementation (MacCormick _et al._, 1995; Platteeuw _et al._, 1996) and to express chromosomally integrated heterologous genes under lactose-induced control, e.g. endolysins (Payne _et al._, 1996; Stentz _et al._, 2010). Previous studies have shown that integration of the lactose operon occurs in a specific region of the chromosome located on a 140 kb _Smal_ fragment (Swindell, 1992), but the exact location and the mechanism behind the lactose operon integration into the chromosome following transduction are unknown.

We hypothesized that a deletion derivative of pLP712 lacking the plasmid replicon had been transduced and subsequently became chromosomally integrated. Based on the pLP712 sequence information, we identified the two IS904 elements flanking the _lac_ operon region on pLP712 as very likely entry points for the genetic exchange with the chromosome. Homologous recombination between IS946 elements A and C or B and E (Fig. 1) can account for a suitably sized replicon-lacking pLP712 derivative carrying the _lac_ operon region flanked by the two IS904 elements. The MG1363 genome contains nine intact and one partial IS904 element, which can be considered target sites. In

**Table 1.** Transduction experiments and transfer of lactose metabolism and antibiotic resistance determinants

<table>
<thead>
<tr>
<th>Source of bacteriophage</th>
<th>Lactose plasmid (size in kb)</th>
<th>Other plasmids</th>
<th>Lactose-utilizing transductants</th>
<th>Chloramphenicol-resistant transductants</th>
<th>Erythromycin-resistant transductants</th>
</tr>
</thead>
<tbody>
<tr>
<td>FI10526</td>
<td>pLP712 (56)</td>
<td>None</td>
<td>$2.1 \times 10^{-7}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FI10527</td>
<td>pSH290 (37) (transduced)</td>
<td>None</td>
<td>$1.1 \times 10^{-3}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FI10528</td>
<td>pLP712 (56) (transduced)</td>
<td>None</td>
<td>$1.7 \times 10^{-7}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FI10529</td>
<td>pSH290 (37) (transduced)</td>
<td>None</td>
<td>$1.2 \times 10^{-3}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>FI10530</td>
<td>pMG393 (38) (transduced)</td>
<td>pCK1*, pIL253†</td>
<td>$6.4 \times 10^{-4}$</td>
<td>$2.1 \times 10^{-3}$</td>
<td>$9.1 \times 10^{-4}$</td>
</tr>
<tr>
<td>FI10531</td>
<td>pMG820 (23) (transduced)</td>
<td>pCK1*, pIL253†</td>
<td>$2.7 \times 10^{-4}$</td>
<td>$6.5 \times 10^{-4}$</td>
<td>$2.0 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

*Plasmid carrying chloramphenicol resistance.
†Plasmid carrying erythromycin resistance.
silico, Smal restriction of the MG1363 genome sequence narrowed the number of possible IS904 insertion sites for the integrated pLP712 plasmid DNA down to two fragments of 126 and 129 kb, the latter of which encompassed the integration hotspot of the MG1363 chromosome. Given the high frequency of integration events in this region of the genome, our efforts to establish the exact nature of the lac operon integration were focussed on this region. We designed PCR experiments based on our hypothesis, which, if proven correct, would generate amplicons covering the left and right junction of the integration site. Following PCR using a pair of primers, one of which hybridized to position 646143–646162 of the L. lactis chromosome and the other to position 3758–3778 of the pLP712 plasmid, a product of 4 kb was obtained for the left junction. Following the same approach, using primers hybridizing to position 649932–649951 of the chromosome and to position 41108–41127 of the pLP712 plasmid, a product of 4 kb could also be obtained for the right junction. Both products were completely sequenced and confirmed our hypothesis for lac operon integration (Fig. 3).

Conclusions

New genomic data can illuminate previous observations and enable them to be interpreted in a detailed molecular manner. The sequence of pLP712 and the analysis of its derivatives have revealed the modular structure of the plasmid. The plasmid has evolved as a collection of metabolic gene cassettes relevant to growth in a milk environment that are interspersed with IS elements. Plasmid pLP712 can be considered a composite of three other lactococcal plasmids, combining their relevant functions for survival in a dairy environment. The difference in the modal codon usage between pLP712 and its host supports the idea that dairy-relevant traits have been acquired through horizontal gene transfer. The processes leading to a more streamlined plasmid carrying the core functions, i.e. lactose metabolism and proteolysis initiation, were driven by IS elements either directly or indirectly through homologous recombination. The relative ease with which derivatives of pLP712 can be isolated can be attributed to the fact that this is still a very active process, whereby the movement of genetic material is not solely between lactococcal plasmids but also occurs between the plasmids and chromosomes in both directions. Our study of the industrially important plasmid pLP712 has reinforced the importance of mobile genetic elements and gene transfer mechanisms in the rearrangement, mixing and flux of genetic information in L. lactis.

ACKNOWLEDGEMENTS

This work was supported by a Biotechnology and Biological Sciences Research Council (BBSRC) Core Strategic Grant. S. J. was supported by EU Marie Curie Fellowship OXYRED O41161. We thank Régis Stentz and Hélène Falentin for critical reading of the manuscript.

REFERENCES


Edited by: P. W. O’Toole