
Charles Darwin University

Contribution of Serotype and Genetic Background to Virulence of Serotype 3 and Serogroup 11 Pneumococcal Isolates

McAllister, Lauren; Ogunniyi, Abiodun; Stroehrer, Uwe; Leach, Amanda; Paton, James

Published in:
Infection and Immunity

DOI:
[10.1128/IAI.05663-11](https://doi.org/10.1128/IAI.05663-11)

Published: 01/01/2011

Document Version
Peer reviewed version

[Link to publication](#)

Citation for published version (APA):

McAllister, L., Ogunniyi, A., Stroehrer, U., Leach, A., & Paton, J. (2011). Contribution of Serotype and Genetic Background to Virulence of Serotype 3 and Serogroup 11 Pneumococcal Isolates. *Infection and Immunity*, 79(12), 4839-4849. <https://doi.org/10.1128/IAI.05663-11>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

1 **Contribution of Serotype and Genetic Background to Virulence of**
2 **Serotype 3 and Serogroup 11 Pneumococcal Isolates**

3
4 **Lauren J. McAllister,¹ Abiodun D. Ogunniyi,¹ Uwe H. Strocher,^{1†} Amanda J.**
5 **Leach,² and James C. Paton^{1*}**

6
7 *Research Centre for Infectious Diseases, School of Molecular and Biomedical Science,*
8 *University of Adelaide, Adelaide 5005, South Australia, Australia¹, Menzies School of Health*
9 *Research, Charles Darwin University, Darwin, 0811, Northern Territory, Australia².*

10

11

12 Running head: CHARACTERIZATION OF PNEUMOCOCCAL OM ISOLATES

13

14

15 † Current address, School of Biological Sciences, Flinders University, Bedford Park 5042, South
16 Australia, Australia.

17

18 *Corresponding author. Mailing address: School of Molecular and Biomedical Science,
19 University of Adelaide, Adelaide, S.A., 5005, Australia. Phone: + 61-8-83035929. Fax: + 61-8-
20 83033262. E-mail: james.paton@adelaide.edu.au.

21

22

23

ABSTRACT

24 Capsular serotype has long been associated with the virulence of *Streptococcus*
25 *pneumoniae*. Here, we present an in-depth study of phenotypic and genetic differences
26 between serotype 3 and serogroup 11 *S. pneumoniae* clinical isolates from both the general
27 and indigenous populations of Australia. Both serotypes/groups included clonally
28 unrelated strains with differences in well-known polymorphic virulence genes, such as
29 *nanA* and *pspA*, as demonstrated by multilocus sequence typing and Western blot analysis,
30 respectively. Nonetheless, the serotype 3 strains were consistently and significantly more
31 virulent in mice than the serogroup 11 strains. Despite extensive genomic analysis, non-
32 capsular genes common to one serotype/group but not the other were not identified.
33 Nevertheless, following the conversion of a serotype 11A isolate to serotype 3 and
34 subsequent analysis in an intranasal infection model, it was evident that both capsular and
35 non-capsular factors determine virulence phenotype in mice. However, it appears that
36 these non-capsular factors vary from strain to strain.

37

INTRODUCTION

38

39 *Streptococcus pneumoniae* (the pneumococcus) is a formidable pathogen, responsible for
40 a broad spectrum of diseases including pneumonia, meningitis, bacteremia and otitis media
41 (OM), and accounts for more deaths worldwide than any other single pathogen (48). In Australia,
42 the overall rate of invasive pneumococcal disease (IPD) in the indigenous population is over 4
43 times the rate seen in non-indigenous Australians (54). Furthermore, chronic suppurative OM,
44 which is normally rare in developed countries, is present in remote Aboriginal communities at 10
45 times the rate defined as a major public health concern by the World Health Organization (9).
46 Children living in these communities not only suffer the highest rates of tympanic membrane
47 perforation in the world, but 80% are estimated to suffer from some degree of OM-related
48 hearing impairment (45).

49 Significant research has already been undertaken on the pneumococcus, but there are still
50 many unanswered questions due to the complex nature of its pathogenicity. The polysaccharide
51 capsule that surrounds the pneumococcus and determines serotype, of which there are more than
52 90, is the most well-established virulence factor. Certain serotypes have been shown to have a
53 higher association with disease or carriage than others (7, 19, 20, 60). In addition, a number of
54 conserved virulence proteins have been widely characterized, including neuraminidase A
55 (NanA), pneumolysin (Ply), autolysin (LytA) and pneumococcal surface proteins A and C (PspA
56 and PspC), and shown to be critical in pathogenesis (29, 50). However, these are only a subset of
57 potential virulence factors. Like other pathogens, such as *Haemophilus influenzae*, another
58 common cause of invasive disease and OM, *S. pneumoniae* has a pangenome (23, 24). Core
59 genes, which are conserved across pneumococcal strains, only represent 70-80% of the genome
60 and a gene pool of over 5000 orthologous clusters is estimated to be available to this naturally
61 transformable organism from other pneumococcal strains and related bacterial species (23). An
62 understanding of the role in virulence of these noncore genes, which are generally found in
63 regions of the genome known as regions of diversity or accessory regions (ARs), is only just
64 beginning to emerge. The majority of studies on pneumococcal ARs to date have focused on IPD

65 (5, 13, 59). In addition, although there have been numerous epidemiological studies highlighting
66 the problem of OM in remote Australian Aboriginal communities (33, 38, 45), very limited
67 molecular analysis of isolates from these communities has been performed. Accordingly, this
68 study aimed to characterize a range of isolates from geographically and temporally diverse
69 Australian Aboriginal communities and the general population using a variety of molecular
70 techniques with a view to identifying common genes or ARs potentially associated with but not
71 necessarily limited to OM.

72

MATERIALS AND METHODS

73

74 **Bacterial strains and growth conditions.** The bacterial strains used in this study are
75 listed in Table 1. Pneumococci were routinely grown in THY (Todd-Hewitt broth supplemented
76 with 1% Bacto™ yeast extract), serum broth (SB; 10% [v/v] donor horse serum in nutrient
77 broth), brain heart infusion (BHI) broth, or on blood agar (BA) plates at 37°C in 95% air and 5%
78 CO₂. Gentamicin and optochin were added at concentrations of 5 µg/ml where appropriate. For
79 storage, *S. pneumoniae* were grown in SB supplemented with glycerol to 30% and stored at -
80 80°C. Pneumococcal opacity phase morphology was determined on THY-catalase plates, as
81 described previously (65). Strains were confirmed as *S. pneumoniae* by optochin sensitivity,
82 while serotype-specific capsule production was confirmed by Quellung reaction using diagnostic
83 antisera obtained from Statens Serum Institut, Copenhagen, Denmark. The level of capsule
84 expression was determined by the uronic acid assay (44). *Escherichia coli* One Shot® TOP10
85 chemically competent cells (Invitrogen, Victoria, Australia) were grown in Luria-Bertani (LB)
86 broth or on LB agar plates, in the presence of ampicillin (Amp) at 50 µg/ml where appropriate.

87 **Bacterial transformation.** Pneumococci were transformed using complete
88 transformation medium (CTM) (17, 40). For cloning and transformation of *E. coli*, the TOPO TA
89 cloning® kit (Invitrogen) with pCR®2.1-TOPO® and One Shot® TOP10 chemically competent
90 cells was used, according to the manufacturer's instructions. LB-Amp plates containing 1.6 mg
91 of X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) were used for selection.

92 **DNA manipulations.** All oligonucleotide primer sequences are available on request.
93 Chromosomal DNA for genome sequencing and PCR was extracted and purified as described
94 previously (55). PCR reactions were carried out in an Eppendorf Mastercycler (Eppendorf,
95 Hamburg, Germany) using a final volume of either 25 or 50 µl. Standard reactions were
96 performed using *Taq* DNA polymerase (Roche Diagnostics, Basel Switzerland), according to the
97 manufacturer's instruction. The Expand™ Long Template or High Fidelity PCR Systems (Roche
98 Diagnostics) were used when high fidelity PCR was required. DNA sequencing reactions were

99 carried out using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems,
100 California, USA).

101 **Multi-locus sequence typing (MLST).** *S. pneumoniae* isolates were subjected to MLST
102 using the method and primer sequences described previously (15) and classified according to the
103 pneumococcal MLST database (<http://spneumoniae.mlst.net/>).

104 **DNA Microarray analysis.** DNA microarray experiments were performed on whole
105 genome *S. pneumoniae* PCR arrays based on TIGR4 and R6 annotations (26, 61). Array slides
106 were obtained from the Bacterial Microarray Group at St George's, University of London. The
107 array design is available in BμG@Sbase (Accession No. A-BUGS-14; [http://bugs.sgu.ac.uk/A-](http://bugs.sgu.ac.uk/A-BUGS-14)
108 [BUGS-14](http://bugs.sgu.ac.uk/A-BUGS-14)) and also ArrayExpress (Accession No. A-BUGS-14). Briefly, *S. pneumoniae* DNA
109 (10.5 μg in 100 μl) was digested with *Sau3A1* or *RsaI* (New England Biolabs [NEB],
110 Massachusetts, USA), and cleaned using a Qiagen MinElute® PCR Purification Kit. Thereafter,
111 20 μl of purified digest was labeled for each dye (Alexa Fluor 555 and Alexa Fluor 647) per
112 slide, using the Genisphere Array 900 DNA™ DNA Labeling Kit for Microarrays (Genisphere,
113 Pennsylvania, USA) according to the manufacturer's instructions. Slides were incubated
114 overnight in a dark, humidified chamber at 65 °C, washed in a 3-step process (15 min with 2×
115 SSC, 0.03% (v/v) SDS, at 65 °C; 15 min with 1× SSC at room temperature [RT]; 15 min with
116 0.2× SSC at RT) and dried. Slides were scanned using a GenePix® 4000B scanner (Molecular
117 Devices) and images were acquired using GenePix® Pro 6.0 software (Axon). Fully annotated
118 microarray data have been deposited in BμG@Sbase and also ArrayExpress databases (accession
119 number E-BUGS-126).

120 **PCR-based subtractive hybridization.** PCR-based subtractive hybridization (3) was
121 performed using the Clontech PCR-Select™ Bacterial Genome Subtraction Kit (Clontech,
122 California, USA), with the exception that the primary hybridization incubation time was
123 increased to 8 h. Ligations were performed following the manufacturer's instructions using T4
124 DNA Ligase and buffer (NEB). PCR reactions were performed with a hot start, using *Taq*

125 polymerase and 1.6× reaction buffer as outlined in the kit manual. Secondary PCR products were
126 used for cloning.

127 **Next generation sequencing.** Sequencing and genome assembly were performed by
128 Geneworks (Thebarton, Adelaide, Australia) using chromosomal DNA prepared as described
129 above, using an Illumina Genome Analyzer *II* (California, USA) and Lasergene® 8 software
130 (DNASTAR Inc, Wisconsin, USA). Sequencing data have been deposited in the NCBI Sequence
131 Read Archive (Accession number SRX026162).

132 **Bioinformatics.** The pneumococcal genomes and BLAST programs available through
133 the Kyoto Encyclopedia of Genes and Genomes (www.genome.jp/kegg/) and National Center for
134 Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>) were used in the analysis of
135 sequencing data. The ClustalW program available through the website of the European
136 Bioinformatics Institute (EBI, <http://ebi.ac.uk/Tools>) was used to align multiple sequences.

137 **Protein analysis.** Bacteria were lysed and subjected to SDS-PAGE, as described
138 previously (41). Separated proteins were electro-blotted onto nitrocellulose (Protran, Melbourne
139 Australia), as described previously (41). After transfer, the membrane was probed with specific
140 polyclonal mouse or rabbit antisera at a dilution of 1/3,000 and then reacted with blotting grade
141 goat anti-mouse or goat anti-rabbit IgG-alkaline phosphatase (AP) conjugates (BioRad
142 Laboratories, USA). The hemolytic activity of pneumococcal lysates was determined as
143 described previously (51), using PBS-washed human erythrocytes.

144 **Virulence and pathogenesis studies.** All animal experiments used outbred 5-6 week old
145 CD-1 (Swiss) mice and were approved by the Animal Ethics Committee of the University of
146 Adelaide. For intraperitoneal (i.p.) and pneumonia/sepsis models, opaque variants were cultured
147 in SB to approximately 1×10^8 CFU/ml and diluted in sterile SB to the appropriate challenge
148 dose. The actual challenge doses were determined retrospectively by plating on BA with or
149 without antibiotic selection, as appropriate. Mice in the pneumonia/sepsis study were
150 anesthetized with Nembutal (pentobarbitone sodium, Rhone-Merieux) and challenged
151 intranasally (i.n.), as previously described (36, 39). For survival studies, mice were monitored

152 closely over a 3-4 week period following challenge for signs of disease and the survival time of
153 each mouse was recorded. For the capsule switch study, mice were sacrificed by CO₂
154 asphyxiation and nasal wash, ear wash, nasal tissue, lung, blood and brain samples were taken.
155 The nasal wash and blood perfusion techniques have been described previously (36). The ear
156 wash was performed by puncturing both tympanic membranes with a pipette tip and washing the
157 ear cavities with a combined 100µl of sterile trypsin buffer (0.5% trypsin in phosphate buffered
158 saline [PBS]). The nasal tissue, lungs and brain were removed and homogenized in 1 ml PBS
159 using a Precellys 24 homogenizer (Bertin Technologies, Pessac, France) according to
160 manufacturer's instructions. Enumeration of bacteria in the various samples was performed as
161 previously described (36, 39). For i.n. colonization studies, transparent phase bacteria were
162 grown in THY to mid-logarithmic phase. I.n. challenge without anesthesia, sample processing
163 and enumeration of bacteria was performed as previously described (41).
164

165

166

RESULTS

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

Strain selection. For the purposes of this study, a number of clinical *S. pneumoniae* isolates belonging to serotypes 3 and 11A were obtained from the Menzies School of Health Research (Menzies), Darwin, Australia, and the Women's and Children's Hospital (WCH), North Adelaide, Australia (Table 1). The Menzies isolates were derived from a variety of indigenous communities in temporally and geographically diverse locations. Serotype 3 is frequently isolated from the ear discharge of indigenous children with acute OM with perforation of the tympanic membrane, yet is less frequently isolated from nasal swabs in asymptomatic members of the same population (34). In addition, this serotype has a high association with OM worldwide (20, 21, 58) and isolation of serotype 3 strains from OM cases has reportedly risen following the introduction of 7-valent pneumococcal conjugate vaccine (PCV7) (42). Serotype 3 is a less frequent cause of IPD, but nevertheless, cases of type 3 IPD have a high fatality rate (19, 22, 60). In Europe, Japan and North America, ST180 is the major clone within serotype 3 (7, 28, 42, 60), whereas in South Africa and Ghana, ST458 is more dominant (35, 46). The second serotype examined, 11A, is a common nasopharyngeal isolate in Aboriginal children but is rarely isolated from the middle ear (34). All serotype 11A isolates used in this study are of nasopharyngeal origin (Table 1). Serotype 11A has been associated with OM in other communities, including in Australia (18, 64), but is generally not as common as serotype 3 (20, 21, 53). Even though it also has a low association with IPD (19), serotype 11A clinical isolates of ST62 have been shown to have a high case-fatality rate in patients with underlying medical conditions (60).

In addition to the Menzies isolates, the serotype 3 and serogroup 11 isolates obtained from WCH are from the general population and have a well-documented clinical history (Table 1). We could not determine whether the WCH serogroup 11 isolates were serotype 11A because the available antiserum used in this study is only serogroup 11-specific. The serotype 2 strain

191 D39 (4), the serotype 4 strain TIGR4 (1), and the serotype 3 strain WU2 (6), which have been
192 used in numerous published studies, were also included in some experiments (Table 1).

193 **Preliminary characterization of serotype 3 and serogroup 11 clinical isolates.** The
194 serotype 3 and serogroup 11 clinical isolates from Menzies and WCH had hitherto been
195 uncharacterized and so it was important to determine their pathogenicity profiles in order to
196 establish any correlation between human disease and our mouse models. To this end, we first
197 compared the expression profiles of well-characterized virulence proteins PspA, PspC, Ply, LytA
198 and NanA with that in the highly virulent serotype 2 *S. pneumoniae* strain D39 by Western
199 blotting, using polyclonal antisera (see Fig. S1 in the supplemental material). Western blot
200 analysis showed no apparent difference in size or intensity using anti-LytA and anti-Ply
201 polyclonal sera for all serotype 3 and serogroup 11 strains. Furthermore, a Ply-specific hemolysis
202 assay using human erythrocytes showed there were no differences in specific hemolytic activity
203 between strains of the same serotype (data not shown). The differences observed in the Western
204 blots for PspA, PspC and NanA were further investigated in selected strains by PspA family
205 typing (63), PCR or sequencing of *pspA*, *pspC* and *nanA* genes, with results summarized in Fig.
206 1. The alleles identified for these three genes have been previously described by Hollingshead et
207 al. (24), Iannelli et al. (26) and King et al. (25, 27, 32), respectively. Interestingly, the serogroup
208 11 strains revealed more variable results than the serotype 3 strains, although some strains, such
209 as Menzies1 and Menzies3 exhibited identical protein profiles. The serotype 3 strains could be
210 divided into two distinct groups (one comprising WCH206 and WCH207, and the other the
211 remainder) based on their Western blot profiles.

212 Selected isolates were subsequently characterized by multi-locus sequence typing
213 (MLST) (14) based on their Western blot profiles. Serotype 3 strains that exhibited identical
214 protein profiles in the Western blot analysis were found to have identical STs; WCH206 and
215 WCH207 were ST180, whereas WCH208, Menzies11, Menzies12, Menzies14 and Menzies17
216 were ST458. WU2 was typed as ST378 (as found by Poolman et al. (52)). As predicted from the
217 divergent protein profiles, the serogroup 11 strains WCH211, WCH213, Menzies1 and Menzies5

218 all belong to unrelated STs (ST3020, ST62, ST3021 and ST662, respectively). WCH211 and
219 Menzies1 are the first reported isolates of their respective STs.

220 The colonies of the ST180 strains WCH206 and WCH207 were larger and more
221 homogenous in size compared to the other serotype 3 strains following overnight growth on BA.
222 No differences in colony morphology were evident between serogroup 11 strains. Capsular
223 polysaccharide production was compared between WCH206, WCH207, Menzies11, Menzies17
224 and WU2 using the uronic acid assay (44). There was no significant difference in capsule
225 production between isolates of the same ST. However, WCH206 and WCH207 produced
226 statistically significantly less capsular polysaccharide than Menzies11 and Menzies17 (Fig. 2).
227 WU2 produced significantly less capsule than Menzies17 but not Menzies11.

228 We subsequently evaluated selected serotype 3 and serogroup 11 strains for their
229 virulence characteristics in mouse pneumonia/sepsis and intraperitoneal (i.p.) (sepsis) models
230 (Fig. 3A and 3B). Statistical analysis of survival data is shown in Table 2. Serotype 3 strains
231 were highly virulent in the pneumonia/sepsis model with the exception of WU2, which killed
232 less than 50% of the mice and was significantly less virulent than all of the other serotype 3
233 strains. Menzies17 was also more virulent than all the other serotype 3 strains, including Menzies
234 11, which belongs to the same ST (Fig. 3A and Table 2). In contrast, serogroup 11 strains were
235 essentially avirulent in the pneumonia/sepsis model and survival times were significantly
236 different from all the serotype 3 strains except WU2. In the i.p. model, they killed less than 50%
237 of mice when injected at very high doses (1×10^7 CFU/mouse), and survival times for three of
238 the strains were significantly greater than those for mice given only 5×10^2 CFU of the serotype
239 3 strains (Fig. 3B and Table 2). Strains were also compared in an intranasal (i.n.) colonization
240 model (Fig. 3C). Interestingly, serogroup 11 strains were isolated more frequently and (for most
241 strains) in significantly greater numbers from the nasopharynx of mice on days 1 and 2 post i.n.
242 challenge than the serotype 3 strains (Fig. 3C and Table 3). Longer term colonization was also
243 explored by performing nasopharyngeal washes on surviving mice in the pneumonia-sepsis study

244 (presented in Fig. 3A) 4-weeks post-inoculation. Both serotypes/groups were found to be
245 capable of longer term colonization of the nasopharynx (Fig. 3D).

246 **Microarray analysis and PCR-based subtractive hybridization of ST180 and ST458**

247 **strains.** The existence of two clonally unrelated groups within the serotype 3 clinical isolates
248 presented an attractive opportunity for assessment of commonality and/or differences in the
249 overall genetic composition between these strains. First, we compared the genomes of WCH208
250 (ST458), Menzies12 (ST458), Menzies13 and Menzies15, which displayed identical protein
251 profiles to one another, by DNA microarray analysis using slides representing the genomes of
252 TIGR4 and R6 (a rough derivative of D39). No differences were detectable, consistent with the
253 protein profiles and ST data. Similarly, microarray and genome sequence analysis of strains from
254 the ST180 lineage have shown them to exhibit few genetic differences, even when isolated from
255 different continents (11, 23). Next, we compared the genomes of selected ST180 (WCH206) and
256 ST458 (Menzies17) strains. The results showed that 3 ORFs were present in WCH206, but
257 absent in Menzies17, while 15 ORFs were detectable in Menzies17 but absent in strain WCH206
258 (see Table S1 in the supplemental material). Similarly, genomic comparisons between serogroup
259 11 strains were made, and several genes were detected to be either present or absent in these
260 comparisons (Table S1). We subsequently carried out genomic comparisons of the
261 serotype/groups 3 and 11 strains to detect any genes that were unique to one serotype. However,
262 no non-capsular gene which was present in all isolates of one serotype but absent in every isolate
263 of the other was identified.

264 One of the limitations of DNA microarray analysis is that it gives no information on
265 genes which may be present in the strains tested, but absent on the array. In order to overcome
266 this, a PCR-based subtractive hybridization method (3) was adapted to detect genes present in
267 both ST180 and ST458, but absent in R6 and TIGR4. Using this procedure, we identified
268 serotype 3 capsule genes (effectively a positive control), and the IIA and IIB components of a
269 putative cellobiose phosphotransferase system (PTS1), which was first sequenced by Shen et al.
270 in a study identifying pneumococcal ORFs absent in the TIGR4 and R6 genomes (57).

271 Interestingly, the pneumococcus has a second cellobiose PTS (PTS2), which is also part of a
272 larger mobile genetic element, and its components were identified during a signature-tagged
273 mutagenesis virulence gene screen using a chinchilla middle ear infection model, suggesting a
274 role in OM (8).

275 Forbes et al. reported that PTS1 belongs to a mobile genetic element consisting of 10
276 genes (16). A putative choline sulfatase was also reported in this region in ATCC 700669 (Spain
277 23F ST81 lineage) by Croucher et al. (10). We compared the genomic organization with those
278 annotated for G54 (12) and Hungary19A-6 (J. Craig Venter Institute). The PTS was found to be
279 associated with a genomic island harboring a ROK (regulator, ORF of unknown function,
280 kinase) family protein and a putative sulfatase with an associated modifying factor, as well as
281 hypothetical proteins (Fig. 4). The ROK family protein does not have the DNA binding motif
282 characteristic of repressors in this family. We agree with the observation of Forbes et al. that the
283 organization of this island is well conserved (16), but G54 has a deletion of about 2.5 kb,
284 including the putative sulfatase gene, as well as parts of the sulfatase modifying factor and the
285 last gene of the putative cellobiose PTS. The same deletion is apparent in the serotype 3 strains
286 Menzies11 and Menzies17 (isolates of ST458), but WCH206 (an isolate of ST180) carries the
287 full genomic island. The sequence of G54 suggests it also has two insertions in two genes of the
288 PTS operon, resulting in frameshifts and premature stop codons, but these insertions are not
289 present in the island sequences of Menzies11 or Menzies17. The absence of the sulfatase in the
290 ST458 isolates explains why genes downstream of the cellobiose PTS were not detected by PCR-
291 based subtractive hybridization. The gene of the ROK family protein is located on a restriction
292 fragment over 3.6 kb, which is well outside the optimal size for PCR-based subtractive
293 hybridization (2).

294 **Distribution of the genomic island carrying a putative cellobiose PTS.** WU2 and the
295 serogroup 11 isolates were analyzed by PCR to determine whether they carried PTS1 or not (see
296 Table S2 in supplemental material). Interestingly, WU2 did not, but all four serogroup 11 strains
297 did. We also examined the distribution of PTS1, as well as the sulfatase and PTS2, by both

298 bioinformatic analysis utilizing publically available pneumococcal genomes and PCR analysis of
299 clinical isolates with a range of serotypes and virulence potential (Table S2). Only one strain
300 appeared to possess neither PTS1 nor PTS2, with the majority of strains possessing both. Only 3
301 out of the 4 serogroup 11 isolates possessed PTS2. Most isolates carrying PTS1 also carried the
302 sulfatase.

303 **Genome sequence comparison.** In order to determine whether there were additional
304 ARs missed by PCR-based subtractive hybridization, the genome of Menzies17 was sequenced
305 and aligned against the published genome of OXC141 (serotype 3, ST180,
306 <ftp://ftp.sanger.ac.uk/pub/pathogens/spn/>). The average coverage depth was 14.68 and sequence
307 gaps over 100 bp in the alignment represented approximately 6.3% of the OXC141 genome.
308 Although serotype 3 ST180 isolates are known to carry additional genes apart from the PTS1 AR
309 which are absent in TIGR4 and D39, none common to Menzies17 were detected (23, 57). This
310 indicates that the PCR-based subtractive hybridization methodology enabled efficient
311 identification and isolation of major genetic differences between isolates.

312 **Impact of capsular serotype and genetic background on pathogenesis.** Previous
313 studies have indicated the importance of additional genetic factors to serotype in pneumococcal
314 disease (31, 43, 56, 59). Indeed the difference in survival between WU2 and the other serotype 3
315 isolates in the pneumonia-sepsis model (Fig. 3A) is indicative of the importance of genotype.
316 However, it has also been demonstrated that serotype influences the behavior of genotype in
317 mouse in vivo models (30, 47, 62). As the isolates of the two groups in this study have distinct
318 behavior in our mouse models, and there are no additional factors to distinguish the two apart
319 from the capsule loci, the role of serotype for our particular isolates was further explored. For
320 this purpose, one of the serotype 11A strains from a remote Aboriginal community (Menzies5)
321 was capsule-switched with a serotype 3 capsule by transformation with a PCR product
322 comprising the capsule locus from WCH206. This strain was designated Menzies5³ and type 3
323 capsule production was confirmed by the Quellung reaction. In addition, loss of reactivity to
324 serogroup 11 antiserum was observed. There was no significant difference in serotype 3 capsular

325 polysaccharide production between Menzies5³ and the donor strain WCH206 as assessed by the
326 uronic acid assay. The mutant also exhibited an identical growth profile to Menzies5 and
327 WCH206 (data not shown). Colonization of various tissues by the three strains in the
328 pneumonia/sepsis model over a 3-day period was investigated (Fig. 5). Acquisition of the
329 serotype 3 capsule by Menzies5³ reduced nasopharyngeal colonization significantly on all three
330 days, but did not increase bacterial numbers recovered in any other niche (ears, lungs, blood or
331 brain) relative to Menzies5 at any time point. WCH206 was isolated at significantly higher
332 numbers than Menzies5³ from the ears (day 3), lungs (day 2), blood (day 2), and brain (day 2 but
333 not day 1). Therefore, although serotype clearly influences the ability of the ST662 isolate to
334 colonize the nasopharynx, the ST180 genotype is fitter than the ST662 genotype in various
335 niches of the pneumonia-sepsis model for serotype 3.
336

337

DISCUSSION

338 Serotype 3 *S. pneumoniae* is an important OM pathogen worldwide, including remote
339 Australian Aboriginal communities (20, 34, 52). Despite the burden of pneumococcal OM in
340 these communities, there has been very little molecular characterization of isolates. In this study
341 we have analyzed isolates from remote communities using a variety of techniques, including
342 Western blot analysis, a hemolysis assay, capsular polysaccharide assay, MLST and murine
343 virulence and colonization models. Although the isolates were acquired from diverse locations,
344 the serotype 3 isolates were found to be from the ST458 clonal cluster, which is unrelated to the
345 ST180 clone that normally dominates in developed countries (7, 28, 42, 60). Interestingly, this
346 ST has also been seen to dominate over ST180 in South Africa and Ghana, which are also
347 socioeconomically disadvantaged (35, 46). Both ST458 and ST180 isolates are highly virulent in
348 our mouse models, in contrast to the serogroup 11 isolates. However, despite observed
349 differences in the epidemiology of serogroup 11 isolates and behavior in our mouse model, the
350 capsule switch experiment performed in this study further underscores the importance of non-
351 capsular factors in disease. This does not mean that serotype is unimportant. In the capsule
352 switch study, Menzies5³ was clearly attenuated in the nasopharynx in comparison to Menzies5.
353 Kadioglu et al. also found their serotype 3 capsule switch variant of D39 and A66 (serotype 3
354 capsule donor) colonized the nasopharynx at lower numbers over a 24 h period in comparison to
355 wild-type D39 (30). Furthermore, although extensive analyses were conducted using DNA
356 microarray, PCR-based subtractive hybridization and Illumina sequencing, we could not identify
357 a specific factor common to the serotype 3 ST180 and ST458 isolates, but absent in the 4
358 unrelated serogroup 11 isolates. In addition, host factors are likely to influence genotype. The
359 prevalence of the ST458 clone over ST180 suggests it may be better adapted for the micro-
360 environmental niches encountered in remote Aboriginal communities.

361 In this study we have shown that PCR-based subtractive hybridization can be adapted to
362 complement DNA microarray hybridization by identifying genes absent on the microarray slide.
363 A previously discovered but uncharacterized putative cellobiose PTS (PTS1) common to ST180

364 and ST458 was identified through this technique, which constitutes part of a genomic island
365 harboring a putative ROK family protein and a putative sulfatase with an associated modifying
366 factor. The putative cellobiose utilization AR had initially been noted to be more common in
367 invasive isolates (16), but as with other pneumococcal ARs (5), when a larger number of isolates
368 were characterized, this relationship could not be confirmed. However, the wide distribution of
369 the AR and the putative role of its genes both suggest potential importance. Furthermore, the
370 apparent involvement of a second pneumococcal cellobiose PTS (PTS2) in OM (8) also suggests
371 that a role for PTS1 in OM is worthy of exploration. It is also interesting to note that the ST662
372 isolate Menzies5, which was less fit than WCH206 (ST180) in various niches of the pneumonia-
373 sepsis model in a serotype 3 background, and WU2 (ST378), which was less virulent than ST180
374 and ST458 isolates in the same model, only possess one of the PTSs, while ST180 and ST458
375 isolates possess both.

376 In this study, despite extensive molecular analyses, we could not identify any ARs that
377 are unique to OM isolates. It is likely that different ARs contribute to virulence in unrelated
378 strains, and that virulence of a given strain is a function of multiple ARs. The impact of any
379 combination of ARs could be influenced by other factors, including redundancy of ARs, and
380 capsular type, as well as host factors. Overall, there appears to be no simple answer to why one
381 strain is more prevalent in disease or carriage than another, and the ability of an isolate to cause
382 disease is likely to reflect a complex mix of serotype, genotype and host factors.

383

384

ACKNOWLEDGEMENTS

385 We thank Andrew Lawrence from the Women's and Children's Hospital, North
 386 Adelaide, South Australia, for the provision of strains. We acknowledge BμG@S (the Bacterial
 387 Microarray Group at St. George's Hospital, University of London) for supply of the microarray
 388 slides and advice. We also acknowledge the use of the pneumococcal MLST database, which is
 389 located at Imperial College London and is funded by the Wellcome Trust.

390 This research was supported by the Channel 7 Children's Research Foundation, as well as
 391 by Program Grant 565526 and Project Grant 627142 from the National Health and Medical
 392 Research Council (NHMRC) of Australia. JCP is a NHMRC Australia Fellow.

393

394

REFERENCES

- 395 1. **Aaberge, I. S., J. Eng, G. Lermark, and M. Lovik.** 1995. Virulence of *Streptococcus*
 396 *pneumoniae* in mice: a standardized method for preparation and frozen storage of the
 397 experimental bacterial inoculum. *Microb. Pathog.* **18**:141-152.
- 398 2. **Agron, P. G., M. Macht, L. Radnedge, E. W. Skowronski, W. Miller, and G. L.**
 399 **Andersen.** 2002. Use of subtractive hybridization for comprehensive surveys of
 400 prokaryotic genome differences. *FEMS Microbiol. Lett.* **211**:175-182.
- 401 3. **Akopyants, N. S., A. Fradkov, L. Diatchenko, J. E. Hill, P. D. Siebert, S. A.**
 402 **Lukyanov, E. D. Sverdlov, and D. E. Berg.** 1998. PCR-based subtractive hybridization
 403 and differences in gene content among strains of *Helicobacter pylori*. *Proc. Natl. Acad.*
 404 *Sci. U.S.A.* **95**:13108-13113.
- 405 4. **Avery, O. T., C. M. MacLeod, and M. McCarty.** 1944. Studies on the chemical nature
 406 of the substance inducing transformation of pneumococcal types. Induction of
 407 transformation by a deoxyribonucleic acid fraction isolated from pneumococcus type III.
 408 *J. Exp. Med.* **79**:137-158.
- 409 5. **Blomberg, C., J. Dagerhamn, S. Dahlberg, S. Browall, J. Fernebro, B. Albiger, E.**
 410 **Morfeldt, S. Normark, and B. Henriques-Normark.** 2009. Pattern of accessory regions
 411 and invasive disease potential in *Streptococcus pneumoniae*. *J. Infect. Dis.* **199**:1032-
 412 1042.
- 413 6. **Briles, D. E., M. Nahm, K. Schroer, J. Davie, P. Baker, J. Kearney, and R. Barletta.**
 414 1981. Antiphosphocholine antibodies found in normal mouse serum are protective
 415 against intravenous infection with type 3 *Streptococcus pneumoniae*. *J. Exp. Med.*
 416 **153**:694-705.
- 417 7. **Brueggemann, A. B., D. T. Griffiths, E. Meats, T. Peto, D. W. Crook, and B. G.**
 418 **Spratt.** 2003. Clonal relationships between invasive and carriage *Streptococcus*
 419 *pneumoniae* and serotype- and clone-specific differences in invasive disease potential. *J.*
 420 *Infect. Dis.* **187**:1424-1432.

- 421 8. **Chen, H., Y. Ma, J. Yang, C. J. O'Brien, S. L. Lee, J. E. Mazurkiewicz, S. Haataja,**
422 **J. H. Yan, G. F. Gao, and J. R. Zhang.** 2007. Genetic requirement for pneumococcal
423 ear infection. *PLoS ONE* **3**:e2950.
- 424 9. **Coates, H. L., P. S. Morris, A. J. Leach, and S. Couzos.** 2002. Otitis media in
425 Aboriginal children: tackling a major health problem. *Med. J. Aust.* **177**:177-178.
- 426 10. **Croucher, N. J., D. Walker, P. Romero, N. Lennard, G. K. Paterson, N. C. Bason, A.**
427 **M. Mitchell, M. A. Quail, P. W. Andrew, J. Parkhill, S. D. Bentley, and T. J.**
428 **Mitchell.** 2009. Role of conjugative elements in the evolution of the multidrug-resistant
429 pandemic clone *Streptococcus pneumoniae* Spain23F ST81. *J. Bacteriol.* **191**:1480-1489.
- 430 11. **Dagerhamn, J., C. Blomberg, S. Browall, K. Sjostrom, E. Morfeldt, and B.**
431 **Henriques-Normark.** 2008. Determination of accessory gene patterns predicts the same
432 relatedness among strains of *Streptococcus pneumoniae* as sequencing of housekeeping
433 genes does and represents a novel approach in molecular epidemiology. *J. Clin.*
434 *Microbiol.* **46**:863-868.
- 435 12. **Dopazo, J., A. Mendoza, J. Herrero, F. Caldara, Y. Humbert, L. Friedli, M.**
436 **Guerrier, E. Grand-Schenk, C. Gandin, M. de Francesco, A. Polissi, G. Buell, G.**
437 **Feger, E. Garcia, M. Peitsch, and J. F. Garcia-Bustos.** 2001. Annotated draft genomic
438 sequence from a *Streptococcus pneumoniae* type 19F clinical isolate. *Microb. Drug*
439 *Resist.* **7**:99-125.
- 440 13. **Embry, A., E. Hinojosa, and C. J. Orihuela.** 2007. Regions of Diversity 8, 9 and 13
441 contribute to *Streptococcus pneumoniae* virulence. *BMC Microbiol.* **7**:80.
- 442 14. **Enright, M. C., and B. G. Spratt.** 1999. Multilocus sequence typing. *Trends Microbiol.*
443 **7**:482-487.
- 444 15. **Enright, M. C., and B. G. Spratt.** 1998. A multilocus sequence typing scheme for
445 *Streptococcus pneumoniae*: identification of clones associated with serious invasive
446 disease. *Microbiology* **144**:3049-3060.
- 447 16. **Forbes, M. L., E. Horsey, N. L. Hiller, F. J. Buchinsky, J. D. Hayes, J. M.**
448 **Compliment, T. Hillman, S. Ezzo, K. Shen, R. Keefe, K. Barbadora, J. C. Post, F. Z.**
449 **Hu, and G. D. Ehrlich.** 2008. Strain-specific virulence phenotypes of *Streptococcus*
450 *pneumoniae* assessed using the *Chinchilla laniger* model of otitis media. *PLoS ONE*
451 **3**:e1969.
- 452 17. **Giammarinaro, P., and J. C. Paton.** 2002. Role of RegM, a homologue of the catabolite
453 repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*. *Infect. Immun.*
454 **70**:5454-5461.
- 455 18. **Hanage, W. P., K. Auranen, R. Syrjanen, E. Herva, P. H. Makela, T. Kilpi, and B.**
456 **G. Spratt.** 2004. Ability of pneumococcal serotypes and clones to cause acute otitis
457 media: implications for the prevention of otitis media by conjugate vaccines. *Infect.*
458 *Immun.* **72**:76-81.
- 459 19. **Hanage, W. P., T. H. Kaijalainen, R. K. Syrjanen, K. Auranen, M. Leinonen, P. H.**
460 **Makela, and B. G. Spratt.** 2005. Invasiveness of serotypes and clones of *Streptococcus*
461 *pneumoniae* among children in Finland. *Infect. Immun.* **73**:431-435.
- 462 20. **Hausdorff, W. P., J. Bryant, C. Kloek, P. R. Paradiso, and G. R. Siber.** 2000. The
463 contribution of specific pneumococcal serogroups to different disease manifestations:
464 implications for conjugate vaccine formulation and use, part II. *Clin. Infect. Dis.* **30**:122-
465 140.

- 466 21. Hausdorff, W. P., G. Yothers, R. Dagan, T. Kilpi, S. I. Pelton, R. Cohen, M. R.
467 Jacobs, S. L. Kaplan, C. Levy, E. L. Lopez, E. O. Mason, Jr., V. Syriopoulou, B.
468 Wynne, and J. Bryant. 2002. Multinational study of pneumococcal serotypes causing
469 acute otitis media in children. *Pediatr. Infect. Dis. J.* **21**:1008-1016.
- 470 22. Henriques, B., M. Kalin, A. Ortqvist, B. Olsson Liljequist, M. Almela, T. J. Marrie,
471 M. A. Mufson, A. Torres, M. A. Woodhead, S. B. Svenson, and G. Kallenius. 2000.
472 Molecular epidemiology of *Streptococcus pneumoniae* causing invasive disease in 5
473 countries. *J. Infect. Dis.* **182**:833-839.
- 474 23. Hiller, N. L., B. Janto, J. S. Hogg, R. Boissy, S. Yu, E. Powell, R. Keefe, N. E.
475 Ehrlich, K. Shen, J. Hayes, K. Barbadora, W. Klimke, D. Dernovoy, T. Tatusova, J.
476 Parkhill, S. D. Bentley, J. C. Post, G. D. Ehrlich, and F. Z. Hu. 2007. Comparative
477 genomic analyses of seventeen *Streptococcus pneumoniae* strains: insights into the
478 pneumococcal supragenome. *J. Bacteriol.* **189**:8186-8195.
- 479 24. Hogg, J. S., F. Z. Hu, B. Janto, R. Boissy, J. Hayes, R. Keefe, J. C. Post, and G. D.
480 Ehrlich. 2007. Characterization and modeling of the *Haemophilus influenzae* core and
481 supragenomes based on the complete genomic sequences of Rd and 12 clinical
482 nontypeable strains. *Genome Biol.* **8**:R103.
- 483 25. Hollingshead, S. K., R. Becker, and D. E. Briles. 2000. Diversity of PspA: mosaic
484 genes and evidence for past recombination in *Streptococcus pneumoniae*. *Infect. Immun.*
485 **68**:5889-5900.
- 486 26. Hoskins, J., W. E. Alborn, Jr., J. Arnold, L. C. Blaszcak, S. Burgett, B. S. DeHoff,
487 S. T. Estrem, L. Fritz, D. J. Fu, W. Fuller, C. Geringer, R. Gilmour, J. S. Glass, H.
488 Khoja, A. R. Kraft, R. E. Lagace, D. J. LeBlanc, L. N. Lee, E. J. Lefkowitz, J. Lu, P.
489 Matsushima, S. M. McAhren, M. McHenney, K. McLeaster, C. W. Mundy, T. I.
490 Nicas, F. H. Norris, M. O'Gara, R. B. Peery, G. T. Robertson, P. Rockey, P. M. Sun,
491 M. E. Winkler, Y. Yang, M. Young-Bellido, G. Zhao, C. A. Zook, R. H. Baltz, S. R.
492 Jaskunas, P. R. Rosteck, Jr., P. L. Skatrud, and J. I. Glass. 2001. Genome of the
493 bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.* **183**:5709-5717.
- 494 27. Iannelli, F., M. R. Oggioni, and G. Pozzi. 2002. Allelic variation in the highly
495 polymorphic locus *pspC* of *Streptococcus pneumoniae*. *Gene* **284**:63-71.
- 496 28. Isozumi, R., Y. Ito, T. Ishida, T. Hirai, I. Ito, K. Maniwa, M. Hayashi, H. Kagioka,
497 M. Hirabayashi, K. Onaru, H. Tomioka, K. Tomii, I. Gohma, M. Osawa, S. Imai, S.
498 Takakura, Y. Iinuma, K. Chin, S. Ichiyama, and M. Mishima. 2008. Molecular
499 characteristics of serotype 3 *Streptococcus pneumoniae* isolates among community-
500 acquired pneumonia patients in Japan. *J. Infect. Chemother.* **14**:258-261.
- 501 29. Jedrzejewski, M. J. 2004. Extracellular virulence factors of *Streptococcus pneumoniae*.
502 *Front. Biosci.* **9**:891-914.
- 503 30. Kadioglu, A., S. Taylor, F. Iannelli, G. Pozzi, T. J. Mitchell, and P. W. Andrew.
504 2002. Upper and lower respiratory tract infection by *Streptococcus pneumoniae* is
505 affected by pneumolysin deficiency and differences in capsule type. *Infect. Immun.*
506 **70**:2886-2890.
- 507 31. Kelly, T., J. P. Dillard, and J. Yother. 1994. Effect of genetic switching of capsular
508 type on virulence of *Streptococcus pneumoniae*. *Infect. Immun.* **62**:1813-1819.
- 509 32. King, S. J., A. M. Whatmore, and C. G. Dowson. 2005. NanA, a neuraminidase from
510 *Streptococcus pneumoniae*, shows high levels of sequence diversity, at least in part
511 through recombination with *Streptococcus oralis*. *J. Bacteriol.* **187**:5376-5386.

- 512 33. **Leach, A. J., and P. S. Morris.** 2007. The burden and outcome of respiratory tract
513 infection in Australian and aboriginal children. *Pediatr. Infect. Dis. J.* **26**:S4-7.
- 514 34. **Leach, A. J., P. S. Morris, G. B. McCallum, C. A. Wilson, L. Stubbs, J. Beissbarth,**
515 **S. Jacups, K. Hare, and H. C. Smith-Vaughan.** 2009. Emerging pneumococcal
516 carriage serotypes in a high-risk population receiving universal 7-valent pneumococcal
517 conjugate vaccine and 23-valent polysaccharide vaccine since 2001. *BMC Infect. Dis.*
518 **9**:121.
- 519 35. **Leimkugel, J., A. Adams Forgor, S. Gagneux, V. Pfluger, C. Flierl, E. Awine, M.**
520 **Naegeli, J. P. Dangy, T. Smith, A. Hodgson, and G. Pluschke.** 2005. An outbreak of
521 serotype 1 *Streptococcus pneumoniae* meningitis in northern Ghana with features that are
522 characteristic of *Neisseria meningitidis* meningitis epidemics. *J. Infect. Dis.* **192**:192-199.
- 523 36. **LeMessurier, K. S., A. D. Ogunniyi, and J. C. Paton.** 2006. Differential expression of
524 key pneumococcal virulence genes *in vivo*. *Microbiology* **152**:305-311.
- 525 37. **Lynch, J. P., 3rd, and G. G. Zhanel.** 2005. Escalation of antimicrobial resistance among
526 *Streptococcus pneumoniae*: implications for therapy. *Semin. Resp. Crit. Care Med.*
527 **26**:575-616.
- 528 38. **Mackenzie, G. A., J. R. Carapetis, A. J. Leach, and P. S. Morris.** 2009. Pneumococcal
529 vaccination and otitis media in Australian Aboriginal infants: comparison of two birth
530 cohorts before and after introduction of vaccination. *BMC Pediatr.* **9**:14.
- 531 39. **Mahdi, L. K., A. D. Ogunniyi, K. S. LeMessurier, and J. C. Paton.** 2008.
532 Pneumococcal virulence gene expression and host cytokine profiles during pathogenesis
533 of invasive disease. *Infect. Immun.* **76**:646-657.
- 534 40. **Martin, B., P. Garcia, M. P. Castanie, and J. P. Claverys.** 1995. The *recA* gene of
535 *Streptococcus pneumoniae* is part of a competence-induced operon and controls
536 lysogenic induction. *Mol. Microbiol.* **15**:367-379.
- 537 41. **McAllister, L. J., H. J. Tseng, A. D. Ogunniyi, M. P. Jennings, A. G. McEwan, and**
538 **J. C. Paton.** 2004. Molecular analysis of the *psa* permease complex of *Streptococcus*
539 *pneumoniae*. *Mol. Microbiol.* **53**:889-901.
- 540 42. **McEllistrem, M. C., J. M. Adams, K. Patel, A. B. Mendelsohn, S. L. Kaplan, J. S.**
541 **Bradley, G. E. Schutze, K. S. Kim, E. O. Mason, and E. R. Wald.** 2005. Acute otitis
542 media due to penicillin-nonsusceptible *Streptococcus pneumoniae* before and after the
543 introduction of the pneumococcal conjugate vaccine. *Clin. Infect. Dis.* **40**:1738-1744.
- 544 43. **Mizrachi Nebenzahl, Y., N. Porat, S. Lifshitz, S. Novick, A. Levi, E. Ling, O. Liron,**
545 **S. Mordechai, R. K. Sahu, and R. Dagan.** 2004. Virulence of *Streptococcus*
546 *pneumoniae* may be determined independently of capsular polysaccharide. *FEMS*
547 *Microbiol. Lett.* **233**:147-152.
- 548 44. **Morona, J. K., R. Morona, and J. C. Paton.** 2006. Attachment of capsular
549 polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for
550 invasive disease. *Proc. Natl. Acad. Sci. U.S.A.* **103**:8505-8510.
- 551 45. **Morris, P. S., A. J. Leach, P. Silberberg, G. Mellon, C. Wilson, E. Hamilton, and J.**
552 **Beissbarth.** 2005. Otitis media in young Aboriginal children from remote communities
553 in Northern and Central Australia: a cross-sectional survey. *BMC Pediatr.* **5**:27.
- 554 46. **Mothibeli, K. M., M. du Plessis, A. von Gottberg, L. de Gouveia, P. Adrian, S. A.**
555 **Madhi, and K. P. Klugman.** 2010. An unusual pneumococcal sequence type is the
556 predominant cause of serotype 3 invasive disease in South Africa. *J. Clin. Microbiol.*
557 **48**:184-191.

- 558 47. **Nesin, M., M. Ramirez, and A. Tomasz.** 1998. Capsular transformation of a multidrug-
559 resistant *Streptococcus pneumoniae* in vivo. *J. Infect. Dis.* **177**:707-713.
- 560 48. **Ostroff, S. M.** 1999. Continuing challenge of pneumococcal disease. *Lancet* **353**:1201-
561 1202.
- 562 49. **Paton, J. C.** 2004. New Pneumococcal Vaccines: Basic Science Developments, p. 382-
563 402. *In* E. I. Tuomanen (ed.), *The Pneumococcus*. ASM Press, Washington, DC.
- 564 50. **Paton, J. C., P. W. Andrew, G. J. Boulnois, and T. J. Mitchell.** 1993. Molecular
565 analysis of the pathogenicity of *Streptococcus pneumoniae*: the role of pneumococcal
566 proteins. *Annu. Rev. Microbiol.* **47**:89-115.
- 567 51. **Paton, J. C., R. A. Lock, and D. J. Hansman.** 1983. Effect of immunization with
568 pneumolysin on survival time of mice challenged with *Streptococcus pneumoniae*. *Infect.*
569 *Immun.* **40**:548-552.
- 570 52. **Poolman, J., P. Kriz, C. Feron, E. Di-Paolo, I. Henckaerts, A. Miseur, D. Wauters,**
571 **R. Prymula, and L. Schuerman.** 2009. Pneumococcal serotype 3 otitis media, limited
572 effect of polysaccharide conjugate immunisation and strain characteristics. *Vaccine*
573 **27**:3213-3222.
- 574 53. **Porat, N., C. Soley, M. M. Marengolciene, D. Greenberg, N. Givon-Lavi, R. Trefler,**
575 **A. Arguedas, and R. Dagan.** 2008. An international serotype 3 clone causing pediatric
576 noninvasive infections in Israel, Costa Rica, and Lithuania. *Pediatr. Infect. Dis. J.*
577 **27**:709-712.
- 578 54. **Roche, P. W., V. Krause, H. Cook, J. Barralet, D. Coleman, A. Sweeny, J. Fielding,**
579 **C. Giele, R. Gilmour, R. Holland, R. Kampen, M. Brown, L. Gilbert, G. Hogg, and**
580 **D. Murphy.** 2008. Invasive pneumococcal disease in Australia, 2006. *Commun. Dis.*
581 *Intell.* **32**:18-30.
- 582 55. **Sambrook, J., E. F. Fritsch, and T. Maniatis (ed.).** 1989. *Molecular cloning: a*
583 *laboratory manual*, 2 ed. Cold Spring Harbor Laboratory Press, New York.
- 584 56. **Sandgren, A., K. Sjostrom, B. Olsson-Liljequist, B. Christensson, A. Samuelsson, G.**
585 **Kronvall, and B. Henriques Normark.** 2004. Effect of clonal and serotype-specific
586 properties on the invasive capacity of *Streptococcus pneumoniae*. *J. Infect. Dis.* **189**:785-
587 796.
- 588 57. **Shen, K., J. Gladitz, P. Antalis, B. Dice, B. Janto, R. Keefe, J. Hayes, A. Ahmed, R.**
589 **Dopico, N. Ehrlich, J. Jocz, L. Kropp, S. Yu, L. Nistico, D. P. Greenberg, K.**
590 **Barbadora, R. A. Preston, J. C. Post, G. D. Ehrlich, and F. Z. Hu.** 2006.
591 Characterization, distribution, and expression of novel genes among eight clinical isolates
592 of *Streptococcus pneumoniae*. *Infect. Immun.* **74**:321-330.
- 593 58. **Shouval, D. S., D. Greenberg, N. Givon-Lavi, N. Porat, and R. Dagan.** 2006. Site-
594 specific disease potential of individual *Streptococcus pneumoniae* serotypes in pediatric
595 invasive disease, acute otitis media and acute conjunctivitis. *Pediatr. Infect. Dis. J.*
596 **25**:602-607.
- 597 59. **Silva, N. A., J. McCluskey, J. M. Jefferies, J. Hinds, A. Smith, S. C. Clarke, T. J.**
598 **Mitchell, and G. K. Paterson.** 2006. Genomic diversity between strains of the same
599 serotype and multilocus sequence type among pneumococcal clinical isolates. *Infect.*
600 *Immun.* **74**:3513-3518.
- 601 60. **Sjostrom, K., C. Spindler, A. Ortqvist, M. Kalin, A. Sandgren, S. Kuhlmann-**
602 **Berenzon, and B. Henriques-Normark.** 2006. Clonal and capsular types decide whether

- 603 pneumococci will act as a primary or opportunistic pathogen. Clin. Infect. Dis. **42**:451-
604 459.
- 605 61. **Tettelin, H., K. E. Nelson, I. T. Paulsen, J. A. Eisen, T. D. Read, S. Peterson, J.**
606 **Heidelberg, R. T. DeBoy, D. H. Haft, R. J. Dodson, A. S. Durkin, M. Gwinn, J. F.**
607 **Kolonay, W. C. Nelson, J. D. Peterson, L. A. Umayam, O. White, S. L. Salzberg, M.**
608 **R. Lewis, D. Radune, E. Holtzapple, H. Khouri, A. M. Wolf, T. R. Utterback, C. L.**
609 **Hansen, L. A. McDonald, T. V. Feldblyum, S. Angiuoli, T. Dickinson, E. K. Hickey,**
610 **I. E. Holt, B. J. Loftus, F. Yang, H. O. Smith, J. C. Venter, B. A. Dougherty, D. A.**
611 **Morrison, S. K. Hollingshead, and C. M. Fraser.** 2001. Complete genome sequence of
612 a virulent isolate of *Streptococcus pneumoniae*. Science **293**:498-506.
- 613 62. **Trzcinski, K., C. M. Thompson, and M. Lipsitch.** 2003. Construction of otherwise
614 isogenic serotype 6B, 7F, 14, and 19F capsular variants of *Streptococcus pneumoniae*
615 strain TIGR4. Appl. Environ. Microbiol. **69**:7364-7370.
- 616 63. **Vela Coral, M. C., N. Fonseca, E. Castaneda, J. L. Di Fabio, S. K. Hollingshead, and**
617 **D. E. Briles.** 2001. Pneumococcal surface protein A of invasive *Streptococcus*
618 *pneumoniae* isolates from Colombian children. Emerg. Infect. Dis. **7**:832-836.
- 619 64. **Watson, M., K. Bayley, J. M. Bell, G. L. Gilbert, G. Hogg, A. D. Keil, V. Krause, D.**
620 **Murphy, P. Roche, H. V. Smith, M. G. Stewart, J. Stylianopoulos, and J. Turnidge.**
621 2003. Laboratory surveillance of invasive pneumococcal disease in Australia in 2001 to
622 2002--implications for vaccine serotype coverage. Commun. Dis. Intell. **27**:478-487.
- 623 65. **Weiser, J. N., R. Austrian, P. K. Sreenivasan, and H. R. Masure.** 1994. Phase
624 variation in pneumococcal opacity: relationship between colonial morphology and
625 nasopharyngeal colonization. Infect. Immun. **62**:2582-2589.
626
627

628
629**Table 1. Pneumococcal strains used in this study.**

Strain	Description	Source (Reference)
D39	serotype 2	(4)
D39-ABΔe5	D39 Δ <i>cpsAB ery</i> ^r	(44)
TIGR4	serotype 4	Tim Mitchell, University of Glasgow, Scotland (1)
WU2	serotype 3	David Briles, University of Alabama, USA (6)
WCH206	serotype 3 (ST180) OM isolate	Women's & Children's Hospital, South Australia (WCH)
WCH207	serotype 3 (ST180) OM isolate	WCH
WCH208	serotype 3 (ST458) OM isolate	WCH
WCH209	serogroup 11 sinusitis isolate	WCH
WCH210	serogroup 11 sinusitis isolate	WCH
WCH211	serogroup 11 (ST3020) sinusitis isolate	WCH
WCH212	serogroup 11 sinusitis isolate	WCH
WCH213	serogroup 11 (ST62) OM isolate	WCH
Menzies1	serotype 11A (ST3021) nasopharyngeal isolate	Menzies School of Health Research, Darwin, Australia)
Menzies2	serotype 11A nasopharyngeal isolate	Menzies
Menzies 3	serotype 11A nasopharyngeal isolate	Menzies
Menzies 4	serotype 11A nasopharyngeal isolate	Menzies
Menzies 5	serotype 11A (ST662) nasopharyngeal isolate	Menzies
Menzies 6	serotype 11A nasopharyngeal isolate	Menzies
Menzies 7	serotype 11A nasopharyngeal isolate	Menzies
Menzies 8	serotype 11A nasopharyngeal isolate	Menzies
Menzies 9	serotype 11A nasopharyngeal isolate	Menzies
Menzies 10	serotype 11A nasopharyngeal isolate	Menzies
Menzies 11	serotype 3 (ST458) nasopharyngeal isolate	Menzies
Menzies 12	serotype 3 (ST458) OM isolate	Menzies
Menzies 13	serotype 3 nasopharyngeal isolate	Menzies
Menzies 14	serotype 3 (ST458) OM isolate	Menzies
Menzies 15	serotype 3 nasopharyngeal isolate	Menzies
Menzies 16	serotype 3 nasopharyngeal isolate	Menzies
Menzies 17	serotype 3 (ST458) OM isolate	Menzies
WCH43	serotype 4 (ST205) blood isolate	WCH
2663	serotype 11A (ST3019) lung isolate	Michael Watson, Path West,

		Perth, Western Australia
3518	serotype 11A (ST62)	Path West
1	serotype 1 (ST304) nasopharyngeal isolate	MSHR
4	serotype 1 (ST227) nasopharyngeal isolate	MSHR
1861	serotype 1 (ST3079) blood isolate	MSHR
4496	serotype 1 (ST3018) blood isolate	WCH
3773	serotype 15B (ST199)	Path West
4104	serotype 19A (ST199)	Path West
WCH16	serotype 6A	WCH
WCH18	serotype 6B	WCH
WCH50	serotype 6	WCH
WCH33	serotype 8	WCH
WCH36	serotype 14	WCH
WCH39	serotype 14	WCH
WCH60	serotype 14	WCH
WCH61	serotype 14	WCH
WCH38	serogroup 19	WCH
WCH64	serogroup 19	WCH
WCH65	serogroup 19	WCH
EF3030	serotype 19F	University of Alabama
WCH67	serogroup 23	WCH
WCH68	serogroup 23	WCH
WCH69	serogroup 23	WCH
WCH101	serogroup 23	WCH

630

631

632
633
634**Table 2. Statistical analysis of survival data.**

Strain	WU2	WCH206	WCH207	Menzies11	Menzies17
Pneumonia Sepsis Model					
WU2	-	0.0007	0.0172	0.0185	0.0003
WCH206	0.0007	-	NS	NS	0.0355
WCH207	0.0172	NS ^a	-	NS	0.003
MSHR11	0.0185	NS	NS	-	0.0039
MSHR17	0.0003	0.0355	0.003	0.0039	-
WCH211	NS	0.0004	0.0023	0.0049	0.0004
WCH213	NS	0.0004	0.0023	0.0049	0.0004
Menzies1	NS	0.0004	0.0023	0.0049	0.0004
Menzies5	NS	0.0004	0.0023	0.0049	0.0004
I.P. Model					
WU2	-	NS	0.0355	NS	0.0015
WCH206	NS	-	0.0147	NS	0.0147
WCH207	0.0355	0.0147	-	NS	NS
MSHR11	NS	NS	NS	-	NS
MSHR17	0.0015	0.0147	NS	NS	-
WCH211	0.0298	0.0302	0.0114	0.0412	0.0123
WCH213	0.0002	0.0004	0.0001	0.0008	0.0002
Menzies1	NS	NS	NS	NS	0.0168
Menzies5	0.0021	0.0028	0.0025	0.0078	0.0027

635
636
637
638
639

Survival time data were analyzed using the Mann Whitney U-test (two tailed).

^a NS indicates not significant ($P > 0.05$).

640
641
642**Table 3. Statistical analysis of nasopharyngeal colonization data.**

Serogroup 11 Strains	WU2	WCH206	WCH207	Menzies11	Menzies17
DAY 1					
WCH211	0.0345	NS ^a	NS	0.0097	NS
WCH213	0.0112	0.0200	0.0159	0.0097	0.0200
Menzies1	NS	NS	NS	0.0259	NS
Menzies5	NS	NS	NS	0.0259	NS
DAY 2					
WCH211	NS	NS	NS	NS	NS
WCH213	0.0112	0.0112	0.0200	0.0112	0.0317
Menzies1	0.0112	0.0112	0.0112	0.0112	0.0079
Menzies5	0.0112	0.0112	0.0112	0.0112	0.0079
DAY 3					
WCH211	NS	0.0311	NS	NS	NS
WCH213	NS	0.0178	NS	NS	NS
Menzies1	NS	0.0097	NS	NS	NS
Menzies5	0.0317	0.0097	NS	NS	NS
DAY 4					
WCH211	NS	NS	NS	NS	NS
WCH213	NS	NS	NS	NS	NS
Menzies1	NS	NS	NS	NS	NS
Menzies5	NS	NS	NS	NS	NS

643
644
645
646

Survival time data were analyzed using the Mann Whitney U-test (two tailed).

^a NS indicates not significant ($P > 0.05$).

647

648

FIGURE LEGENDS

649

650 FIG. 1. Alleles of *nanaA*, *pspA* and *pspC* detected in serogroup 3 and 11 strains. Alleles
651 were identified for *nanaA*, *pspA* and *pspC* by sequencing or PspA family typing, in accordance
652 with King et al. (32), Hollingshead et al. (25) and Iannelli et al. (27), respectively. Proteins are
653 shown in proportion to their approximate size. NanA is represented in lilac. The NanA identified
654 in WCH206 is identical to the NanA present in D39. The other *nanaA* alleles have a 15 bp
655 insertion (indicated by a dark purple box) and a 180 bp deletion (indicated by a black line),
656 which have been previously described in the *nanaA* alleles of other pneumococcal strains (32). In
657 addition, one of the alleles, which is found in both serotype 3 and serogroup 11, has a 500 bp
658 region of diversity at the 5'-terminus, which is indicated by pink and has also been previously
659 described (32). PspA and PspC proteins are categorized as described by Hollingshead et al. (25)
660 and Iannelli et al. (27), respectively. The colour scheme of Iannelli et al. has been used.
661 Homologous sequences are represented by boxes of the same colour. Dark grey/black = signal
662 peptide, other shades of grey = random coil, green = proline rich region, blue = anchor, other
663 colours = α -helix. The clade determining regions of PspA (25) are indicated by white dots. Lines
664 indicate absence of DNA.

665

666 FIG. 2. Capsular polysaccharide production by type 3 isolates. Capsular polysaccharide
667 produced by the indicated strains was quantitated by uronic acid assay, as described in the
668 Materials and Methods. Data are the mean \pm SD of five independent experiments; *, $P < 0.05$;
669 **, $P < 0.01$, unpaired *t*-test).

670

671 FIG. 3. Virulence and nasopharyngeal colonization of selected serotype 3 and serogroup 11
672 isolates. (A) Survival of CD-1 mice in the pneumonia/sepsis model. Groups of 10 mice were
673 inoculated with $1.5 - 4 \times 10^7$ CFU of the indicated strains. The median survival time for each

674 group is indicated by a horizontal bar. (B) Survival of CD-1 mice following i.p. injection of
675 groups of 10 mice with either 5×10^2 CFU (serotype 3) or 1×10^7 CFU (serogroup 11) of
676 bacteria. The median survival time for each group is indicated by a horizontal bar. (C) The
677 number of bacteria in the nasopharynx (nasopharyngeal tissue and nasal wash) over a 7 day
678 period following intranasal inoculation with $1-2 \times 10^7$ CFU/mouse. Groups of 20 mice were
679 inoculated without anaesthesia and 5 mice were euthanased per timepoint, survival permitting.
680 The median number of bacteria in the nasopharynx for each strain is represented by a horizontal
681 bar. The broken line represents the limit of detection (40 CFU). (D). Long term colonization after
682 intranasal challenge was determined by plating nasopharyngeal washes from the mice that were
683 still alive 4 weeks after intranasal challenge in (A) above.

684

685 FIG. 4. The genomic islands in *S. pneumoniae* Hungary^{19A}-6 and G54 containing the
686 putative cellobiose PTS system. The figure is based on the genome maps of Hungary^{19A}-6 (nt
687 1784704-1794633, Refseq [NC_010380](#)) and G54 (nt 1630647-1637715, Refseq [NC_011072](#)).
688 The indicated ORFs are annotated as follows: sph1921/spg1693, hypothetical protein;
689 sph1922/spg1693, hypothetical; sph1923, sulfatase-modifying factor 1 (spg1694, hypothetical);
690 sph1924, sulfatase; sph1925, PTS system cellobiose-specific IIC component (spg1995,
691 hypothetical; spg1696, hypothetical); sph1926/spg1697, lichenan-specific phosphotransferase
692 enzyme IIA component; sph1927/spg1698, lichenan-specific phosphotransferase enzyme IIB
693 component; sph1928, hypothetical protein (spg1699, hypothetical; spg1700, hypothetical);
694 sph1929/spg1701, lichenan permease IIC component; sph1930/spg1702, putative ROK family
695 protein. Genes were shaded according to their KEGG pathway function with black indicating
696 membrane transport and white indicating an unassigned pathway. WCH206 possesses the same
697 island as Hungary19A-6. Menzies17 and Menzies11 have the same deletion as G54, but do not
698 have the two single bp insertions causing frame-shifts and premature stop codons.
699

700 FIG. 5. Ability of the capsule switch mutant to colonize various niches of the mouse.
701 Groups of mice were anesthetized and inoculated i.n. with approximately 10^7 CFU/mouse of
702 WCH206, Menzies5³ or Menzies5. Numbers of pneumococci in various host niches were
703 determined for 5 mice from each group after 1, 2 or 3 days, as described in the Materials and
704 Methods. The median number of bacteria is indicated by a black, horizontal bar, and the limit of
705 detection is indicated by a broken line, which is 40 CFU in the nasopharynx, lungs, blood and
706 brain, and 4 CFU in the ears. The Mann Whitney *U* test (two tailed) was used to analyse data. *,
707 $P < 0.05$; **, $P < 0.01$.
708
709

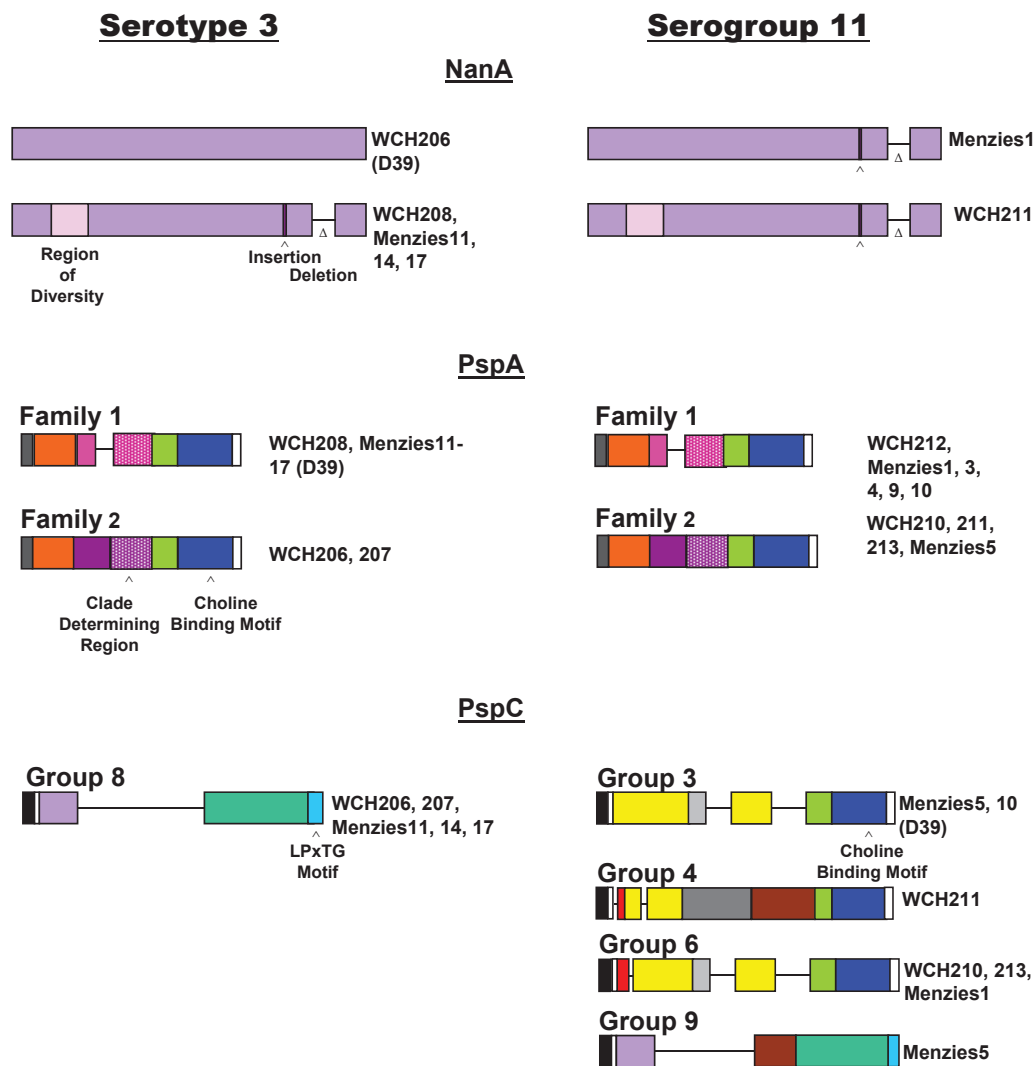


Fig. 1

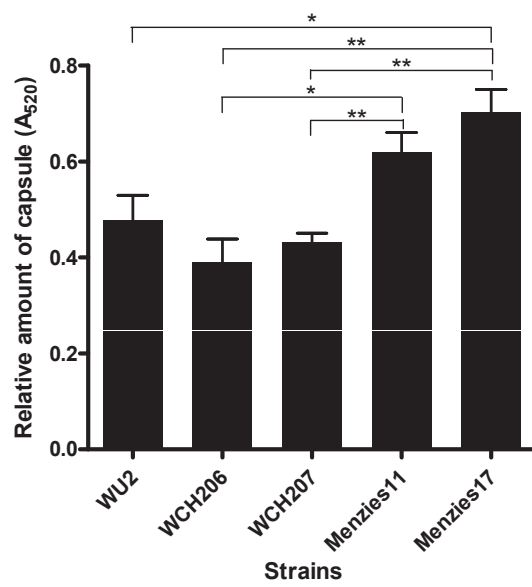


Fig. 2

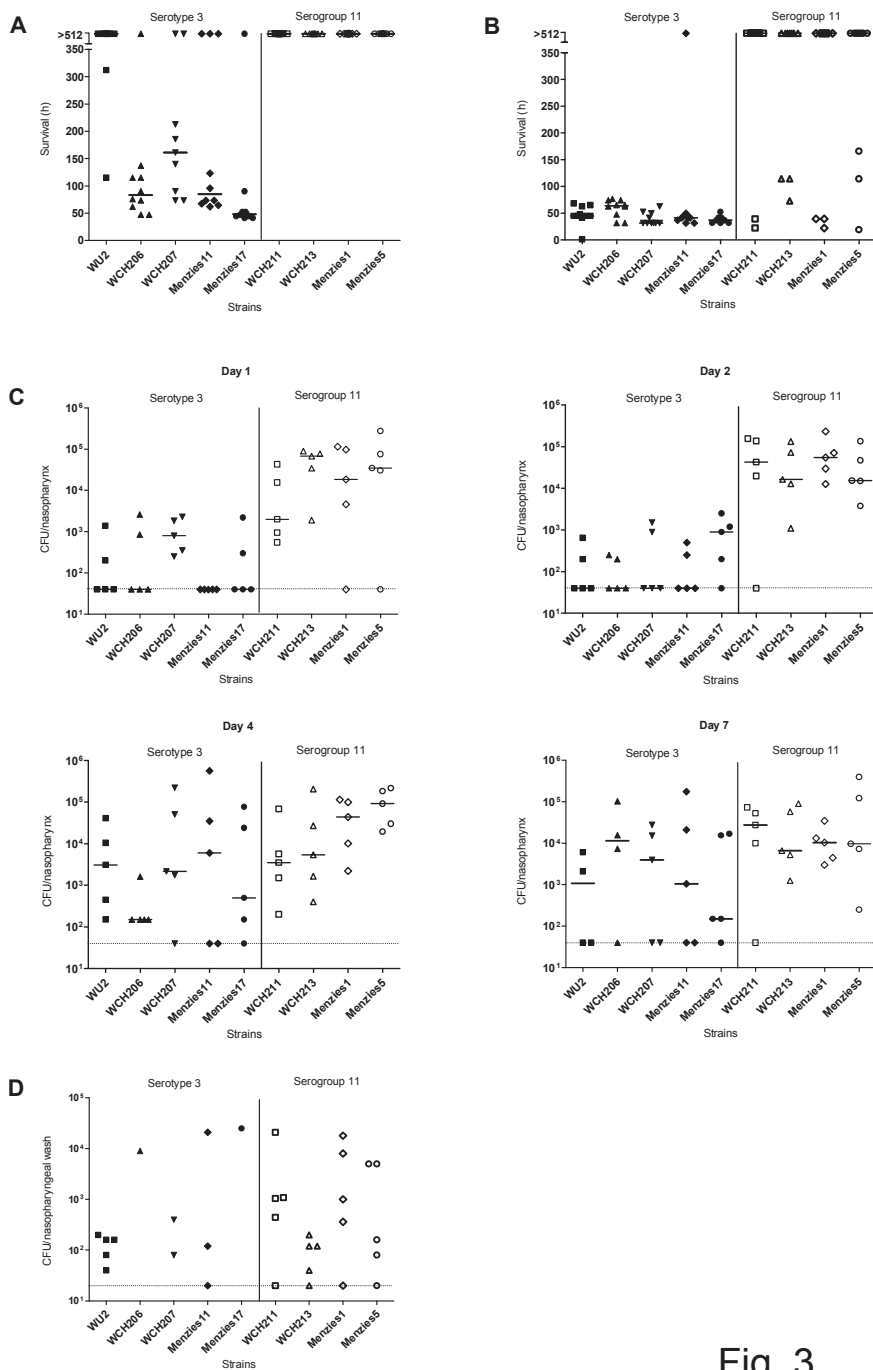


Fig. 3

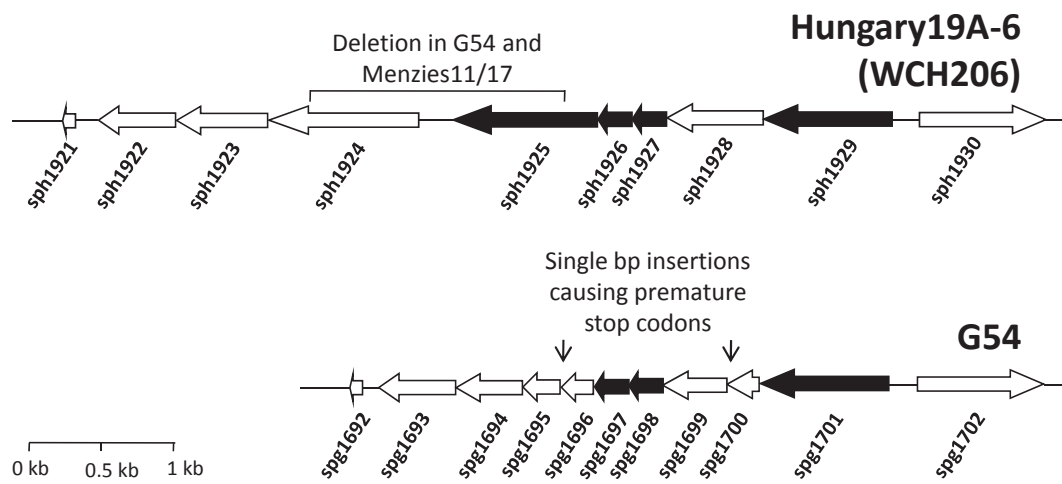


Fig. 4

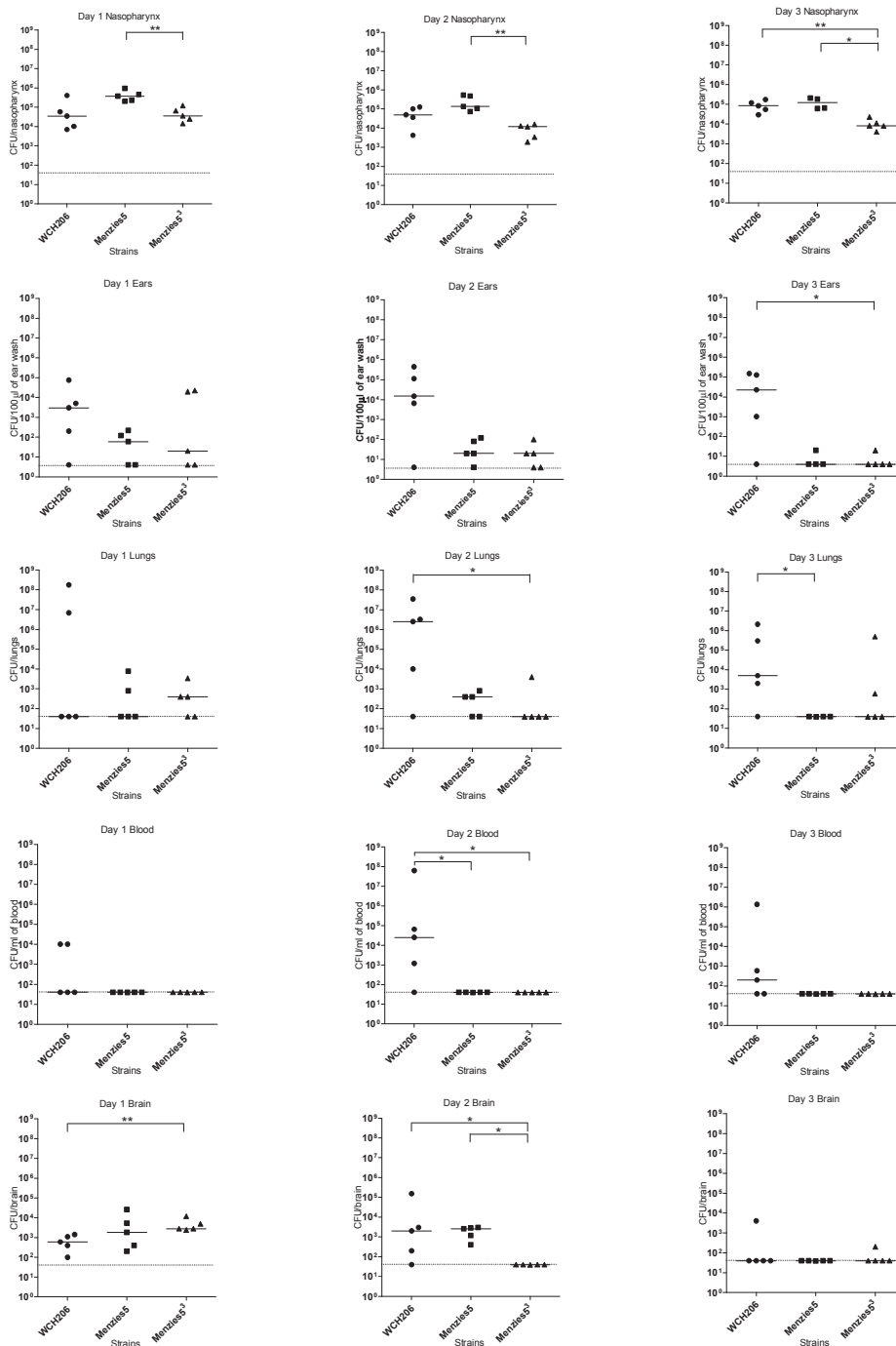


Fig. 5