Table S1. Bacterial strains and plasmids

Strain or plasmid	Description	Reference
S. pyogenes		
MGAS2221	Wild-type, M1-serotype strain	(1)
SF370	Wild-type, M1-serotype strain	(2)
JRS4-PolHis	The JRS4 strain has the plasmid integrated into the chromosome by a single cross-over event at the *rpoC* gene which encodes the β subunit of RNA polymerase, creating a His_6 tag at its C-terminal end.	(3)
MGAS8232	Wild-type, M18-serotype strain	(4)
87-282	Wild-type, M18-serotype strain	(5)
Manfredo	Wild-type, M5-serotype strain	(6)
Vaughn	Wild-type, M5-serotype strain	(5)
Alab49	Wild-type, M53-serotype strain	(7)
NZ131	Wild-type, M49-serotype strain	(8)
Bisno9	M53-serotype clinical isolate	This study
Bisno21	M3-serotype clinical isolate	This study
2221Δ*covR*	*covR* deletion mutant of the MGAS2221strain	(9)
2221ΔP2Δ*hasS*	The P2 promoter and the entire *hasS* sequence are deleted in MGAS2221	This study
2221ΔP2	The P2 promoter is deleted in MGAS2221	This study
2221P1$^-$	The –10 box of the P1 promoter is inactivated by mutations in MGAS2221	This study
2221ΔP2P1$^-$	In MGAS2221 P2 promoter is deleted and the –10 box of the P1 promoter is inactivated by mutations	This study
2221Δ42T	MGAS2221 has a 42-nt deletion which includes the *hasS* terminator sequence and the T-rich region located immediately downstream	This study
2221Δ63T	MGAS2221 has a 63-nt deletion which includes the *hasS* terminator sequence, the T-rich region and 25 nts located immediately downstream	This study
2221Δ*covR*ΔP2Δ*hasS*	The P2 promoter and the entire *hasS* sequence are deleted in 2221Δ*cov*R	This study
2221Δ*covR*Δ42T	2221Δ*covR* has a 42-nt deletion which includes the *hasS* terminator sequence and the T-rich region located immediately downstream	This study
2221Δ*covR*P1$^-$	In 2221Δ*covR* the –10 box of the P1 promoter is inactivated by mutations	This study
2221Δ*covR*Δ42TP1$^-$	2221Δ*cov*R has mutations in the –10 box of the P1 promoter and a 42-nt deletion which includes the *hasS* terminator sequence and the T-rich region located immediately downstream	This study
2221Δ*covR*ΔP2P1$^-$	In 2221Δ*cov*R the P2 promoter is deleted and the –10 box of the P1 promoter is inactivated by mutations	This study
2221/pKSM*hasS*	The MGAS2221 strain carrying the pBBL740*hasS* plasmid	This study
SF370Δ63T[a]	The SF370 strain has a 63-nt deletion which includes the *hasS* terminator sequence and the T-rich region located immediately downstream	This study
E. coli		
DH5α	Cloning host	Invitrogen
Rosetta (DE3)	Protein expression host	Novagen

Plasmids		
pKSM720	*E. coli* - *Streptococcus* vector containing a promoterless luciferase (*luc*) gene with its ribosomal-binding site	(10)
pBBL740	*S. pyogenes* integrational plasmid	(11)
pKSM720P2	pKSM720 containing transcriptional fusion of P2 with *luc*	This study
pKSM720P2T	pKSM720 containing transcriptional fusion of *luc* with a 0.184 kb DNA fragment harboring P2, *hasS* and the terminator	This study
pKSM*hasS*	pKSM720 containing a 305 bp DNA fragment carrying the P2 promoter and *hasS*	This study
pBBL740*hasS*	pBBL740 containing a 1.4 kbp DNA fragment carrying *hasS* locus and its flanking either side regions	This study
pBBL740ΔP2	pBBL740*hasS* derived plasmid containing a deletion of the P2 promoter	This study
pBBL740P1	pBBL740*hasS* derived plasmid containing mutations in the –10 box of the P1 promoter	This study
pBBL740ΔP2P1	pBBL740*hasS* derived plasmid containing a deletion of the P2 promoter and mutations in the –10 box of the P1 promoter	This study
pBBL740ΔP2Δ*hasS*	pBBL740*hasS* derived plasmid containing a deletion of the P2 promoter and the entire *hasS* sequence including the terminator	This study
pBBL740Δ42TP1	pBBL740*hasS* derived plasmid containing mutations in the –10 box of the P1 promoter and a 42-nt deletion which includes the *hasS* terminator sequence and the T-rich region located immediately downstream	This study
pBBL740Δ42T	pBBL740*hasS* derived plasmid containing a 42-nt deletion which includes the *hasS* terminator sequence and the T-rich region located immediately downstream	This study
pBBL740Δ63T	pBBL740*hasS* derived plasmid containing a 63-nt deletion which includes the *hasS* terminator sequence, the T-rich region and 25 nts located immediately downstream	This study
pET21d	*E. coli* protein expression vector harboring T7 promoter	Novagen
pETCovR	pET21d derived plasmid containing *covR* with TEV cleavage site followed by His_6 tag at its C-terminus	This study
pETCovRD53E	pETCovR derived plasmid containing a D53E mutation in CovR	This study
pEU7534	pET15b derived vector containing *rpoD*	(12)

[a] - The whole genome of the mutant was sequenced (Virginia Bioinformatics Institute at Virginia Tech). We found that the strain has a deletion in the *spyB* gene in addition to a deletion of the terminator. The *spyB* deletion was rescued through insertion of the WT copy of *spyB* in the mutant chromosome. The "rescued" strain had the phenotype similar to the SF370Δ63T mutant indicating that the phenotype of the mutant is due to a deletion of the terminator.

Table S2. General primers

Primer	Sequence[a]
hasR-BamHI-f	GATCGGATCCGTTCTGGGCTACCATATTCTGAC
hasR-XhoI-r	CAGTCTCGAGCTGTTGCAGCATAAACTGTCTCATC
P2-BglII-f	CGTGAGATCTCTACAATATATCCTTTACCAG
P2-XhoI-r	CAGTCTCGAGTATCTCAAAATAATTGAATTG
term-XhoI-r	ACGTCTCGAGCCCCCAAAAAAAGAAAAATCAGG
RPA-f	GCGGTTATTTTGAGATAATTTATATTAA
RPA-r	TTCCCCCAAAAAAAGAAAAATCAGG
hasART-f	CGCGGAGTAATATTAATAACCTATC
hasS-f	CCGTGAGTCTCCTTTACCAGTTATCATATTTCTTG
hasS-r	CCGTGAGTCTGGCACAATTACACCTCTTTC
hasART-r	CAGCAGCAACTTTATAGTCATG
plr-f	CGTCTTGCATTCCGCCGTATTC
plr-r	GATCTGTAAGGTCATTGATACG
hasSRT-f	CCGCGCCGCAATTCAATTATTTTGAGATAAT
hasSRT-r	GCGCGCGGAAAAGAAAAATCAGGATAATTTATT
1	GCGGCTTTTTTTGACTTAATTGTCGGG
2	CCATTTTTATTCAGGTTAGGTG
covR-NcoI-f	CGCACCATGGCCACAAAGAAAATTTTAATTATTGAAGATG
covR-XbaI-r	GCGCTCTAGATTATTTCTCACGAATAACGTATCCC
RNA-has-r1	GCGCGTTCCCCCAAAAAAAGAAAAATCAGG
mga-f	TGCGTGATCGTGACTGTTTTTTG
mga-r	GTTCTCTCCACTGTTGACTTGTAAAC
F1	CAAGGAAATTAAAAAAGAAAG
F2	TTATTAAAAATATTTCTATGACTAGTTGAC
F3	GATAAAACTATTTGAAAAGTTTGC
F4	CTACAATATATCCTTTACCAGTTATCATATTTCTTG
R	CCATAAATTCCTACAGTTGATGTTCC
rtqRT-f	GCGCGCGTTATTAAAAATATTTCTATGAC
rtqRT-r	GACAGGGACCTCGATATTC
P1-f	GCGCCAATTCAATTATTTTGAG
P1-r	CACGTCTCGAGCATAAATTCCTACAGTTGATGTTC

[a] - Restriction sites are underlined.

Table S3. Primers for site-directed mutagenesis

primer	sequence	Introduced mutations
ΔhasS-f	CACAAAATTAGATAAAACTATTTGGGGGAAATTTTTTAATTTAAATAC	The P2 promoter and the entire *hasS* sequence are deleted
ΔhasS-r	GTATTTAAATTAAAAAATTTCCCCCAAATAGTTTTATCTAATTTTGTG	
ΔP2-f	CACAAAATTAGATAAAACTATTTGTAACAATTCAATTATTTTGAG	The P2 promoter is deleted
ΔP2-r	CTCAAAATAATTGAATTGTTACAAATAGTTTTATCTAATTTTGTG	
mP1-f	GACATTACCCCTTATTTATATTCTAGAATCGAGGTCCCTGTCTTTCAAGG	The –10 box of the P1 promoter is mutated
mP1-r	CCTTGAAAGACAGGGACCTCGATTCTAGAATATAAATAAGGGGTAATGTC	
Δ42T-f	CAATTCAATTATTTTGAGGGGGAAATTTTTTAATTTAAATAC	a 42-nt deletion which includes the *hasS* terminator and the T-rich region
Δ42T-r	GTATTTAAATTAAAAAATTTCCCCCTCAAAATAATTGAATTG	
Δ63T-f	CAATTCAATTATTTTGAGATAATTTTTATTAAAAATATTTCTATGAC	a 63-nt deletion which includes the *hasS* terminator, the T-rich region and 25 nts located immediately downstream
Δ63T-r	GTCATAGAAATATTTTTAATAAAAATTATCTCAAAATAATTGAATTG	
CovR-D53E-f	GATTTAATCCTGCTTGAATTAATGTTACCAGAGATG	a D53E mutation in CovR
CovR-D53E-r	CATCTCTGGTAACATTAATTCAAGCAGGATTAAATC	

Table S4. Genetic diversity in the *hasS* region.

GAS strain	serotype	*hasABC* locus	*hasS* terminator				P2 promoter mutation	IS*1239*	Repeats in *hasS*
			42 bp stem-loop	34 bp stem-loop	25 bp stem-loop	no stem-loop			
Strains analyzed in this study									
MGAS2221[a]	M1	+[b]	–[c]	–	+	–	–	–	–
SF370[a]	M1	+	–	–	+	–	–	–	–
Bisno21	M3	+	–	–	–	+	+	–	–
Manfredo[a]	M5	+	–	–	–	+	+	+	–
Vaughn	M5	+	–	–	–	+	+	+	–
MGAS8232[a]	M18	+	–	–	–	+	–	–	–
87-282	M18	+	–	–	–	+	–	–	–
NZ131[a]	M49	+	–	+	–	–	–	–	–
Alab49[a]	M53	+	+	–	–	–	+	–	–
Bisno9	M53	+	+	–	–	–	+	–	–
NCBI's complete genomes									
M1 GAS	M1	+	–	–	+	–	–	–	–
MGAS5005	M1	+	–	–	+	–	–	–	–
M1 476	M1	+	–	–	+	–	–	–	–
A20	M1	+	–	–	+	–	–	–	–
MGAS10270	M2	+	–	–	–	+	–	–	–
SSI-1	M3	+	–	–	–	+	+	–	–
MGAS315	M3	+	–	–	–	+	+	–	–
MGAS10750	M4	–	–	–	–	–	–	–	–
MGAS10394	M6	+	+	–	–	–	+	+	+
MGAS9429	M12	+	–	+	–	–	–	–	–
MGAS2096	M12	+	–	+	–	–	–	–	–
HKU QMN11M0907901	M12	+	–	+	–	–	–	–	–
HSC5	M14	+	–	–	–	+	+	–	–
MGAS6180	M28	+	–	+	–	–	–	–	–
MGAS1881	M59	+	–	+	–	–	–	–	–
MGAS15252	M59	+	–	+	–	–	–	–	–
NCBI's whole genome shotgun sequencing [d]									
GA41345	M1	+	–	–	+	–	–	–	–
DSM 20565	M1	+	–	–	+	–	–	–	–
GA40634	M4	–	–	–	–	–	–	–	–
GA19700	M6	+	+	–	–	–	+	+	+
GA19681	M6	+	+	–	–	–	+	+	+
BJCYGAS15	M12	+	–	–	–	+	–	–	–
HLJGAS12011	M12	+	–	–	–	+	–	–	–
SP1-LAU	M12	+	–	+	–	–	–	–	–
GA41394	M18	+	–	–	–	+	–	–	–
ATCC 10782	M24	+	–	+	–	–	–	–	–
GA40377	M28	+	–	+	–	–	–	–	–
06BA18369	M41.2	+	–	+	–	–	+	–	–
M49 591	M49	+	–	+	–	–	–	–	–
GA40884	M59	+	–	+	–	–	–	–	–
GA03747	M66	+	–	+	–	–	–	–	–
GA03455	n.a.[e]	+	+	–	–	–	+	–	–
GA03805	n.a.	+	+	–	–	–	+	–	–
GA41039	n.a.	+	+	–	–	–	+	–	–
UTMEM-1	n.a.	+	–	+	–	–	–	–	–
UTSW-2	n.a.	+	–	+	–	–	–	–	–
GA16797	n.a.	+	–	+	–	–	–	–	–
GA03799	n.a.	+	–	+	–	–	–	–	–

GA40056	n.a.	+	–	+	–	–	–	–	–
GA19702	n.a.	+	–	+	–	–	–	–	–
GA41046	n.a.	+	–	–	–	+	–	–	–

[a] – complete genomes; [b] – present; [c] – absent; [d] - GAS genomes with the partial sequence of the *hasS* locus were not selected for analysis; [e] - serotype information is not available.

Figure S1

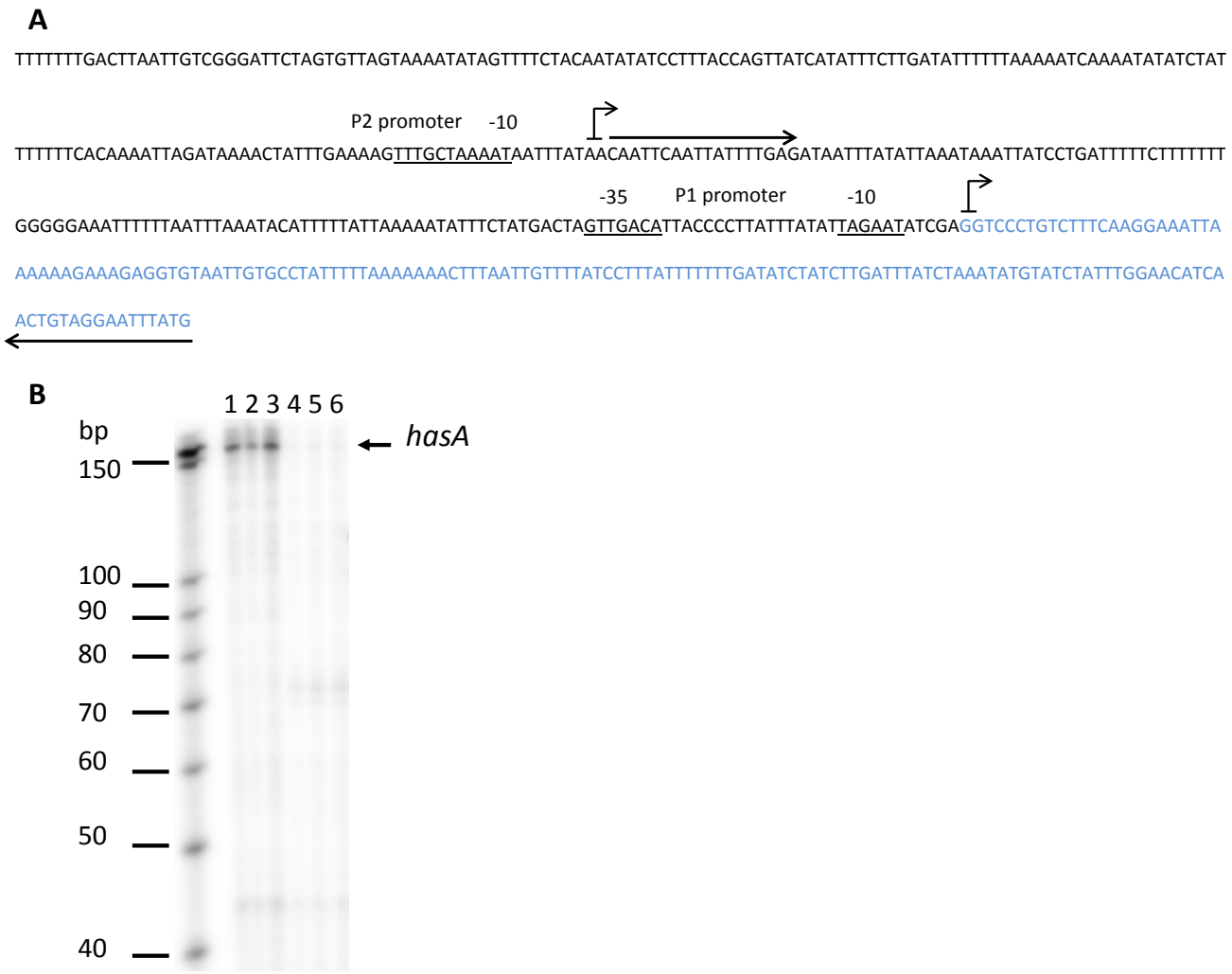

Fig. S1. The *hasA* upstream region encodes the P1 promoter. **A**. Nucleotide sequence of the *hasA* upstream region in GAS M1-serotype. –35, –10 and extended −10 promoter regions are underlined. Bent arrows denote the transcriptional start sites of *hasS* identified by 5' RACE analysis and *hasA* according to the published results of primer extension analysis (13). Horizontal arrows indicate the location of the primers, P1-f and P1-r (Table S2), used for *in vitro* transcription of the P1 promoter region. The transcriptional product is shown in blue font. **B.** Time-course of *in vitro* RNA synthesis from a PCR product carrying the P1 promoter. Shown is a 10% denaturing urea polyacrylamide gel containing *in vitro* transcribed RNA products using GAS RNA polymerase and σ^{A}. DNA template for transcriptional reactions loaded in the lanes 1, 2 and 3 was PCR amplified from MGAS2221 genomic DNA using the primers pair of P1-f and P1-r. DNA template for transcriptional reactions loaded in the lanes 4, 5 and 6 was PCR amplified from 2221P1^{-} strain genomic DNA using the same primer pair. Lane 1 and 4, 30 min; lane 2 and 5, 60 min; lane 3 and 6, 90 min. The RNA ladder is labeled in bases.

Figure S2

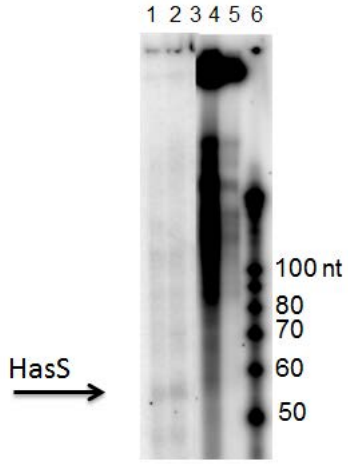

Fig. S2. RPA identifying transcript in the intergenic region upstream of *hasA* and downstream of M5005_Spy_1850. Lane 1, RPA with RNA isolated from 2221Δ*covR*; Lane 2, RPA with RNA isolated from 2221/pKSM*hasS*; Lane 3, RPA with RNA isolated from MGAS2221; Lane 4, no-RNAse control; Lane 5, the probe, Lane 6, RNA ladder

Figure S3

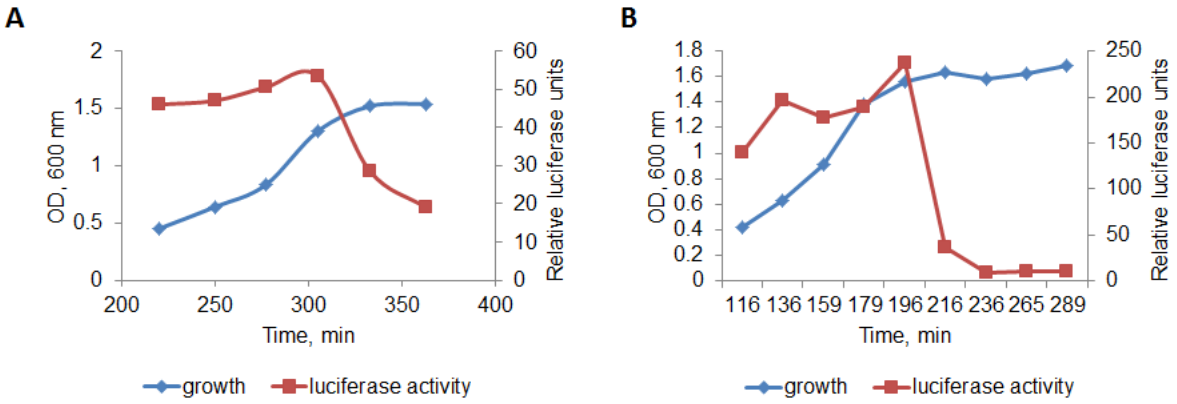

Fig. S3. Transcriptional analysis of the P2 promoter. Time courses of luciferase activity was carried out on *S. pyogenes* strains harboring the P2::*luc* fusion. Panel **A** shows the results obtained with the WT strain and panel **B** shows the results ontained with the 2221Δ*covR* mutant. The results of one representative experiment out of three independent experiments are shown.

Figure S4

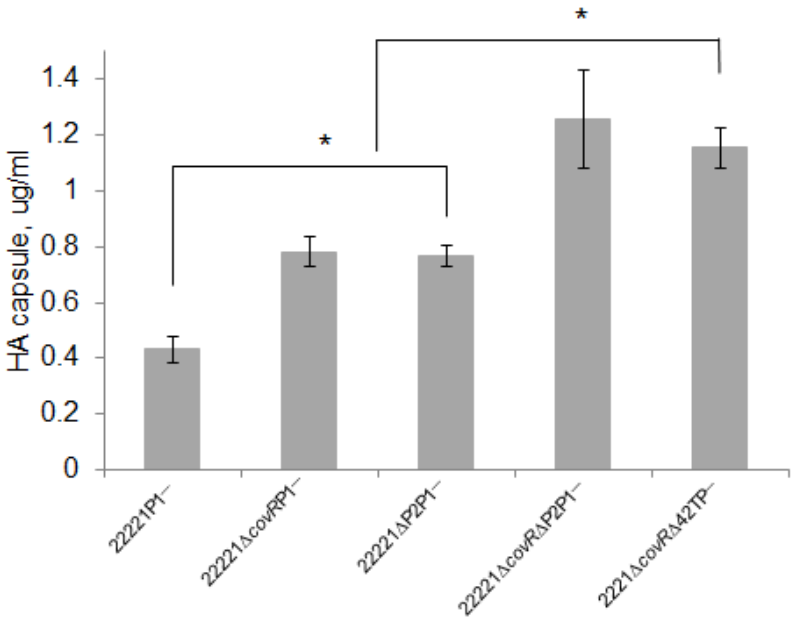

Fig. S4. The RNA product from *hasS* region negatively regulates capsule expression. The strains harvested at the mid-exponential growth phase. The experiments were performed in triplicate, and mean values ± standard deviations are shown. * denotes statistically significant differences calculated using the Student's test.

Figure S5

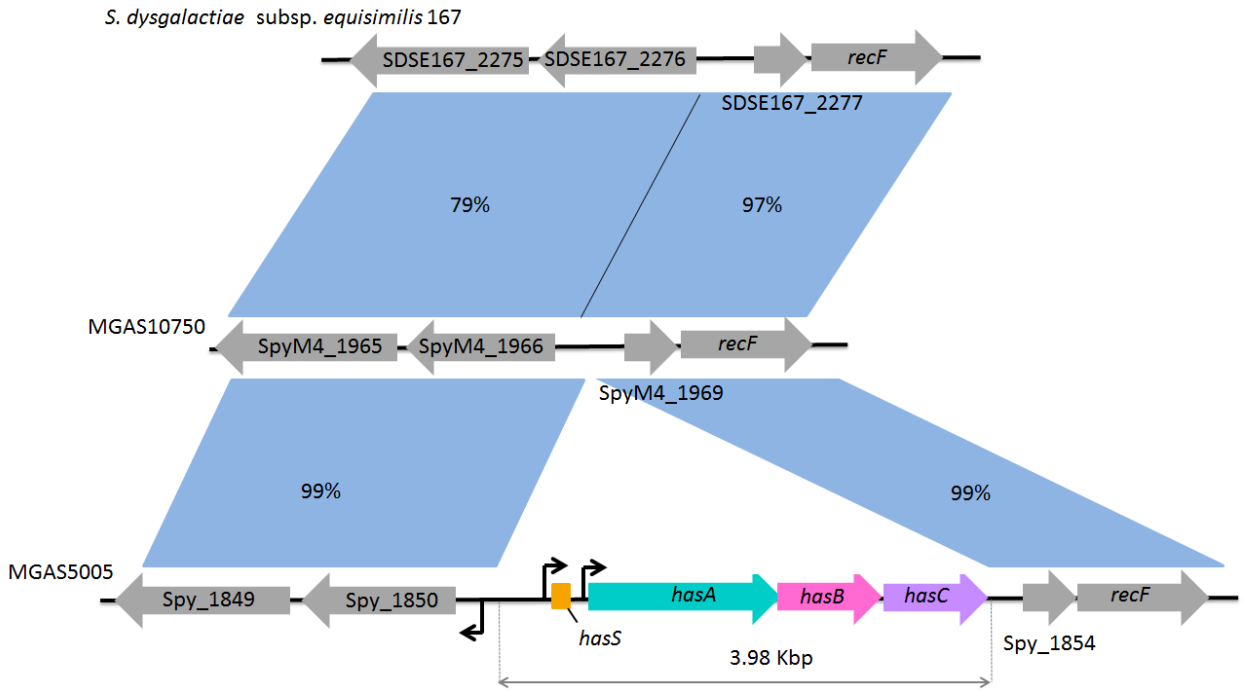

Fig. S5. Comparative genome alignment of the *hasABC* operon and the flanking DNA regions from *S. dysgalactiae* subsp. *equisimilis* 167 and *S. pyogenes* strains — MGAS5005 (M1-serotype) and MGAS10750 (M4-serotype). Bent arrows indicate the promoters in the *hasA* upstream region. Blue bars separating each genome represent similarity matches.

Figure S6.

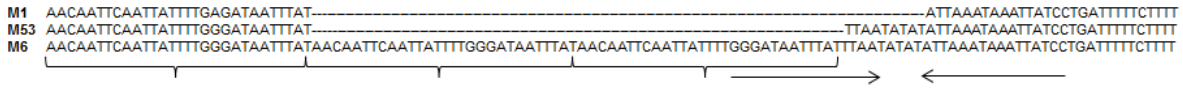

Fig. S6. Sequence alignment of *hasS* from serotypes M1, M53 and M6. The terminator hairpin is marked by inverted arrows. Brackets denote DNA repeats in *hasS* from serotype M6.

References

1. **Sumby P, Porcella SF, Madrigal AG, Barbian KD, Virtaneva K, Ricklefs SM, Sturdevant DE, Graham MR, Vuopio-Varkila J, Hoe NP, Musser JM.** 2005. Evolutionary origin and emergence of a highly successful clone of serotype M1 group a Streptococcus involved multiple horizontal gene transfer events. J Infect Dis **192:**771-782.
2. **Ferretti JJ, McShan WM, Ajdic D, Savic DJ, Savic G, Lyon K, Primeaux C, Sezate S, Suvorov AN, Kenton S, Lai HS, Lin SP, Qian Y, Jia HG, Najar FZ, Ren Q, Zhu H, Song L, White J, Yuan X, Clifton SW, Roe BA, McLaughlin R.** 2001. Complete genome sequence of an M1 strain of Streptococcus pyogenes. Proceedings of the National Academy of Sciences of the United States of America **98:**4658-4663.
3. **Opdyke JA, Scott JR, Moran CP, Jr.** 2001. A secondary RNA polymerase sigma factor from Streptococcus pyogenes. Molecular microbiology **42:**495-502.
4. **Smoot JC, Barbian KD, Van Gompel JJ, Smoot LM, Chaussee MS, Sylva GL, Sturdevant DE, Ricklefs SM, Porcella SF, Parkins LD, Beres SB, Campbell DS, Smith TM, Zhang Q, Kapur V, Daly JA, Veasy LG, Musser JM.** 2002. Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks. Proceedings of the National Academy of Sciences of the United States of America **99:**4668-4673.
5. **Ashbaugh CD, Alberti S, Wessels MR.** 1998. Molecular analysis of the capsule gene region of group A Streptococcus: the hasAB genes are sufficient for capsule expression. Journal of bacteriology **180:**4955-4959.
6. **Holden MT, Scott A, Cherevach I, Chillingworth T, Churcher C, Cronin A, Dowd L, Feltwell T, Hamlin N, Holroyd S, Jagels K, Moule S, Mungall K, Quail MA, Price C, Rabbinowitsch E, Sharp S, Skelton J, Whitehead S, Barrell BG, Kehoe M, Parkhill J.** 2007. Complete genome of acute rheumatic fever-associated serotype M5 Streptococcus pyogenes strain manfredo. Journal of bacteriology **189:**1473-1477.
7. **Lizano S, Luo F, Tengra FK, Bessen DE.** 2008. Impact of orthologous gene replacement on the circuitry governing pilus gene transcription in streptococci. PLoS One **3:**e3450.
8. **McShan WM, Ferretti JJ, Karasawa T, Suvorov AN, Lin S, Qin B, Jia H, Kenton S, Najar F, Wu H, Scott J, Roe BA, Savic DJ.** 2008. Genome sequence of a nephritogenic and highly transformable M49 strain of Streptococcus pyogenes. Journal of bacteriology **190:**7773-7785.
9. **Trevino J, Perez N, Ramirez-Pena E, Liu Z, Shelburne SA, 3rd, Musser JM, Sumby P.** 2009. CovS simultaneously activates and inhibits the CovR-mediated repression of distinct subsets of group A Streptococcus virulence factor-encoding genes. Infection and immunity **77:**3141-3149.
10. **Kinkel TL, McIver KS.** 2008. CcpA-mediated repression of streptolysin S expression and virulence in the group A streptococcus. Infection and immunity **76:**3451-3463.
11. **Zhu H, Liu M, Sumby P, Lei B.** 2009. The secreted esterase of group a streptococcus is important for invasive skin infection and dissemination in mice. Infection and immunity **77:**5225-5232.
12. **Gusa AA, Scott JR.** 2005. The CovR response regulator of group A streptococcus (GAS) acts directly to repress its own promoter. Molecular microbiology **56:**1195-1207.
13. **Crater DL, van de Rijn I.** 1995. Hyaluronic acid synthesis operon (has) expression in group A streptococci. The Journal of biological chemistry **270:**18452-18458.