Brief Report:

Tricontinental detection of Streptococcus pyogenes $M1_{UK}$: A call for wider research and active surveillance.

Ana Vieira[1,2], Valerie WC Soo[1], Ho Kwong Li[1,2], Xiangyun Zhi[1,2], Lucy Reeves[1], Kristin K. Huse[1,2], Kai Yi Mok[1], Oscar Cowen[1], Elita Jauneikaite[3,4], Juliana Coelho[3,5], Shiranee Sriskandan[1,2,3]

[1]Department of Infectious Disease, Faculty of Medicine, Imperial College London, London, UK.

[2]Centre for Bacterial Resistance Biology, Imperial College London, London, UK.

[3]NIHR Health Protection Research Unit in Healthcare-associated Infections and AMR, Imperial College London, London, UK.

[4]School of Public Health, Faculty of Medicine, Imperial College London, London, UK.

[5]Centre for Infections, UK Health Security Agency, London, UK,

Correspondence: Shiranee Sriskandan s.sriskandan@imperial.ac.uk

Word count 1441

Summary

The *Streptococcus pyogenes* $M1_{UK}$ lineage, characterised by an intrinsic ability to express SpeA toxin and defined by 27 single nucleotide polymorphisms in the core genome, dominates the population of *emm*1 *S. pyogenes* isolates in England. The lineage has been identified elsewhere in Europe, North America, and, most recently, Australia. SpeA however may not be the sole deterministic factor underlying success of the $M1_{UK}$ lineage. Production of SpeA by strains belonging to an intermediate *emm*1 sublineage, $M1_{23SNP}$, is indistinguishable from $M1_{UK}$ strains. Despite this, in England at least, $M1_{UK}$ has outcompeted strains from the $M1_{23SNP}$ sublineage. We infer that the fitness of $M1_{UK}$ resides in additional properties that confer an advantage to *S. pyogenes*, underlining a need for further research. A single nucleotide polymorphism (SNP) in the ssrA leader sequence upstream of *speA* is one of a limited number of SNPs that distinguish intermediate sublineages that differ in SpeA production. Introduction of the ssrA SNP into representative isolates of the widely disseminated $M1_{global}$ clone and the intermediate $M1_{13SNP}$ lineage (that cannot otherwise produce readily-detectable SpeA in culture) resulted in SpeA expression, confirming the importance of the ssrA SNP to SpeA phenotype. However, RNAseq analysis of clinical strains showed that presence of the SNP was not invariably linked to read-through from the ssrA leader sequence or SpeA expression. Literature review suggests that read through and *speA* mRNA transcript length may be impacted by the two component regulator CovRS, pointing to a complex regulatory network interaction between the bacterial chromosome and phage-encoded superantigens.

Introduction

A marked increase in serious *Streptococcus pyogenes* infections has been reported in several countries following relaxation of public health interventions designed to limit the spread of COVID-19 (1). In England, notifications of both scarlet fever and invasive *S. pyogenes* infections increased throughout 2022; infections peaked at the end of the calendar year but remained high thereafter (2). In the 30 weeks from mid September 2022-mid April 2023 there were 54,394 notifications of scarlet fever and 2965 invasive *S. pyogenes* infections in England alone (2). The overall case fatality rate of 14.2% (2) underlines the major public health impact of the observed increase, and the importance

of understanding the relative roles of strain pathogenicity, population immune susceptibility, co-infections, season, and access to healthcare.

*Emm*1 *S. pyogenes* strains are known to be inherently invasive and normally account for one quarter of invasive infections; currently however, *emm*1 isolates account for over half of invasive infections, rising to over 60% in children in England (2). The population of *emm*1 *S. pyogenes* isolates in England is dominated by a novel sublineage, $M1_{UK,}$ that has an intrinsic ability to express the phage-encoded superantigenic toxin, streptococcal pyrogenic exotoxin (Spe) A (3, 4). $M1_{UK}$ strains can be differentiated from strains belonging to the highly successful *emm*1 *S. pyogenes* globally-disseminated clone that emerged in the 1980's ($M1_{global}$) by just 27 single nucleotide polymorphisms in the core genome (3). The $M1_{UK}$ lineage has also been identified elsewhere in Europe (5), North America (6, 7), and most recently in Australia, in isolates from Queensland and Victoria dating from 2013 (8), just 3 years after its first detection in England (3). While $M1_{UK}$ represented ~60% of Australian *emm*1 isolates by 2020, in England, this proportion reached 91% by 2020 (4,8). The detection of $M1_{UK}$ in Australia, including Queensland, underlines the ability of this lineage to spread in temperate and more tropical climates. Of added concern are reports of strains infected with a ΦHKU488.vir-like phage, that carries both *speC* and *ssa* superantigen genes in Australia (8). To date, this phage has not been identified in *emm*1 strains from England, although ~10% of *emm*1 strains in England do possess a phage encoding *speC* and *spd*.

Previous genomic analysis of *emm*1 *S. pyogenes* strains in England identified two intermediate sublineages in addition to $M1_{global}$ and $M1_{UK,}$ with just 13 and 23 of the 27 SNPs that define $M1_{UK}$ (3). We recently reported that strains from the $M1_{13SNP}$ intermediate sublineage produce negligible amounts of SpeA, similar to $M1_{global}$ (9). Phylogenetic analysis of Australian sequences (9) shows the $M1_{13SNP}$ sublineage to be present in Australia too (Figure 1A). Strains from the $M1_{23SNP}$ sublineage, in contrast, produce SpeA at an increased level that is indistinguishable from $M1_{UK}$ isolates (9). Despite increased SpeA production, by 2020, we were unable to detect $M1_{23SNP}$ isolates when screening all *emm*1 isolates submitted to the reference laboratory in England (4). The $M1_{23SNP}$

75 intermediate sublineage appears to have been outcompeted by $M1_{UK}$, suggesting an added fitness
76 advantage in $M1_{UK}$ strains beyond production of SpeA alone. We postulate this further adaptation
77 may be conferred by the additional 4 SNPs that distinguish $M1_{UK}$ from $M1_{23SNP}$ isolates (3 non-
78 synonymous SNPs and an intergenic SNP 39bp upstream of *glpF2*, aquaporin, expression of which
79 is significantly reduced in $M1_{UK}$) (8, 9). In this report we seek to further understand the mechanisms
80 that differentiate SpeA-producers from non-producers among *emm*1 *S. pyogenes* strains that carry
81 the phage that encodes *speA*.

82 **Results.**

83 The $M1_{UK}$ SNP at position 983438 of the MGAS5005 reference genome, within the leader sequence
84 of *ssrA* and upstream of the *speA* start site (referred to hereafter as the 'ssrA SNP'), is present in
85 both $M1_{23SNP}$ and $M1_{UK}$ sublineages, but is absent in $M1_{global}$ and $M1_{13SNP}$ sublineages. Through
86 screening of natural mutants arising during outbreaks, the ssrA SNP was determined to be
87 specifically associated with increased SpeA production (9). We introduced the ssrA SNP into an
88 $M1_{global}$ strain, and into an intermediate $M1_{13SNP}$ strain, to determine if this would be sufficient to
89 enhance SpeA production. Consistent with the findings of Davies et al (8), SpeA production in culture
90 supernatants increased in both transformants (Figure 1B). Enhancement of SpeA was similar in both
91 the $M1_{global}$ and $M1_{13SNP}$ backgrounds, suggesting that the accumulated SNPs in the $M1_{13SNP}$ lineage,
92 which include almost half of the SNPs that characterize $M1_{UK}$, do not appreciably contribute to the
93 SpeA phenotype. We considered if our mutagenesis strategy itself might lead to increased SpeA
94 expression but this was ruled out by use of a control construct in each strain (Figure 1B).

95

96 To understand the basis for difference in *speA* expression between the lineages, we compared
97 RNAseq reads from four randomly selected $M1_{global}$ strains, which do not have the ssrA SNP, and
98 four randomly selected $M1_{UK}$ strains, which all have this SNP. None of the strains possessed
99 regulatory gene mutations other than those that characterize $M1_{UK}$. Similar to Davies et al (8), we
100 found that there was indeed evidence of read-through in the intergenic regions between the end of
101 *ssrA* and the start of *spea*. (Figure 1C). However, we did not see clear differentiation between the
102 four $M1_{global}$ strains and the four $M1_{UK}$ strains within the intergenic region, where read abundance in

103 two of the $M1_{global}$ strains was almost identical to $M1_{UK}$ strains (Figure 1C). Regardless, transcription
104 of *speA* was significantly reduced in all four $M1_{global}$ strains compared with $M1_{UK}$. Interestingly, *speA*
105 read abundance was somewhat higher in the two $M1_{global}$ strains that had more intergenic reads, but
106 not as high as the $M1_{UK}$ strains. These data provide support for intergenic read-through affecting
107 *speA* transcription. However, based on short read RNA sequencing data, we cannot definitively
108 ascribe read abundance in the intergenic region to the ssrA SNP in all cases, noting that all four of
109 the $M1_{global}$ strains studied lack any SNP in this region. Comparative WGS analysis between the four
110 $M1_{global}$ strains failed to detect genetic features associated with higher levels of read-through. Taken
111 together, transcriptional read-through was observed in this region, but the association with the ssrA
112 SNP and SpeA over-expression was not absolute. Despite not possessing the ssrA SNP, two of four
113 $M1_{global}$ strains tested showed similar read-through to $M1_{UK}$ strains, yet made less SpeA, pointing to
114 additional factors that influence SpeA release. The ssrA SNP however does indeed appear
115 important, at least in the strains examined.
116
117 **Comment**
118 A polycistronic transcript for speA was originally reported ~25 years ago by Cleary *et al*, who
119 identified both ~2kB and ~900bp *speA* mRNA transcripts in a single *emm*1 *S. pyogenes* strain (10).
120 In that case, the longer 2kB transcript was however associated with colonies that failed to produce
121 SpeA, while the shorter 900bp transcript, corresponding to the expected size for the *speA* transcript,
122 was associated with mucoid *emm*1 colonies that produced abundant SpeA. Based on the date of
123 publication, we infer the *emm*1 strain in that study (10) to be from an $M1_{global}$ background. The within-
124 strain variation in *speA* transcript length coupled with mucoid phenotype points to involvement of the
125 two-component regulator covRS (csrRS) which is now known to repress capsule and *speA*
126 transcription in *S. pyogenes* (11-13). Although $M1_{global}$ strains produce little detectable SpeA protein,
127 SpeA production can be markedly upregulated following mutation of *covS* (11,12) or *covR* (12, 13)
128 (Figure 1D), likely accounting for previous reports of upregulation of SpeA following *in vivo* passage
129 of *emm*1 *S. pyogenes* isolates (14).

130 Consistent with the above experimental observations, contemporary $M1_{global}$ invasive strains with
131 *covRS* mutations arising in patients produce SpeA at amounts equivalent to $M1_{UK}$ strains (9). A
132 historic *emm*1 isolate NCTC8198 (SF130; H250; H305) (15) produces 5-10 fold more SpeA than
133 $M1_{UK}$ strains (Figure 1E) yet does not have the $M1_{UK}$ ssrA SNP. It does however possess a stop
134 mutation in *covS* at residue 318. Northern blotting has previously identified only a single dominant
135 ~750bp *speA* mRNA transcript for this strain (16), potentially consistent with abundant SpeA
136 production. We postulate that release from indirect or direct covRS repression may allow *speA*
137 transcription to start at a native promoter site proximal to *speA*, triggered by growth-phase related
138 factors. Alternatively, this may result in more rapid processing of any longer transcript. Whether
139 there is an interaction between the ssrA SNP and covRS, other signals, or RNAseY is unknown.
140
141 $M1_{UK}$ is associated with a recent upsurge in invasive infections in Europe (17,18). The two
142 intermediate sublineages, $M1_{13SNP}$ and $M1_{23SNP,}$ appear to be no longer present in England (4).
143 Despite the $M1_{23SNP}$ sublineage being able to produce SpeA at a level that is indistinguishable from
144 $M1_{UK}$ isolates (9), it has been fully outcompeted by $M1_{UK}$ in England (4). Although increased SpeA
145 production is predicted to perpetuate community transmission, the fitness advantage(s) of $M1_{UK}$ may
146 yet require additional features that promote *emm*1 survival in the human population, underlining a
147 need for further research and surveillance.
148
149 **Author contributions.**
150 AV, VWCS, HKL, and SS conceived the work described. AV, VS, HKL, XZ, EJ, LR, KKH, KYM, OC,
151 JC (UKHSA) undertook data collection, analysis, and data presentation. AV, VWCS, and SS wrote
152 the first draft of the report and all authors contributed to its revision.
153
154 **The authors declare no competing interests**
155
156 **Acknowledgements.** Funding was from UKRI (MR/P022669/1); Wellcome Trust Collaborative
157 award (215539/Z/19/Z); the NIHR Imperial Biomedical Research Centre (BRC). VWCS is supported

158 by an MRC Career Development Award (MR/X007421/1) and HKL was supported by an MRC CMBI
159 Clinical Training Fellowship. The authors acknowledge the support of the NIHR Health Protection
160 Research Unit in Healthcare-associated infections and AMR at Imperial and the Imperial BRC
161 Leonard and Dora Colebrook Laboratory. The authors thank all those who contribute strains to the
162 national and northwest London collections of *S. pyogenes* and the diagnostic and reference
163 laboratories that process these.

164

165

166 **References**

167 1. World Health Organization. Increase in invasive Group A streptococcal infections among children
168 in Europe, including fatalities. https://www.who.int/europe/news/item/12-12-2022-increase-in-
169 invasive-group-a-streptococcal-infections-among-children-in-europe--including-
170 fatalities#:~:text=The%20observed%20increases%20reported%20to,during%20the%20COVID%2
171 D19%20pandemic. Accessed, April 30th 2023.

172

173 2. UK Health Security Agency. Group A Streptococcal infections: 13th update on seasonal activity in
174 England. April 23rd 2023 https://www.gov.uk/government/publications/group-a-streptococcal-
175 infections-activity-during-the-2022-to-2023-season/group-a-streptococcal-infections-13th-update-
176 on-seasonal-activity-in-england. Accessed April 30th 2023.

177

178 3. Lynskey NN, Jauneikaite E, Li HK, Zhi X, Turner CE, Mosavie M, Pearson M, Asai M, Lobkowicz
179 L, Chow JY, Parkhill J, Lamagni T, Chalker VJ, Sriskandan S. Emergence of dominant toxigenic
180 M1T1 *Streptococcus pyogenes* clone during increased scarlet fever activity in England: a population-
181 based molecular epidemiological study. Lancet Infect Dis. 2019 Nov;19(11):1209-1218. doi:
182 10.1016/S1473-3099(19)30446-3.

183

184 4. Zhi X, Li HK, Li H, Loboda Z, Charles S, Vieira A, Huse K, Jauneikaite E, Reeves L, Mok KY,
185 Coelho J, Lamagni T, Sriskandan S. Emerging invasive group A Streptococcus $M1_{UK}$ lineage

186 detected by allele-specific PCR, England, 2020. Emerg Infect Dis. 2023 Apr 5;29(5). doi:
187 10.3201/eid2905.221887.
188
189 5. Rümke LW, de Gier B, Vestjens SMT, van der Ende A, van Sorge NM, Vlaminckx BJM,
190 Witteveen S, van Santen M, Schouls LM, Kuijper EJ. Dominance of $M1_{UK}$ clade among Dutch M1
191 *Streptococcus pyogenes*. Lancet Infect Dis. 2020 May;20(5):539-540. doi: 10.1016/S1473-
192 3099(20)30278-4.
193
194 6. Demczuk W, Martin I, Domingo FR, MacDonald D, Mulvey MR. Identification of *Streptococcus*
195 *pyogenes* $M1_{UK}$ clone in Canada. Lancet Infect Dis. 2019 Dec;19(12):1284-1285. doi:
196 10.1016/S1473-3099(19)30622-X.
197
198 7. Li Y, Nanduri SA, Van Beneden CA, Beall BW. $M1_{UK}$ lineage in invasive group A streptococcus
199 isolates from the USA. Lancet Infect Dis. 2020 May;20(5):538-539. doi: 10.1016/S1473-
200 3099(20)30279-6.
201
202 8.Davies MR, Keller N, Brouwer S, Jespersen MG, Cork AJ, Hayes AJ, Pitt ME, De Oliveira DMP,
203 Harbison-Price N, Bertolla OM, Mediati DG, Curren BF, Taiaroa G, Lacey JA, Smith HV, Fang NX,
204 Coin LJM, Stevens K, Tong SYC, Sanderson-Smith M, Tree JJ, Irwin AD, Grimwood K, Howden BP,
205 Jennison AV, Walker MJ. Detection of *Streptococcus pyogenes* $M1_{UK}$ in Australia and
206 characterization of the mutation driving enhanced expression of superantigen SpeA. Nat Commun.
207 2023;14(1):1051. doi: 10.1038/s41467-023-36717-4.

208 9. Li HK, Zhi X, Vieira A, Whitwell HJ, Schricker A, Jauneikaite E, Li H, Yosef A, Andrew I, Game L,
209 Turner CE, Lamagni T, Coelho J, Sriskandan S. Characterization of emergent toxigenic $M1_{UK}$
210 *Streptococcus pyogenes* and associated sublineages. Microb Genom. 2023 Apr;9(4). doi:
211 10.1099/mgen.0.000994. PMID: 37093716.

212 10. Cleary PP, McLandsborough L, Ikeda L, Cue D, Krawczak J, Lam H. High-frequency intracellular
213 infection and erythrogenic toxin A expression undergo phase variation in M1 group A streptococci.
214 Mol Microbiol. 1998 Apr;28(1):157-67. doi: 10.1046/j.1365-2958.1998.00786.x.

215 11. Sumby P, Whitney AR, Graviss EA, DeLeo FR, Musser JM. Genome-wide analysis of group A
216 streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS
217 Pathog. 2006 Jan;2(1):e5. doi: 10.1371/journal.ppat.0020005.

218 12. Treviño J, Perez N, Ramirez-Peña E, Liu Z, Shelburne SA 3rd, Musser JM, Sumby P. CovS
219 simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A
220 Streptococcus virulence factor-encoding genes. Infect Immun. 2009 Aug;77(8):3141-9. doi:
221 10.1128/IAI.01560-08.

222 13. Horstmann N, Myers KS, Tran CN, Flores AR, Shelburne Iii SA. CovS inactivation reduces CovR
223 promoter binding at diverse virulence factor encoding genes in group A Streptococcus. PLoS Pathog.
224 2022 Feb 18;18(2):e1010341. doi: 10.1371/journal.ppat.1010341

225 14. Kazmi SU, Kansal R, Aziz RK, Hooshdaran M, Norrby-Teglund A, Low DE, Halim AB, Kotb M.
226 Reciprocal, temporal expression of SpeA and SpeB by invasive M1T1 group A streptococcal isolates
227 in vivo. Infect Immun. 2001 Aug;69(8):4988-95. doi: 10.1128/IAI.69.8.4988-4995.2001

228 15. Griffith F. Types of Haemolytic Streptococci in Relation to Scarlet Fever. The Journal of Hygiene, Vol.
229 25, No. 4 (Nov., 1926), pp. 385-397
230

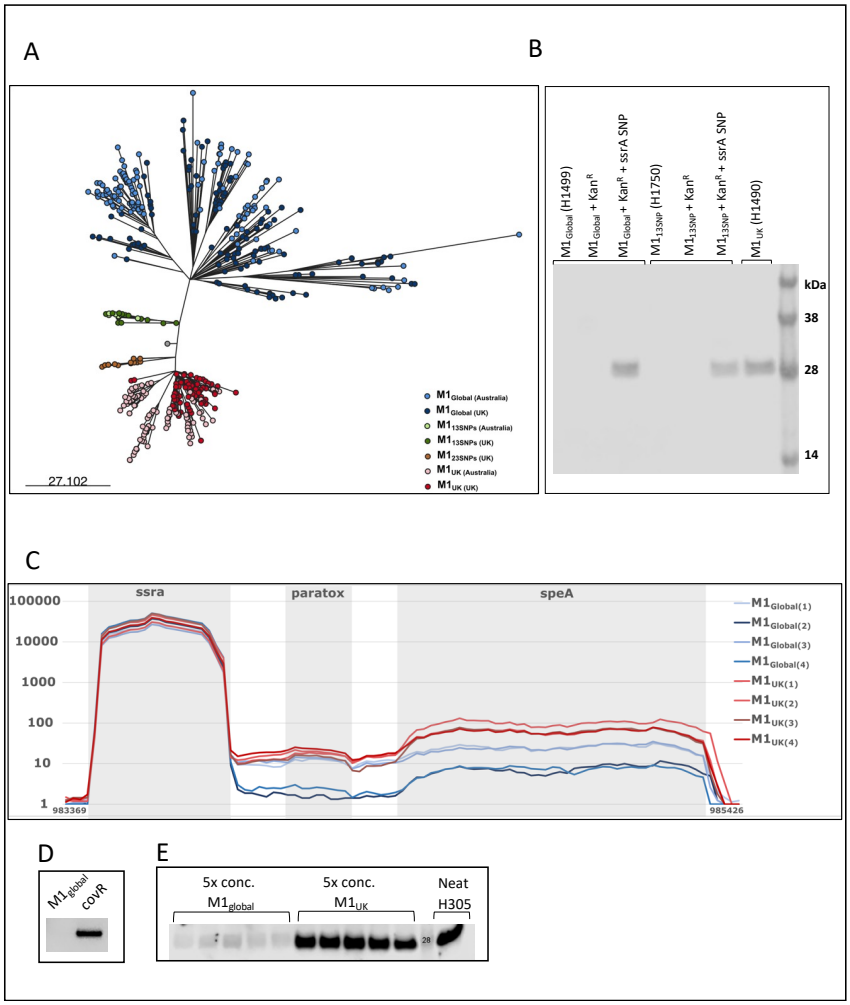
231 16. Unnikrishnan M, Cohen J, Sriskandan S. Growth-phase-dependent expression of virulence
232 factors in an M1T1 clinical isolate of *Streptococcus pyogenes*. Infect Immun. 1999 Oct;67(10):5495-
233 9. doi: 10.1128/IAI.67.10.5495-5499.1999
234

235 17. Guy R, Henderson KL, Coelho J, Hughes H, Mason EL, Gerver SM, Demirjian A, Watson C,
236 Sharp A, Brown CS, Lamagni T. Increase in invasive group A streptococcal infection notifications,

England, 2022. Euro Surveill. 2023 Jan;28(1):2200942. doi: 10.2807/1560-7917.ES.2023.28.1.2200942.

18. van der Putten BCL, Vlaminckx BJM, de Gier B, Freudenburg-de Graaf W, van Sorge NM. Group A streptococcal meningitis with the $M1_{UK}$ variant in the Netherlands. JAMA. 2023 Apr 7:e235927. doi: 10.1001/jama.2023.5927.

244 Figure 1



245
246

247 **Figure 1. M1UK and sublineages; complex regulation of SpeA production**

248 Phylogenetic tree constructed with 413 core SNPs (without recombination regions) obtained by mapping
249 269 *emm*1 *S. pyogenes* isolates selected from UK to demonstrate sublineages (ref. 7) with 319 isolates
250 from Australia (ref. 1) to the reference genome MGAS5005 **(A)**; coloured tips indicate the various
251 sublineages, with Australian strains indicated by lighter colours and a UK strain with 19SNPs in grey. **(B)**
252 Introduction of ssrA SNP into M1global strain H1499 and M113SNP strain H1750 results in detectable SpeA
253 in supernatants when tested using SpeA immunoblot. AphA3 (kanR) was introduced upstream of the ssrA
254 leader sequence as a selectable marker in all transformants. Only transformants with the ssrA SNP
255 introduced expressed detectable SpeA. For comparison, M1UK strain H1490 also shown. **C.** RNA
256 sequencing read abundance across *ssrA*; an intergenic region which includes *ptx*; and *speA* comparing
257 four M1global strains (blue lines) and four M1UK strains (red lines). All were non-invasive throat isolates.
258 M1global isolates were BHS0674 (H1499); BHS0162 (H1489); BHS0130 (H1504); and BHS0151 (H1488).
259 M1UK isolates were BHS0170 (H1490); BHS0128 (H1503); BHS0258 (H1491); and BHS0581 (H1496).
260 **D**. SpeA immunoblot of supernatant from M1global strain H584 and an isogenic covR T65P mutant (H1565).
261 **E.** SpeA immunoblot of concentrated *emm*1 supernatants from M1global strains (H1488; H1489; H1492;
262 H1495; H1499) and M1UK strains (H1490; H1491; H1493; H1539; H1496) (all non-invasive throat strains)
263 compared with non-concentrated supernatant from *emm*1 strain H305 (NCTC8198).