

Supplementary file for manuscript: Characterisation of emergent toxigenic M1_{UK} *Streptococcus pyogenes* and associated sublineages

By Ho Kwong Li*, Xiangyun Zhi*, Ana Vieira, Harry Whitwell, Amelia Schricker, Elita Jauneikaite, Hanqi Li, Ahmed Yosef, Ivan Andrew, Laurence Game, Claire E. Turner, Theresa Lamagni, Juliana Coelho, Shiranee Sriskandan

CONTENTS

Supplementary Figure S1 Real-time PCR measurement of *glpF2* transcription

Supplementary Figure S2 SpeA production by non-invasive *emm1* *S. pyogenes*

Supplementary Figure S3 String analysis of proteomic data: M1_{global} vs. M1_{UK} cellular fractions

Supplementary Figure S4 String analysis of proteomic data: four M1 sublineages compared

Supplementary Table S1 Strains used in experimental work

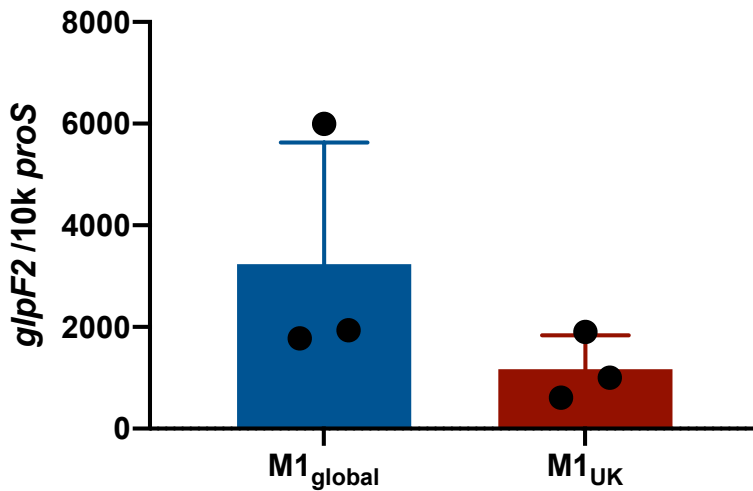
Supplementary Table S2 Strains and WGS used for phylogenetic tree and SpeA quantification

Supplementary Table S3 Chemically-defined medium composition

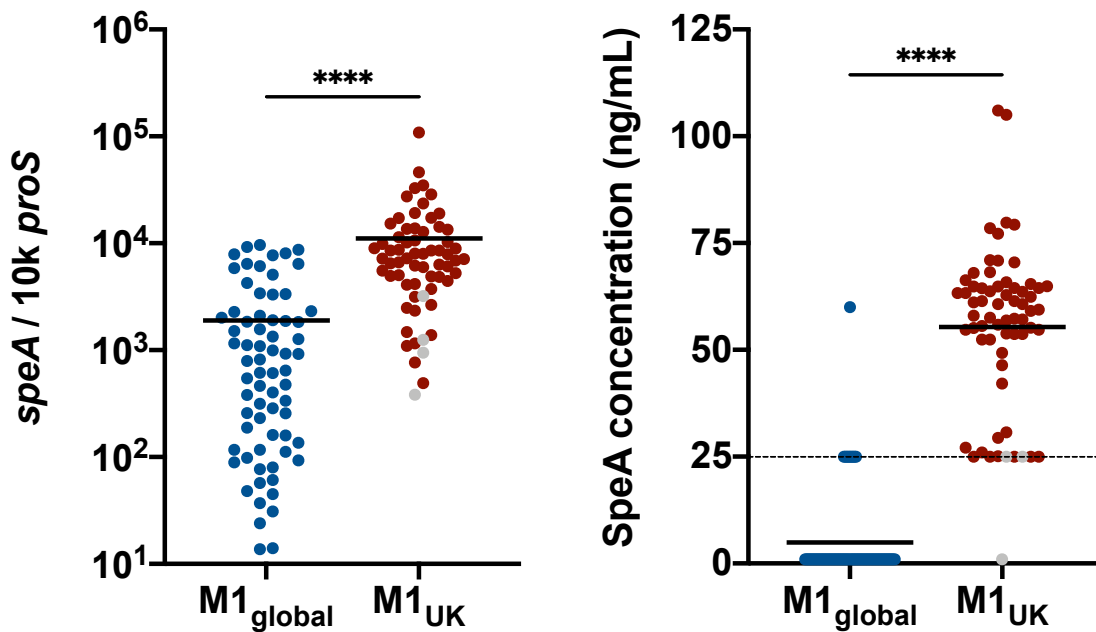
Supplementary Table S4 Primers used for RT-qPCR

Excel Suppl. Files Proteomic comparisons
(supplied separately)

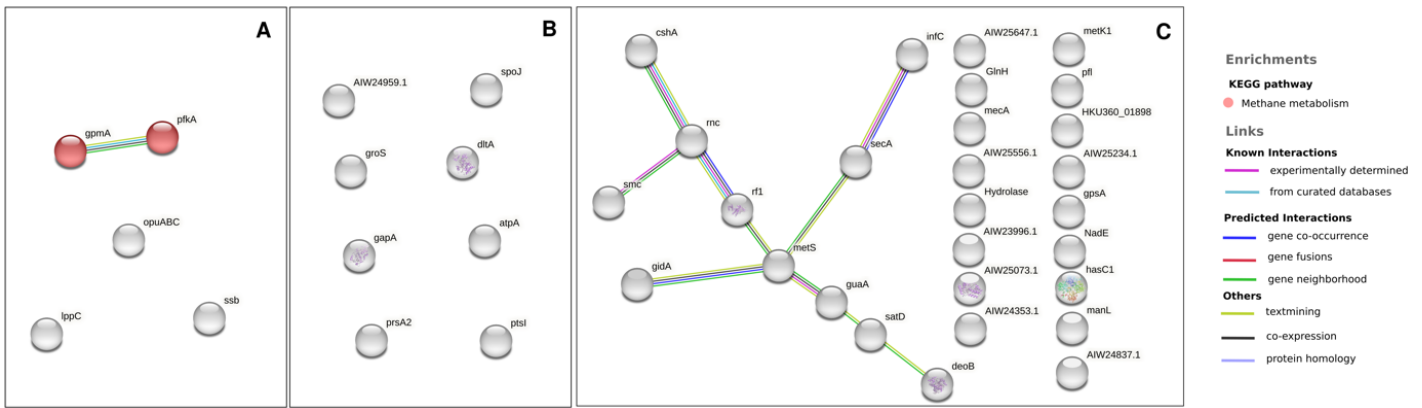
SUPPLEMENTARY FIGURES



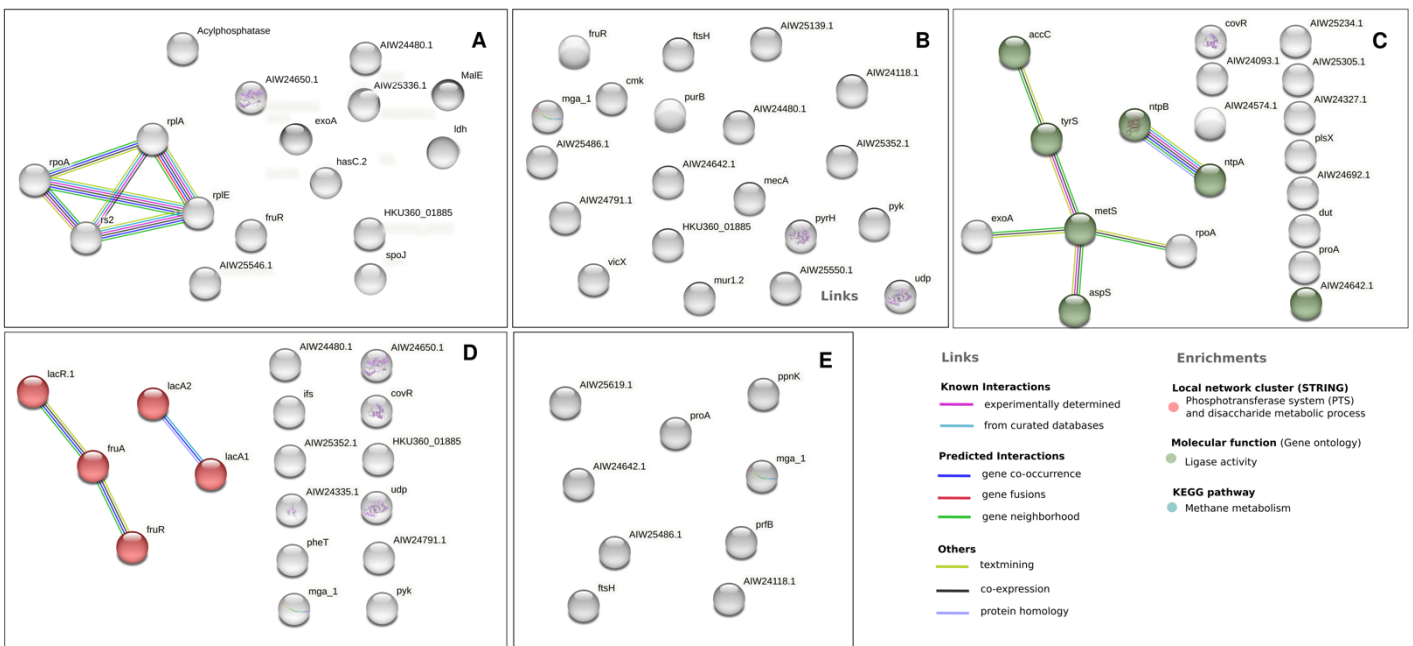
Supplementary Figure S1. Quantitative real-time PCR measurement of *glpF2* transcription in n=3 strains of *S. pyogenes* per lineage. Individual data points represent the average of 3 technical replicates for each strain. Strains used were: M1_{global} (blue bar) (BHS0162; BHS0130; BHS0674) and M1_{UK} (red bar) (BHS0258; BHS0170; BHS0128)



Supplementary Figure S2. Measurement of SpeA production by non-invasive *emm1* *S. pyogenes* isolates, (A), Quantitative real-time PCR measurement of *speA* transcription in M1_{global} and M1_{UK} strains of *S. pyogenes* per lineage (previously published in Lynskey N. & Jauneikaite E. et al, Lancet Infect Dis, 2019 Nov; 19(11): 1209-1218, shown for reference only). (B) Semi quantitative measurement of SpeA expression by western blotting. N=135 strains. Solid line represents the mean, dashed line represents the limit of detection. Grey dots indicate an intermediate sublineage of M1. Unpaired T-test, **** is p<0.0001.



Supplementary Figure S3. Enrichment and protein-protein interactions of differentially expressed proteins (p -value < 0.05) in different bacterial fractions using String database. Coloured lines represent physical associations as described in the legend above. Coloured bubbles represent the type of enrichment observed. Data compare M1_{UK} and M1_{global} Supernatants (A); Cell wall extracts (B); Cytosol extracts (C) obtained from *S. pyogenes* strains cultured in chemically defined medium (CDM). (N=5 different strains per group).



Supplementary Figure S4. Enrichment and protein-protein interactions of differentially expressed proteins (cytosol only, p -value < 0.05) using String database comparing strains from four sublineages M1_{UK}; M1_{23SNPs}; M1_{13SNPs}; and M1_{global} cultured in CDM (N=5 per group). Coloured lines represent physical associations as described in the legend above. Coloured bubbles represent the type of enrichment observed. Associations shown represent: (A) All 4 groups compared, (B) M1_{global} vs M1_{UK}, (C) [M1_{UK}+M1_{23SNPs}] vs [M1_{13SNPs} + M1_{global}], (D) M1_{global} vs all others, (E) M1_{UK} vs all others.

SUPPLEMENTARY TABLES

Supplementary Table S1. Bacterial strains used in RNA sequencing and RT-PCR

	Accession¶	Strain identifier	Lineage	Year	RNAseq	gldA operon	glpF2
M1_{global}	ERS1020620	BHS0674	M1 _{global}	2015	Y	Y	Y
	ERS1020158	BHS0162	M1 _{global}	2012	Y	Y	Y
	ERS1020095	BHS0130	M1 _{global}	2011	Y	Y	Y
	ERS1020136	BHS0151 or H1488	M1 _{global}	2012	Y		
	ERS1020472	BHS0503	M1 _{global}	2010		Y	
	ERS1020045	BHS0013	M1 _{global}	2009		Y	
	ERS1020385	BHS0368	M1 _{global}	2009		Y	
	ERS1020523	BHS0448	M1 _{global}	2009		Y	
M1_{UK}	ERS1020174	BHS0170	M1 _{UK}	2012	Y	Y	Y
	ERS1020090	BHS0128	M1 _{UK}	2011	Y	Y	Y
	ERS1020341	BHS0258	M1 _{UK}	2014	Y	Y	Y
	ERS1020603	BHS0581	M1 _{UK}	2015	Y		
	ERS1020508	BHS0521	M1 _{UK}	2015		Y	
	ERS1020714	BHS0643	M1 _{UK}	2015		Y	
	ERS1463088	BHS0762	M1 _{UK}	2016		Y	

¶Accession numbers relate to genome sequences deposited in the European nucleotide archive

Supplementary Table S2. *S. pyogenes* strains and genome sequences used to create phylogenetic tree (Figure 2) and used to quantify SpeA expression by non-invasive and invasive strains.

ENA ^{††}	STUDY	DISEASE	LINEAGE	LOCATION	YEAR
ERR2864947	Sharma <i>et al.</i>	Unknown	M1global	UK	2017
ERR2864949	Sharma <i>et al.</i>	Unknown	M1global	UK	2017
ERR2864951	Sharma <i>et al.</i>	Unknown	M1global	UK	2017
ERR2864964	Sharma <i>et al.</i>	Unknown	M1global	UK	2017
ERR2864950	Sharma <i>et al.</i>	Outbreak	23 SNPs	UK	2017
ERR2864953	Sharma <i>et al.</i>	Outbreak	23 SNPs	UK	2017
ERR2864957	Sharma <i>et al.</i>	Outbreak	22 SNPs	UK	2017
ERR2864962	Sharma <i>et al.</i>	Outbreak	23 SNPs	UK	2017
ERR2864966	Sharma <i>et al.</i>	Outbreak	23 SNPs	UK	2017
ERR2864969	Sharma <i>et al.</i>	Outbreak	23 SNPs	UK	2017
ERR2864948	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864952	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864954	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864955	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864956	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864958	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864959	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864960	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864961	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864963	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864965	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864967	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERR2864968	Sharma <i>et al.</i>	Unknown	M1uk	UK	2017
ERS4267588	This Study	Outbreak	26 SNPs	UK	2018
ERS4267589	This Study	Outbreak	27 SNPs	UK	2018
CP000017.2	Sumby <i>et al.</i>	Reference	M1global	x	x
ERS1448799	Kapatai <i>et al.</i>	Invasive	13 SNPs	UK	2014
ERS1450651	Kapatai <i>et al.</i>	Invasive	13 SNPs	UK	2015
ERS1450822	Kapatai <i>et al.</i>	Invasive	13 SNPs	UK	2015
ERS1448193	Kapatai <i>et al.</i>	Invasive	23 SNPs	UK	2014
ERS1450815	Kapatai <i>et al.</i>	Invasive	23 SNPs	UK	2015
ERS1448173	Kapatai <i>et al.</i>	Invasive	M1global	UK	2014
ERS1448879	Kapatai <i>et al.</i>	Invasive	M1global	UK	2014
ERS1450607	Kapatai <i>et al.</i>	Invasive	M1global	UK	2015
ERS1450839	Kapatai <i>et al.</i>	Invasive	M1global	UK	2015
ERS1448481	Kapatai <i>et al.</i>	Invasive	M1uk	UK	2014
ERS1449006	Kapatai <i>et al.</i>	Invasive	M1uk	UK	2014
ERS1450390	Kapatai <i>et al.</i>	Invasive	M1uk	UK	2015
ERS1450879	Kapatai <i>et al.</i>	Invasive	M1uk	UK	2015
ERS1594714	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2013
ERS1594852	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2013
ERS1594863	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2013
ERS1594843	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2013
ERS1594990	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2016
ERS1594904	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2016
ERS1594914	Lynskey & Jauneikaite <i>et al.</i>	Invasive	13 SNPs	UK	2016
ERS1594824	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2013
ERS1594734	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2013
ERS1594744	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2013
ERS1594757	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2013
ERS1594864	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2013
ERS1594882	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2013
ERS1594950	Lynskey & Jauneikaite <i>et al.</i>	Invasive	23 SNPs	UK	2016

ERS362145	Turner <i>et al.</i>	Invasive	M1global	UK	2011
ERS362146	Turner <i>et al.</i>	Invasive	M1uk	UK	2011
ERS362147	Turner <i>et al.</i>	Invasive	M1uk	UK	2011
ERS362165	Turner <i>et al.</i>	Invasive	M1global	UK	2011
ERS362031	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362033	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362040	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362047	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362051	Turner <i>et al.</i>	Invasive	M1global	UK	2006
ERS362058	Turner <i>et al.</i>	Invasive	M1global	UK	2006
ERS362063	Turner <i>et al.</i>	Invasive	M1global	UK	2006
ERS362066	Turner <i>et al.</i>	Invasive	M1global	UK	2006
ERS362067	Turner <i>et al.</i>	Invasive	M1global	UK	2006
ERS362071	Turner <i>et al.</i>	Invasive	13 SNPs	UK	2006
ERS362110	Turner <i>et al.</i>	Invasive	M1global	UK	2003
ERS362122	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362029	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS361855	Turner <i>et al.</i>	Invasive	M1global	UK	2001
ERS361867	Turner <i>et al.</i>	Invasive	M1global	UK	2002
ERS361913	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361937	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361941	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361950	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361954	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361960	Turner <i>et al.</i>	Invasive	13 SNPs	UK	2005
ERS361966	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361929	Turner <i>et al.</i>	Invasive	M1global	UK	2004
ERS361826	Turner <i>et al.</i>	Invasive	M1global	UK	2001
ERS361827	Turner <i>et al.</i>	Invasive	M1global	UK	2001
ERS361841	Turner <i>et al.</i>	Invasive	M1global	UK	2001
ERS361865	Turner <i>et al.</i>	Invasive	M1global	UK	2002
ERS362030	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362032	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362049	Turner <i>et al.</i>	Invasive	M1global	UK	2005
ERS362078	Turner <i>et al.</i>	Invasive	M1global	UK	2006
ERS379360	Turner <i>et al.</i>	Invasive	M1global	UK	2006

¶ Non-invasive strains that were used in SpeA quantification (Supplementary figure 2) are highlighted by grey shading.

† Invasive strains representing four sublineages used for SpeA quantification are highlighted in blue shading

Supplementary Table S3. Components of chemically defined media (CDM) –

COMPONENT ¹	CONSTITUENT	STOCK CONC	VOLUME IN 500mL CDM
Water	n/a	n/a	350
Iron	Fe(NO ₃) ₃ .9H ₂ O - 50mg	1000X	0.5
	F ₂ SO ₄ .7H ₂ O - 250mg		
	made up to 50mL ddH ₂ O		
Phosphate	K ₂ HPO ₄ - 5g	25X	20
	KH ₂ PO ₄ - 25g		
	NaH ₂ PO ₄ .H ₂ O - 79.9g		
	Na ₂ HPO ₄ .7H ₂ O - 346.8g		
	made up to 1L ddH ₂ O		
MgSO ₄ .7H ₂ O	MgSO ₄ .7H ₂ O - 87.5g	500X	1
	made up to 250mL ddH ₂ O		
MnSO ₄ .H ₂ O	MnSO ₄ .H ₂ O - 1.25g	1000X	0.5
	made up to 250mL ddH ₂ O		
Sodium acetate	NaC ₂ H ₃ O ₂ .3H ₂ O - 112.5g	50X	10
	made up to 500mL ddH ₂ O		
Calcium chloride	CaCl ₂ .2H ₂ O - 338mg	1000X	0.5
	made up to 50mL ddH ₂ O		
Sodium bicarbonate	NaHCO ₃ - 50g	20X	25
	made up to 1L dd H ₂ O		
L-Cysteine HCl	Cysteine HCl - 16.25mg	500X	1
	made up to 50mL ddH ₂ O		
Bases	2N HCl - 50mL	100X	5
	adenine - 0.5g		
	guanine hydrochloride - 0.5g		
	uracil - 0.5g		
	ddH ₂ O - 200ml		
Vitamins	p-aminobenzoic acid - 10mg	1000X	0.5
	biotin - 10mg		
	folic acid - 40mg		
	nicotinamide - 50mg		
	b-nicotinamide adenine dinucleotide phosphate - 125mg		
	pantothenate calcium salt - 100mg		
	pyridoxal - 50mg		
	pyridoxamine dihydrochloride - 50mg		
	riboflavin - 100mg		
	thiamine hydrochloride - 50mg		
	vitamin B ₁₂ - 5mg		
	dissolve in 50mL ddH ₂ O		
Amino acids	DL- alanine - 2.5g	50X	10

	L-arginine - 2.5g		
	L-aspartic acid - 2.5g		
	L-asparagine - 2.5g		
	L-cystine 1.25g		
	L-glutamic acid - 2.5g		
	L-glutamine - 5.0g		
	Glycine - 2.5g		
	L-histidine - 2.5g		
	L-isoleucine - 2.5g		
	L-leucine - 2.5g		
	L-lysine - 2.5g		
	L-methionine - 2.5g		
	L-phenylalanine - 2.5g		
	L-proline - 2.5g		
	hydroxy-L-proline - 2.5g		
	L-serine - 2.5g		
	L-threonine - 5.0g		
	L-tryptophan - 2.5g		
	L-tyrosine - 2.5g		
	L-valine - 2.5g		
	made up to 500mL ddH2O		
Carbon source	e.g., glucose	n/a	5g

[¶]adapted from van de Rijn and Kessler, *Infect Immun*, 27(2):444

Supplementary Table S4. Primers used for RT-qPCR

GENE	PRIMER	SEQUENCE	PRODUCT (bp)
gldA	gldA_RTf	GCCTCAGATAATGAAATCAGCC	121
	gldA_RTR	CAAGTAAGTCAGCGATAGCC	
mipB	mipB_RTf	TAGGAGCACAAGCCATCAC	100
	mipB_RTR	TCCAATCCTTGCCAAAATC	
pflD	pflD_RTf	CAAAAACGGCTAAACCAGAAC	118
	pflD_RTR	ATGAACCAGACAGATTGCAC	
ptsIIC	PTS_RTf	GCAAAACATCATCAAGCCAATC	121
	PTS_RTR	CCCAGCAATCAGGAAAAGAC	
glpF2	glpF_RTf	GCTATGGCTTAGGAGTTATG	120
	glpF_RTR	AGACGTGAGCCCATGGGAAC	
speA	speA_Fwd	GAGGGGTAACAAATCATGAAGG	94
	speA_Rv	TCAAATGATAGGCTTTGGATACC	
proS	proS_Fwd	TGAGTTTATTATGAAAGACGGCTATAGTTTC	93
	proS_Rv	AATAGCTTCGTAAGCTTGACGATAATC	