Analysis of Genomes of Bacterial Isolates from Lameness Outbreaks in Broilers

N. Simon Ekesi,[a]* Beata Dolka,[ab] Adnan A.K. Alrubaye,[a] Douglas D. Rhoads[a#]

[a]Cell and Molecular Biology program, University of Arkansas, Fayetteville, AR, USA

[b]Department of Pathology and Veterinary Diagnostics, Institute of Veterinary Medicine, Warsaw University of Life Sciences – SGGW, Warsaw, Poland.

Running Head: Broiler Lameness Bacterial Genomes

[#]Address correspondence to Douglas D. Rhoads, drhoads@uark.edu

*Present address: N. Simon Ekesi, Arkansas Regional Laboratory, Division of Microbiology, Food and Drug Administration, Jefferson, AR, USA

Abstract:

We investigated lameness outbreaks at commercial broiler farms in Arkansas. From Bacterial Chondronecrosis with Osteomyelitis (BCO) lesions, we isolated several distinct bacterial species. The results show that BCO-lameness pathogens on particular farms can differ significantly. Genomes for *Staphylococcus aureus* were highly related to chicken isolates from Europe, but present in the Arkansas area for at least a decade. Phylogenomics suggest that this *S. aureus* has been restricted to poultry for more than 40 years. Detailed analysis of genomes from two neighboring clades of *S. aureus* human and chicken isolates, identifies the acquisition of a particular pathogenicity island in the transition from human to chicken pathogen and that pathogenesis in chickens may depend on this mobile element. Genomes assembled from *Escherichia coli* isolates were quite different between farms, and more similar to genomes from different geographical locations. The *E. coli* phylogenomics suggests frequent host shifts for this species, in contrast to *S. aureus*. Isolate-specific genome characterizations will help further our understanding of the disease mechanisms of BCO-lameness, a significant animal welfare issue.

Importance:

Detailed inspection of the genome sequences of different bacterial species associated with causing lameness in broiler chickens reveals that one species, *E. coli*, appears to easily switch hosts from humans to chickens and other host species. Conversely, isolates of *S. aureus* appear to be restricted to specific hosts. One potential mobile DNA element has been identified that may be critical for causing disease in chickens for *S. aureus.*

Keywords: *E. coli*; *S. aureus*; lameness; broiler; genome

35 **Introduction**

36 Lameness poses animal health welfare issues which results in significant losses in poultry
37 production. Modern broilers selectively bred for rapid growth are particularly prone to leg
38 problems (1). Bacterial chondronecrosis with osteomyelitis (BCO) is the leading cause of
39 lameness in broiler and broiler breeder flocks (1-3). In birds that develop lameness, bacteria
40 translocate into the bloodstream via the integument, respiratory system or gastrointestinal
41 tract (1, 4, 5). Bacteria may have come from the immediate environment, or vertical transfer
42 though the egg (6). Bacteria that survive in the blood may colonize the proximal growth plate of
43 the rapidly growing femorae and tibiae inducing necrosis leading to BCO-lameness (1, 5, 7).
44 Stressors, or other factors contributing to immunosuppression, can facilitate bacterial
45 colonization and BCO spread in commercial poultry flocks (1, 8-13). In our research facility,
46 *Staphylococcus agnetis* is the primary bacterial isolate from lame broilers induced by growth on
47 raised-wire-flooring (2, 5). Genomic analysis of *S. agnetis* isolates from chickens and dairy cattle
48 demonstrate that the chicken isolates appear to be a clade arising from one branch of the cattle
49 isolates (14). Multiple bacterial species have been historically identified from BCO lesions in
50 different studies (1, 2, 11, 15-22) but there are few reports on characterization of genetic
51 relatedness within a species of chicken BCO isolates (20). Genomic analysis of a chicken isolate
52 of *Staphylococcus aureus* suggested a genetic basis for the jump from human to chicken for an
53 outbreak in the United Kingdom (23). Surveys of 20 broiler flock farms in Australia suggested
54 that avian pathogenic *E. coli* was the main BCO isolate (24). Genetic analysis by multilocus
55 sequence type (MLST), pulsed field gel electrophoresis and PCR phylogenetic grouping, of 15 *E.*
56 *coli* isolates from 8 flocks in Brazil indicated significant diversity for vertebral osteomyelitis, and
57 arthritis isolates, even within the same flock (25). The aim of this study was to characterize the
58 genomes of BCO isolates from three different commercial broiler farms in Arkansas.

59 **Results and Discussion**

60 *Diagnosis and Microbiological Sampling*

61 In June of 2016 we surveyed two commercial broiler houses on separate farms experiencing
62 outbreaks of BCO-lameness. Both houses had experienced a loss of cooling a week earlier,
63 causing heat stress for several hours. The farms were in rural, central Arkansas separated by 6.3

km, operated by the same integrator, and stocked from the same hatchery. The company veterinarian reported that samples from lame birds had been routinely submitted to a poultry health diagnostics laboratory and were primarily diagnosed as *E. coli.* For both farms we randomly collected lame birds for necropsy for BCO lesions. Blood and lesion swabs were collected from these birds, and house air was sampled, for bacterial species surveys.

In Farm 1 the birds were 31 days old. We diagnosed and necropsied six lame birds (Table 1). KB1 and KB2 were symptomatic of spondylolisthesis/kinky-back (KB). KB1 had BCO lesions in T4, left tibia, and both femora. We obtained thousands (TNTC; too numerous to count) of small green colonies from the T4 sample that were determined to be *Enterococcus cecorum*. KB2 had BCO of only the left tibia, but no colonies were recovered from sampling from this site. We did recover approximately 50 green colonies from what appeared to be a normal T4 that were *E. cecorum*. Lame3 and Lame4 both had bilateral BCO of the femora and tibiae. Lame3 had TNTC white colonies from microbiological sampling of T4 that were *S. agnetis*. Lame5 had bilateral FHN, bilateral tibial dyschondroplasia (TD), and pericarditis. We recovered green colonies (20 from left and TNTC from right) from the TD lesions that were *E. cecorum*. Due to limited supplies there was no microbiological sampling for Lame4 and Lame6.

For Farm 2 the birds were 41 days of age. We diagnosed and necropsied seven lame birds (Table 1). Lame7 was diagnosed with BCO lesions of the left femur, and right tibia. We obtained 30 purple colonies that were *E. coli* from the blood sample. We diagnosed Lame8 with bilateral BCO of tibia and femur. No colonies were obtained from sampling this bird. Lame9 was diagnosed with BCO of bilateral tibia, and the left femur, with evident pericarditis. Approximately 100 purple colonies were produced from sampling of the left femur that were determined to be *E. coli*. Lame10 was diagnosed with BCO of both tibia and right femur. Lame10 also had pericarditis. We got 40 white colonies from the left tibia, and 70 white colonies from the right femur that were *Salmonella enterica*. Lame11 had BCO lesions on both tibiae and femora. We got TNTC purple colonies from the left tibia, three from the left femur, and 15 from blood that were determined to be *E. coli.* Lame12 was diagnosed with bilateral BCO of tibia and femur. We recovered approximately 500 purple colonies from blood, 40 purple

92 colonies from the left tibia, and TNTC purple colonies from the left femur that were *E. coli*. Due
93 to limited supplies there was no microbiological sampling for Lame13.

94 Average plate counts for air sampling were 80 and 125 for Farm 1 and Farm 2, respectively. The
95 predominant species was *Staphylococcus cohnii* (~95%) with 3-4% *Staphylococcus lentus* and 1-
96 2% *E. coli*. Air sampling during our broiler experiments at the University of Arkansas Poultry
97 Research farm typically identify predominantly S. cohnii but not *E. coli* (26, 27).

98 In July 2019, we sampled a third commercial broiler farm (Farm 3) in Northwest Arkansas, more
99 than 88 km from Farms 1 & 2 and operated by a different integrator supplied from a different
100 hatchery. We sampled 11 lame birds at 35 days of age (Table 1). Lame14, was diagnosed with
101 BCO of bilateral tibiae and right femur. We recovered numerous white colonies from the left
102 femur, right tibia, and left tibia that were determined to be *Staphylococcus aureus*. Lame15 had
103 BCO of all femorae and tibiae. Culture plates had numerous white colonies from all four
104 sampled sites that were *S. aureus*. Lame16 had BCO of all tibiae and femorae. We recovered 10
105 green colonies from the right femur that were not analyzed, and 10 white colonies from the left
106 tibia that were determined to be *S. aureus*. We diagnosed Lame17 with BCO of the right femur
107 and both tibiae. We recovered numerous white colonies from all three sites that were *S.*
108 *aureus*. Lame18 was diagnosed with bilateral BCO of the femorae and the tibiae. We got TNTC
109 purple colonies from both femorae, and a few purple colonies from both tibiae, that were *E.*
110 *coli*. Lame19 had BCO of all femorae and tibiae. Swabs gave only a few colonies of *S.*
111 *epidermidis* that we assumed were contaminants during sampling. Lame20 had bilateral BCO of
112 the tibiae. We recovered a few green colonies of *Staphylococcus cohnii* which were presumed
113 contaminants during sampling. Lame21 had BCO of all femorae and tibiae. We isolated TNTC
114 white colonies from all four sites that were *S. aureus*. Lame22 was diagnosed with bilateral BCO
115 of femorae and tibiae. Microbial sampling only yielded only 10 green colonies from the right
116 femur that were determined to be *Staphylococcus simulans.* Lame23 was diagnosed with
117 bilateral BCO of femorae and tibiae. Culture plates had only white colonies, TNTC from both
118 femora and tibiae, that were *S. aureus*. Lame24 had BCO of all femorae and tibiae. We
119 recovered more than 100 white colonies from all four BCO lesions that were *S. aureus*.

120 *BCO Genome Assemblies*

121 We chose to characterize genomes for four representative *E. coli* isolates: 1409 for Farm 1,
122 1413 from Farm 2, and 1512 and 1527 from one bird on Farm 3 (Table 2). A hybrid assembly for
123 1409 produced 5.05 Mbp in 23 contigs that organized into 4 DNA assembly graphs. We resolved
124 the replicons using the long reads for contiguity analysis of the assembly graphs using the
125 Bandage software. The resolved genome contains a 4.84 Mbp chromosome, with episomes of
126 113.6, 108.7, and 2.3. The predicted serotype is O16. The hybrid assembly for 1413 produced
127 5.37 Mbp in 59 contigs and 3 DNA assembly graphs. Unfortunately, the Nanopore reads were
128 not of sufficient quality or length to complete a contiguity analysis of the entire genome, but
129 does identify at least two episomes of 98.8 kbp and 2257 bp. The predicted serotype is O78.
130 Draft assemblies were generated for *E. coli* 1512 and 1527. The assembly of 1512 contained
131 4.96 Mbp in 152 contigs with a N50 of 150 kbp. The assembly of 1527 was 4.90 Mbp in 179
132 contigs. The N50 was 97 Kbp with the largest contig of 258 Kbp. Both 1512 and 1527 are
133 predicted to be serotype O78, like 1413. We generated draft assemblies for 14 *S. aureus*
134 isolates from Farm 3 to examine genome diversity within a farm and within individual birds
135 (Table 2). Two separate colonies from lame bird were used for draft genome assembly (1510 &
136 1511, 1513 & 1514, 1515 & 1516, 1517 & 1518, 1519 & 1520, 1521 & 1522, 1523 & 1524). The
137 assemblies ranged from 2.79 to 2.82 Mbp in 60 to 96 contigs (excluding contigs < 300 bp). The
138 largest contigs were between 279 and 284 Kbp. N50 values ranged from 58 to 113 kbp. The L50
139 values ranged from 7 to 14 contigs. Each of the *S. aureus* assemblies had at least 3 circular
140 contigs (episomes). Table 2 summarizes all BCO isolate assemblies.

141 *Phylogenetic Comparison*

142 To examine the phylogenetic relationships between *E. coli* isolates from the three farms, we
143 identified the most closely related genomes according to PATRIC for 1409 (Farm 1), 1413 (Farm
144 2), 1512, and 1527 (Farm 3). We then identified the clades to which the closest related
145 genomes are assigned in the NCBI dendrogram for all *E. coli* genomes (Figure 1). The *E. coli* from
146 each farm clearly map to very different clades. We downloaded 57 genomes representing the
147 most closely related genomes identified by PATRIC for our four new *E. coli* genomes. We
148 generated a phylogenetic tree for all 61 genomes using Average Nucleotide Identity (ANI), with
149 isolates identified by region and host/source (Figure 2 and Table S1). Isolate 1409 grouped with

150 one chicken isolate from Pakistan and 4 isolates from chickens in China. The other closely
151 related isolates in that same branch were from cows, pigs, dogs and humans from either the
152 USA, China, France or Mexico. Isolate 1413 clustered with one *E. coli* isolate from layer
153 peritonitis in the USA, four Denmark chickens, and one USA ground turkey meat. Other closely
154 related isolates in the same branch came from USA chicken breast, thigh, brain pericardium and
155 two from feces, and from Bolivia human feces, UK chicken feces, and Latvia human clinical
156 sample. Isolate 1512 and 1527 are virtually identical, with an ANI of 0.99995, which is not
157 surprising since they were isolated from different sites in the same lame bird. Isolates 1512 and
158 1527 grouped with a pig isolate from China, a human isolate from France, and one chicken
159 isolate from the UK, with ANI of 0.9996. The other closely related isolates came from human in
160 Japan, sick chickens in Poland, pigs from South Dakota, a FDA water contamination project in
161 Arizona, USA chicken eggs, human clinical samples, a coliseptic turkey in Israel, US citizens
162 afflicted with haemolytic uraemic syndrome during a 2011 outbreak in Germany, Switzerland
163 chicken meat, and US deer feces. The ANI for 1512/5127 relative to 1409 and 1413, is 0.98,
164 while the ANI between 1409 and 1413 is 0.97, so the three genomes are equally distant from
165 each other.

166 The PATRIC closest-genome report for our new *S. aureus* chicken assemblies listed 11 *S. aureus*
167 genomes for isolates obtained from chickens. Ten of these isolates cluster exclusively on one
168 branch of a subtree in the NCBI genome dendrogram, and this cluster is next to a sister cluster
169 of 13 human isolates (Figure 3). The close relationship of these chicken isolates had been
170 previously demonstrated by Lowder *et al. (23)* using MLST to characterize *S. aureus* isolates
171 from avian hosts. Those analyses suggested a host switch from humans to poultry and a close
172 relationship among *S. aureus* isolated from the global chicken industry from 1970 to 2000. The
173 genome of isolate ED98 was assembled as the type strain representing a BCO isolate from *S.*
174 *aureus* from 1986 or 1987 in Ireland. We generated an ANI-based phylogenomic tree (Figure 4)
175 for our 14 new *S. aureus* chicken BCO isolate genomes including the 11 chicken isolates from
176 the ED98 clade, and included the two most closely related human isolates (based on PATRIC)
177 from the sister clade. Our 14 BCO *S. aureus* genomes grouped together with an average ANI of
178 0.999915, indicating a clonally-derived population. The 14 BCO *S. aureus* clustered with
179 genomes for four isolates (B4-59C, B3-17D, B2-15A, B8-13D) from retail chicken meat from

180 Tulsa, Oklahoma, in 2010 (ANI=0.9997) and a deep-wound lesion from a broiler from Poland in
181 2008 (ch23). Further distant are two additional Poland 2008 isolates (ch21 and ch22) from
182 chicken wounds and lesions, and ED98 from the Ireland 1980s chicken BCO outbreak. Near the
183 base of the chicken clade are ch3 and ch5 which are recorded as chicken commensals from
184 Belgium in 1976. The closest human isolates are CFBR-171 from a sputum sample from the USA
185 in 2012 and GHA2 from a 2018 patient in Ghana. Additional details on these isolates are
186 provided in Table S1. Our ANI based phylogenomics are in agreement with the work of Louder
187 *et al.* (23), where they used mutational analysis of bi-allelic polymorphisms to demonstrate the
188 tight relationship of poultry isolates of *S. aureus*. They had concluded that the "jump" to
189 chickens likely was associated with the same chicken isolates from Belgium, and likely derived
190 from human isolates in Poland. The first instance of association with pathogenesis in chickens is
191 ED98 in 1986 or 1987. Lowder *et al.* (23) had suggested that the host switch by ED98 was
192 associated with "acquisition of novel mobile genetic elements from an avian-specific accessory
193 gene pool, and by the inactivation of several proteins important for human disease
194 pathogenesis." This was evidenced by their demonstration of enhanced resistance to killing by
195 chicken heterophils.

196 We used the RAST SEED Viewer proteome comparison tool to determine whether we could
197 more precisely identify the genes associated with the progression from human pathogen (GHA2
198 and CFBR-171), to either chicken commensal (ch3 and ch5), or chicken pathogen (ED98, B4-59C,
199 and ch21). We used the ED98 predicted proteins as the reference, to identify polypeptides
200 conserved at >80% identity in chicken pathogens, while the chicken commensals and human
201 pathogens, encode a polypeptide of <80% identity. This filtering identified 33 protein encoding
202 genes (PEG) from the main chromosome and none from the three plasmids in ED98 (Table 3
203 marked with *). Two clusters were identified associated with two mobile elements: a
204 transposon and a *S. aureus* pathogenicity island (SaPI). There was also an additional short PEG
205 (PEG 48). To further define any association of the 33 PEGs with the jump from humans to
206 chickens we used tBLASTn to individually query the genomes of 11 *S. aureus* chicken isolates; 2
207 chicken commensals and 9 pathogens. We also queried a BLAST database created from 29
208 human *S. aureus* isolates including 10 from the sister clade in Figure 3 and an additional 19
209 representing clades nearby in the NCBI genome tree. The tBLASTn included all 33 PEGs as well

210 as immediate flanking or intervening PEGs (Table 3). PEG 48 is a hypothetical 56 residue
211 polypeptide that appears only in the chicken pathogens. The transposon-related region (PEG
212 1641-1659) contained no PEG that was highly conserved in all the chicken isolates and less
213 conserved in all the human isolates and the chicken commensals. On the other hand, the SaPI
214 region (PEG 755-773) contained a number of PEGs that appear highly conserved in all the
215 chicken pathogens. Some of these PEGs are also highly conserved in some of the human
216 pathogen isolates, but lacking in the chicken commensals (i.e., PEG 756-760, 767-771). There
217 are an additional 4 hypothetical PEGs (774-778) downstream of the SaPI terminase which
218 appear to be specific to the chicken pathogens. Close inspection of the results in Table 3
219 suggest that isolate ch9 appears to be intermediate between the chicken commensals, ch3 and
220 ch5, and all of the other chicken pathogens in the ED98 clade. This is concordant with our ANI
221 phylogenomic analysis (Figure 4) where ch9 is basal relative to the other chicken pathogens.
222 Thus, ch9 may represent a transitional state between the human pathogens as the bacterial
223 genome evolved increased virulence in chickens. The data is consistent with the *S. aureus*
224 genome evolving in two directions during the host switch from human to chickens: loss of the
225 SaPI to become commensals, and acquisition of new PEGs in the SaPI to become chicken-
226 specialists. Louder *et al.* (23) had originally identified this SaPI region as associated with the
227 jump to chickens but they surveyed other avian isolates by PCR and did not examine at the
228 resolution of individual PEGs.

229 SaPI are mobile elements that are packaged by "helper" virus assembly systems and integrate
230 at a specific location (28, 29). The SaPI in question is integrated between genes for a
231 Methionine ABC transporter (PEG 754) and a CsbD stress response protein (PEG 779). We used
232 tBLASTn of the *S. aureus* accessions in NCBI to survey for the number of entries with an
233 integrase (PEG 755) homolog directly downstream of the Met ABC transporter homolog; with
234 homolog threshold set at ≥80% identity. Out of 1012 PEG 754 homologs (Table 3) there were
235 307 entries with a PEG 755 homolog directly downstream. Conversely, there were 568 entries
236 where PEG 779 directly followed PEG 754, consistent with no SaPI inserted in those genomes.
237 Our tBLASTn searches also were consistent with the SaPI mobile element extending from the
238 integrase (PEG 755) through PEG 778, as we never found homologs of PEGs 777, 778, 780 or
239 781 directly downstream of PEG 754 (data not shown). Surveys by tBLASTn using the PEG

240 sequences from 754 to 779 demonstrate a high variability across the SaPI for whether the
241 coding sequence is conserved in the genomes of other *S. aureus* isolates (see Count for NCBI S.
242 aureus Desc column in Table 3). PEG 755, 759, 760, 764-773 appear to be present in many
243 (Count range: 199-839) genomes, consistent with many having defined functions for SaPI
244 mobilization and therefore conserved in many SaPIs. The regions for PEG 756-758, 761-763,
245 and 774-778 are found in far fewer (Count range: 10-89) genomes. We then used tBLASTn to
246 identify those *S. aureus* accessions in NCBI with homologs to these less conserved PEGs. Only 8
247 accessions contained all 11: ED98, ch21, ch22, B2-15A, B8-13D, B4-59C, B3-17D and X22 (Table
248 3). All are chicken pathogens from the same clade as ED98 with the exception of X22, which is a
249 genome deposited by the China Animal Disease Control Center but the host is not specified.
250 Notably, isolates ch3 and ch5 are chicken isolates but are listed as commensals, and not
251 necessarily pathogens. Both of these genomes have no SaPI downstream of the MetABC
252 transporter (PEG 754). Therefore, this SaPI does not specify chicken colonization, but rather
253 likely contributes to pathogenicity in chickens. Clearly, more work needs to be focused on the
254 actual functions of many of the genes in this SaPI and whether the chicken pathogenicity results
255 from the combination of PEG 48 and the SaPI. Some of the PEGs in this SaPI have homologs in
256 SaPIs in human isolates, however, the actual functions of the predicted polypeptides are not
257 well understood (Table 3). One of the hypothetical PEGs, PEG 777, in this SaPI appears to be
258 restricted to only chicken pathogens

259 We next explored whether there were signatures of selection in the evolution of the chicken
260 pathogen genome progressing from ED98 to the present. We used the RAST SEED Viewer
261 proteome comparison tool to further analyze the evolution of this *S. aureus* chicken clade since
262 1987 in the UK (Table S2). We selected our assembly for 1519 as it was the largest assembly
263 with the fewest contigs to represent the 2019 isolates from Farm 3. ED98 represents a 1986-
264 1987 isolate, ch21 is from Poland in 2008, and B4-59C is from 2010 in Tulsa, Oklahoma retail
265 poultry meat. The SEED Viewer filter was set with the ED98 proteome as reference, to identify
266 predicted proteins absent (<50% identity) in one or more of the other three proteomes (Table
267 S2). The analysis suggests that 32 proteins (31 phage and hypothetical proteins, and a efflux
268 pump for Tetracycline resistance) were lost between 1996 and 2008. Eight phage and
269 hypothetical proteins in ED98 and Ch21, were lost from the genome before appearance in Tulsa

270 in 2010, and only 4 hypothetical proteins in ED98, ch21 and B4-59C, are absent in 1519 in
271 Northwest Arkansas in 2019. We then reversed the analysis with 1519 as the reference to
272 identify new proteins that appeared in the lineage from ED98, through Ch21 to B4-59C to 1519.
273 The analysis identified 35 proteins present in 1519 for which the other 3 genomes lack a protein
274 with 50% or greater identity. Twenty-eight are phage, hypothetical or plasmid-maintenance
275 related. The remaining seven include a partial coding sequence for phosphoglycerate kinase, an
276 aminoglycoside N6'-acetyltransferase, a DUF1541 domain-containing protein, and a
277 lead/cadmium/zinc/mercury/copper transporting ATPase. Two open reading frames (genes
278 1916 and 1917) are also novel to the 1519 genome that have overlapping open reading frames
279 (ORFs) and may represent a frame shifted LPXTG cell wall anchor protein with a SdrC adhesin of
280 unknown specificity. The adjacent gene (1915) is for a SdrD adhesin of unknown specificity. We
281 reexamined this particular 4512 bp contig for assembly errors and could find none; templated
282 alignments of the Illumina HiSeq reads identified no frame shifts. BLASTn searches at NCBI
283 found identical regions in a few other *S. aureus* genomes from human isolates. Therefore, the
284 assembly appears correct, but that does not preclude ribosomal frame-shifting. The Unicycler
285 assembly predicted this contig to be circular, so this may be a plasmid that encodes one or
286 more adhesin functions. DUF1541 proteins of similar size are found in a wide range of different
287 bacterial species. The divalent cation transporter is found in many different *Staphylococcus*
288 species. The aminoglycoside-N6'-acetyltransferase (gene 1919) has no significant BLASTp
289 homologs in any *S. aureus* genome in NCBI, and the best homologs are 70% identical in isolates
290 of *Staphylococcus sciuri*, *Staphylococcus lentus,* and *Staphylococcus fleurettii.* This gene is
291 present in a 4357 bp contig that Unicycler could not circularize. However, the contig termini
292 each contain portions of a plasmid recombination MobE mobilization protein that likely could
293 be fused into one ORF using long read sequence data. The other genes in this contig are two
294 hypothetical proteins, a tetracycline resistance predicted region, and an ArsR-family
295 transcriptional regulator. However, this contig appears to possibly contain a mobile element
296 affecting antibiotic resistance with the aminoglycoside transferase and the tetracycline
297 resistance marker.

298 Further evolution of this genome is evidenced by proteins highly conserved (>80%) in 1519 and
299 B4-59C, but not (<50%) in ED98 and ch21. The Tulsa 2010 and Arkansas 2019 isolate genomes

300 contain 16 proteins not found in ED98 and ch21; including a toxic shock syndrome toxin 1 (PEG
301 327), and a phage associated exotoxin superantigen (gene 329). Interestingly this region
302 appears to be a probable SaPI, as inspection of neighboring genes identifies an integrase (PEG
303 335), terminase (PEG 326) and SaPI associated homologs (PEGs 325 and 330). PEG 327 and 329
304 are in a 97,219 bp contig predicted as circular that also encode a number of genes for exotoxins
305 and SaPI functions. The contig only contains two phage predicted proteins, so it may be a large
306 plasmid containing many virulence determinants. Additional PEGs found in B4-59C and 1519
307 genomes but not ED98 and ch21 include a cluster (PEG 2179, 2180, and 2181) of homologs to
308 hypothetical proteins found in SaPIs. However these PEGs are on a 3446 bp contig. Inspection
309 of the SEED annotation of the B4-59C assembly identify flanking integrase and terminase
310 homologs and other SaPI associated homologs. So, this region also may be a functional, mobile
311 SaPI.

312 There were only 3 proteins identified in 1519, B4-59C and Ch21, but not in ED98; two are
313 hypothetical (PEG 1197 and 1421) and the other a secretory antigen SsaA-like protein (PEG 49).
314 This secretory antigen has been associated with transposons and also annotates as a CHAP
315 domain protein, or putative cell wall lysis protein.

316 Our conclusions on the evolution of this clade of *S. aureus* that infects_chickens is that the
317 genome continues to evolve as it adapts from 1986 to 2020, with the acquisition of additional
318 adhesins and virulence determinants, possibly facilitated by transduction of new virulence
319 determinants in mobile SaPI. As such, this clade seems to be restricted, or specialized, for
320 infecting chickens. The clade appears to have been in the Oklahoma, Arkansas region for more
321 than a decade, but how it is transmitted to different farms or flocks is not clear. It could be
322 vertically transmitted from hen to chicks. Alternatively, chicks could be exposed at the
323 hatchery, or workers could spread the bacterium to farms through breakdowns in biosecurity.

324 In contrast the four *E. coli* genomes we characterized from three different farms show a very
325 different pattern. Isolates 1512 and 1527 are highly related as they came from different BCO
326 lesions in the same lame bird. We have previously reported that, for individual lame birds, we
327 recover the same species from multiple BCO lesions, and sometimes from the blood (2, 5). Our
328 surveys of the three farms also demonstrate that there may be a predominant BCO pathogen

329 within each farm, while in some broiler houses multiple species can be causing BCO lameness.
330 *E. coli* 1409, 1413, and 1512/1527 genomes are very distinct and come from very different
331 clades, which is very different from the patterns for *S. aureus,* or as we have reported for *S.*
332 *agnetis* infecting chickens (14). Reports from Brazil using virulence genes or MLST reported
333 distinct *E. coli* genotypes within a flock (25). However, their data could not place the *E. coli*
334 relative to those from non-chicken sources. The pattern we report from *E. coli* phylogenomics
335 is most consistent with a generalist pathogen that easily jumps to different host species.
336 Remarkably, two neighboring farms (Farm1 and Farm2) supplied by the same hatchery and
337 operated by the same integrator, had very different *E. coli* (1409 and 1413) involved in BCO
338 lameness outbreaks. This is more consistent with the *E. coli* on each farm originating from other
339 hosts (zoonoses) or each farm could have "evolved" it's own *E. coli* BCO pathogen over many
340 flocks and years.

341 **Conclusions**

342 Overall, the *E. coli* isolates from BCO lesions in Arkansas appear to be highly diverse, as they
343 derive from different clades that contain *E. coli* closely related to isolates from non-chicken
344 hosts. Conversely, the *S. aureus* isolates appear to come from a clade of chicken-specific
345 isolates associated exclusively with chicken hosts for at least five decades. Thus, the
346 phylogenomics suggest that *E. coli* appears to be a generalist and can switch hosts much more
347 easily than can *S. aureus*, which appears to be more of a specialist. This distinction may likely
348 derive from a difference in genome size as the *E. coli* genomes are roughly twice the size of the
349 *S. aureus* genomes. The larger genome size would allow *E. coli* to retain a greater diversity of
350 host-specificity virulence genes. Chicken pathogenicity of *S. aureus* appears to depend, in part,
351 on a specific, mobile pathogenicity island in which the function of several genes are not
352 defined. There is genomic evidence that the chicken pathogen clade continues to evolve,
353 possibly driven by integration of additional different pathogenicity islands.

354

355 **Materials and Methods**

356 Microbiological Sampling and Bacterial Species Identification

357 Diagnosis of and sampling of BCO lesions and blood have been described (1, 2, 7, 30). Initial
358 characterization of bacterial diversity by number of colonies of a particular color, was on
359 CHROMagar Orientation (CO; DRG International, Springfield, NJ), and further refined by
360 restreaking on CHROMagar Staphylococcus (CS; DRG International) (2, 5). Representative
361 colonies were then diagnosed to species by 16S rRNA gene sequencing (2, 5).

362 Air sampling was by waving open CO plates within the building. CO and then further evaluated
363 on CS plates. Representative colonies were typed to species as above.

364 *Genomic DNA Isolation and Sequencing*

365 Cultures were preserved in 40% glycerol at -80°C. Working stocks were maintained on tryptic
366 soy agar slants at 4°C. For DNA extraction, staphylococci were grown in tryptic soy broth to log
367 phase and DNA was isolated using as described (14). DNA isolation from *E. coli* used lysozyme
368 treatment, followed by organic extractions (31). DNA was quantified using a GloMax® Multi Jr
369 Detection System (Promega Biosystems Sunnyvale Inc., CA, USA) and purity evaluated with a
370 Nanovue spectrophotometer (Healthcare Biosciences AB Uppsala, Sweden). DNA size was
371 verified by agarose gel (1.5%) electrophoresis.

372 Library construction and Illumina MiSeq 2 x 250 sequencing were at the Michigan State
373 University Genomics Core Facility. Libraries for Illumina HiSeqX 2 x 125 sequencing were
374 prepared using a RipTide kit (iGenomX, Carlsbad, CA) and sequenced by Admera Health (South
375 Plainfield, NJ). Long reads were generated using Oxford Nanopore-MinION bar-code kit, as
376 described (14).

377 *Genome Assembly and Analysis*

378 *De novo* genome assemblies from short reads were generated as described (14). For hybrid
379 assemblies the long reads were phase corrected using the Illumina short reads and Ratatosk
380 v0.3 (https://github.com/DecodeGenetics/Ratatosk), before hybrid assembly with Unicycler
381 v0.4.8 (32). Unicycler hybrid assembly graphs were further analyzed for contiguity in Bandage
382 0.8.1 (33) to discern and export replicons. The PATRIC (Pathosystems Resource Integration
383 Center) webserver (34) was used for some Unicycler assemblies, assembly annotation, and
384 identification of similar genomes. Average Nucleotide Identity (ANI) values were determined

385 using pyANI 0.2.9 (35). ANI values were subtracted from 1 to generate distance matrices which
386 were submitted to FastME 2.0 (36) to generate Newick trees. Archaeoptryx 0.9928 beta (37)
387 was used to transform Newick trees into graphic representations. Assemblies were annotated
388 and compared using the Rapid Annotation using Subsystem Technologies (RAST) and SEED
389 viewer (38, 39). Serotype prediction was using the ECTyper module at GalaxyTrakr.org.
390 Chromosome-level genomes were obtained from NCBI using genome_updater
391 (https://github.com/pirovc/genome_updater).

392 **Acknowledgements**

393 Support has been provided in part by grants from Cobb-Vantress, Inc., Zinpro LLC, and the
394 Arkansas Biosciences Institute, the major research component of the Arkansas Tobacco
395 Settlement Proceeds Act of 2000. BD was supported by KNOW (Leading National Research
396 Centre), Scientific Consortium "Healthy Animal – Safe Food", under of Ministry of Science and
397 Higher Education No. 05-1/KNOW2/2015, and Department of Pathology and Veterinary
398 Diagnostics, Institute of Veterinary Medicine, Warsaw University of Life Sciences–SGGW,
399 Poland. The funders had no role in the design of this study, the interpretation of the results, or
400 the contents of the manuscript. Thank you to Dr. Mark Hart and Dr. Karen Christensen for
401 thoughtful comments regarding this manuscript.

402 **References**

403 1. Wideman RF. 2016. Bacterial chondronecrosis with osteomyelitis and lameness in broilers: a
404 review. Poult Sci 95 325-344.

405 2. Al-Rubaye AAK, Couger MB, Ojha S, Pummill JF, Koon JA, II, Wideman RF, Jr., Rhoads DD. 2015.
406 Genome analysis of *Staphylococcus agnetis*, an agent of lameness in broiler chickens. PLoS One
407 10:e0143336.

408 3. Thøfner ICN, Poulsen LL, Bisgaard M, Christensen H, Olsen RH, Christensen JP. 2019.
409 Longitudinal Study on Causes of Mortality in Danish Broiler Breeders. Avian Dis 63:400-410.

410 4. Jiang T, Mandal RK, Jr. RFW, Khatiwara A, Pevzner I, Kwon YM. 2015. Molecular survey of
411 bacterial communities associated with bacterial chondronecrosis with osteomyelitis (BCO) in
412 broilers. PLoS One 10:e0124403.
413 5. Al-Rubaye AAK, Ekesi NS, Zaki S, Emami NK, Wideman RF, Rhoads DD. 2017. Chondronecrosis
414 with osteomyelitis in broilers: Further defining a bacterial challenge model using the wire
415 flooring model. Poult Sci 96:332-340.
416 6. Anonymous. !!! INVALID CITATION !!! (2, 7).
417 7. Wideman RF, Prisby RD. 2013. Bone circulatory disturbances in the development of
418 spontaneous bacterial chondronecrosis with osteomyelitis: A translational model for the
419 pathogenesis of femoral head necrosis. Front Endocrinol (Lausanne) 3:183.
420 8. Andreasen JR, Andreasen CB, Anwer M, Sonn AE. 1993. Heterophil chemotaxis in chickens with
421 natural Staphylococcal infections. Avian Dis 37:284-289.
422 9. Butterworth A. 1999. Infectious components of broiler lameness: a review. . Worlds Poult Sci J
423 55:327-352.
424 10. McNamee P, McCullagh J, Thorp B, Ball H, Graham D, McCullough S, McConaghy D, Smyth J.
425 1998. Study of leg weakness in two commercial broiler flocks. Vet Rec 143:131-135.
426 11. McNamee PT, Smyth JA. 2000. Bacterial chondronecrosis with osteomyelitis ('femoral head
427 necrosis') of broiler chickens: a review. Avian Pathol 29:477-495.
428 12. Mutalib A, Riddell C, Osborne AD. 1983. Studies on the pathogenesis of Staphylococcal
429 osteomyelitis in chickens. II. Role of the respiratory tract as a route of infection. Avian Dis
430 27:157-160.
431 13. El-Lethey H, Huber-Eicher B, Jungi TW. 2003. Exploration of stress-induced immunosuppression
432 in chickens reveals both stress-resistant and stress-susceptible antigen responses. Vet Immunol
433 Immunopathol 95:91-101.

434 14. Shwani A, Adkins PRF, Ekesi NS, Alrubaye A, Calcutt MJ, Middleton JR, Rhoads DD. 2020. Whole
435 genome comparisons of *Staphylococcus agnetis* isolates from cattle and chickens. Appl Environ
436 Microbiol 86:e00484-20.
437 15. Kibenge FSB, Wilcox GE, Perret D. 1982. Staphylococcus aureus isolated from poultry in Australia
438 I. Phage typing and cultural characteristics. Vet Microbiol 7:471-483.
439 16. Emslie KR, Nade S. 1983. Acute hematogenous Staphylococcal osteomyelitis: a description of the
440 natural history in an avian model. Am J Pathol 110:333-345.
441 17. Hocking PM. 1992. Musculo-skeletal disease in heavy breeding birds., p 297-309. *In* Whitehead
442 CC (ed), Bone Biology and Skeletal Disorders in Poultry. Carfax Publishing Company, Abingdon,
443 United Kingdom.
444 18. Thorp BH, Whitehead CC, Dick L, Bradbury JM, Jones RC, Wood A. 1993. Proximal femoral
445 degeneration in growing broiler fowl. Avian Pathol 22:325-342.
446 19. Thorp BH. 1994. Skeletal disorders in the fowl: a review. Avian Pathol 23:203-236.
447 20. Butterworth A, Reeves NA, Harbour D, Werrett G, Kestin SC. 2001. Molecular typing of strains of
448 *Staphylococcus aureus* isolated from bone and joint lesions in lame broilers by random
449 amplification of polymorphic DNA. Poult Sci 80:1339-1343.
450 21. Tarr PE, Sakoulas G, Ganesan A, Smith MA, Lucey DR. 2004. Hematogenous enterococcal
451 vertebral osteomyelitis: report of 2 cases and review of the literature. J Infect 48:354-362.
452 22. Stalker MJ, Brash ML, Weisz A, Ouckama RM, Slavic D. 2010. Arthritis and osteomyelitis
453 associated with *Enterococcus cecorum* infection in broiler and broiler breeder chickens in
454 Ontario, Canada. J Vet Diagn Invest 22:643-645.
455 23. Lowder BV, Guinane CM, Ben Zakour NL, Weinert LA, Conway-Morris A, Cartwright RA, Simpson
456 AJ, Rambaut A, Nübel U, Fitzgerald JR. 2009. Recent human-to-poultry host jump, adaptation,
457 and pandemic spread of *Staphylococcus aureus*. PNAS USA 106:19545-19550.

458 24. Wijesurendra DS, Chamings AN, Bushell RN, Rourke DO, Stevenson M, Marenda MS,
459 Noormohammadi AH, Stent A. 2017. Pathological and microbiological investigations into cases
460 of bacterial chondronecrosis and osteomyelitis in broiler poultry. Avian Pathol 46:683-694.
461 25. Braga JFV, Chanteloup NK, Trotereau A, Baucheron S, Guabiraba R, Ecco R, Schouler C. 2016.
462 Diversity of Escherichia coli strains involved in vertebral osteomyelitis and arthritis in broilers in
463 Brazil. BMC Vet Res 12:140.
464 26. Alrubaye A, Ekesi NS, Hasan A, Koltes DA, Wideman Jr R, Rhoads D. 2020. Chondronecrosis with
465 osteomyelitis in broilers: Further defining a bacterial challenge model using standard litter
466 flooring and protection with probiotics. Poult Sci In Press.
467 27. Alrubaye AAK, Ekesi NS, Hasan A, Elkins E, Ojha S, Zaki S, Dridi S, Wideman RF, Rebollo MA,
468 Rhoads DD. 2020. Chondronecrosis with Osteomyelitis in Broilers: Further Defining Lameness-
469 Inducing Models with Wire or Litter Flooring, to Evaluate Protection with Organic Trace
470 Minerals. Poult Sci 99:5422-5429.
471 28. Tormo MÁ, Ferrer MD, Maiques E, Úbeda C, Selva L, Lasa Í, Calvete JJ, Novick RP, Penadés JR.
472 2008. *Staphylococcus aureus* Pathogenicity Island DNA Is Packaged in Particles Composed of
473 Phage Proteins. J Bact 190:2434.
474 29. Dearborn AD, Dokland T. 2012. Mobilization of pathogenicity islands by *Staphylococcus aureus*
475 strain Newman bacteriophages. Bacteriophage 2:70-78.
476 30. Wideman RF, Al-Rubaye A, Gilley A, Reynolds D, Lester H, Yoho D, Hughes JM, Pevzner I. 2013.
477 Susceptibility of 4 commercial broiler crosses to lameness attributable to bacterial
478 chondronecrosis with osteomyelitis. Poult Sci 92:2311-2325.
479 31. Sambrook J, Fritsch EF, Maniatis T. 1989. Molecular cloning: a laboratory manual. Cold Spring
480 Harbor Laboratory Press, Cold Spring Harbor, NY, USA.
481 32. Wick RR, Judd LM, Gorrie CL, Holt KE. 2017. Unicycler: Resolving bacterial genome assemblies
482 from short and long sequencing reads. PLoS Comput Biol 13:e1005595.

483 33. Wick RR, Schultz MB, Zobel J, Holt KE. 2015. Bandage: interactive visualization of de novo
484 genome assemblies. Bioinformatics 31:3350-3352.
485 34. Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, Chlenski P, Conrad N, Dickerman A,
486 Dietrich EM, Gabbard JL, Gerdes S, Guard A, Kenyon RW, Machi D, Mao C, Murphy-Olson D,
487 Nguyen M, Nordberg EK, Olsen GJ, Olson RD, Overbeek JC, Overbeek R, Parrello B, Pusch GD,
488 Shukla M, Thomas C, VanOeffelen M, Vonstein V, Warren AS, Xia F, Xie D, Yoo H, Stevens R.
489 2020. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities.
490 Nucleic Acids Res 48:D606-d612.
491 35. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2016. Genomics and taxonomy in
492 diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods 8:12-
493 24.
494 36. Lefort V, Desper R, Gascuel O. 2015. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-
495 Based Phylogeny Inference Program. Mol Biol Evol 32:2798-2800.
496 37. Han MV, Zmasek CM. 2009. phyloXML: XML for evolutionary biology and comparative genomics.
497 BMC Bioinformatics 10:356.
498 38. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM,
499 Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D,
500 Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O.
501 2008. The RAST server: Rapid annotations using subsystems technology. BMC Genomics 9:75.
502 39. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B,
503 Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of
504 microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206-14.

505

506 Table 1. Microbiological sampling of bone and blood samples from two commercial broiler
507 farms experiencing BCO outbreaks. BCO diagnoses are listed along with necropsy comments;
508 RF- right femur, LF- left femur, RT-right tibia, LT- the left tibia, T4- vertebral joint, N- normal,
509 THN- tibial head necrosis, THNS- THN severe, FHN- femoral head necrosis, KB- kinky back, TD-
510 tibial dyschondroplasia, FHS- femoral head separation. High Colony Counts lists the Species
511 diagnosis from 16S rRNA gene sequencing where a large number (>50) of same color colonies
512 were recovered from one or more Site(s) indicating high probability of infection.

		BCO Diagnoses					**High Colony Counts**	
Bird	**Farm**	**LT**	**RT**	**LF**	**RF**	**T4**	**Species**	**Site**
KB1	1	THN	N	FHN	FHN	KB	*E. cecorum*	T4
KB2	1	THN	N	N	N	N	*E. cecorum*	T4
Lame3	1	THN	THN	FHN	FHN	KB	*E. coli*	RT
							S. agnetis	T4
Lame4	1	THN	THN	FHN	FHN	KB	-	-
Lame5	1	TD	TD	FHN	FHN		*E. cecorum*	LT RT
Lame6	1	THN	THN	FHN	FHN	KB	-	-
Lame7	2	N	THN	FHN	N	N	*E. coli*	Blood
Lame8	2	THN	THN	FHS	FHN	N	-	
Lame9	2	THN	THN	FHN	N	KB	*E. coli*	blood LT LF
Lame10	2	THN	THN	N	FHN	KB	*S. enterica*	LT RF
Lame11	2	THNS	THNC	FHS	FHN	N	*E. coli*	blood LT LF
Lame12	2	THN	THN	FHT	FHT	N	*E. coli*	blood LT LF
Lame13	2	-	-	FHN	FHN	-	-	-
Lame14	3	THNS	THNS	FHN	N	-	*S. aureus*	LT LF RT
Lame15	3	THN	THNS	FHN	FHN	-	*S. aureus*	LT LF RT RF
Lame16	3	THNS	THNS	FHS	FHN	-	*S. aureus*	LT
Lame17	3	THNS	THNS	N	FHN	-	*S. aureus*	LT RT RF
Lame18	3	THNS	THNS	FHN	FHN	-	*E. coli*	LT LF RT RF
Lame19	3	THNS	THNS	FHN	FHS	-	*S. epidermidis*	LT RF
Lame20	3	THNS	THNS	N	N	-	*S. cohnii*	LT RT
Lame21	3	THNS	THNS	FHN	FHN	-	*S. aureus*	LT LF RT RF
Lame22	3	THNS	THNS	FHN	FHS	-	*S. simulans*	RF
Lame23	3	THNS	THN	FHN	FHN	-	*S. aureus*	LT LF RT RF
Lame24	3	THNS	THN	FHT	FHT	-	*S. aureus*	LT LF RT RF

513

514 Table 2. Bacterial genome assemblies produced in these analyses are listed by species, Isolate
515 designation, host source, assembly genome Status and statistics (Mbp, Contig count), NCBI
516 Biosample. Abbreviations are as in Table 1.

Isolate	Source	Status	Mbp	Contigs	Biosample
E. coli					
1409	RT Lame3	Finished	5.063	4	SAMN12285857
1413	Blood Lame12	Finished	5.375	59	SAMN12285859
1512	LF Lame18	Draft	4.962	152	SAMN13245724
1527	RF Lame18	Draft	4.904	179	SAMN13245725
S. aureus					
1510	LT Lame14	Draft	2.794	96	SAMN13245722
1511	RT Lame14	Draft	2.804	87	SAMN15589960
1513	LF Lame15	Draft	2.822	78	SAMN15589961
1514	RF Lame15	Draft	2.821	78	SAMN15589962
1515	RF Lame16	Draft	2.820	84	SAMN15589963
1516	LT Lame16	Draft	2.827	79	SAMN13245723
1517	LT Lame17	Draft	2.827	78	SAMN15589964
1518	RF Lame17	Draft	2.817	85	SAMN15589965
1519	LT Lame21	Draft	2.846	60	SAMN15589966
1520	RF Lame21	Draft	2.827	78	SAMN15589967
1521	RF Lame23	Draft	2.820	89	SAMN15589968
1522	RT Lame23	Draft	2.816	89	SAMN15589969
1523	RF Lame24	Draft	2.821	87	SAMN15589970
1524	LT Lame24	Draft	2.820	83	SAMN15589971

517

Table 3. Percent identity from tBLASTn searches of *S. aureus* genomes using the polypeptide sequences for protein encoding genes from ED98. PEG originally identified as possibly chicken pathogen specific are indicated (*). Annotations are derived from SEED augmented by BLASTp queries at NCBI. The 29 human isolates are from the clades surrounding the ED98 clade as described in the text and identified in Table S1. Avian isolates include 2 chicken commensals (ch3 and ch5); 10 chicken pathogens from the ED98 clade; a partridge (pa3), pheasant (ph3) and unknown host (X22) from the X22 clade; and chicken isolate ch24 (not in the ED98 or X22 clade). Count for NCBI *S. aureus* Desc is the number of descriptions returned (max 5000) with >80% identity for an NCBI tBLASTn search of *S. aureus* (Taxonomy ID: 1280), which roughly approximates the number of genome entries encoding a highly related polypeptide. Additional genome details are in the text and Table S1.

Protein Encoding Gene			tBLASTn percent identity for isolate genome(s)																	Count for NCBI *S. aureus* Desc
Protein Encoding Gene				Commensals		Pathogens										X22clade				
PEG	Polypeptide Annotation	Length	29 human	ch3	ch5	ch9	ED98	ch21	ch22	ch23	B2-15A	B3-17D	B4-59C	B8-13D	1519	pa3	ph2	X22	ch24	
47	hypothetical	638	54	93	93	100	100	100	100	100	100	100	67	100	100	93	93	0	0	25
*48	hypothetical	56	0	0	0	100	100	100	100	100	100	100	95	100	100	0	0	0	0	28
49	hypothetical	452	81	96	96	100	100	100	100	100	100	100	93	100	100	96	96	29	29	381
754	Methionine ABC transporter substrate-binding	274	100	100	100	100	100	100	100	100	100	100	100	100	100	99	99	99	99	1012
*755	Integrase SaPI	407	99	45	45	100	100	100	100	100	100	100	100	100	100	99	99	100	45	307
*756	SAP domain (DNA binding domain)	463	100	0	0	100	100	100	100	100	100	100	100	100	100	100	100	100	0	53
*757	hypothetical in SaPI	191	100	0	0	100	100	100	100	100	100	100	100	100	100	98	98	100	30	53
*758	hypothetical in SaPI	74	100	0	0	100	100	100	100	100	100	100	100	100	100	100	100	100	0	53
*759	winged helix-turn-helix domain-containing	91	100	0	0	100	100	100	100	100	100	100	100	100	100	100	100	100	41	394
*760	hypothetical in SaPI	49	100	0	0	100	100	100	100	100	100	100	100	100	100	100	100	100	96	368
*761	hypothetical in SaPI	71	41	0	0	45	100	100	100	100	100	100	100	100	100	42	42	100	0	89
*762	hypothetical in SaPI	120	0	0	0	0	100	100	100	100	100	100	100	100	100	0	0	100	0	24
*763	DUF1474 family	106	84	0	0	49	100	100	100	100	100	100	100	100	100	63	63	100	60	84
*764	primase alpha helix C-terminal domain-containing	290	83	0	0	94	100	100	100	100	100	100	100	100	100	97	97	100	26	200
*765	DNA helicase	570	21	0	0	99	100	100	100	100	100	100	100	100	100	99	99	100	22	200
*766	hypothetical in SaPI	127	30	0	0	99	100	100	100	100	100	100	100	100	100	98	98	100	0	199
*767	hypothetical in SaPI	214	97	0	0	100	100	100	100	100	100	100	100	100	99	95	95	100	92	405
*768	hypothetical in SaPI	114	97	0	0	100	100	100	100	100	100	100	100	100	100	99	99	100	98	353
*769	hypothetical in SaPI	193	98	0	0	100	100	100	100	100	100	100	100	100	100	97	97	100	0	312
*770	capsid morphogenesis B	43	97	0	0	100	100	100	100	100	100	100	100	100	100	97	97	100	0	312

Protein Encoding Gene			tBLASTn percent identity for isolate genome(s)																	Count for NCBI *S. aureus* Desc
Protein Encoding Gene				Commensals		Pathogens										X22clade				
PEG	Polypeptide Annotation	Length	29 human	ch3	ch5	ch9	ED98	ch21	ch22	ch23	B2-15A	B3-17D	B4-59C	B8-13D	1519	pa3	ph2	X22	ch24	
*771	spore coat	176	97	0	0	100	100	100	100	100	100	100	100	100	98	96	96	100	94	406
*772	hypothetical in SaPI	114	86	36	36	100	100	100	100	100	100	100	100	100	100	85	85	100	85	405
*773	terminase SaPI	190	97	46	46	96	100	100	100	100	100	100	100	100	100	99	99	100	96	839
*774	hypothetical	59	0	0	0	87	100	100	100	100	100	100	100	100	100	100	100	100	0	42
*775	hypothetical	130	0	0	0	95	100	100	100	100	100	100	100	100	100	100	100	100	0	48
*776	hypothetical	403	0	0	0	89	100	100	100	100	100	100	100	100	100	91	91	100	0	48
*777	hypothetical	109	44	44	44	52	100	100	100	100	100	99	100	100	100	58	58	100	55	10
*778	hypothetical	164	0	0	0	100	100	100	100	100	100	100	100	100	100	0	0	100	0	37
779	CsbD stress response	65	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	1011
1641	hypothetical	121	45	60	60	64	100	100	96	64	64	84	84	64	84	60	60	38	38	7
*1642	hypothetical	198	50	75	75	75	100	99	99	75	75	100	100	75	100	75	75	0	0	7
*1643	Secretory antigen SsaA-like transposon-related	343	61	79	79	79	100	100	99	79	79	99	99	79	99	79	79	46	46	7
*1644	hypothetical	644	51	71	71	70	100	100	100	70	70	98	96	70	98	71	71	0	0	7
1645	hypothetical	78	0	72	82	0	100	100	100	0	0	70	70	0	70	72	72	0	0	3
1646	hypothetical	56	0	0	0	85	100	100	85	85	85	85	87	85	85	0	0	0	0	2
1647	hypothetical	454	81	95	95	94	100	100	99	94	94	100	100	94	100	95	95	27	27	27
*1648	hypothetical	111	0	69	69	0	100	100	94	0	0	94	94	0	94	69	69	0	0	7
*1649	hypothetical	77	0	0	0	0	100	100	99	0	0	96	96	0	96	0	0	0	0	4
*1650	hypothetical	73	0	0	0	0	100	100	100	0	0	94	94	0	94	0	0	0	0	7
1651	hypothetical	832	85	91	91	90	100	90	100	90	90	100	97	90	100	91	91	0	0	387
*1652	hypothetical transposon-related	130	62	78	78	78	100	100	100	78	78	98	93	78	98	78	78	0	0	7
1653	hypothetical	87	60	72	72	72	100	100	100	72	72	97	82	72	97	72	72	0	0	7
1654	hypothetical transposon-related	359	65	86	86	87	100	96	100	86	86	98	91	86	98	86	86	0	0	6
1655	hypothetical	359	67	92	92	91	100	100	100	91	91	99	96	91	99	92	91	0	0	26
1656	hypothetical	101	63	92	92	92	100	100	100	92	92	98	94	92	92	92	92	0	0	27
1657	hypothetical	112	55	87	87	87	100	100	100	87	87	100	95	87	100	87	87	0	0	27
1658	conserved hypothetical transposon-related	102	26	85	85	85	100	100	100	85	85	100	85	85	100	85	85	0	0	27
*1659	hypothetical protein	95	71	78	80	87	100	100	100	87	87	99	91	87	99	78	78	0	0	22
1660	Acetyl-coenzyme A carboxyl transferase alpha	315	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	99	100	1014

Protein Encoding Gene			tBLASTn percent identity for isolate genome(s)																		Count for NCBI *S. aureus* Desc
Protein Encoding Gene				Commensals		Pathogens										X22clade					
PEG	Polypeptide Annotation	Length	29 human	ch3	ch5	ch9	ED98	ch21	ch22	ch23	B2-15A	B3-17D	B4-59C	B8-13D	1519	pa3	ph2	X22	ch24		
	chain																				

525

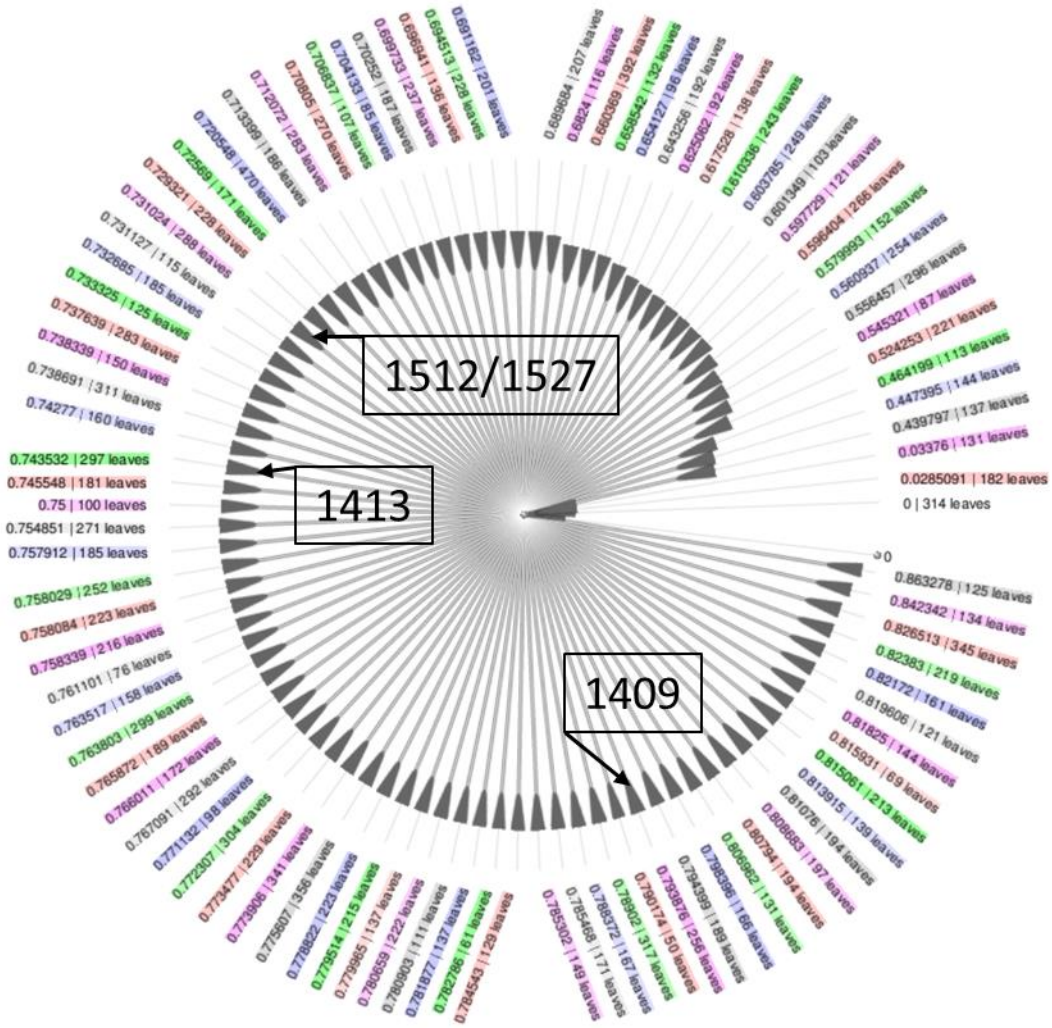


Figure 1. NCBI dendrogram based on genomic BLAST for 21,187 *E. coli* genomes (https://www.ncbi.nlm.nih.gov/genome/167). The locations of the nearest genomes for the *E. coli* isolates presented in this work are indicated on the dendrogram. Each node of the tree is indicated by the distance from the root and the number of leaves in that node

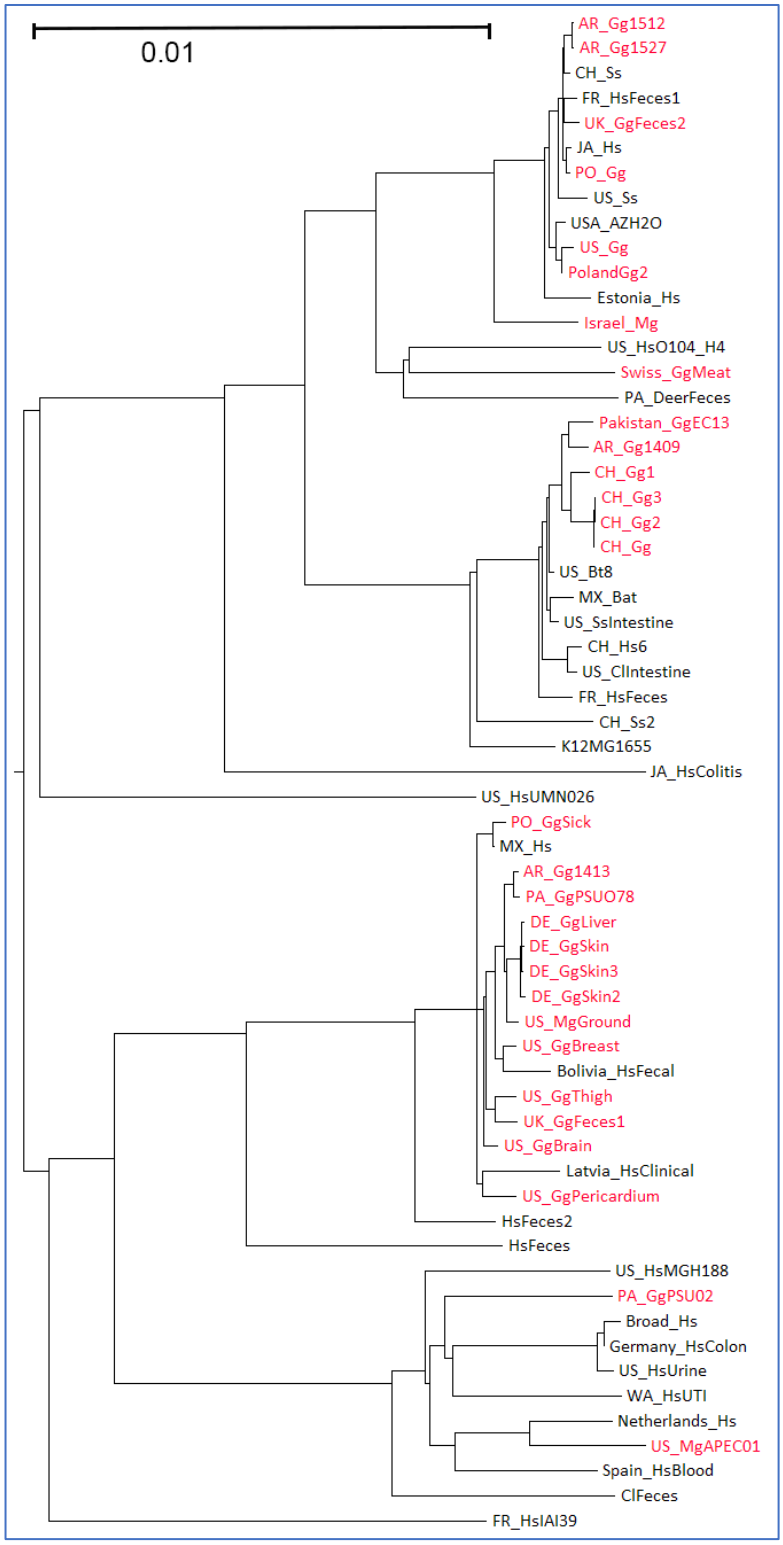
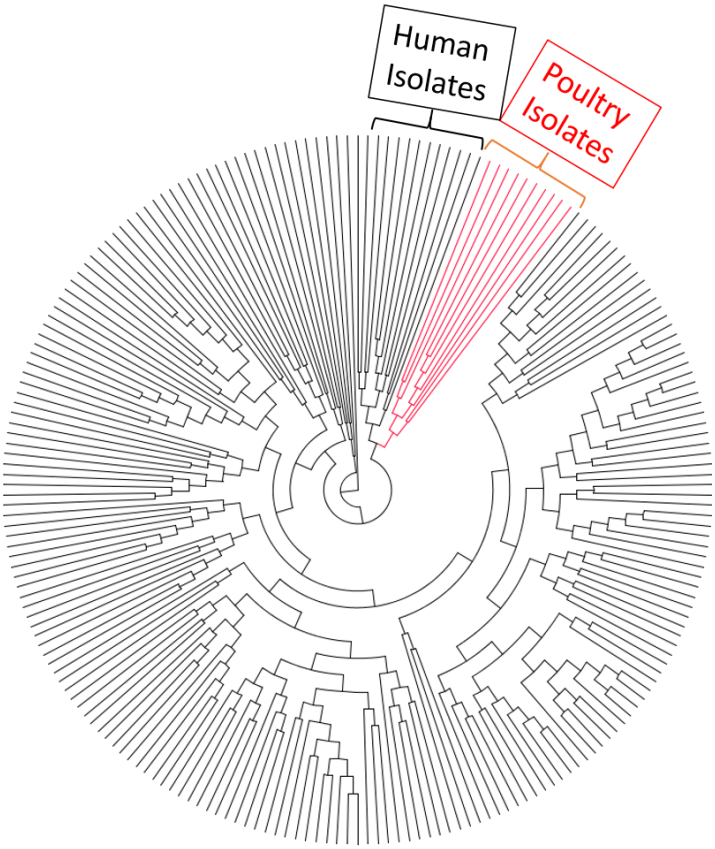


Figure 2. Phylogenetic tree for 61 *E. coli* genomes based on Average Nucleotide Identity. Key for isolate genomes is in Table S1. In brief, first two characters indicate location, first two characters after the underline indicate host, and remaining characters indicate source or isolate. Isolates in red are from poultry or poultry products.

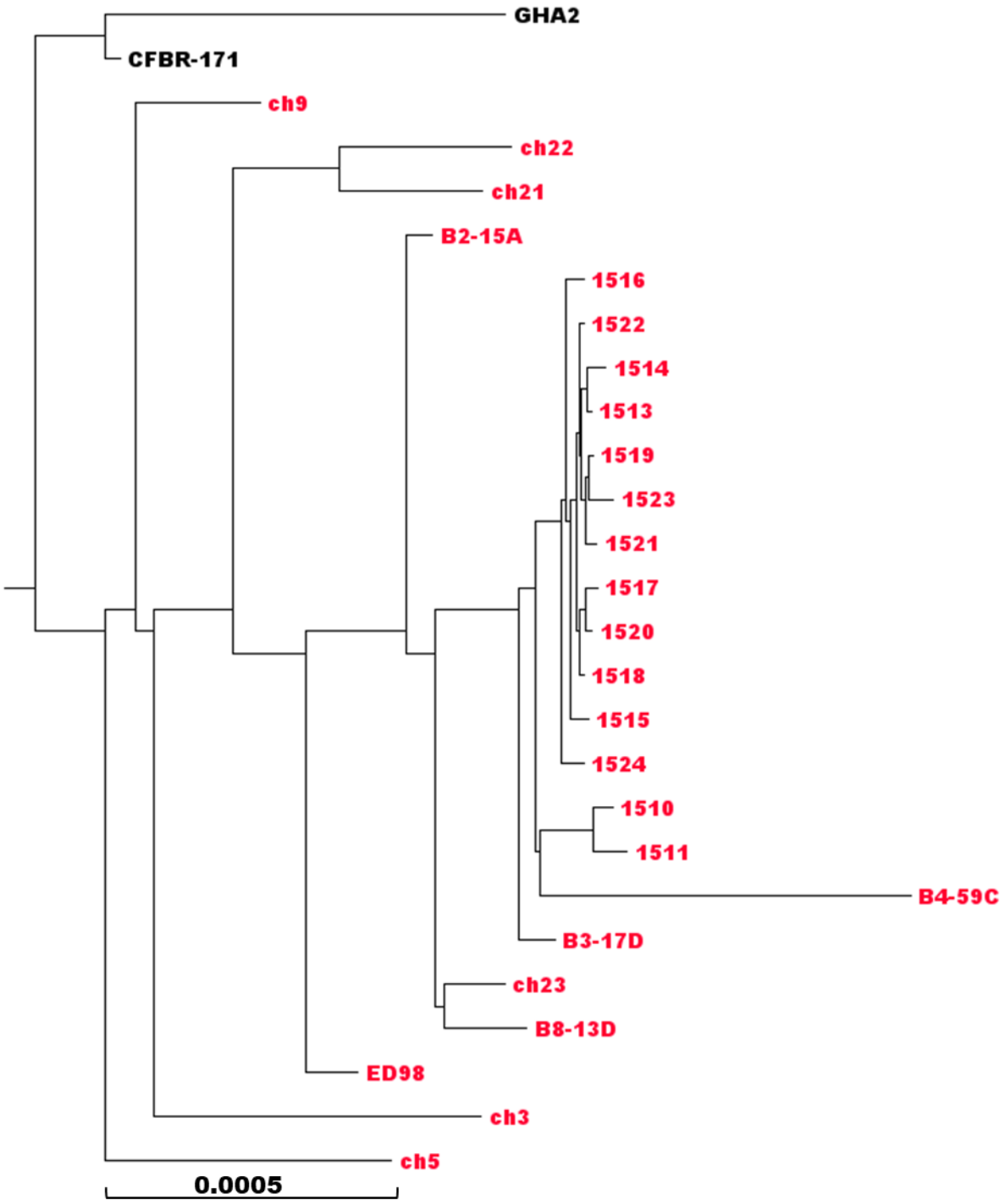
535



536

537 Figure 3. Phylogenetic subtree for 427 *S. aureus* genomes from NCBI based on whole genome BLAST
538 comparisons. This subtree was an expanded view for the branch containing known poultry isolate genomes.
539 The poultry isolate clade and a sister clade of human isolates are indicated.

540



541

542 Figure 4. Phylogenetic tree for 27 *S. aureus* genomes based on Average Nucleotide Identity. Key for isolate
543 genomes is in Table S1. The tree compares 25 known isolates from poultry or poultry products (in red) to the
544 closest related human isolates (in black).

545 Table S1. Genomes used for genomic analyses. Designation is the coding used in the trees or text, Assembly is the NCBI accession, Country is the
546 source of the isolate (if known), State is indicated for some USA isolates, Host is genus species and Source is the tissue/sampling from which the
547 isolate was obtained. BioProject is the NCBI registry for the isolate characterization.

Designation	Assembly	Country	State	Host	Source	Strain	BioProject
S. aureus							
1510	JACEHY000000000	USA	AR	*Gallus gallus*	Bone	1510	PRJNA554887
1511	JACEHW000000000	USA	AR	*G. gallus*	Bone	1511	PRJNA554887
1513	JACEHV000000000	USA	AR	*G. gallus*	Bone	1513	PRJNA554887
1514	JACEHU000000000	USA	AR	*G. gallus*	Bone	1514	PRJNA554887
1515	JACEHT000000000	USA	AR	*G. gallus*	Bone	1515	PRJNA554887
1516	JACEHX000000000	USA	AR	*G. gallus*	Bone	1516	PRJNA554887
1517	JACEHS000000000	USA	AR	*G. gallus*	Bone	1517	PRJNA554887
1518	JACEHR000000000	USA	AR	*G. gallus*	Bone	1518	PRJNA554887
1519	JACEHQ000000000	USA	AR	*G. gallus*	Bone	1519	PRJNA554887
1520	JACEHP000000000	USA	AR	*G. gallus*	Bone	1520	PRJNA554887
1521	JACEHO000000000	USA	AR	*G. gallus*	Bone	1521	PRJNA554887
1522	JACEHN000000000	USA	AR	*G. gallus*	Bone	1522	PRJNA554887
1523	JACEHM000000000	USA	AR	*G. gallus*	Bone	1523	PRJNA554887
1524	JACEHL000000000	USA	AR	*G. gallus*	Bone	1524	PRJNA554887
B2-15A	GCA_007726495.1	USA	OK	*G. gallus*	Retail Meat	B2-15A	PRJNA555718
B3-17D	GCA_007726565.1	USA	OK	*G. gallus*	Retail Meat	B3-17D	PRJNA555718
B4-59C	GCA_007726525.1	USA	OK	*G. gallus*	Retail Meat	B4-59C	PRJNA555718
B8-13D	GCA_007726545.1	USA	OK	*G. gallus*	Retail Meat	B8-13D	PRJNA555718
ch21	GCA_003343155.1	Poland		*G. gallus*	Deep Wound	ch21	PRJNA344860
ch22	GCA_003350605.1	Poland		*G. gallus*	Deep Wound	ch22	PRJNA344860
ch23	GCA_003336545.1	Poland		*G. gallus*	Deep Wound	ch23	PRJNA344860
ch3	GCA_003336625.1	Belgium		*G. gallus*	Commensal	ch3	PRJNA344860
ch5	GCA_003336635.1	Belgium		*G. gallus*	Commensal	ch5	PRJNA344860
ch9	GCA_003336495.1	USA		*G. gallus*	Hock	ch9	PRJNA344860
ED98	GCA_000024585.1	Ireland		*G. gallus*	Bone	ED98	PRJNA39547
X22	GCA_007998145.1	China		*Unknown*	Unknown	X22	PRJNA558859
ph2	GCA_003336565.1	Scotland		Pheasant	Tenosynovitis	ph2	PRJNA344860
pa3	GCA_003336575.1	Scotland		Partridge	Liver	pa3	PRJNA344860

Designation	Assembly	Country	State	Host	Source	Strain	BioProject
3688STDY6124906*	GCA_900126025.1	Thailand		*Homo sapiens*	Unknown	3688STDY6124906	PRJEB9575
3688STDY6125000*	GCA_900124915.1	Thailand		*H. sapiens*	Unknown	3688STDY6125000	PRJEB9575
502A*	GCA_000597965.1	USA	NY	*H. sapiens*	Newborn Nursery	RN6607	PRJNA240093
BCH-SA-12*	GCA_003720885.1	USA		*H. sapiens*	Throat	BCH-SA-12	PRJNA480016
CFBR-102*	GCA_003720115.1	USA		*H. sapiens*	Sputum	CFBR-102	PRJNA480016
CFBR-122*	GCA_003720355.1	USA		*H. sapiens*	Sputum	CFBR-122	PRJNA480016
CFBR-149*	GCA_003720135.1	USA		*H. sapiens*	Sputum	CFBR-149	PRJNA480016
CFBR-171*	GCA_003719905.1	USA		*H. sapiens*	Sputum	CFBR-171	PRJNA480016
CFSA134*	GCA_002123885.1	USA		*H. sapiens*	Sputum	CFSA134	PRJNA380429
DAR1813*	GCA_000609765.1	USA		*H. sapiens*	Blood	DAR1813	PRJNA228330
DAR1890*	GCA_000609945.1	USA		*H. sapiens*	Fluid	DAR1890	PRJNA228339
DAR3156*	GCA_000610525.1	Argentina		*H. sapiens*	Bone	DAR3156	PRJNA228372
DAR3157*	GCA_000610545.1	Argentina		*H. sapiens*	Sepsis	DAR3157	PRJNA228373
DAR3166*	GCA_000610605.1	Argentina		*H. sapiens*	Bone	DAR3166	PRJNA228376
DAR3178*	GCA_000610685.1	Argentina		*H. sapiens*	Bone	DAR3178	PRJNA228381
F29982*	GCA_000571715.1	USA		*H. sapiens*	Nares	F29982	PRJNA225050
F41882*	GCA_000559345.1	USA		*H. sapiens*	Nares	F41882	PRJNA224323
FDAARGOS_16*	GCA_001019355.2	USA	NY	*H. sapiens*	Unknown	FDAARGOS_16	PRJNA231221
FDAARGOS_359*	GCA_002554295.1	USA	MD	*H. sapiens*	Wound	FDAARGOS_359	PRJNA231221
GHA2*	GCA_008630855.1	Ghana		*H. sapiens*	Hospital	GHA2	PRJNA564764
H27777*	GCA_000561365.1	USA		*H. sapiens*	Nares	H27777	PRJNA224442
M0455*	GCA_000361205.1	USA		*H. sapiens*	ICU	M0455	PRJNA173479
M6K089*	GCA_003421965.1	Japan		*H. sapiens*	Skin	M6K089	PRJDB5246
M6K136*	GCA_003422345.1	Japan		*H. sapiens*	Skin	M6K136	PRJDB5246
NCTC10804*	GCA_900458375.1	UK		unknown	Unknown	NCTC10804	PRJEB6403
NCTC9393*	GCA_900458005.1	UK		*H. sapiens*	Unknown	NCTC9393	PRJEB6403
NRS149*	GCA_001018775.2	USA	NY	*H. sapiens*	Unknown	FDAARGOS_14	PRJNA231221
SAM-7*	GCA_003038745.1	Lebanon		*H. sapiens*	Throat	SAM-7	PRJNA437720
W76127*	GCA_000562505.1	USA		*H. sapiens*	Nares	W76127	PRJNA224506
E. coli							
Africa_Bt	GCA_001419785.1	Tanzania		*B. taurus*		33	PRJNA293513
AR_Gg1409	JACGTG000000000	USA	AR	*G. gallus*	Bone	1409	PRJNA554886

Designation	Assembly	Country	State	Host	Source	Strain	BioProject
AR_Gg1413	JACGTF000000000	USA	AR	*G. gallus*	Bone	1413	PRJNA554886
AR_Gg1512	JACGTE000000000	USA	AR	*G. gallus*	Bone	1512	PRJNA554886
AR_Gg1527	JACGTD000000000	USA	AR	*G. gallus*	Bone	1527	PRJNA554886
Bolivia_HsFecal	GCA_003850735.1	Bolivia		*H. sapiens*	feces	286A	PRJNA427943
Broad_Hs	GCA_002244745.1	USA	WA	*H. sapiens*	Urine	136-1758	PRJNA269984
CH_Gg	GCA_003009015.1	China		*G. gallus*		12c7	PRJNA417344
CH_Gg1	GCA_002959165.1	China		*G. gallus*		YH17134	PRJNA434044
CH_Gg2	GCA_003008775.1	China		*G. gallus*		12c8	PRJNA417344
CH_Gg3	GCA_003009715.1	China		*G. gallus*		12c5	PRJNA417344
CH_Hs6	GCA_003302635.1	China		*H. sapiens*	feces	A61	PRJNA400107
CH_Ss	GCA_003328175.1	China		*Sus scrofa*		E565	PRJNA450836
CH_Ss2	GCA_000987875.1	China		*S. scrofa*	Feces	SEC470	PRJNA244370
ClFeces	GCA_003043915.1			*Canis lupus*	Feces	RM14723	PRJNA341281
DE_GgLiver	GCA_001652345.1	Denmark		*G. gallus*	liver	E44	PRJNA321591
DE_GgSkin	GCA_003015065.1	Denmark		*G. gallus*	skin	L7S7	PRJNA438734
DE_GgSkin2	GCA_003013555.1	Denmark		*G. gallus*	skin	L3S3	PRJNA438662
DE_GgSkin3	GCA_003015075.1	Denmark		*G. gallus*	skin	L5S5	PRJNA438735
Estonia_Hs	GCA_006238365.1	Estonia		*H. sapiens*	clinical sample	EEIVKB55	PRJNA528606
FR_HsFeces	GCA_900536595.1	France		*H. sapiens*	feces	CEREMI_E32	PRJEB28341
FR_HsFeces1	GCA_900499885.1	France		*H. sapiens*	feces	884A	PRJEB28020
FR_HsIAI39	GCA_000026345.1	France		*H. sapiens*		IAI39	PRJNA33411
Germany_HsColon	GCA_000183345.1	Germany		*H. sapiens*	Ileum	NRC857c	PRJNA41221
HsFeces	GCA_003018255.1			*H. sapiens*	Feces	2012C-4502	PRJNA218110
HsFeces2	GCA_003018095.1			*H. sapiens*		2014C-3338	PRJNA218110
Israel_Mg	GCA_000819645.1	Israel		*Meleagris gallopavo*	Blood	789	PRJNA262513
JA_Hs	GCA_006535915.1	Japan		*H. sapiens*		SMEc189	PRJDB8148
JA_HsColitis	GCA_000008875.1	Japan		*H. sapiens*	Hemorrhagic colitis	SAKAI (EHEC)	PRJNA226
K12MG1655	GCA_000005845.2					MG1655	PRJNA603343
Latvia_HsClinical	GCA_006236595.1	Latvia		*H. sapiens*	Clinical sample	LVSTRB103	PRJNA528606
MX_Bat	GCA_002456375.1	Mexico		*Tadarida brasiliensis*	Feces	MOD1-EC908	PRJNA230969

Designation	Assembly	Country	State	Host	Source	Strain	BioProject
MX_Hs	GCA_002485345.1	Mexico		*H. sapiens*	Feces	MOD1-EC6621	PRJNA230969
Netherlands_Hs	GCA_002888415.1	Netherlands		*H. sapiens*	Meninges	SP-16	PRJNA429029
PA_DeerFeces	GCA_002215155.1	USA	PA	*Odocoileus virginianus*	Feces	PSUO103	PRJNA314794
PA_GgPSU02	GCA_002215095.1	USA	PA	*G. gallus*	Peritoneum	PSUO2	PRJNA287563
PA_GgPSUO2	GCA_002215095.1	USA	PA	*G. gallus*	Peritoneum	PSUO2	PRJNA287563
PA_GgPSUO78	GCA_002215115.1	USA		*G. gallus*	Peritoneum	PSUO78	PRJNA287566
Pakistan_GgEC13	GCA_004284075.1	Pakistan		*G. gallus*	infection	EC_13	PRJNA522294
PO_Gg	GCA_001709145.1	Poland		*G. gallus*		019PP2015	PRJNA319144
PO_GgSick	GCA_001696335.1	Poland		*G. gallus*	sick domestic birds	012PP2015	PRJNA319144
PolandGg2	GCA_001758245.1	Poland		*G. gallus*		022PP2016	PRJNA319144
Spain_HsBlood	GCA_000500875.1	Spain		*H. sapiens*	Blood	GSK2024	PRJNA221787
Swiss_GgMeat	GCA_001660565.1	Switzerland		*G. gallus*	Meat	S51	PRJNA323827
UK_GgFeces1	GCA_900490165.1	United Kingdom		*G. gallus*	feces	VREC0540	PRJEB8774
UK_GgFeces2	GCA_900482085.1	United Kingdom		*G. gallus*	feces	VREC0637	PRJEB8774
US_Bt8	GCA_004792865.1	USA		*B. taurus*	feces	KCJK8229	PRJNA420036
Germany_BtMastitis	GCA_000829985.1	Germany		*B. taurus*	Mastitis	1303	PRJNA46529
US_ClIntestine	GCA_002232435.1	USA	NY	*C. lupus*	Intestine	MOD1-EC5097	PRJNA230969
US_Gg	GCA_002512585.1	USA	AL	*G. gallus*	Egg	MOD1-EC6339	PRJNA230969
US_GgBrain	GCA_002537555.1	USA		*G. gallus*	brain	MOD1-EC6094	PRJNA230969
US_GgBreast	GCA_003793955.1	USA		*G. gallus*	Chicken Breasts	CVM N17EC0744	PRJNA292663
US_GgPericardium	GCA_002231405.1	USA	PA	*G. gallus*	Pericardial Sac	MOD1-EC5115	PRJNA230969
US_GgThigh	GCA_003794735.1	USA		*G. gallus*	Chicken Thighs	CVM N17EC0412	PRJNA292663
US_HsMGH188	GCA_002152225.1	USA	MA	*H. sapiens*		MGH188	PRJNA271899
US_HsO104_H4	GCA_000299455.1	USA		*H. sapiens*	Feces	2011C-3493	PRJNA81095
US_HsUMN026	GCA_000026325.2	USA		*H. sapiens*		UMN026	PRJNA33415
US_HsUrine	GCA_002245075.1	USA	WA	*H. sapiens*	Urine	225-2935	PRJNA269984
US_MgAPEC01	GCA_000014845.1	USA	IA	*Meleagris gallopavo*	Lung	APEC 01	PRJNA16718

Designation	Assembly	Country	State	Host	Source	Strain	BioProject
US_MgGround	GCA_003774815.1	USA		*M. gallopavo*	ground meat	CVM N17EC1100	PRJNA292663
US_Ss	GCA_002474525.1	USA	SD	*S. scrofa*	Ileum	MOD1-EC5757	PRJNA230969
US_SsIntestine	GCA_002464015.1	USA		*S. scrofa*	Jejunum	MOD1-EC6458	PRJNA230969
USA_AZH2O	GCA_002534895.1	USA	AZ		water	MOD1-EC5915	PRJNA230969
WA_HsUTI	GCA_000778565.1	USA	WA	*H. sapiens*	Urine	UPEC-208	PRJNA248737

548 *- 29 *S. aureus* isolates used in tBLASTn analyses in Table 3.

549

550 Table S2. Evolution of proteomes in four genomes of *S. aureus* infecting chickens. The RAST SEED viewer was used to identify genes present in
551 1519 where the predicted polypeptide had a % identity less than 50% in one or more of the genomes for the indicated isolates. Gene is the RAST
552 1519 annotation gene number, Length is for the predicted polypeptide in 1519, and function is the annotation from RAST and/or NCBI.

% Identity			1519		
ED98	Ch21	B4-59C	Gene	Length	function
100.0	100.0	0.0	6	148	hypothetical protein
100.0	100.0	0.0	7	83	hypothetical protein
100.0	100.0	0.0	16	112	hypothetical protein
0.0	0.0	0.0	32	182	DUF1541 domain-containing protein
36.4	35.9	36.4	33	688	Lead, cadmium, zinc and mercury transporting ATPase; Copper-translocating P-type ATPase
0.0	0.0	0.0	117	43	hypothetical protein
0.0	0.0	0.0	153	38	hypothetical protein
0.0	0.0	100.0	296	97	hypothetical protein
0.0	34.4	100.0	327	235	Toxic shock syndrome toxin 1 (TSST-1)
0.0	0.0	100.0	328	55	hypothetical protein
0.0	55.0	100.0	329	41	Exotoxin, phage associated, enterotoxin SEB (partial)
0.0	0.0	100.0	331	140	Phage protein
0.0	0.0	100.0	332	390	Phage protein
0.0	0.0	0.0	421	40	Phosphoglycerate kinase (partial)
45.9	45.9	100.0	500	48	hypothetical protein
0.0	100.0	100.0	562	49	Secretory antigen SsaA-like protein transposon-related
0.0	0.0	0.0	595	44	hypothetical protein
36.0	36.0	100.0	729	222	hypothetical protein
26.1	26.1	100.0	732	242	hypothetical protein
0.0	0.0	100.0	733	71	Phage protein
100.0	100.0	0.0	1021	128	hypothetical protein
0.0	100.0	100.0	1197	45	hypothetical protein
100.0	100.0	0.0	1210	39	hypothetical protein
0.0	0.0	0.0	1289	64	hypothetical protein
0.0	100.0	100.0	1421	44	hypothetical protein
100.0	98.6	0.0	1553	74	Hypothetical protein, phi-ETA orf24 homolog

% Identity			1519		
ED98	Ch21	B4-59C	Gene	Length	function
0.0	0.0	0.0	1554	72	Phage protein
37.4	36.1	37.4	1555	414	Phage DNA helicase
0.0	0.0	0.0	1556	119	Phage protein
0.0	0.0	0.0	1557	255	Phage replication initiation protein
95.9	97.7	0.0	1558	225	Hypothetical protein, PV83 orf19 homolog
100.0	99.5	0.0	1560	213	Phage-associated recombinase
100.0	100.0	0.0	1561	160	ORF027
98.8	97.7	39.5	1563	87	Hypothetical protein, PVL orf39 homolog
97.3	98.6	0.0	1565	74	Hypothetical protein, PV83 orf12 homolog
98.2	98.2	0.0	1566	57	Uncharacterized protein pCM2_0059
100.0	100.0	0.0	1567	97	phage protein
100.0	81.3	0.0	1568	251	Phage antirepressor protein
100.0	100.0	0.0	1569	62	hypothetical protein
100.0	100.0	0.0	1570	201	Phage protein
0.0	0.0	0.0	1572	65	Phage protein
0.0	0.0	0.0	1573	257	Phage antirepressor protein
0.0	0.0	0.0	1574	79	hypothetical protein
100.0	93.4	0.0	1576	110	Phage protein
99.7	100.0	0.0	1577	553	DNA adenine methylase
100.0	100.0	25.7	1578	350	Phage integrase
0.0	0.0	0.0	1719	41	hypothetical protein
0.0	0.0	100.0	1760	52	hypothetical protein transposon-related
0.0	0.0	0.0	1862	158	hypothetical protein
0.0	0.0	0.0	1863	76	hypothetical protein
0.0	0.0	0.0	1864	47	hypothetical protein
0.0	0.0	0.0	1865	65	hypothetical protein
0.0	0.0	0.0	1866	56	hypothetical protein
0.0	0.0	0.0	1867	88	hypothetical protein
0.0	0.0	0.0	1868	77	hypothetical protein
0.0	0.0	0.0	1869	157	hypothetical protein
0.0	0.0	0.0	1870	99	hypothetical protein

% Identity			1519		
ED98	Ch21	B4-59C	Gene	Length	function
0.0	0.0	0.0	1871	186	hypothetical protein
0.0	0.0	0.0	1872	54	hypothetical protein
0.0	0.0	0.0	1916	52	Adhesin of unknown specificity SdrC
0.0	0.0	0.0	1917	205	Type I restriction-modification system, DNA-methyltransferase subunit M / Type I restriction-modification system, specificity subunit S
24.3	24.3	24.3	1919	178	Aminoglycoside N6'-acetyltransferase
0.0	0.0	0.0	1921	144	hypothetical protein
31.8	31.8	31.8	1922	110	Transcriptional regulator, ArsR family
45.0	0.0	0.0	1923	173	Plasmid recombination, MobE mobilization protein
0.0	0.0	100.0	2179	95	Hypothetical protein in SaPI
30.1	30.1	100.0	2180	121	Hypothetical protein in SaPI
0.0	0.0	100.0	2181	486	Virulence-associated protein E in SaPI
0.0	0.0	100.0	2189	45	hypothetical protein
29.4	29.4	100.0	2191	193	hypothetical protein
0.0	0.0	100.0	2192	59	hypothetical protein
0.0	0.0	0.0	2358	280	hypothetical protein
0.0	0.0	0.0	2369	205	hypothetical protein
0.0	0.0	0.0	2384	146	Phage holin
100.0	100.0	0.0	2688	39	hypothetical protein

553