J Mol Evol (2009) 68:498–505
DOI 10.1007/s00239-009-9221-x

Prevalence and Acquisition of the Genes for Zoocin A and Zoocin A Resistance in *Streptococcus equi* subsp. *zooepidemicus*

Amy S. Gargis · Anna-Lee D. O'Rourke · Gary L. Sloan · Robin S. Simmonds

Received: 25 September 2008 / Accepted: 27 February 2009 / Published online: 9 April 2009

Abstract Zoocin A is a streptococcolytic enzyme produced by *Streptococcus equi* subsp. *zooepidemicus* strain 4881. The zoocin A gene (*zooA*) and the gene specifying resistance to zoocin A (*zif*) are adjacent on the chromosome and are divergently transcribed. Twenty-four *S. equi* subsp. *zooepidemicus* strains were analyzed to determine the genetic difference among three previously characterized as zoocin A producers (strains 4881, 9g, and 9h) and the 21 nonproducers. LT-PCR and Southern hybridization studies revealed that none of the nonproducer strains possessed *zooA* or *zif*. RAPD and PFGE showed that the 24 strains were a genetically diverse population with eight RAPD profiles. *S. equi* subsp. *zooepidemicus* strains 9g and 9h appeared to be genetically identical to each other but quite different from strain 4881. Sequences derived from 4881 and 9g showed that *zooA* and *zif* were integrated into the chromosome adjacent to the gene *flaR*. A comparison of these sequences with the genome sequences of *S. equi* subsp. *zooepidemicus* strains H70 and MGCS10565 and *S. equi* subsp. *equi* strain 4047 suggests that *flaR* flanks a region of genome plasticity in this species.

Keywords Zoocin A Peptidoglycan hydrolase Horizontal gene transfer RAPD PFGE

Electronic supplementary material The online version of this article (doi:10.1007/s00239-009-9221-x) contains supplementary material, which is available to authorized users.

A. S. Gargis · G. L. Sloan
Department of Biological Sciences, The University of Alabama, Box 870334, Tuscaloosa, AL 35487-0334, USA

A.-L. D. O'Rourke · R. S. Simmonds (✉)
Microbiology Department, University of Otago, P.O. Box 56, 9001 Dunedin, New Zealand
e-mail: robin.simmonds@otago.ac.nz

Introduction

Streptococcus equi are important pathogens in equine respiratory disease. Originally regarded as a single species, *S. equi* have recently been divided into two subspecies, *equi* and *zooepidemicus* (Alber et al. 2004). These two subspecies share 98% DNA homology but exhibit distinctly different phenotypes. *S. equi* subsp. *equi* (*S. equi*) is the cause of the equine disease known as strangles and is considered far more pathogenic than *S. equi* subsp. *zooepidemicus* (*S. zooepidemicus*) (Skinner and Quesnel 1978; Timoney 2004). *S. zooepidemicus* is a normal commensal organism and an opportunistic pathogen in the equids (Timoney 2004).

S. zooepidemicus strain 4881 was isolated from the aborted fetus of a horse in New Zealand and was identified to produce the Class III bacteriocin zoocin A (Schofield and Tagg 1983). Subsequent studies identified additional zoocin A producing *S. zooepidemicus* strains 9g and 9h, which yielded the same specific inhibitory spectrum as strain 4881 (Barnham et al. 1987). Zoocin A is a D-alanyl–L-alanyl endopeptidase that hydrolyzes the peptidoglycan cross bridges of susceptible streptococci (Gargis et al. 2009). Producer cell immunity to zoocin A is encoded by *zif*, the zoocin A immunity factor (Beatson et al. 1998). *zooA* and *zif* are chromosomally encoded and are transcribed in opposite directions. The *zooA-zif* locus is flanked by two transposon-like sequences with significant similarity to the plasmid-encoded genes for lysostaphin and lysostaphin resistance (*end-epr* [*lss-lif*]) in *Staphylococcus simulans* biovar *staphylolyticus*) (Beatson et al. 1998). The similarity in genetic arrangement between the *zooA-zif* and the *end-epr* loci raised the possibility that zoocin A was acquired via horizontal gene transfer (Beatson et al. 1998). In recent years, endopeptidase genes with homology to *zooA* have

been discovered in other streptococcal species including *Streptococcus milleri* NMSCC 061 (Beukes and Hastings 2001) and *Streptococcus constellatus* subsp. *constellatus* (Heng et al. 2006). The *S. milleri* endopeptidase millericin B gene (*milB*) and its immunity factor gene (*milF*), and the *S. constellatus* subsp. *constellatus* endopeptidase gene (*stlA*) and its immunity factor gene (*stlI*), are both chromosomally located like *zooA* and *zif*. It is presumed that bacteria produce bacteriolytic endopeptidases, like zoocin A, for the purpose of killing closely related bacterial species. Therefore, the acquisition of *zooA* and *zif* would give an organism a competitive advantage in environments where nutrients are limited (Tagg et al. 1976).

In the present study, 24 *S. zooepidemicus* strains were examined in order to determine the prevalence of the *zooA* and *zif* genes and the genetic differences between the zoocin A producing and the zoocin A nonproducing strains. Marked differences in the RAPD profiles, PFGE patterns, and DNA sequence proximal to the *zooA-zif* locus supports the hypothesis that *zif* and *zooA* were acquired by horizontal gene transfer.

Materials and Methods

Bacterial Strains and Growth Conditions

All strains used and their associated epidemiological data are listed in Supplementary Table S1. Bacterial strains were maintained on Columbia agar base (CAB) (Difco, Becton Dickinson, Sparks, MD, USA), grown at 37°C in 5% CO_2 for 24 h, and then stored at 4°C. All broth cultures were grown in either Todd Hewitt broth (THB) or M17 broth (Difco) at 37°C in 5% CO_2 for 24 h.

Biochemical Analysis and Deferred Antagonism of *S. zooepidemicus*

Biochemical tests were carried out on the 24 strains of *S. zooepidemicus* to confirm the identity of each strain in terms of genus and species. Hemolysis reactions were determined by growing strains on CAB containing 5% (v/v) horse blood. The ability to produce proteolytic enzymes was tested by growing each isolate on CAB containing 5% skim milk powder. Fermentation tests were carried out using the sugars sorbitol, lactose, and trehalose as described in *Bergey's Manual of Systematic Bacteriology* (Holt et al. 1994). The 24 strains were tested for production of zoocin A and immunity to zoocin A using the deferred antagonism test, as described by Tagg and Bannister (1979). The *S. zooepidemicus* strains were also typed according to their Lancefield grouping using a Slidex Streptokit (BioMérieux, Marcy-l'Etoile, France).

Chromosomal DNA Extractions

Chromosomal DNA was extracted from the streptococcal species using two different methods. The first method was a phenol/chloroform extraction, as described by Simpson and Cleary (1987). The second method used was the GeneElute Mammalian DNA kit (Sigma-Aldrich, St. Louis, MO, USA) and was carried out according to the manufacturer's instructions with the following modifications. Cultures were grown in 20 ml of THB at 37°C in 5% CO_2 for 16–24 h, then centrifuged at 5000*g* for 10 min at 4°C, and the supernatant was removed by aspiration. The cell pellet was resuspended and washed twice in 20 ml and once in 1 ml of TE buffer, pH 8. The cell pellet was then resuspended in 140 µl enzymatic lysis buffer (20 mM Tris-HCl, 2 mM EDTA, 1.2 g Triton X-100), containing 18 µl lysozyme (200 mg/ml; Sigma-Aldrich), and incubated at 37°C for 30 min, with occasional gentle shaking. Following incubation, 10 µl RNase (10 mg/ml) and 8 µl mutanolysin (500 U/ml) were added and the tubes were incubated for a further 15 min at 37°C. Twenty-five microliters of proteinase K (10 mg/ml; Sigma-Aldrich) and 200 µl of solution SL (Sigma GeneElute kit lysis solution) were added to the cells, mixed by gentle inversion, and then incubated at 70°C for 30 min. Following this final incubation step, 200 µl of 100% ethanol was added to the tube, which was then mixed by gentle inversion. This mixture was added to a GeneElute nucleic acid binding column and eluted according to the manufacturer's instructions.

Random Amplification of Polymorphic DNA

The 24 strains were typed using RAPD analysis. The RAPD primers P6 (5′-CTGTGACACCGGGATACGA-3′) and H2 (5′- CCTCCCGCCACC-3′) (Bert et al. 1996) were used for analysis of group A, B, C, and G streptococci. The RAPD PCR was carried out in a 50-µl reaction mixture containing 50 ng of template DNA, 25 µl of master mix, 50 pM of primer P6, 50 pM of primer H2, and 23 µl of PCR-quality H_2O (*Taq*PCR master mix kit; Qiagen). PCR amplification conditions were as follows: one cycle of denaturation at 94°C for 5 min, followed by 35 cycles of denaturation at 94°C for 1 min, primer annealing at 36°C for 1 min, and template extension at 72°C for 3 min. A negative control incorporating all PCR components except genomic DNA was carried out for each RAPD-PCR conducted. Following completion of the RAPD-PCR, duplicate 10-µl aliquots of each reaction mixture were loaded onto a 1× TAE, 1% (w/v) agarose gel and electrophoresed at 100 V for 1 h in 1× TAE buffer. After electrophoresis, the gel was stained with ethidium bromide. Strict adherence to the guidelines set out in Power's (1996) review of RAPD

typing was employed. RAPD patterns were analyzed both visually and using GelQuest and ClusterVis Software (SequentiX; Klein Raden, Germany). After fingerprints were normalized and the background was subtracted, the levels of similarity between all RAPD profiles were calculated using the Dice correlation coefficient (Nei and Li 1979). A dendrogram was generated from the distance matrix using the unweighted pair group method with arithmetic averages.

Pulsed Field Gel Electrophoresis

PFGE was performed on *S. zooepidemicus* strains 4881, 9g, 9h, 9d, 9i, and 9j using a protocol based on that of Keis et al. (1995) for the preparation of clostridial genomic DNA, with the following modifications. Bacterial cultures were grown to an OD_{600nm} of 0.5–0.7 in 10 ml of THB. Once the bacterial cells were embedded in the agarose plugs, the plugs were placed in 1 ml of lysis buffer containing 100 µl of lysozyme (10 mg/ml; Sigma), 2 µl of RNase A (10 mg/ml; Sigma) and incubated at 37°C for 16–24 h. Following lysis the protocol according to Keis et al. was resumed, with *Sma*I used to digest genomic DNA. The digests were electrophoresed at 6 V/cm in 0.5× TBE cooled to 14°C, at an angle of 120°, on a 1.2% (w/v) agarose gel (USB Corp., Cleveland, OH, USA), with pulse times of 5 to 25 s, for 22 h. The gel was stained with ethidium bromide for 1 h and analyzed following the criteria proposed by Tenover et al. (1995). These criteria dictate that strains are indistinguishable if they give the same PFGE pattern, closely related if the PFGE patterns differ by no more than two to three fragment shifts, possibly related when four to six fragment differences are observed, and unrelated when two PFGE patterns share less than 50% of the same well-defined fragments (usually seven or more fragment differences).

Southern Hybridization

Genomic DNA of all 24 *S. zooepidemicus* strains was digested with *Eco*RI in the presence of buffer H (Roche Diagnostics, Indianopolis, IN, USA). The DNA was transferred to a Hybond–N^+ membrane (Amersham Pharmacia Biotech, Buckinghamshire, UK). DNA transfer was carried out using the Southern capillary blotting protocol, and alkali fixation procedure, in accordance with the manufacturer's instructions (Amersham Pharmacia Biotech). Radiolabeling of PCR probes to *zooA* (zooA 3840F, 5′-CATAATATTCGAGAAGGCGCT-3′; zooA 4838R, 5′-CCTACTACAGTTGACAAAGAGCCA-3′) and *zif* (zif 2492F, 5′-TTTCTAGGGTTTGGGTAGT-3′; zif 3537R, 5′-GGAGGAAATAAAGGTTGTCGC-3′) was carried out using Amersham Ready-to-Go DNA Labeling Beads (-dCTP), following the manufacturer's instructions with the following modifications. Briefly, 25–50 ng of PCR amplicon was made up to 50 µl with sterile water. Probe DNA was then denatured by boiling for 8 min and then immediately cooled on ice for 2 min. The probe was then added to a tube containing a Ready-to-Go DNA labeling bead and mixed by pipetting until the bead dissolved. A 2- to 5-µl aliquot of [α-^{32}P]dCTP was then added to the probe and incubated at 37°C for 15 min and the reaction stopped by the addition of 2 µl of 0.5 M EDTA (pH 8.0). Prior to the addition of radiolabeled probe DNA, membranes were prehybridized. The membrane was placed in a hybridization tube with the DNA side of the membrane exposed. Ten to twenty milliliters of Church and Gilbert (1984) solution was added and the tube was rotated in a Hybaid hybridization oven (Thermo Electron Corp., Waltham, MA, USA) for a minimum of 1 h at 65°C. Following prehybridization, 50 µl of the radiolabeled DNA probe was added to the hybridization tube and rotation continued at 65°C for 16–24 h. The membrane was washed twice in 300 ml of 5× SSC (20× SSC: 3 M NaCl, 0.3 M sodium citrate, 1 mM EDTA) plus 0.5% SDS with gentle agitation at room temperature for 2 min, followed by two washes in 2× SSC + 0.1% SDS at 65°C for 20 min with gentle agitation. Detection of membrane-bound radiolabeled probes was performed by exposing Agfa CP-BU new medical X-ray film (Agfa-Gevaert N.V., Belgium) to the hybridized membranes. The membrane was placed in an autoradiograph cassette (Eastman Kodak Co., Rochester, NY, USA) with the probe bound DNA side facing upward. Two Agfa X-ray films were placed on top and the cassette was placed at –70°C for 4–24 h. Film was developed for 5 min in Kodak D-19 developer (Eastman Kodak Co.), prepared according to the manufacturer's instructions, fixed in Ilford Hypam rapid paper and film fixer (Ilford Imaging Australia Pty Ltd., Victoria, Australia) for 4 min, and soaked in water for 10 min.

Long Template-PCR (LT-PCR) and Sequencing of *S. zooepidemicus* Isolates

Three strains have been isolated that contain *zif* and *zooA*: *S. zooepidemicus* 4881 (Beatson et al. 1998) and *S. zooepidemicus* strains 9g and 9h (Barnham et al. 1987). PCR primers *rgg69* (5′-GCAAGCCGTTTCACGCGGAAGC-3′) and *flaR6684* (5′-GCGTCCGTCTAACAGAATCC-3′) were designed to the 5′ and 3′ ends of a 6.8-kb *Eco*RI contig containing *zif* and *zooA* from *S. zooepidemicus* 4881 and used in LT-PCR analysis of the 24 strains of *S. zooepidemicus*. Primers ORFA4177 (5′-ATTTAAGCACCA TATCCCTCAAG-3′) and ORFAR5158 (5′-GATTTAGC ATTACTGTGGTGGTCT-3′) were used for the LT-PCR screening for *dsg*, a putative permease gene that occurred

near *flaR* in the nonproducer strains but not in the producer strains. Each LT-PCR was carried out using either the Expand LT-PCR kit (Roche Diagnostics) or the KOD hot start DNA polymerase kit (Novagen, San Diego, CA, USA) according to the manufacturer's instructions. Reactions were set up in 50-µl volumes using 50 ng of template DNA and 25 pM of each primer. Amplification conditions were as follows: one cycle of denaturation at 94°C for 2 min, followed by 10 cycles of denaturation at 94°C for 10 s, primer annealing at 60°C for 30 s, and primer extension at 68°C for 7 min. This was followed by 20 subsequent cycles using a 7-min primer extension time with an incremental addition of 20 s per cycle. The other cycle parameters were as above, with a final extension carried out at 68°C for 7 min. The resulting products were visualized by gel electrophoresis.

BLAST Search Analysis and Sequencing of the Regions Flanking the *zooA-zif* loci in *S. zooepidemicus* Strains 4881 and 9g

In order to identify genes flanking the *zooA-zif* loci, primer walking with primers designed against the conserved regions of strains *S. zooepidemicus* 4881 and 9g and the draft genome sequence of *S. zooepidemicus* H70 and completed genome sequence of *S. equi* 4047 (available at www.sanger.ac.uk/Projects/S_zooepidemicus/ and www.sanger.ac.uk/Projects/S_equi/, respectively) (Supplementary Table S2) was undertaken. All sequencing was performed at the Allan Wilson Centre Genome Service (Massy University, Palmerston North, NZ) and analyzed with Bioedit (Ibis Biosciences, Carlsbad, CA), BLASTN, and BLASTP (available at http://www.ncbi.nlm.nih.gov/BLAST/).

Results

Phenotypic Characterization of Strains

All 24 *S. zooepidemicus* strains were Gram positive, protease negative, and β hemolytic; fermented lactose and sorbitol but not trehalose; and belonged to Lancefield group C. The deferred antagonism test revealed that only three of the strains (4881, 9g, and 9h) produced zoocin A and that all three strains displayed cross-immunity to each other.

Genetic Characterization of Strains

Southern hybridization analysis showed that *zif* and *zooA* were only present in *S. zooepidemicus* strains 4881, 9g, and 9h. In strain 4881, the probes reacted with an *Eco*RI fragment of ~6.8 kb, whereas in strains 9g and 9h, the probes reacted with an *Eco*RI fragment of ~8 kb (Supplementary Fig. S1).

An UPGMA dendrogram of genetic relationships among the 24 strains of *S. zooepidemicus*, using the Dice coefficient for an estimation of genetic similarity, was constructed from the RAPD data (Fig. 1a). The 24 strains grouped into eight different RAPD profiles. RAPD profile 1 is the zoocin A type strain *S. zooepidemicus* 4881, while RAPD profile 6 represents the other zoocin A producing isolates, 9g and 9h. This cluster analysis revealed that strains 9g and 9h are genetically distant from strain 4881. Amplification reactions with primers P6 and H2 produced PCR amplicons ranging from ~400 to 2650 bp (Fig. 1b). The zoocin A producer organisms showed strain-specific banding patterns in their RAPD profiles, with *S. zooepidemicus* strains 9g and 9h yielding a 1250-bp band and

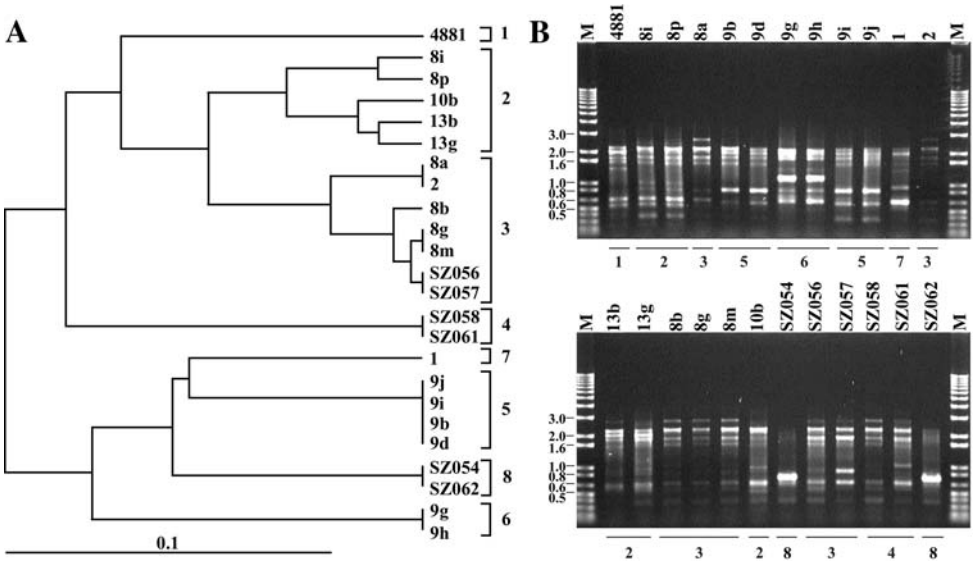

Fig. 1 **a** UPGMA dendrogram of genetic relationships among the 24 strains of *S. zooepidemicus*, using the Dice coefficient for an estimation of genetic similarity. Scale bar represents genetic distance (dissimilarity). **b** RAPD profiles of *S. zooepidemicus* isolates. Bars and numbers under the gel indicate group profiles and profile numbers respectively. *M* 1-kb Plus DNA ladder (Invitrogen). Molecular size markers are indicated as kilobases

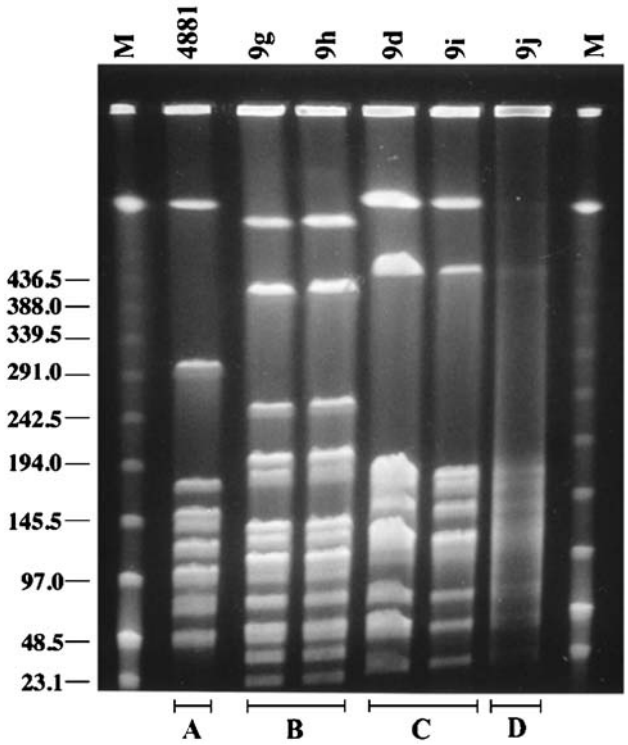

Fig. 2 PFGE patterns of selected *S. zooepidemicus* isolates. Bars and letters underneath the gel indicate group patterns and PFGE pattern designations, respectively. *M* lambda DNA ladder (New England Biolabs). Molecular size markers are indicated as kilobases

strain 4881 lacking this band. These RAPD data suggested that 9g and 9h were two isolates of the same strain obtained from independent samples.

PFGE (based on *Sma*I digests) was used to further investigate the genetic heterogeneity of the three zoocin A producer strains (4881, 9g, and 9h) and of three nonproducer strains (9d, 9i, and 9j). The three nonproducer strains were chosen for analysis because they were isolated from the same outbreak as the producer strains 9g and 9h. Four distinct patterns (Fig. 2) were identified: pattern A, strain 4881; pattern B, strains 9g and 9h; pattern C, strains 9d and 9i; and pattern D, 9j. Based on the criteria of Tenover et al. (1995), strain 4881 and the other two producer strains, 9g and 9h, appear to be unrelated, sharing only four common bands and differing by seven well-defined fragments. By contrast, strains 9g and 9h appear to be identical; the same is true for strains 9d and 9i. Strain 9j did not group with the other strains.

LT-PCR resulted in the generation of five different length amplicons, ranging from ~850 bp to 8 kb (Supplementary Fig. S2). *S. zooepidemicus* 4881 generated a 6.6-kb product, as did strains 9b, 9d, 9i, and 9j. An 8-kb product was generated from producer strains 9g and 9h, which was in agreement with the Southern hybridization results.

Sequencing of *S. zooepidemicus* Strains 4881 and 9g

GenBank accession numbers for sequences derived from *S. zooepidemicus* strains 4881 and 9g are U50357 and EU636005, respectively. Previously, *zooA-zif*, present in *S. zooepidemicus* 4881, was shown to be flanked by two potential transposase genes, ORF1 and ORF2 (Beatson et al. 1998). BLASTX search analysis confirmed that ORF1 encoded a protein related to the IS200-like transposase family (77% identity and 88% similarity to an IS200-like transposase present in *Lactobacillus reuteri* 100-23) and that ORF2 encoded a protein related to the IS1165 and IS1167 transposases (73% identity and 84% similarity to the transposase IS204/IS1001/IS096/IS1165 present in *Streptococcus suis* 89/1591 and 69% identity and 83% similarity to IS1167 present in *Streptococcus pneumoniae* TIGR4). ORFs 1 and 2 were therefore annotated *tnpA* and *tnpB*, respectively (Fig. 3). Sequencing to the RHS (right-hand side) of *tnpB* identified an ORF (*flaR*) encoding a protein having 95% identity and 95% similarity to a topology modulation protein from *S. zooepidemicus* MGCS10565 and 74% identity and 82% similarity to the DNA topology modulation protein FlaR of *Streptococcus thermophilus*. Sequencing of a further 6 kb on the RHS of *flaR* (block D) identified ORFs corresponding to housekeeping genes encoding a cell wall surface anchor and proteins involved in carbon metabolism, including what appears to be a carbon phosphotransferase system (Fig. 3).

Sequencing to the RHS of *zooA* in *S. zooepidemicus* 9g revealed two copies of the *tnpB* gene, *flaR*, and the same set of housekeeping genes (block D) identified in *S. zooepidemicus* 4881 (Fig. 3). BLAST searching the genome sequences of *S. equi* 4047 and *S. zooepidemicus* strains H70 and MGCS10565 revealed that they also contained genes very similar to *flaR* and block D (Fig. 3) and that their genome sequences remained very similar for ~50 kb beyond block D.

Sequencing to the LHS (left-hand side, as shown in Fig. 3) of *tnpA* in *S. zooepidemicus* strains 4881 and 9g revealed that both had an ORF (*rgg*) followed by blocks C (323 bp), B (5 kb), and E (5 kb). The ORF designated *rgg* encoded a protein having 100% identity and 100% similarity to the *S. zooepidemicus* MGCS1065 Rgg/GadR/MutR family of transcriptional activators. The sequence designated C was short (323 bp) and did not contain any identifiable ORFs. Block B (5 kb) contained ORFs with genes encoding proteins including an AAA family ATPase, components for the GATase1 anthranilate synthase and para-amino benzoate synthase. Block E (5 kb) contained ORFs with genes encoding proteins including transcriptional regulators, ribosomal methyltransferases, and proteins involved in the phosphotransferase system.

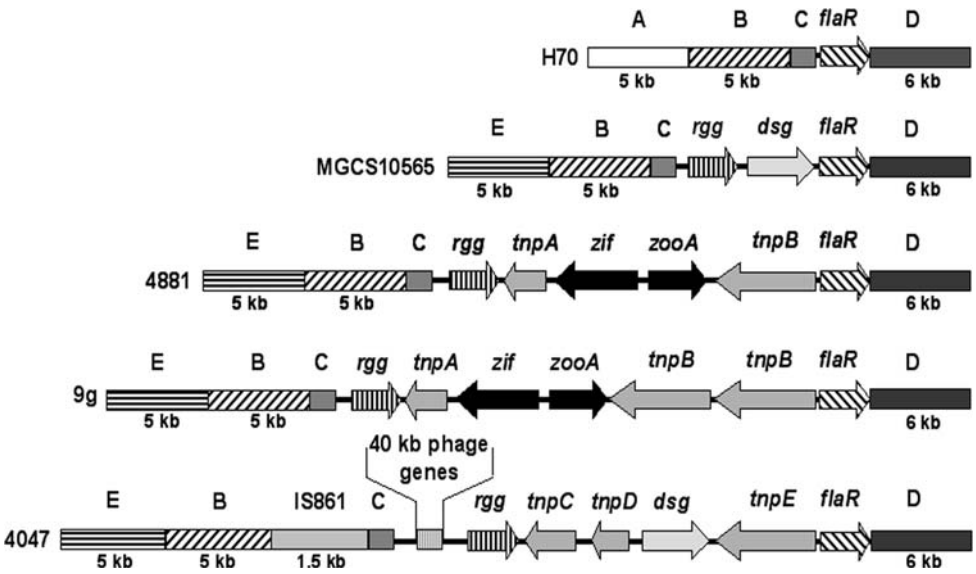

Fig. 3 Sequence analysis of *S. zooepidemicus* strains. GenBank accession numbers for sequence derived from *S. zooepidemicus* strains 4881 and 9g are U50357 and EU636005, respectively. Putative ORFs were identified using BioEdit software and the function of each ORF was deduced either by functional analysis (*zif* and *zooA* encode fully functional and expressed proteins) or from BLAST comparisons. Sequence blocks A–E were assigned on the basis of a high similarity, >99%

BLAST searching the *S. zooepidemicus* H70 genome sequence identified sequences corresponding to blocks C and B (Fig. 3). However, to the LHS of block B, the sequence differed from those of *S. zooepidemicus* strains 4881 and 9g and appeared to contain genes associated with the broad host range conjugative transposon Tn5252, including genes involved in the transfer of conjugative elements from one bacterium to another. This has been designated block A (Fig. 3).

Searching to the LHS of *flaR* in the *S. zooepidemicus* MGCS10565 genome sequence identified the presence of *rgg*. The region spanning *rgg* to *flaR* contained an ORF (*dsg*) encoding a protein with 39% identity and 65% similarity to a putative permease designated Dsg present in *S. gordonii* DL-1. Searching further to the LHS of *rgg* in the *S. zooepidemicus* MGCS10565 genome sequence identified block C, block B, and block E (Fig. 3).

Searching to the LHS of *flaR* in the *S. equi* 4047 genome sequence identified the presence of *rgg*. The region spanning *rgg* to *flaR* was 7.3 kb and contained six ORFs. ORFs labeled *tnpC*, *tnpD*, and *tnpE* encoded proteins with similarity to transposase genes from various bacteria. Transposase similarities were as follows: *tnpC*, encoding a protein with 89% identity and 94% similarity to the transposase IS861 from *S. pyogenes* M1 GAS; *tnpD*, encoding a protein with 81% identity and 91% similarity to a transposase from *S. pyogenes* MGAS10750; and *tnpE*, encoding a protein with 30% identity and 53% similarity to a transposase from *Caldicellulosiruptor saccharolyticus* DSM8903. An ORF (*dsg*) encoding the same Dsg protein as seen in *S. zooepidemicus* MGCS10565 was also identified. Screening of the 24 *S. zooepidemicus* strains by PCR showed that *dsg* was not present in any of the zoocin A producers and was present in only 13 of the 21 nonproducer strains. Sequencing of the *rgg69-flaR6684* LT-PCR products showed that, when present in the nonproducer strains, *dsg* occurred at the same position as *zooA* and *zif* in the producer strains.

Searching further to the LHS of *rgg* in the *S. equi* 4047 genome sequence identified sequences of DNA labeled as phage-associated genes (40 kb), including a putative bacteriophage replication repressor, a bacteriophage excisionase, and a bacteriophage-related DNA polymerase, as well as genes with similarity to proteins involved in bacteriophage assembly and packaging of DNA into prophage heads. Four other regions were also identified: block C, IS861 (79% identical to the IS861 transposase OrfA and Orf B from *Streptococcus agalactiae* A909), block B, and block E (Fig. 3).

There do not appear to be any additional highly conserved regions between *S. equi* 4047 and *S. zooepidemicus* H70 for at least 100 kb to the LHS of the sequence depicted in Fig. 3. BLAST searching the *S. zooepidemicus* H70 genome for sequences similar to that of block E revealed that block E is located 1.8 million bp to the LHS of block A (or 370 kb to the RHS of block A). The sequences flanking the *zooA* and *zif* region in 4881 and 9g are more similar to the sequences found in *S. zooepidemicus* strain MGAS10750 than those of *S. zooepidemicus* strain H70.

Discussion

The presence of both *zif* and *zooA* in strains 4881, 9g, and 9h was confirmed by Southern hybridization analysis. The probable genetic relatedness of strains 9g and 9h to each other and their dissimilarity to 4881 was suggested by the

size of the *Eco*RI fragments that reacted with the probes. RAPD analysis and subsequent PFGE confirmed that strains 9g and 9h were isogenic and differed significantly from strain 4881.

LT-PCR confirmed that the sequences between *rgg69* and *flaR6684* were present among the majority of the *S. zooepidemicus* strains, but the various sizes of the PCR amplicons generated suggested considerable genetic variability within this region. The *dsg* sequence occurred in the chromosomes of non-zoocin A producers at the same position as *zooA* and *zif* in the zoocin A producers. If *dsg* were a conserved sequence that was replaced by *zooA* and *zif*, then we would expect to see it in all of the non-zoocin A producers. However, *dsg* was present in only 13 of the 21 nonproducer strains examined and was present among strains that displayed different RAPD profiles. Also, the presence of *dsg* within a given RAPD profile was variable; for example, all members of RAPD profile 5 had *dsg* but only some members of RAPD profiles 2, 3, and 4 were positive for *dsg* (data not shown). These observations confirmed that this region of the *S. zooepidemicus* chromosome was one of considerable genetic variability.

The marked differences among the *S. zooepidemicus* strains (the presence or absence of *dsg, zif* and *zooA* genes) and the different genetic lineages of the producer strains suggest that *zif* and *zooA* may have been acquired by horizontal gene transfer. Based on these results, sequencing of 4881 and 9g in the region of the *zif* and *zooA* genes was undertaken. The only significant difference in producer strains 4881 and 9g between block C and *flaR* was the presence of a second copy of *tnpB* in 9g. The *zif* and *zooA* intergenic regions, which contain the promoter and gene regulator sequences (data not shown), were identical. Thus, the observed twofold higher expression level of zoocin A in 4881 is not likely due to changes to these sequences. Continuous sequencing for 6 kb and a further 4 kb of partial sequence (data not shown) on the RHS of the *flaR* gene revealed that the RHS sequences in *S. zooepidemicus* strains H70, MGCS10565, 4881, and 9g and *S. equi* 4047 were very similar, suggesting that insertion or deletion of genetic material occurred to the LHS of *flaR*. In *S. equi* 4047, blocks B and C are separated by an insertion sequence but, otherwise, could be regarded as a single block common to all five strains. In *S. equi* 4047, block C is separated from *rgg* by a 40-kb cluster of bacteriophage genes, within which the exotoxins SeeH and SeeI (SePE-H and SePE-I) are encoded. SeeH and SeeI are superantigenic toxins unique to *S. equi*, and their genes are located adjacent to a gene encoding a phage muramidase (Artiushin et al. 2002; Alber et al. 2004). SeeH and SeeI are absent in *S. zooepidemicus* and it has been suggested that their acquisition by phage-mediated transfer was an important event in the formation of the more virulent *S. equi* from the putative *S. zooepidemicus* ancestor (Artiushin et al. 2002). The commonality of the regions block C and *flaR* suggests that the target sites for insertion of the *zooA-zif* locus lies between these two sequences.

The original hypothesis regarding the horizontal acquisition of the *zooA-zif* locus developed from the similarity between the chromosomally encoded *zooA-zif* locus and the plasmid-encoded lysostaphin and lysostaphin immunity factor locus (*end-epr*) (Beatson et al. 1998). The mechanism of *zooA-zif* transfer is not known, but such transfer appears to be a rare event. Bacteria can share genetic information in three general ways: natural competence, transduction, and conjugation (Chen and Dubnau 2004). The genetic variability of the *S. equi* subspecies was recently highlighted with respect to the hyaluronic acid synthesis (*has*) operon. Intragene rearrangements, duplication events, and horizontal gene transfer are thought to have played a role in the evolution of the *has* operon (Blank et al. 2008). The recently completed genome sequence of *S. zooepidemicus* MGCS10565 revealed the presence of 22 of the 23 genes determined to be essential for streptococci to display natural competence (as determined for *Streptococcus pneumoniae*) (Beres et al. 2008). Much like the naturally transformable *S. pneumoniae, S. gordonii, S. sanguis*, and *S. mutans*, the MGCS10565 genome lacks prophages (streptococci that are not naturally competent are known to harbor prophages). These similarities between MGCS10565 and other naturally competent strains suggest that *S. zooepidemicus* may be naturally competent (Beres et al. 2008). It is probable that the *zooA-zif* locus was acquired by natural transformation in select strains of *S. zooepidemicus*. The results of this study confirm that the *zooA-zif* locus is not common within the *S. zooepidemicus* gene pool and that, when it is present, it integrates into the chromosome between block C and *flaR*. The chromosomal region in which *zooA* and *zif* are located appears to be one of considerable plasticity and a key differentiator of the *equi* and *zooepidemicus* subspecies.

Acknowledgments This work was supported in part by Grant Number R03AI073412 (to G.L.S.) from the National Institute of Allergy and Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Allergy and Infectious Diseases or the National Institutes of Health.

References

Alber J, El-Sayed A, Lämmler C, Hassan AA, Weiss R, Zschöck M (2004) Multiplex polymerase chain reaction for identification and differentiation of *Streptococcus equi* subsp. *zooepidemicus* and *Streptococcus equi* subsp. *equi*. J Vet Med B Infect Dis Vet Public Health 51:455–458

Artiushin SC, Timoney JF, Sheoran AS, Muthupalani SK (2002) Characterization and immunogenicity of pyrogenic mitogens

SePE-H and SePE-I of *Streptococcus equi*. Microb Pathog 32:71–85

Barnham M, Cole G, Efstratiou A, Tagg JR, Skjold SA (1987) Characterization of *Streptococcus zooepidemicus* (Lancefield group C) from human and selected animal infections. Epidemiol Infect 98:171–182

Beatson SA, Sloan GL, Simmonds RS (1998) Zoocin A immunity factor: a *femA*-like gene found in a group C streptococcus. FEMS Microbiol Lett 163:73–77

Beres SP, Sesso R, Pinto SWL, Hoe NP, Porcella SF, DeLeo FR, Musser JM (2008) Genome sequence of a lancefield group C *Streptococcus zooepidemicus* strain causing epidemic nephritis: new information about an old disease. PLoS ONE 3:e3026

Bert F, Branger C, Lambert-Zechovsky N (1996) Analysis of genetic relationships among strains of groups A, C and G streptococci by random amplified polymorphic DNA analysis. J Med Microbiol 45:278–284

Beukes M, Hastings JW (2001) Self-protection against cell wall hydrolysis in *Streptococcus milleri* NMSCC 061 and analysis of the millericin B operon. Appl Environ Microbiol 67:3888–3896

Blank LM, Hugenholtz P, Nielsen LK (2008) Evolution of the Hyaluronic Acid Synthesis (*has*) operon in *Streptococcus zooepidemicus* and other pathogenic streptococci. J Mol Evol 67:13–22

Chen I, Dubnau D (2004) DNA uptake during bacterial transformation. Nat Rev Microbiol 2:241–249

Church GM, Gilbert W (1984) Genomic sequencing. Proc Natl Acad Sci USA 81:1991–1995

Gargis SR, Heath HE, Heath LS, LeBlanc PA, Simmonds RS, Abbott BD, Timkovich R, Sloan GL (2009) Use of 4-sulfophenyl isothiocyanate labeling and mass spectrometry to determine the site of action of the streptococcolytic peptidoglycan hydrolase zoocin. Appl Environ Microbiol 75:72–77

Heng NCK, Swe PM, Ting Y, Dufour M, Baird HJ, Ragland NL, Burtenshaw GA, Jack RW, Tagg JR (2006) The large antimicrobial proteins (bacteriocins) of streptococci. Int Congr Ser 1289:351–354

Holt JG, Krieg NR, Sneath PHA, Staley JT, Williams ST (eds) (1994) Bergey's manual of determinative bacteriology, 9th edn. Williams and Wilkins, Baltimore, MD

Keis S, Bennett CF, Ward VK, Jones DT (1995) Taxonomy and phylogeny of industrial solvent-producing clostridia. Int J Syst Bacteriol 45:693–705

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Power EGM (1996) RAPD typing in microbiology—a technical review. J Hosp Infect 34:247–265

Schofield CR, Tagg JR (1983) Bacteriocin-like activity of group B and group C streptococci of human and of animal origin. J Hyg (Lond) 90:7–18

Simpson WJ, Cleary PP (1987) Expression of M type 12 protein by a group A streptococcus exhibits phaselike variation: evidence for coregulation of colony opacity determination and M protein. Infect Immun 55:2448–2455

Skinner FA, Quesnel LB (1978) Streptococci. Academic Press, New York

Tagg JR, Bannister LV (1979) "Fingerprinting" beta-haemolytic streptococci by their production of and sensitivity to bacteriocine-like inhibitors. J Med Microbiol 12:397–411

Tagg JR, Dajani AS, Wannamaker LM (1976) Bacteriocins of gram-positive bacteria. Microbiol Rev 40:722–756

Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, Swaminathan B (1995) Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. J Clin Microbiol 33:2233–2239

Timoney JF (2004) The pathogenic equine streptococci. Vet Res 35:397–409