PinR mediates the generation of reversible population diversity in *Streptococcus zooepidemicus*

Karen F. Steward, Tihana Harrison, Carl Robinson, Josh Slater, Duncan J. Maskell, Simon R. Harris, Matthew T. G. Holden and Andrew S. Waller

1Animal Health Trust, Kentford, Newmarket CB8 7UU, UK
2Royal Veterinary College, Hawkshead Lane, Hatfield AL9 7TA, UK
3Department of Veterinary Medicine, University of Cambridge, Madingley Road, Cambridge CB3 0ES, UK
4Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton CB10 1SA, UK

Opportunistic pathogens must adapt to and survive in a wide range of complex ecosystems. *Streptococcus zooepidemicus* is an opportunistic pathogen of horses and many other animals, including humans. The assembly of different surface architecture phenotypes from one genotype is likely to be crucial to the successful exploitation of such an opportunistic lifestyle. Construction of a series of mutants revealed that a serine recombinase, PinR, inverts 114 bp of the promoter of SZO_08560, which is bordered by GTAGACTTTA and TAAAGTCTAC inverted repeats. Inversion acts as a switch, controlling the transcription of this sortase-processed protein, which may enhance the attachment of *S. zooepidemicus* to equine trachea. The genome of a recently sequenced strain of *S. zooepidemicus*, 2329 (Sz2329), was found to contain a disruptive internal inversion of 7 kb of the FimIV pilus locus, which is bordered by TAGAAA and TTTCTA inverted repeats. This strain lacks pinR and this inversion may have become irreversible following the loss of this recombinase. Active inversion of FimIV was detected in three strains of *S. zooepidemicus*, 1770 (Sz1770), B260863 (SzB260863) and H050840501 (SzH050840501), all of which encoded pinR. A deletion mutant of Sz1770 that lacked pinR was no longer capable of inverting its internal region of FimIV. The data highlight redundancy in the PinR sequence recognition motif around a short TAGA consensus and suggest that PinR can reversibly influence the wider surface architecture of *S. zooepidemicus*, providing this organism with a bet-hedging solution to survival in fluctuating environments.

**INTRODUCTION**

The Gram-positive organism *Streptococcus equi* subsp. *zooepidemicus* (*S. zooepidemicus*) is the most frequently isolated opportunistic pathogen of horses, associated with respiratory disease in young horses (Lindahl et al., 2013; Velineni et al., 2014; Wood et al., 1993, 2005) and uterine infections in mares (Hong et al., 1993; Rasmussen et al., 2013; Smith et al., 2003). The bacterium is also associated with disease in a wide range of other animal hosts including dogs (Abbott et al., 2010; Chalker et al., 2003; Pesavento et al., 2008) and humans (Abbott et al., 2010; Baiter et al., 2000). The *S. zooepidemicus* group contains a wide variety of strain types, reflecting the diverse array of hosts and tissues that this species of bacteria can infect, and there are 324 distinct sequence types (STs) currently listed on the MLST online database http://pubmlst.org/szooepidemicus/ (last accessed 24 November 2014; Webb et al., 2008). However, *S. zooepidemicus* strains of the same ST are frequently isolated from several host species, highlighting that at least some strains are equipped to exploit new pathogenic niches as and when the opportunity arises.

Within the *S. zooepidemicus* group, *S. equi* subsp. *equi* (*S. equi*) is the causative agent of strangles, which is the most frequently diagnosed infectious disease of horses worldwide. *S. equi* is host-restricted and only causes strangles,
which is characterized by abscessation of the lymph nodes of the head and neck. Comparison of the genomes of S. zooepidemicus strain H70 (SzH70) and S. equi strain 4047 (Se4047) provided evidence of functional loss in the genome of Se4047 due to mutation and deletion, coupled with pathogenic specialization through the acquisition of mobile genetic elements (Heather et al., 2008; Holden et al., 2009). The majority of S. zooepidemicus isolates (101 of 140 isolates tested), including SzH70, encode a 131 kDa putative sortase-processed surface protein, SZO_08560, which contains a C-terminal LPXTG motif (Holden et al., 2009). SZO_08560 contains four Listeria–Bacteroides repeat Pfam domains (PF09479) with structural similarity to mucin-binding proteins (Ebbes et al., 2011), but the function of this protein remains unknown. The Se4047 genome encodes only the final 112 aa of the orthologous protein (SEQ_1307a) and lacks an orthologue of an adjacent gene, SZO_08550, which is predicted to encode a serine recombinase (pfam00239), named PinR (COG1961). Examination of the SzH70 genome sequencing data revealed five of 50 sequence reads that positioned 114 bp of the promoter region of SZO_08560 (−170 to −55 bp) in the inverted ‘B’ orientation as opposed to the annotated reference ‘A’ orientation. This sequence is bordered by GTAGACTTTA and TAAAGTCTAC inverted repeats and it has been proposed that inversion of this sequence by PinR switches transcription of SZO_08560 on or off, thereby modulating the expression of this protein. The SZO_08560+ allele encodes a 131 kDa protein of unknown function, while the SZO_08560− allele encodes a truncated 106 kDa protein (SzH70 strain) or a truncated 108 kDa protein (Se4047 strain) (SzH70 strain). To generate each modified strain, SzH70 or Se1770 was transformed with the relevant pG’+host9 plasmid and transformants were subjected to two rounds of homologous recombination as described previously (Hamilton et al., 2006). The first recombination event, leading to the integration of the plasmid into the bacterial chromosome, was achieved by growing transformants in THB containing erythromycin at 0.5 μg ml⁻¹ (THBE) at 28 °C overnight and then increasing the temperature to 37 °C for 3 h. Integrants were selected following growth on THA containing erythromycin at 0.5 μg ml⁻¹ (THAE) overnight at 37 °C. Integrants were inoculated into THB and grown at 37 °C overnight followed by dilution into THB and incubation at 28 °C for a further 48 h. Incubation at the permissive temperature (28 °C) allowed plasmid replication and facilitated the second recombination event. Bacteria were plated on THA and grown at 37 °C to promote the loss of free plasmid. Putative mutant colonies were subcultured onto fresh THA and THAE plates to confirm their erythromycin sensitivity. The presence of the relevant mutant allele in the chromosome of putative mutants was determined by PCR using the primers listed in Table S2 followed by DNA sequencing on an ABI3100 DNA sequencer with BigDye fluorescent terminators. A schematic of the mutants generated in this study is shown in Fig. 1.

We constructed a series of S. zooepidemicus deletion mutants to determine if PinR mediates the inversion of the SZO_08560 promoter and investigate the wider recombinase-mediated regulation of protein production in S. zooepidemicus.

METHODS

Bacterial isolates. Full details of all the isolates examined in this study are available in Table S1 (available in the online Supplementary Material) and on the MLST database (http://pubmlst.org/szeoepidemicus/). SzH70 was isolated from a nasopharyngeal swab taken from a healthy thoroughbred racehorse in Newmarket, UK, during 2000 and is an ST-1 strain (Holden et al., 2009). S. zooepidemicus strain 2329 (Sz2329) is an ST-118 strain that was isolated from a tracheal wash recovered from a case of acute fatal haemorrhagic pneumonia in a greyhound from Kent in 2008 (SzH70) was transformed according to the manufacturer’s instructions (Sigma).

Preparation of total bacterial RNA and preparation of cDNA. An overnight culture was diluted 1 : 20 in fresh THB and grown to an OD₆₀₀ of 0.3. The culture was mixed with two volumes of RNA protect (Qiagen) and cells were harvested by centrifugation at 4 °C at 5000 g for 10 min followed by 8000 g for 10 min. Supernatant was poured off and the pellet resuspended in 200 μl Tris-EDTA buffer (Fluka), 3 mg lysozyme (Sigma) and 500 U mutanolysin (Sigma). The cells were vortexed repeatedly for 45 min, 700 μl of RLT buffer (Qiagen) was added and the sample was vortexed for 10 s. Then, 0.05 g of acid-washed glass beads (Sigma) was added and the sample vortexed for 5 min to complete cell lysis. The sample was centrifuged at 16 100 g and RNA was extracted from the supernatant using an RNeasy midi kit (Qiagen) with the inclusion of two on-column DNase 1 treatment steps according to the manufacturer’s instructions. RNA was quantified using a NanoDrop 1000 V3.7.1 spectrophotometer and reverse transcribed using a Verso cDNA kit according to the manufacturer’s instructions (Thermo Scientific).

Quantitative PCR (qPCR) of transcripts and the orientation of the invertible region. The number of copies of DNA or cDNA of interest were quantified by qPCR with the primers listed in Table S2. Reactions contained 10 μl Kapa SYBR fast (Kapa Biosystems), 0.3 μM forward primer, 0.3 μM reverse primer and 6 μl of a 1 : 10 dilution of DNA or cDNA. Reactions were made up to 20 μl with water and thermocycled on an ABI StepOnePlus instrument at 95 °C for 3 min followed by 40 cycles of 95 °C for 30 s and 60 °C for 10 s with an SYBR
read taken at the end of each cycle, then 95 °C for 15 s. A melt curve was generated from 60 to 95 °C with SYBR reads every 0.3 °C to differentiate potential non-specific amplification products and data were analysed using StepOnePlus Software v2.1. No template and no reverse transcription controls were used as negative controls and standard curves with a DNA reference were created for each primer pair. The experiments were repeated in triplicate and data were normalized by comparison with the housekeeping gene gyrA. Amplified FimIV DNA fragments were purified using a PCR purification kit (Qiagen), and the sequences were obtained on both strands using an ABI3100 DNA sequencer with BigDye fluorescent terminators using the original PCR primers. Sequence data were assembled using SeqMan 5.03 (DNASTAR).

**Quantification of in vitro growth rate.** Mutant strains were inoculated into THB containing 10% fetal calf serum (THBS) in triplicate and the growth of each strain was monitored by measuring the OD600.

**Air-interface infection model.** Air-interface respiratory tract organ cultures were constructed using explants of equine trachea as described previously (Hamilton et al., 2006). The trachea used in this study were recovered from six ponies that were euthanized for reasons unrelated to this project and processed on the same day to maximize cell viability. Trachea were washed in Dulbecco’s modified Eagle’s medium supplemented with 2 mM l-glutamine (DMEM) containing penicillin (100 U ml⁻¹), streptomycin (50 μg ml⁻¹), gentamicin (100 μg ml⁻¹) and amphotericin-B (2.5 μg ml⁻¹) (PAA) for 4 h to remove commensal flora. Following further washing in DMEM to remove residual antibiotics and amphotericin-B, the trachea were dissected into pieces approximately 5 mm² and mounted on agarose platforms surrounded by 3 ml DMEM supplemented with 2 mM l-glutamine, in six-well cell culture plates. Organ cultures were maintained in a humidified 5% CO₂ incubator at 37 °C. The viability of the air-interface organ cultures was assessed using 1 μm polystyrene bead (Park Scientific) clearance. Contamination was monitored by running a bacteriology loop around all four edges of the culture pieces.
and streaking onto strep select plates. Any tissue pieces in which contamination was detected were discarded. Organ culture pieces were infected with a 10 μl suspension containing 1 × 10^6 c.f.u. of SzH70 mutants, or were mock-infected with THB. Attachment of bacteria to the organ culture pieces was quantified by measuring viable counts (six organ culture pieces per time point) of adherent bacteria at 2 h post-infection. Organ culture pieces were vortexed for 5 s in PBS to remove non-adherent bacteria and then homogenized before plating serial ten-fold dilutions onto THA and enumerating colonies. Data are presented from six independent experiments.

Whole genome sequencing. Strain Sz2329 has previously been shown to lack pinR and SZO_08560 by PCR screening of a diverse population of S. zooepidemicus (Holden et al., 2009). Lack of pinR raised the possibility that novel invertible sequences could be fixed in the genome, facilitating their identification. Therefore, the genome of Sz2329 was sequenced to 25-fold coverage using a Genome Sequencer-FLX (454 Life Sciences, Roche Applied Sciences). Two sequencing libraries were prepared from genomic DNA, the first a fragment (~250 bp read length) and a second 3000 bp insert, long-tag paired end library (~100 bp) to provide scaffolding. The reads were assembled with Newbler (v2.0.0.11.14) using default assembly parameters. Comparison with the genome sequence of SzH70 (FM204884) (Holden et al., 2009) was facilitated by using the Artemis Comparison Tool (ACT) (Carver et al., 2005). The sequence and annotation of the Sz2329 genome have been deposited in the EMBL database under accession number JTH0000000.

Statistical analysis. A two-sided Student’s t-test was used to compare continuous data where assumptions of a normal distribution and equal variance were satisfied. A Kruskal–Wallis test was performed to compare continuous data where assumptions of a normal distribution were not satisfied. A paired Student’s t-test was used to compare the attachment of wild-type and mutant strains of SzH70 to explants of equine trachea, accounting for variation between the six different strains.

RESULTS

PinR is responsible for inversion of the region upstream of SZO_08560

We generated a series of mutant and complementation strains in SzH70 (Fig. 1) to determine if PinR mediates inversion of the promoter of SZO_08560. Each deletion was confirmed by PCR and sequencing across the deletion site. The amount of the invertible promoter region in both the A and B orientations was quantified by qPCR. Wild-type SzH70 contained 96% of SZO_08560 promoter copies in the A orientation and 4% in the B orientation (Fig. 2). Deletion of pinR (ΔpinR A) fixed the promoter in the A orientation; no copies of the promoter in the B orientation were identified in this mutant. Both pinR and the SZO_08560 promoter were deleted (ΔpinR) and then the SZO_08560 promoter was reintroduced in the B orientation to produce a mutant strain (ΔpinR B) that contained only the SZO_08560 promoter in the B orientation. Complementation of the pinR deletion in strains ΔpinR A and ΔpinR B by insertion of a copy of pinR under the control of its native promoter into the pseudogene SZO_07770 (ΔpinR A c and ΔpinR B c, respectively) restored inversion of the promoter of SZO_08560 in the ΔpinR A c strain such that 0.8% of promoter copies were in the B orientation. However, inversion of the promoter of SZO_08560 was not restored in the ΔpinR B c strain (Fig. 2). The ΔpinR A and ΔpinR A c strains had a significantly reduced growth rate when compared with SzH70 and the other mutant strains (P<0.006) (Fig. S1).

The orientation of the invertible region determines SZO_08560 transcription

To determine if the promoter of SZO_08560 was more active in the A or B orientation, total RNA isolated from each mutant strain was reverse transcribed and used to quantify the transcription of SZO_08560 by qPCR. Data were normalized based on the number of gyrA transcripts in each triplicate sample. The transcription of SZO_08560 in wild-type SzH70 was equivalent to that of gyrA in this strain (Fig. 3). Deletion of pinR such that the promoter of SZO_08560 was fixed in the A orientation (ΔpinR A) caused a reduction of SZO_08560 transcription to 0.7% of wild-type levels (P<0.0001). However, fixation of the promoter of SZO_08560 in the B orientation increased SZO_08560 transcription to 189% of wild-type levels (P<0.0001). Deletion of SZO_08560 abolished its transcription, which was restored to 50% of wild-type transcription levels by complementation through the insertion of a copy of SZO_08560 under the control of its native promoter in the B orientation into the pseudogene SZO_07770 (strain ΔO8560 c).

Deletion or increased transcription of SZO_08560 alone did not significantly affect attachment of S. zooepidemicus to equine tissues

The number of bacteria recovered from explants of equine trachea 2 h post-infection with the wild-type SzH70 strain did not significantly differ from the number recovered from those infected with the mutant strains (Fig. 4).
However, the reduction in the amount of D08560 and D08560 c strains recovered relative to SzH70 approached statistical significance (P=0.0859 and P=0.0883, respectively). A higher number of bacteria were recovered from those explants infected with the DpinRB mutant, which transcribes the most SZO_08560 relative to SzH70, although this was also not statistically significant (P=0.67).

**PinR inverts sequences distant to the promoter of SZO_08560**

Analysis of the draft genome sequence of Sz2329 using the ACT confirmed that this strain contained a deletion of pinR and the majority of its SZO_08560 homologue, which was identical to that previously identified in Se4047 (Holden et al., 2009). The assembled Sz2329 draft genome contained one example of altered locus architecture consisting of an inversion of a 7137 bp region containing the major and minor pilin genes, but not the AraC-like regulator or associated sortases of FimIV (Fig. 5). The inversion occurred in 100% of the sequencing reads covering this region, which was represented in a single contig. The inverted region of FimIV is flanked by a six-base inverted repeat (TAGAAA), which partially (TAGA) matches the 10-base inverted repeat (GATACTTTA) that flanks the invertible promoter region upstream of SZO_08560 in SzH70 (Holden et al., 2009).

To determine if inversion of the FimIV locus was actively occurring in other strains of *S. zooepidemicus*, PCR primers were designed to amplify a product when the FimIV region was in either the original orientation, as annotated in the SzH70 genome, or the inverted orientation. A collection of 10 FimIV-containing strains were screened by PCR for the occurrence of amplification products suggesting the presence of DNA in both orientations (Table S1). Active inversion of the FimIV sequence was identified in strains Sz1770, SzB260863 and SzH050840501. Only the inverted FimIV PCR product was amplified from strain Sz2329. The PCR products were purified and sequenced, confirming that the inverted region in FimIV was flanked by the same inverted repeat (TAGAAA) in all strains.

**Fig. 3.** Transcription of SZO_08560 in the mutant strains. The number of transcript copies of SZO_08560 was quantified by qPCR and normalized relative to the amount of gyrA. Error bars indicate sd.

**Fig. 4.** Attachment of SzH70 and mutant strains to explants of equine trachea. Error bars indicate 95% confidence intervals.

**Fig. 5.** Partial inversion of the FimIV locus in strain Sz2329 relative to the SzH70 reference genome visualized using the ACT (Carver et al., 2005). The coloured bars separating each genome (blue and red) represent similarity matches identified by reciprocal TBLASTX analysis, with a score cut-off of 100. Blue lines link matches in the same orientation; red lines link matches in the reverse orientation.
To determine if PinR was mediating FimIV inversion, pinR was deleted from Sz1770 by allelic replacement mutagenesis. Deletion of pinR was confirmed by PCR and sequencing across the deletion site. The numbers of original and inverted copies of FimIV in the ΔpinR mutant, wild-type Sz1770 and Sz2329 were quantified by qPCR and normalized to gyrA. Wild-type Sz1770 contained 0.01 % (1 : 10 000) of FimIV copies in the inverted orientation (Fig. 6). Deletion of pinR from strain Sz1770 prevented inversion of the FimIV region, yielding 100 % of qPCR products in the original orientation.

DISCUSSION

The surface architecture of S. zooepidemicus is likely to be crucial to its ability to adapt and interact with mammalian hosts and the wider environment in order to fulfil the requirements of its opportunistic lifestyle. The organism must survive outside a host, in drinking water or on soil, grass and other surfaces in competition with a vast array of other micro-organisms, whilst remaining in a state of readiness to infect a susceptible new host should the opportunity arise. The population of S. zooepidemicus infects many different mammalian hosts and tissues. Indeed, individual strains are themselves capable of infecting multiple hosts, and zoonotic transmission, for example from an infected dog to a veterinary nurse, has been demonstrated (Abbott et al., 2010). S. zooepidemicus persists in the tonsils or on the mucosal surfaces of recovered horses in the face of a mature immune response, increasing the likelihood of onward transmission (Lindahl et al., 2013). Therefore, the ability of S. zooepidemicus to modulate its surface is likely to be essential to its long-term survival.

Here we present evidence that the inversion of the promoter of SZO_08560 is performed by PinR and demonstrate that inversion acts as a switch, controlling transcription of SZO_08560. SZO_08560 contains an N'-terminal signal sequence, C’-terminal LPXTG sortase-processing motif and four Listeria–Bacteroides repeat Pfam domains (PF09479) with structural similarity to mucin-binding proteins (Ebbes et al., 2011). Whilst the exact receptor bound by SZO_08560 remains unknown, the reduced ability of SZO_08560 mutants to attach to explants of equine trachea, which approached statistical significance, suggests that SZO_08560 is likely to play a role in the attachment of S. zooepidemicus to host tissue.

The ΔpinR A mutant lacks pinR with the SZO_08560 promoter orientated in the A direction and shows a slow growth rate. One explanation for the slow growth of this strain is interference of the transcription of SZO_08540 or SZO_08530 by the SZO_08560 promoter, which could be enhanced by its closer proximity to these coding sequences following the deletion of pinR. SZO_08540 encodes a conserved hypothetical protein, whilst SZO_08530 encodes RpsP, the 30S ribosomal protein S16. Interestingly, the ΔpinR B mutant, which lacks pinR, whilst orientating the SZO_08560 promoter in the B direction had a normal growth rate, as did the ΔpinR mutant, which lacks both pinR and the SZO_08560 promoter (Fig. S1). Therefore, the inversion of the SZO_08560 promoter from the A to the B orientation in strain ΔpinR A c, which contains a complementing copy of pinR, may be preferred as it is likely to yield strains with a normal growth rate. However, the inversion of the SZO_08560 promoter from the B to the A orientation in strain ΔpinR B c was not detected, probably as the resultant mutants would have a slower growth rate.

Analysis of the Sz2329 genome sequence, which lacks pinR, identified a disruptive internal inversion of the FimIV locus, which was bordered by short inverted repeats that shared a four-base motif (TAGA) with the SZO_08560 promoter. The FimIV locus encodes an AraC-like regulator, three putative sortase enzymes, a putative exported protein and three putative sortase-processed proteins that are predicted to form a surface pilus structure (Holden et al., 2009). Screening of a panel of S. zooepidemicus isolates identified three strains with active FimIV inversion. The deletion of pinR in one of these strains, Sz1770, stopped FimIV inversion revealing a wider role for PinR in the global regulation of bacterial surface components and highlighting redundancy in the DNA sequences of the inverted repeats. Our data suggest that the ancestor of Sz2329 contained a functional copy of pinR and was actively inverting the FimIV region until the loss of pinR fixed this region in the position that it was in at the time. FimIV was present in 81 (58 %) of 140 isolates of S. zooepidemicus that were tested and is missing from the S.
equi genome (Holden et al., 2009), indicating that its loss from the genomes of some strains may be beneficial in the particular environments that they occupy. It is interesting that inversion of the FinIV region was not observed in strain S2H70, despite this strain actively inverting the SZO_08560 promoter via PinR, suggesting that co-factors may assist PinR to invert alternative substrates. Variation in the sequence of inverted repeats and size of the inverted regions of DNA confounds the in silico identification of substrate sites and further research is required to identify the range of PinR substrates and the consequences of inversion on the properties of the variants produced.

The data presented here suggest that PinR plays an important role in modulating the surface architecture of S. zooepidemicus forming a mixture of distinct phenotypes, which provides this organism with a bet-hedging solution to survival in fluctuating environments (Stewart & Cookson, 2012). PinR of S. zooepidemicus shares >60% predicted amino acid identity with putative resolvases including those encoded by strains of Streptococcus anginosus, Streptococcus pneumonia (strain GA17545), Streptococcus constellatus, Streptococcus suis, Streptococcus mitis, Streptococcus ovis, Streptococcus pseudopneumoniae, Peptoniphilus indolicus, Eubacterium sphenum, Parvimonas micra, Eggerthia cateniformis, Gemella bergeri, Gemella cuniculi, Bulteldia extructa, Enterococcus faecium, Erysipelotrichaceae bacterium, Clostridiales bacterium, Gardnerella vaginalis, Coprobacillus sp., Catenibacterium sp. and Mogibacterium sp., suggesting that serine recombinase-mediated modulation of surface architecture is a mechanism that is widely adopted by other Gram-positive bacteria. It is intriguing to note that PinR shares 27% amino acid sequence identity and conserved serine residue with the site-specific recombinase of Bacteroides fragilis, FinA (also known as Mpi), which modulates the production of several surface components in this Gram-negative bacterium by inversion of promoter sequences (Cerdeño-Tárraga et al., 2005; Coyne et al., 2003). To our knowledge, our data provide the first evidence to suggest that the reversible ON–OFF phenotype known as phase variation can be mediated by a recombinase in streptococci.

ACKNOWLEDGEMENTS

This project was funded by a project grant from the Horserace Betting Levy Board (vet/prj/751). S2H050840501 was kindly provided by Dr Androuilla Efstratiou of Public Health England.

REFERENCES


Edited by: S. Jörg