Examination of Australian *Streptococcus suis* Isolates From Clinically Affected Pigs in a Global Context and the Genomic Characterisation of ST1 as a Predictor of Virulence

Mark A O'Dea[a*], Tanya Laird[a], Rebecca Abraham[a], David Jordan[b], Kittitat Lugsomya[c], Laura Fitt[d], Marcello Gottschalk[e], Alec Truswell[a], Sam Abraham[a*]

[a] Antimicrobial Resistance and Infectious Disease Laboratory, School of Veterinary and Life Sciences, Murdoch University, Perth, Western Australia

[b] Wollongbar Primary Industries Institute, NSW Department of Primary Industries, NSW, Australia

[c] Department of Veterinary Microbiology, Faculty of Veterinary Science, Chulalongkorn University , Bangkok, Thailand

[d] ACE Laboratory Services, Bendigo, Victoria, Australia

[e] Laboratory of Research on *Streptococcus suis*, Faculty of Veterinary Medicine, University of Montreal, Saint-Hyacinthe, QC, Canada

Running title

Streptococcus suis in Australian pigs

*address correspondence to Mark O'Dea m.odea@murdoch.edu.au; Sam Abraham s.abraham@murdoch.edu.au

Abstract

Streptococcus suis is a major zoonotic pathogen that causes severe disease in both humans and pigs. In this study, we investigated *S. suis* from 148 cases of clinical disease in pigs from 46 pig herds over a period of seven years. These isolates underwent whole genome sequencing, genome analysis and antimicrobial susceptibility testing. Genome sequence data of Australian isolates was compared at the core genome level to clinical isolates from overseas. Results demonstrated eight predominant multi-locus sequence types and two major *cps* gene types (cps2 and 3). At the core genome level Australian isolates clustered predominantly within one large clade consisting of isolates from the UK, Canada and North America. In particular, serotype 2 MLST25 strains were very closely associated with Canadian and North American strains. A very small proportion of Australian swine isolates (5%) were phylogenetically associated with south-east Asian and UK isolates, many of which were classified as causing systemic disease, and derived from cases of human and swine disease. In addition, we show that ST1 clones carry a constellation of putative virulence genes not present in other Australian STs, and that this is mirrored in overseas ST1 clones. Based on this dataset we provide a comprehensive outline of the current *S. suis* clones associated with disease in Australian pigs and their global context, and discuss the implications this has on antimicrobial therapy, potential vaccine candidates and public health.

Importance:

In this study, we examine in detail, the genomic characteristics of 148 *Streptococcus suis* isolates from clinically diseased Australian pigs. We report the antimicrobial susceptibility profiles, virulence gene analysis and relationship to isolates from other regions of the world. We also demonstrate that ST1 clones, regardless of serotype, carry a large array of putative virulence genes while maintaining a small total gene content. This compilation of data has major ramifications for vaccine development, and refines the understanding of the distribution of various strains of this potentially-fatal zoonotic agent in the global pig industry

Introduction

Streptococcus suis is an agent of serious concern for the global swine industry, and an emerging zoonotic agent causing meningitis, sepsis, arthritis, endocarditis and endopthalmitis in humans. Human disease is particularly prevalent in the western Pacific and south-east Asian regions, where risk is predominantly attributable to poor sanitation at slaughter and ingestion of undercooked pork, followed by Europe and then significantly less in the Americas (1). Zoonotic infection with *S. suis* has occurred in Australia, although there are few reports in the literature. The most recently published cases occurred in 2007 and 2008 in an abattoir worker from Victoria and two piggery workers from New South Wales respectively (2, 3) with each case attributed to serotype 2, which, along with serotype 14, is one of two globally-dominant serotypes detected in human cases.

In swine, *S. suis* is predominantly carried on the tonsils, but also within nasal cavities, genital and gastrointestinal tracts of many clinically healthy pigs, a feature which complicates management strategies in herds with disease outbreaks. Despite this carriage in healthy pigs, *S suis* is a cause of a wide range of clinical syndromes ranging from sudden death in the peracute manifestation of disease, to meningitis, septicaemia, endocarditis, arthritis and pneumonia. The clinical manifestations generally occur in the post-weaning period, although can occur less frequently in suckers and in adult pigs (4). Isolates obtained from areas indicative of invasive disease such as joints, brain tissue, heart and abdomen may be referred to as systemic, or causative of systemic infection, in contrast to isolates obtained from lung tissue.

Treatment for affected animals is usually reliant on administration of β-lactams such as penicillin and amoxicillin, and, where permitted, farms may prophylactically treat all pigs at the peri-weaning stage. Another management option employed by some producers is vaccination, usually in the form of bacterins produced as autogenous vaccines from on-farm isolates (5). Due to the highly variable antigenicity of the capsular polysaccharides (CPS), *S. suis* is currently classified into 29 serotypes (six previously classified serotypes have been reassigned to different bacterial species) (6), and it is considered that protection (if any) is only provided by homologous vaccine serotypes. In addition to

this, virulence factors and yet to be determined genetic factors may play a role in poor vaccine efficacy (7), although virulence factors are not necessarily protective antigens and the current uncertainty surrounding the role of virulence factors makes this difficult to assess.

A recent review on the worldwide distribution and typing of *S. suis* by Goyette-Desjardins et al. (2014) provided a comprehensive analysis of the predominant serotypes and MLSTs in swine production systems across the US, South America, Europe and South-East Asia (8). Of note in this review was the lack of information on Australian *S. suis* types, with no published surveillance since 1994 (9), a gap of 20 years. Following this a study published in 2015 on 45 Australian *S. suis* isolates (encompassing 3 human isolates from 2006-2008, and 42 swine isolates from 1981-2011) described 4 MLST's (1, 25, 369 and 28) (10).

Given the paucity of data available on *S. suis* associated with disease in Australian pigs, the aim of this study was to analyse and characterise a significant number of *S. suis* isolates from diseased pigs obtained across multiple production sites and spanning a seven year period. In addition to serotype and MLST, we report on virulence factors and antimicrobial susceptibility, and compare the core genome to international isolates to assess evolution in a global context.

Results

Samples

A total of 148 swabs from archived clinical isolates were transferred from ACE Laboratory Services (a major swine-industry referral lab) to the Murdoch University Antimicrobial Resistance and Infectious Diseases Laboratory for detailed molecular analysis. The majority of isolates were acquired from the lungs of diseased pigs and were classified as potentially virulent, followed by isolates from heart, brain, abdomen and joints which were classified as virulent isolates, and miscellaneous sites including upper respiratory tract, abscesses and lymphoid tissue (Table 1). Isolates spanned a seven year period with 2, 1, 29, 27, 9, 49 and 29 isolates from years 2010, 2011, 2013, 2014, 2015, 2016 and 2017 respectively, and two isolates for which the year was unknown. Isolates were originally

99 obtained from at least 10 of Australia's major pig production enterprises (17 samples were of
100 unknown origin) and encompassed 46 known individual farms (7 isolates had unavailable farm data).
101 Of the available farm data, 39 farms accounted for between one and five isolates each, five farms
102 accounted for between six and eight isolates each and one farm accounted for 33 isolates.
103

Isolation Site	Number of Isolates
Abdomen	2
Brain	12
Heart	12
Joint	6
Lung	93
Lymph node	1
Neck abscess	3
Upper respiratory tract	3
Unknown	16

104

105 **Table 1** Number of isolates and site of origin used in the study

106

107 **Genomic and phylogenetic analysis**

108 Genomic analysis of the 148 isolates revealed that 110 isolates belonged to 11 previously identified
109 MLSTs with 38 isolates belonging to one of 26 newly identified MLSTs. The most prominent MLSTs
110 were ST 27 (27/148), ST 25 (26/148), ST 28 (11/148), ST 483 (11/148) and ST 1 (10/148). The new
111 MLSTs assigned as a result of this study were ST 1031 – ST 1056 inclusive.

Analysis of the capsular genes resulted in two major serotypes being detected; serotype 2 (39/148) and serotype 3 (37/148). The other main serotypes were 1/2 (11/148), 16 (8/148), 19 (8/148), 8 (7/148) and 4 (7/148). Analysis of serotypes and MLST combined showed a high proportion of isolates as serotype 2 ST25 (17.6%), serotype 2 ST28 (6.1%) and serotype 3 ST27 (18.2%). Of serotypes where more than one isolate was present, serotype 15 (n=2), 16 (n=8), 21 (n=4) and 31 (n=5) all had previously unidentified STs.

The proportion of isolates carrying the main virulence factors *mrp*, *epf* or *sly* across all isolates was 7.4% (cps 1/2, 2, 14 and 19), 6.8% (cps 1/2, 2 and 14) and 33.1% (caps1/2, 2, 3, 4, 5, 7, 8, 11, 12, 14, 15, 16, 18, 19, 21, and 23) respectively. Only a single isolate carried both *mrp* and *sly*, while 6.8% of isolates carried both *mrp* and *epf*, and 6.8% carried *mrp, epf* and *sly*. The isolates carrying all three genes belonged to serotypes 1/2 (n=7), 2 (n=2) and 14 and all belonged to MLST 1, indicating a significant association between ST 1 and the carriage of all three genes (Fisher's exact test P-value <0.0001). Although predominantly isolated from lung, two of these isolates originated from brain and two from joint fluid and all four of these isolates were serotype 1/2. These serotype 1/2 ST1 isolates also carried more virulence genes with an average of 46 compared to all other isolates carrying 36 virulence genes and were clustered in virulence gene group 2 (Figure 1).

The main putative virulence genes present across all isolates included *enolase* (100%), *GAPDH* (100%), *gdh* (100%), *luxS* (100%), *pgdA* (100%), *stk* (100%), *divIVA* (99.3%), *dltA* (99.3%), *dnaJ* (99.3%), *ef-tu* (99.3%), *ldh* (99.3%), *oatA* (99.3%), *srtA* (99.3%) and *ypfJ* (99.3%) and this conserved group was the only common feature found in isolates associated with systemic disease (brain, joint, heart, abdomen isolates). Detailed examination of the virulence gene profiles demonstrates that the Australian isolates can be grouped according to four main blocks of virulence genes (Figure 1). Aside from the core virulence genes present in most isolates, group 1 can be broken into subgroups based on small clusters of genes, however no discernible pattern relating to potential virulence is present. Group 2 contains a conserved block of 45 virulence genes and is the only group to contain the genes

137 *dppIV*, *epf*, *mrp*, *sbp2*, *oppA* and *lde*. This group consists predominantly of serotype 1/2 and serotype
138 2 isolates, and all are ST1. Groups 3 and 4 are similar in gene composition with the exception that
139 group 3 isolates are lacking a combination of *fba*, *gidA*, *ciaR* and *hylA*, and group 4 isolates have a
140 combination of *iga*, *sly* and in some cases *fbps* and *dnaK*. However group 3 isolates have a
141 combination of *iga*, *sly* and *Tran* which is not present in Group 4.

142

143 **Figure 1**: Presence of putative virulence factors in all 148 isolates from this study. Blocks indicate
144 presence of individual genes. Genes are identified by columns and isolate number, MLST and
145 serotype are identified in rows.

146

147 Analysis of the amino acid sequence of GDH across the Australian strains showed that all ST1 isolates
148 had amino acid substitutions of A, S, K and K at positions 296, 299, 305 and 330 respectively,
149 whereas non-ST1 isolates had the substitutions A,A,E,E, or T, A, E, E at the corresponding positions.

150 On average, the total gene content of systemic isolates from brain and joint, was 2057, significantly
151 lower than that from respiratory isolates which carried 2109 (Mann-Whitney *U* test p-value = 0.025).
152 When isolates from brain, joint, heart and abdomen were analysed, the average total gene content
153 was 2095, which was not significantly different from respiratory isolates (p-value=0.682). When the
154 total gene content of isolates was compared to the number of virulence genes a trend was observed
155 based upon ST (Figure 2a). This was most apparent in the ST1 isolates which had a significantly
156 higher number of virulence genes compared to the other ST's, and the lowest number of total genes
157 with an average total gene content of 1918 (p-value<0.00001). ST25 isolates also formed a tight
158 cluster with a mean of 35.7 virulence genes and a relatively low total mean gene content of 2050.
159 Despite having a mean virulence gene content of 36, ST28 isolates clustered with some of the
160 highest total gene contents with all isolates having greater than 2100 genes and 9/11 having greater

than 2200 genes. These findings were mirrored by an assessment of global *S. suis* virulence genes against total gene content (Figure 2b). Most striking is the consistent clustering of ST1 isolates with a mean virulence gene content of 45.8 and a mean gene content of 1936. The six non-ST1 isolates within this cluster consist of ST107 (n=1), ST105 (n=2) and ST144 (n=3) which are single locus variants (SLV) of ST1 at the *dpr*, *gki* and *cpn*60 loci respectively. Additionally, for the published systemic isolates from Weinert et al. (2015), from which data was analysed and MLST able to be determined using our pipeline, 51/108 (47.2%) were ST1, with the next highest proportion being ST28 at 17/108. When SLVs were analysed, another seven isolates were found to be variants at the *gki* locus, such that when ST1 and ST1 SLV isolates were combined they accounted for 53.7% of systemic isolates.

Figure 2. a) Total gene content of individual Australian isolates plotted against virulence factor content based on MLST. **b)** Total gene content of international isolates for which MLST could be determined, plotted against virulence factor content.

Phylogenetic comparison of the *S. suis* isolated from Australian pigs against a global collection of *S. suis* strains resulted in the identification of four major clades (Figure 3). The strains from this study were present in all four clades. Clade 1 was highly divergent from the other clades, and contained 20 Australian strains, with strain 74-1662 (serotype 12 ST87), an isolate from lung tissue, being the most divergent. The majority of the Australian strains (89/148) clustered in clade 2, which was comprised entirely of strains from the UK, North America and Canada. Only eight strains were present in clade 4, which was made up of Vietnamese and UK isolates. Australian serotype 2 ST25 isolates clustered with Canadian and North American serotype 2 ST25 isolates within clade 2. To explore this relationship, further analysis of the core genome was performed on these isolates (Figure 4) demonstrating that despite the similarity, the Australian isolates can be distinguished from North America and Canadian strains.

Figure 3. Core genome SNP phylogeny of Australian and international *S. suis* isolates. Pink labels encompass clade 1, purple labels encompass clade 2, blue labels encompass clade 3 and green labels encompass clade 4. Yellow labels indicate Australian isolates. Orange fill encompasses Canadian/Nth American Serotype 2 ST25 and within this red fill encompasses Australian serotype 2 ST25.

Figure 4. Serotype 2 ST25 isolate core genome SNP phylogeny. Red clade is Canadian/Nth American isolates, blue clade is Australian clade. Remainder are Australian divergent isolates. Scale indicates SNP differences.

Antimicrobial susceptibility testing

All 148 isolates were subjected to micro-broth dilution to determine the MIC value against ten antibiotics belonging to eight classes (Table 2). A high proportion of the isolates were resistant to both tetracycline (99.3%) and erythromycin (83.8%). Low levels of resistance was observed for florfenicol (14.9%), penicillin G (8.1%), ampicillin (0.7%) and trimethoprim/ sulphamethoxazole (0.7%). None of the isolates were resistant to enrofloxacin.

202

drug	n	0.13	0.25	0.5	1	2	4	8	16	32	64	128	256	ns_ci	cr_ci
Ampicillin	148		99.3 (147)						.7 (1)					0.7(0,3.7)	0.7(0,3.7)
Ceftriaxone	148		88.5 (131)	7.4 (11)	2 (3)	1.4 (2)					.7 (1)			.	.
Ciprofloxacin	148		37.2 (55)	37.2 (55)	22.3 (33)	3.4 (5)								.	.
Enrofloxacin	148	14.2 (21)	44.6 (66)	40.5 (60)	.7 (1)									0.7(0,3.7)	0(0,2.5)
Erythromycin	148	11.5 (17)	4.1 (6)	.7 (1)	1.4 (2)			2 (3)	1.4 (2)	79.1 (117)				84.5(77.6,89.9)	83.8(76.8,89.3)
Florfenicol	148				2 (3)	4.1 (6)	79.1 (117)	11.5 (17)		.7 (1)	1.4 (2)	.7 (1)	.7 (1)	93.9(88.8,97.2)	14.9(9.6,21.6)
Gentamicin	148					6.1 (9)	40.5 (60)	43.2 (64)	9.5 (14)				.7 (1)	.	.
PenicillinG	148	81.8 (121)	4.7 (7)	4.7 (7)	3.4 (5)	3.4 (5)		.7 (1)		.7 (1)				12.8(7.9,19.3)	8.1(4.3,13.7)
Tetracycline	148				.7 (1)		.7 (1)	1.4 (2)	2 (3)	1.4 (2)	93.9 (139)			100 (97.5,100)	99.3(96.3,100)
TMS	148			91.9 (136)	5.4 (8)	2 (3)		.7 (1)						8.1(4.3,13.7)	0.7(0,3.7)

203

204 **Table 2** Distribution of minimum inhibitory concentrations for *S.* suis. Percentage of isolates classified as non-susceptible (ns ci) and/or clinically resistant (cr
205 ci) with corresponding 95% confidence intervals for those where breakpoints are available. Shaded areas indicate the range of dilutions evaluated. Vertical
206 bars indicate clinical breakpoint (where available). TMS; trimethoprim/sulfamethoxazole

Among the *S. suis* isolates, 15.6 % were multi-drug resistant (MDR; resistance to ≥3 classes of antimicrobial) (Table 3). Commonly identified MDR phenotypes included macrolide/phenicol/tetracycline (7.4%) and β-lactam/macrolide/tetracycline (6.1%). All isolates carried resistance against at least one antimicrobial class with 73.7% resistant to two classes of antimicrobial classes (2). The most common phenotypic profile was resistance to a combination of macrolides and tetracyclines with 67.6% of isolates determined to carry this resistance. When analysed by clonal type or serotype there was no discernible pattern in AMR phenotype.

The two dominant antimicrobial resistance genes detected in these isolates were *tetO* and *ermB*. The percentage of isolates carrying identified resistance genes for tetracyclines (95.3%) and macrolides (80.5%) support the MIC data with 96.3% and 83.8% of isolates being resistant to tetracyclines and macrolides respectively. Other resistant genes present included *aph*(3') (n=6), *fexA* (n=1), *optrA* (n=3), *lnuB* (n=6) and *spc* (n=5).

phenotype	n	%
1: mac	1	0.7
1: tet	15	10.1
2: mac_tet	100	67.6
2: phe_tet	9	6.1
3: bla_mac_tet	9	6.1
3: mac_phe_tet	11	7.4
4: bla_fpi_mac_tet	1	0.7
4: bla_mac_phe_tet	2	1.4

221 **Table 3** Phenotypic AMR profiles of *Streptococcus suis* isolated from Australian pigs. The breakpoints
222 used to classify isolates in this case are clinical breakpoints. mac; macrolides, tet; tetracyclines, phe;
223 phenicols, bla; beta lactams, fpi; folate pathway inhibitors.

224

225 **Discussion**

226 *Streptococcus suis* is a major cause of disease in pigs, with the potential for significant public health
227 impact. Along with antimicrobial therapy, control measures in swine herds rely on the production of
228 autogenous vaccines, with variable, and often poor, success rates (7, 11). Given the range of
229 serotypes and STs, in conjunction with proven and putative virulence factors and markers, choosing
230 appropriate clones for vaccine stock is difficult, and reliant on having adequate data. In this study we
231 present a detailed assessment of *S.suis* isolates from clinically affected Australian pigs, and compare
232 these isolates to internationally available genome sequence data. Key findings include the virulence
233 potential of ST1 clones, the zoonotic disease potential of Australian ST1 clones which cluster with
234 Vietnamese isolates from cases of human disease, and the limited evolution of Australian clones
235 from their global seed strains.

236 Analysis of 148 Australian isolates determined that the majority were serotype 2 (26%) or serotype 3
237 (25%), followed by serotype 1/2 (7.4%). This is consistent with global *S. suis* strains isolated from
238 cases of disease, particularly for serotypes 2 and 3 in Canada, North America and China (8). While
239 the majority of isolates were obtained from lungs, 32 isolates were obtained from brain, abdomen,
240 heart or joint, suggesting invasive or systemic isolates. Of these, 13 were serotype 2 and six were
241 serotype 1/2, however given this relatively small sample size of systemic isolates, an increased
242 pathogenicity cannot necessarily be ascribed to these serotypes. Significantly, even within this
243 sample set, there were 20 individual serotypes detected (Supplementary Table 1). Given that the
244 capsule is a recognised virulence factor and an immunogenic target (12), a minimum requirement
245 for vaccine production should be serotyping from an outbreak strain.

246 While ST27 and ST25 were the most prevalent MLSTs, 37 individual STs were identified, including 26
247 new MLSTs. This greatly expands upon previously available Australian data which documented four
248 STs including 1, 25, 28 and 369 (10). The prevalence of isolates as serotype 2 ST25 (17.6%) and
249 serotype 2 ST28 (6.1%) is consistent with those from North America, Canada, Germany, Spain, the
250 United Kingdom and China, and serotype 3 ST27 (18.2%) is consistent with reports from Spain and
251 the United Kingdom (8). These findings indicate that capsular and ST combinations are stable and
252 conserved, and it is likely that the circulating strains present in Australia are reflective of the seed
253 stock which was imported from the United Kingdom, Canada and New Zealand (13).

254 Despite conservation of serotype/ST combinations, the use of whole genome sequencing
255 demonstrates geographic clustering at the level of the core genome. Analysis of Australian and
256 international isolates demonstrates four clades, although more than 80% of the isolates are present
257 in clades 2 and 4. Clade 2 consists of isolates from the United Kingdom, Canada and Australia. It can
258 be seen that the Australian isolates form distinct clusters, which correlate with serotype/MLST
259 combinations. The most interesting and distinct cluster was seen in Clade 2, whereby Australian
260 serotype 2 ST25 isolates (Fig 4 red) clustered within Canadian serotype 2 ST25 isolates (Fig 4 yellow).
261 Further analysis by re-deriving the core genome for only these isolates showed two distinct
262 populations separated according to geographic origin. This further supports the hypothesis that
263 serotype/ST combinations are stable, however once introduced into a region these will then form
264 divergent populations which over time become phylogenetically distinct from the original clone.
265 Only eight Australian isolates were present in clade 4, which otherwise consisted entirely of
266 Vietnamese and UK isolates. These Australian isolates were all serotype 1/2 ST1: aside from a single
267 isolate which was serotype 14 ST1. This appears to indicate that the MLST of *S. suis* is a better
268 predictor of core genome phylogeny than the serotype. The presence of these isolates and their core
269 genome similarity to the south-east Asian isolates could be due to two reasons. Historically, assays
270 did not differentiate between serotypes 2 and 1/2, with isolates being reported as serotype 2 or
271 serotype 2 (plus 1/2) (14). For this reason it may be that these isolates are derived from European

serotype 1/2 isolates which were initially labelled as serotype 2. Another scenario is that a clinically normal swine worker from south-east Asia transmitted a precursor serotype 1/2 clone from the Asian region to Australian pigs. It is difficult to assess the likelihood of this pathway however, given that most human *S. suis* documentation refers to clinically affected humans, clinically affected humans have not been documented to infect other people or pigs, and that the epidemiology and carriage in clinically unaffected humans is not well defined (15).

There are an increasing number of virulence factors termed critical to virulent strains of *S. suis*, which arguably are not critical but in various combinations may affect the virulence of different clones (16). Historically there has been a focus on three particular virulence factors; muraminidase-released protein (*mrp*), extracellular protein factor (*epf*) and suilysin (*sly*), shown to be associated with highly virulent strains, with *cps*2 strains carrying all three genes considered the most virulent globally (17). Analysis of Australian isolates revealed that serotype 2 and 3 isolates were almost exclusively $mrp^{-}/epf^{-}/sly^{-}$. This is in agreement with early studies demonstrating that the presence of all three of these factors is not necessary for a clone to exhibit high levels of virulence (18), with 9/18 (50%) of the isolates obtained from joints or brain having this gene combination. The presence of the $mrp^{+}/epf^{+}/sly^{+}$ combination in 7/11 (63.6%) of serotype 1/2 isolates, along with 4/7 of these isolates being from joints or brain, indicates that Australian serotype 1/2 isolates have particular virulence potential, in contrast to earlier studies which concluded clinical disease in Canadian pigs with serotype 1/2 was likely due to inherent herd factors (19). It was notable that serotype 1/2 isolates were obtained from six separate farms and six separate production enterprises, indicating that this serotype is not confined to a single company or nucleus herd. The $mrp^{+}/epf^{+}/sly^{+}$ combination was also carried by all ST1 isolates, and these were consistently the isolates with the largest array of putative virulence factors. Like serotype 1/2 isolates, ST1 isolates were obtained from six farms and production enterprises, indicating this is a widespread, potentially virulent clone.

296 The clustering of isolates by MLST is generally reflected in the virulence gene content of isolates as
297 seen in Figure 2, providing further evidence that MLST is a significant indicator of clonal virulence
298 potential when investigating a disease outbreak. In the case of Australian ST1 isolates, a conserved
299 constellation of virulence genes which was not present in other MLSTs provides some insight into
300 the virulence of ST1 clones. While the majority of these genes have their potential virulence
301 characteristics defined in vitro, there is a distinct aggregation of bacterial adhesion factors such as
302 fibronectin binding factors *dpp IV*, *mrp, epf* and *sbp2* and oligopeptide binding protein *oppA* (20-23).
303 The presence of *ide*, an IgM protease would also be involved in early phase infection and attachment
304 to cells by cleaving IgM blocking cellular binding factors (24).

305 When this gene block was investigated in ST1 clones from overseas, it was also found to be highly
306 conserved across 226 isolates, with the exception of the *oppA* gene which was absent in 21.2% of
307 isolates. Taken together these factors suggest that both Australian and international ST1 clones have
308 a distinct fitness advantage in terms of binding ability towards host target cells. In order to
309 determine if the potential virulence of Australian ST1 isolates could be classified by analysis of the
310 GDH amino acid sequence, as had been reported in an overseas study, we compared of the amino
311 acid sequence of GDH across the Australian strains (25). We found that all ST1 isolates had amino
312 acid substitutions of A, S, K and K in positions 296, 299, 305 and 330 respectively, the same
313 combination of substitutions reported by Kutz et al. (2008) in highly virulent serotype 2 clones. None
314 of the other Australian STs had this sequence, further demonstrating that virulence may be
315 significantly linked to ST, in this case ST1.

316 Studies have shown that virulent *S. suis* clones have smaller genome sizes combined with a larger
317 number of virulence genes when compared with less virulent isolates (26). While the sample size of
318 Australian isolates from sites other than the respiratory tract was too small to confirm this, the
319 clustering of ST1 isolates as seen in Figure 2 clearly shows a high number of virulence genes in
320 association with the lowest total gene content of all STs examined. To see if this held true on a global

scale, we mapped the same data from 443 isolates which clearly demonstrated the same pattern, inclusive of six ST1 single locus variants. A similar pattern, with clustering in the mid genome size and virulence gene range could be seen with ST25 isolates, and again with a larger total gene number in ST28 isolates. This confers with reports suggesting ST28 clones are non-virulent and ST25 clones are of medium virulence (27) in mouse models. This data provides further evidence that ST1 clones worldwide carry high virulence gene content in combination with a low total number of genes and that virulence is more closely associated with MLST than cps type (28).

Antimicrobial resistance of Australian strains was similar to levels reported overseas with regards to tetracycline (99.3%), erythromycin (83.8 %) and trimethoprim/sulfamethoxazole (0.7%). Resistance to florfenicol was 14.9%, while all isolates were clinically susceptible to enrofloxacin, likely due to this being banned from use in food producing animals in Australia, and evidence of the success of this programme. Of concern was the observed clinical resistance to penicillin G, albeit at a relatively low level in 8.1% of isolates. This is a first line therapy for *S. suis,* and indeed the beta-lactams are used in human therapy (29). Penicillin resistance levels of 5% have been reported in isolates from England (30), and 0% in isolates from China (31). Therefore this is an aspect of *S suis* in Australia that must be carefully monitored from both an animal and public health point of view. It should also be noted that our analysis did not detect any common beta-lactamase genes, as was the case in a recent study of Spanish and Canadian serotype 9 isolates (32), indicating a potentially unknown mechanism of resistance in these isolates.

This study has greatly increased the data available on the circulating strains of *S. suis* in the Australian pig herd. Comparative genome analysing using Australian and overseas data revealed that production of autogenous vaccine stock should not be based totally on serotype of circulating strains. In fact it is our view that it is equally important to take into account MLST, as we have demonstrated that there is a strong correlation between MLST and virulence gene content. It is possible that the core genome, or specific gene constellations associated with MLST groupings as

346 seen with ST1 isolates in this study, may uncover suitable vaccine targets for development of pan-
347 MLST vaccines. Therefore, we recommend that the MLST of isolates from clinical cases be regularly
348 monitored to ensure they match vaccine stocks. With the rapidly expanding range of whole genome
349 sequence analysis pipelines such as that used in this study, clonal typing and virulence marker
350 determination of clones associated with disease outbreaks can be achieved within a matter of days.
351 Utilisation of this technology can allow for highly targeted autogenous vaccine use. Additionally, the
352 presence of bacterial adhesion factors associated with ST1 clones which were not present in other
353 clonal groups also presents a target for future vaccine studies, potentially in the form of multivalent
354 subunit vaccines.

355 Additionally, this study demonstrates the phylogenetic relationship of *S. suis* clones evolving from
356 single clonal types imported into a country. Australia's animal import policies have prevented live
357 pigs from being imported since 1987 (13), meaning that circulating clones are a derivation of those
358 from at least 30 years ago and likely introduced with stock from the United Kingdom, Canada and
359 New Zealand. As can be seen in clade 1 of Figure 3, approximately 13% (20/148) of Australian clones,
360 only one of which was from brain tissue, have diverged significantly from the majority of those
361 studied, and consist predominantly of serotypes 16 and 31, with all except isolate 74-1662 having
362 newly assigned MLSTs. Despite this, it can be seen that for the majority of isolates the core genome
363 has remained consistent with European and North American isolates, such that while Australian
364 clones evolved in isolation, they have continued clustering closely with their hypothesised genetic
365 predecessors.

366 In conclusion, Australian clones associated with disease in pigs consist predominantly of serotypes 2,
367 3 and 1/2, which is consistent with reports from other pig producing countries (8). Despite the
368 limited number, the characterisation of serotype 1/2 ST1 clones is significant, as all displayed
369 distinctive factors associated with highly virulent *S. suis*, along with grouping separately to other
370 Australian isolates in the Vietnamese/UK clade, potentially indicating a separate source of

371 introduction. In addition to this, these strains have only low levels of divergence from Vietnamese
372 and UK isolates from cases of swine and human systemic disease, making this a sequence type which
373 requires further investigation.

374

375 **Methods**

376 **Isolates.**

377 Swabs or fresh tissue from diseased pigs were submitted by consultant veterinarians to ACE
378 Laboratory Service, Victoria, Australia. Samples were plated onto Sheep Blood Agar (SBA),
379 MacConkey Agar and Chocolate Agar (Oxoid, Thermo Fisher Scientific), and incubated at 37°C for 24
380 hours. Suspect *S. suis* isolates, which showed alpha haemolysis on SBA and that were assessed to
381 demonstrate morphology consistent with *S. suis* were sub-cultured onto fresh plates for isolation.
382 Identification was carried out using matrix assisted laser desorption ionization-time of flight mass
383 spectrometry (MALDI-TOF) typing (Bruker). Following confirmation, isolates were harvested from
384 SBA plates into 1ml of Tryptic Soy Broth (TSB)+10 % glycerol and stored at -80°C in sterile
385 microcentrifuge tubes. A random subset of isolates was thawed and streaked on to SBA for
386 resurrection at 37˚C. Purity of each isolate was confirmed prior to harvesting plates onto transport
387 swabs for postage to Murdoch University.

388 **MIC Testing.**

389 All isolates were subjected to antimicrobial susceptibility testing via broth microdilution according to
390 the Clinical Laboratory Standards Institute (CLSI) Performance Standards. MIC results were
391 categorized as susceptible, intermediate and resistant using the clinical interpretative criteria
392 specified in CLSI performance standard VET01-S3 (33). If interpretive criteria was not present in
393 VET01-S3, CLSI performance standard M100-S25 was used (34).

394 **Whole-genome sequencing.**

DNA extractions were performed on the 148 isolates using a MagMax DNA multi-sample kit (ThermoFisher Scientific) according to the manufacturer's instructions, with the modification to omit the RNAse treatment step. Library preparation was performed with a Nextera XT kit with the only change from the manufacturer's instructions being an increased tagmentation time of 7 minutes. Sequencing was performed on an Illumina Nextseq 500 platform using a mid-output V2 (2 x 150 cycles) reagent kit.

Sequence analysis

All sequencing files were parsed through the Nullarbor bioinformatics pipeline (v1.20) (35) to determine MLST and antimicrobial resistance genes. A database was manually created containing all capsular serotypes and 40 previously described putative virulence genes (36). The Abricate programme was used to query contig files against the database to determine capsular type and virulence genes, with cutoffs of ≥95% coverage and ≥99% identity used to determine gene presence. Distinguishing serotype 2 from serotype 1/2, and serotype 1 from serotype 14 was performed as outlined by Athey et al. (2016) (37).

Phylogenetic trees were based on single nucleotide polymorphisms in the core genome. For genomic comparison against international *S. suis* isolates, 383 previously published sequences from Vietnamese, UK, North American and Canadian isolates (26, 38)were downloaded from NCBI and ENA. Genome annotation was performed using Prokka (v1.12) (39) and outputs were processed using Roary (v3.8.0) (40) for core genome determination and Gubbins (v2.2.3) (41) for recombination removal and alignment. Manual annotation of trees was performed in iTOL (v4.2) (42). Sequences for new MLST allele variations were uploaded to the *Streptococcus suis* MLST Databases site (https://pubmlst.org/ssuis/) for assignment of allele identification, and final allele combinations were then uploaded for assignment of new MLSTs.

Accession number(s). All sequence read data generated in this study has been deposited in the NCBI Sequence Read Archive under accession number SRP150885.

Supplementary Table 1. Summary table of isolate metadata from study.

1. Huong VT, Ha N, Huy NT, Horby P, Nghia HD, Thiem VD, Zhu X, Hoa NT, Hien TT, Zamora J, Schultsz C, Wertheim HF, Hirayama K. 2014. Epidemiology, clinical manifestations, and outcomes of *Streptococcus suis* infection in humans. Emerg Infect Dis 20:1105-14.
2. Tramontana AR, Graham M, Sinickas V, Bak N. 2008. An Australian case of *Streptococcus suis* toxic shock syndrome associated with occupational exposure to animal carcasses. Med J Aust 188:538-9.
3. Kennedy KJ, Jadeer AA, Ong CW, Senanayake SN, Collignon PJ. 2008. Two cases of *Streptococcus suis* endocarditis in Australian piggery workers. Med J Aust 189:413.
4. Gottschalk M. 2011. p 841-855. *In* Zimmerman JJ RA, Schwartz JK, Stevenson G (ed), Diseases of Swine. Wiley Publishers, Iowa.
5. Varela NP, Gadbois P, Thibault C, Gottschalk M, Dick P, Wilson J. 2013. Antimicrobial resistance and prudent drug use for *Streptococcus suis*. Anim Health Res Rev 14:68-77.
6. Tien le HT, Nishibori T, Nishitani Y, Nomoto R, Osawa R. 2013. Reappraisal of the taxonomy of *Streptococcus suis* serotypes 20, 22, 26, and 33 based on DNA-DNA homology and *sodA* and *recN* phylogenies. Vet Microbiol 162:842-9.
7. Segura M. 2015. *Streptococcus suis* vaccines: candidate antigens and progress. Expert Rev Vaccines 14:1587-608.
8. Goyette-Desjardins G, Auger JP, Xu J, Segura M, Gottschalk M. 2014. *Streptococcus suis*, an important pig pathogen and emerging zoonotic agent-an update on the worldwide distribution based on serotyping and sequence typing. Emerg Microbes Infect 3:e45.

444 9. Mwaniki CG, Robertson ID, Trott DJ, Atyeo RF, Lee BJ, Hampson DJ. 1994. Clonal analysis and
445 virulence of Australian isolates of *Streptococcus suis* type 2. Epidemiol Infect 113:321-34.
446 10. Groves MD, Jordan D, Chapman TA, Jassim RA. 2015. Multilocus sequence typing of
447 Australian *Streptococcus suis* type 2 by MALDI-TOF mass spectrometry analysis of PCR
448 amplicons. Vet Microbiol 177:394-7.
449 11. Torremorell M PC, Trigo E. 1997. Vaccination against *Streptococcus suis*: Effect on nursery
450 mortality. Swine Health and Production 5:139-143.
451 12. Smith HE, Damman M, van der Velde J, Wagenaar F, Wisselink HJ, Stockhofe-Zurwieden N,
452 Smits MA. 1999. Identification and characterization of the cps locus of *Streptococcus suis*
453 serotype 2: the capsule protects against phagocytosis and is an important virulence factor.
454 Infect Immun 67:1750-6.
455 13. Todd A. The Australian Pig Industry-200 years of development p69. *In* The 1988 Australian
456 National Pig Fair.
457 14. Smith HE, Veenbergen V, van der Velde J, Damman M, Wisselink HJ, Smits MA. 1999. The cps
458 genes of *Streptococcus suis* serotypes 1, 2, and 9: development of rapid serotype-specific
459 PCR assays. J Clin Microbiol 37:3146-52.
460 15. Segura M. 2009. *Streptococcus suis*: an emerging human threat. J Infect Dis 199:4-6.
461 16. Segura M, Fittipaldi N, Calzas C, Gottschalk M. 2017. Critical *Streptococcus suis* virulence
462 factors: Are they all really critical? Trends Microbiol 25:585-599.
463 17. Oh SI, Jeon AB, Jung BY, Byun JW, Gottschalk M, Kim A, Kim JW, Kim HY. 2017. Capsular
464 serotypes, virulence-associated genes and antimicrobial susceptibility of *Streptococcus suis*
465 isolates from pigs in Korea. J Vet Med Sci 79:780-787.
466 18. Smith HE, Vecht U, Wisselink HJ, Stockhofe-Zurwieden N, Biermann Y, Smits MA. 1996.
467 Mutants of *Streptococcus suis* types 1 and 2 impaired in expression of muramidase-released
468 protein and extracellular protein induce disease in newborn germfree pigs. Infect Immun
469 64:4409-12.

19. Martinez G, Harel J, Lacouture S, Gottschalk M. 2002. Genetic diversity of *Streptococcus suis* serotypes 2 and 1/2 isolates recovered from carrier pigs in closed herds. Can J Vet Res 66:240-8.

20. Ge J, Feng Y, Ji H, Zhang H, Zheng F, Wang C, Yin Z, Pan X, Tang J. 2009. Inactivation of dipeptidyl peptidase IV attenuates the virulence of *Streptococcus suis* serotype 2 that causes streptococcal toxic shock syndrome. Curr Microbiol 59:248-55.

21. Yu Y, Qian Y, Du D, Xu C, Dai C, Li Q, Liu H, Shao J, Wu Z, Zhang W. 2016. SBP2 plays an important role in the virulence changes of different artificial mutants of *Streptococcus suis*. Mol Biosyst 12:1948-62.

22. Zheng F, Shao ZQ, Hao X, Wu Q, Li C, Hou H, Hu D, Wang C, Pan X. 2018. Identification of oligopeptide-binding protein (OppA) and its role in the virulence of *Streptococcus suis* serotype 2. Microb Pathog 118:322-329.

23. Li Q, Fu Y, Ma C, He Y, Yu Y, Du D, Yao H, Lu C, Zhang W. 2017. The non-conserved region of MRP is involved in the virulence of *Streptococcus suis* serotype 2. Virulence 8:1274-1289.

24. Seele J, Singpiel A, Spoerry C, von Pawel-Rammingen U, Valentin-Weigand P, Baums CG. 2013. Identification of a novel host-specific IgM protease in *Streptococcus suis*. J Bacteriol 195:930-40.

25. Kutz R, Okwumabua O. 2008. Differentiation of highly virulent strains of *Streptococcus suis* serotype 2 according to glutamate dehydrogenase electrophoretic and sequence type. J Clin Microbiol 46:3201-7.

26. Weinert LA, Chaudhuri RR, Wang J, Peters SE, Corander J, Jombart T, Baig A, Howell KJ, Vehkala M, Valimaki N, Harris D, Chieu TT, Van Vinh Chau N, Campbell J, Schultsz C, Parkhill J, Bentley SD, Langford PR, Rycroft AN, Wren BW, Farrar J, Baker S, Hoa NT, Holden MT, Tucker AW, Maskell DJ, Consortium BRT. 2015. Genomic signatures of human and animal disease in the zoonotic pathogen *Streptococcus suis*. Nat Commun 6:6740.

27. Fittipaldi N, Xu J, Lacouture S, Tharavichitkul P, Osaki M, Sekizaki T, Takamatsu D, Gottschalk M. 2011. Lineage and virulence of *Streptococcus suis* serotype 2 isolates from North America. Emerg Infect Dis 17:2239-44.

28. Kerdsin A, Akeda Y, Takeuchi D, Dejsirilert S, Gottschalk M, Oishi K. 2018. Genotypic diversity of *Streptococcus suis* strains isolated from humans in Thailand. Eur J Clin Microbiol Infect Dis 37:917-925.

29. Wertheim HFL, Nghia HDT, Taylor W, Schultsz C. 2009. *Streptococcus suis*: An emerging human pathogen. Clin Infect Dis 48:617-625.

30. Hernandez-Garcia J, Wang JH, Restif O, Holmes MA, Mather AE, Weinert LA, Wileman TM, Thomson JR, Langfor PR, Wren BW, Rycroft A, Maskell DJ, Tucker AW, Consortium BT. 2017. Patterns of antimicrobial resistance in *Streptococcus suis* isolates from pigs with or without streptococcal disease in England between 2009 and 2014. Vet Microbiol 207:117-124.

31. Chen L, Song YJ, Wei ZG, Hei HK, Zhang AD, Jin ML. 2013. Antimicrobial Susceptibility, Tetracycline and Erythromycin resistance genes, and Multilocus Sequence Typing of *Streptococcus suis* isolates from diseased pigs in China. J Vet Med Sci 75:583-587.

32. Zheng H, Du P, Qiu X, Kerdsin A, Roy D, Bai X, Xu J, Vela AI, Gottschalk M. 2018. Genomic comparisons of *Streptococcus suis* serotype 9 strains recovered from diseased pigs in Spain and Canada. Vet Res 49:1.

33. CLSI. 2015. Antimicrobial Disk and Dilution Susceptibility Tests for Bacteria Isolated From Animals (VET01-S3) 3rd Edition. Clinical and Laboratory Standards Institute.

34. CLSI. 2015. Performance Standards for Antimicrobial Susceptibility Testing; (M100-S25). 25th Informational Supplement. .

35. Seeman T, Bullach D, Schultz M, Kwong J, Howden B. Nullarbor, https://github.com/tseemann/nullarbor.

519 36. Segura M, Calzas C, Grenier D, Gottschalk M. 2016. Initial steps of the pathogenesis of the
520 infection caused by *Streptococcus suis*: fighting against nonspecific defenses. FEBS Lett
521 590:3772-3799.
522 37. Athey TB, Teatero S, Lacouture S, Takamatsu D, Gottschalk M, Fittipaldi N. 2016.
523 Determining *Streptococcus suis* serotype from short-read whole-genome sequencing data.
524 BMC Microbiol 16:162.
525 38. Athey TB, Teatero S, Takamatsu D, Wasserscheid J, Dewar K, Gottschalk M, Fittipaldi N.
526 2016. Population Structure and Antimicrobial Resistance Profiles of *Streptococcus suis*
527 Serotype 2 Sequence Type 25 Strains. PLoS One 11:e0150908.
528 39. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068-9.
529 40. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane
530 JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics
531 31:3691-3.
532 41. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J, Harris SR.
533 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome
534 sequences using Gubbins. Nucleic Acids Res 43:e15.
535 42. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and
536 annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-5.

537

Figure 1: Presence of putative virulence factors in all 148 isolates from this study. Blocks indicate presence of individual genes. Genes are identified by columns and isolate number, MLST and serotype are identified in rows.

Figure 2. a) Total gene content of individual Australian isolates plotted against virulence factor content based on MLST. **b)** Total gene content of international isolates for which MLST could be determined, plotted against virulence factor content.

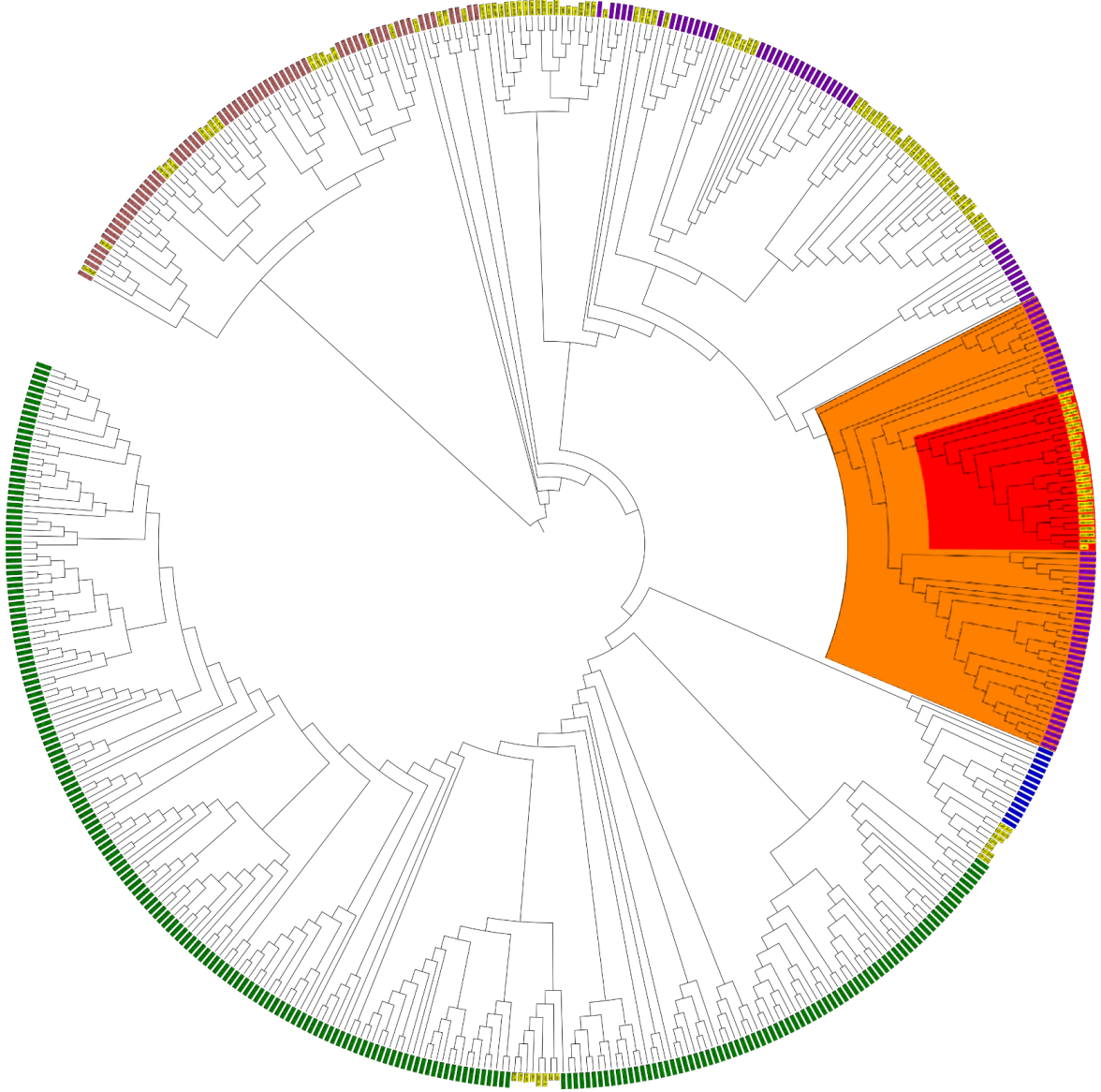

Figure 3. Core genome SNP phylogeny of Australian and international S. suis isolates. Pink labels encompass clade 1, purple labels encompass clade 2, blue labels encompass clade 3 and green labels encompass clade 4. Yellow labels indicate Australian isolates. Orange fill encompasses Canadian/Nth American Serotype 2 ST25 and within this red fill encompasses Australian serotype 2 ST25.

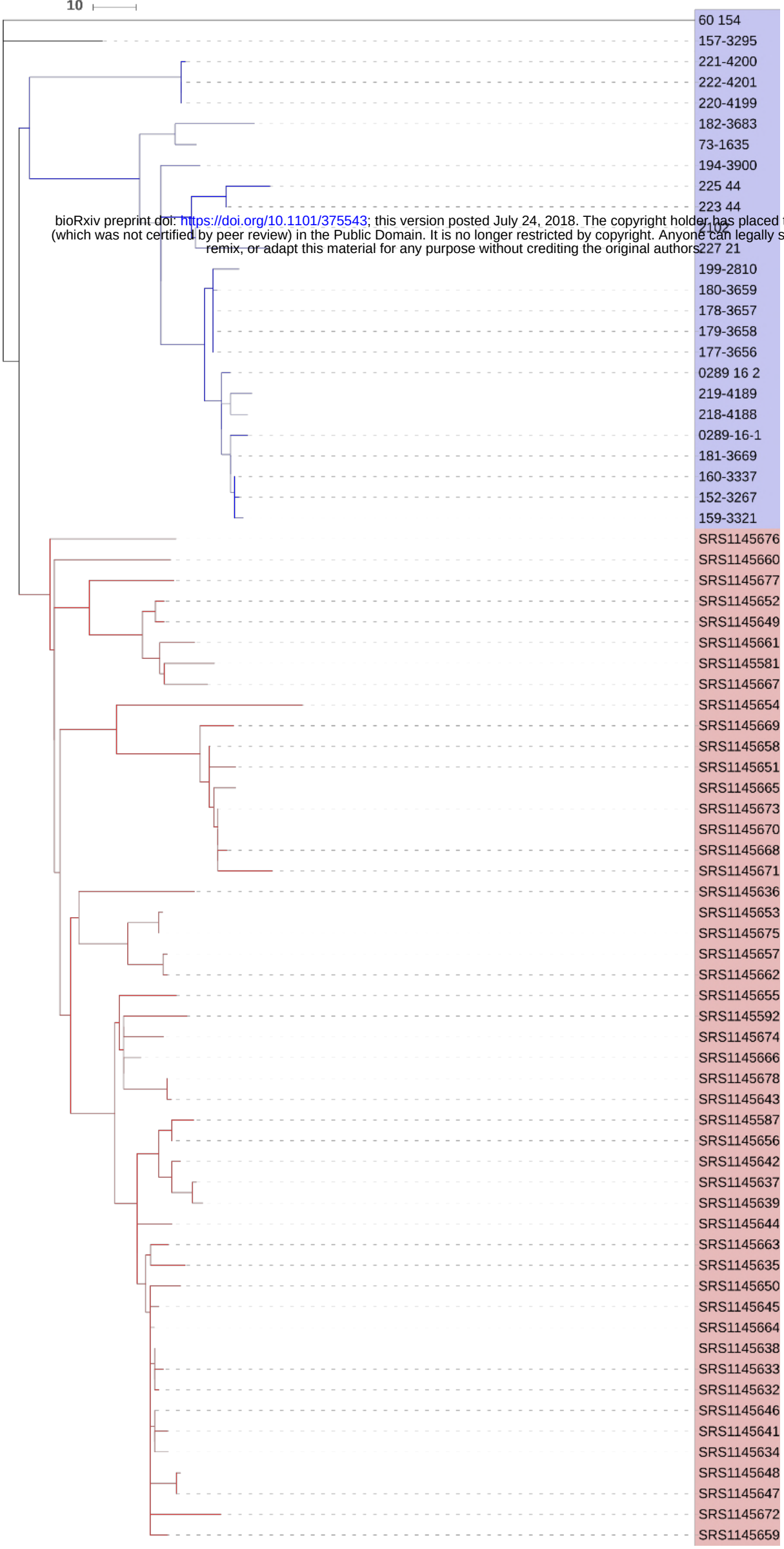

Figure 4. Serotype 2 ST25 isolate core genome SNP phylogeny. Red clade is Canadian/Nth American isolates, blue clade is Australian clade. Remainder are Australian divergent isolates. Scale indicates SNP differences