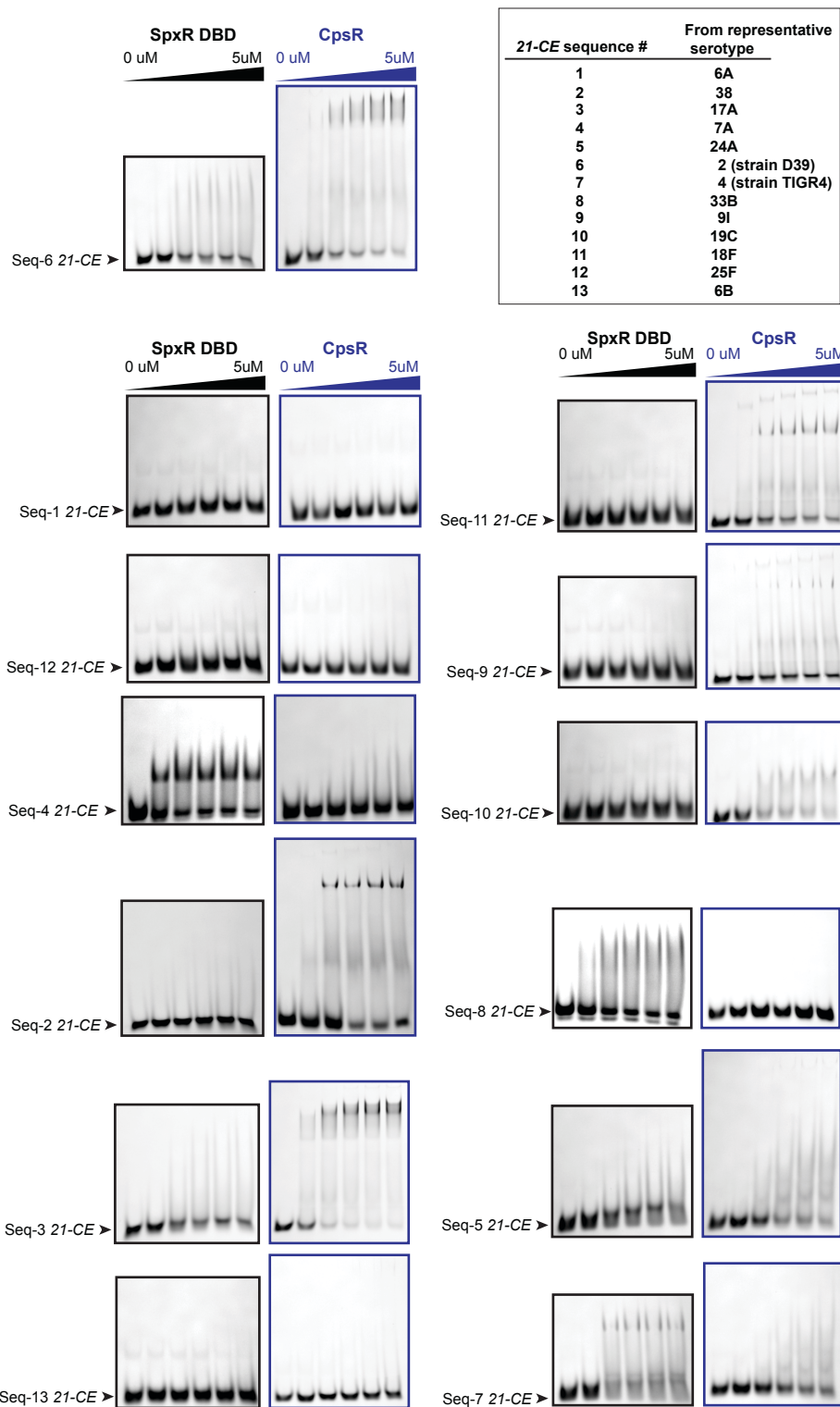
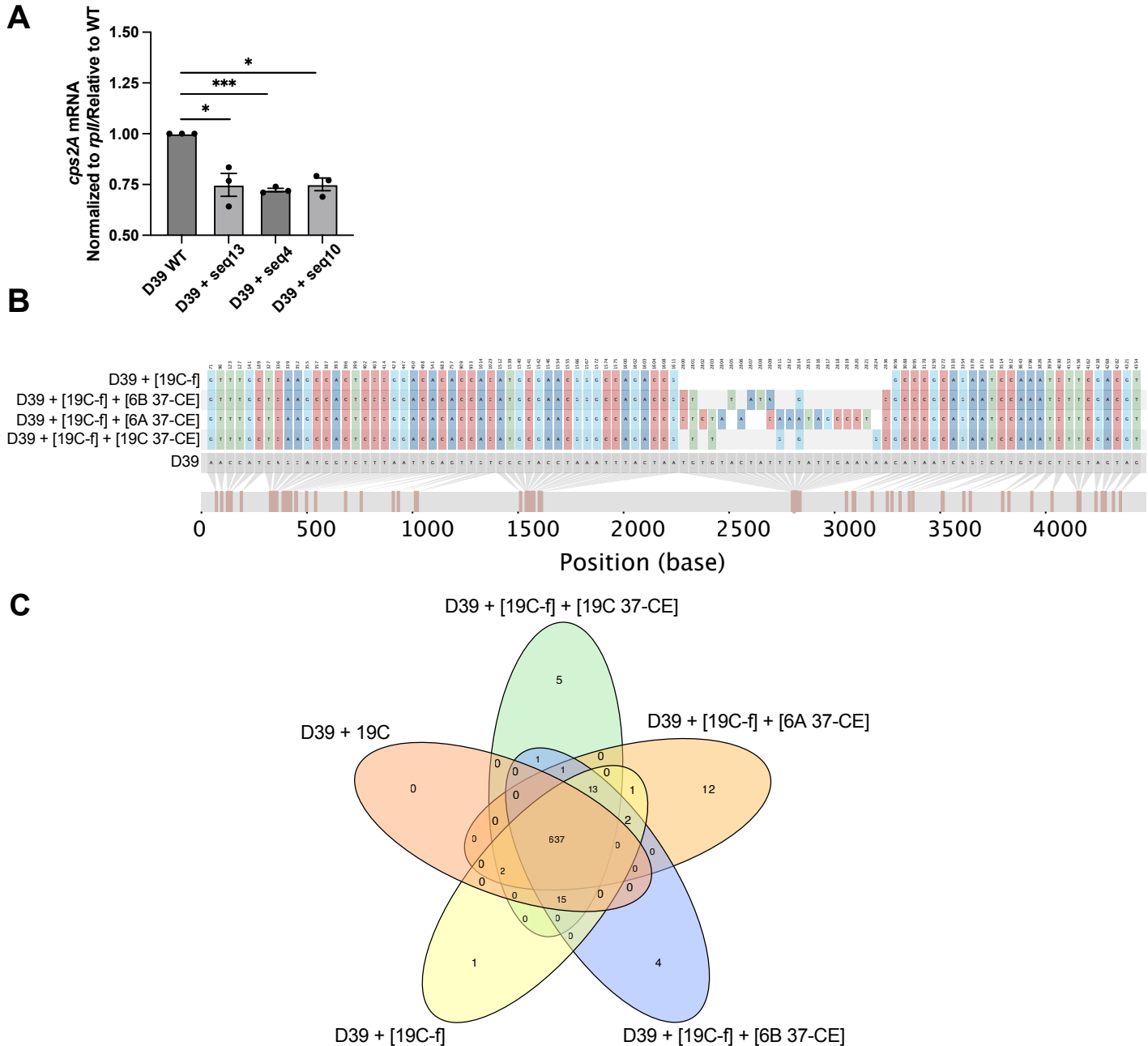


Supplementary Figure 2



Supplementary Figure 2. EMSA titrations of all 13 different 21-CE sequences with SpxR (DBD) and full-length CpsR protein. Increasing concentrations (0 to 5 μ M) of protein were added to fluorescent DNA probe and imaged after electrophoresis using a non-denaturing gel. Source data are provided as a Source Data file.

Supplementary Figure 3



Supplementary Figure 3. Comparison of mutations within the 37-CE switched strains.

(A) qRT-PCR results from aerated D39 WT (sequence #6), D39 + seq13, D39 + seq4 and D39 + seq10 37-CE strains. The mean and standard error of the mean (SEM) consisting of 3 biological replicates are plotted. Statistical differences were determined using a one sample t-test. *, $P < 0.05$; ***, $P < 0.001$. Source data are provided as a Source Data file. (B) Alignment and visualization of the 37-CE locus with all identified single nucleotide polymorphisms (SNPs) for the 4 switched strains relative to the parent D39 WT. Grey bar at the bottom represents the entire alignment region inclusive of *dexB* and *cpsA*, with orange vertical lines representing each SNP displayed, with their exact positions printed above the respective SNP. The 37-CE locus extends from positions 2,800-2,837. (C) Venn diagram of all detected structural variants (SNPs and insertions and deletions (Indels)) within switch strain genomes relative to the parent wild-type switch strain (D39).

Supplementary Figure 4

Serotype	Repeat unit structure
2 (D39)	
6A	
6B	
19C	

Glucopyranose

Galactopyranose

Rhamnose

N-acetylmannosamine

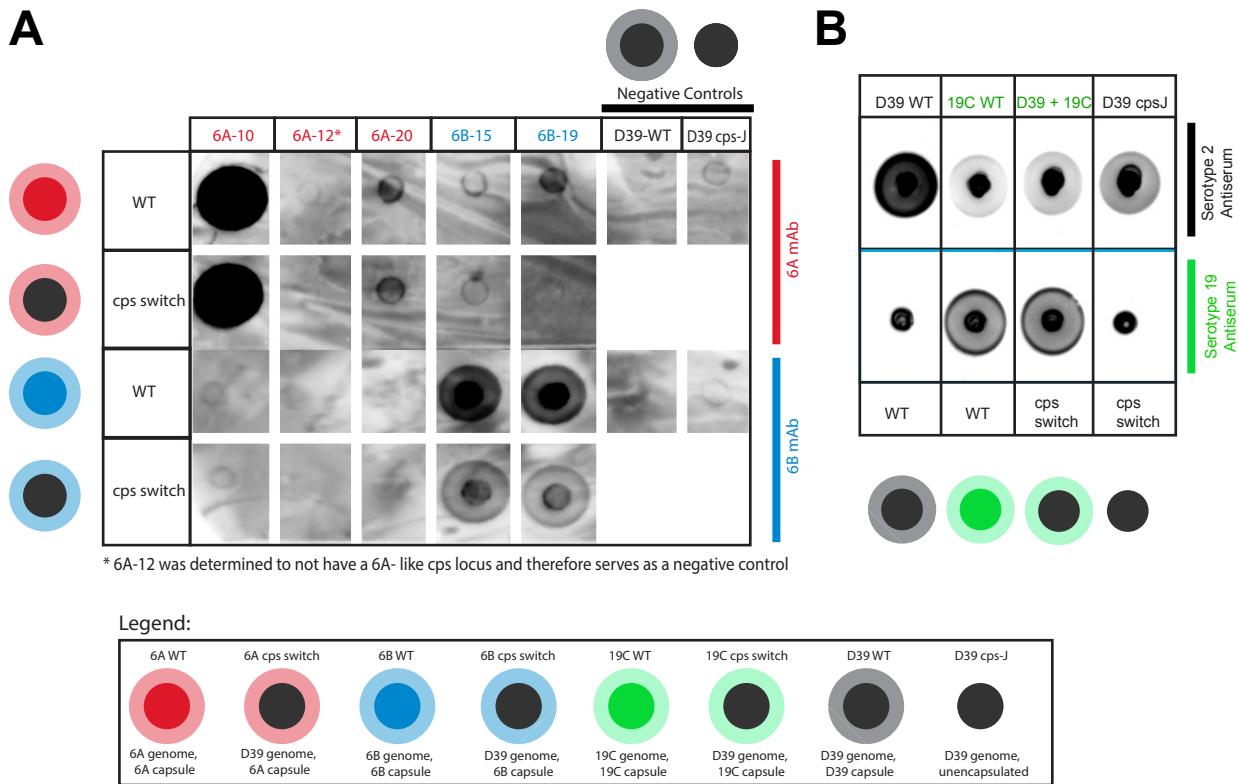
Ribitol

Ribofuranose

Glucuronic acid

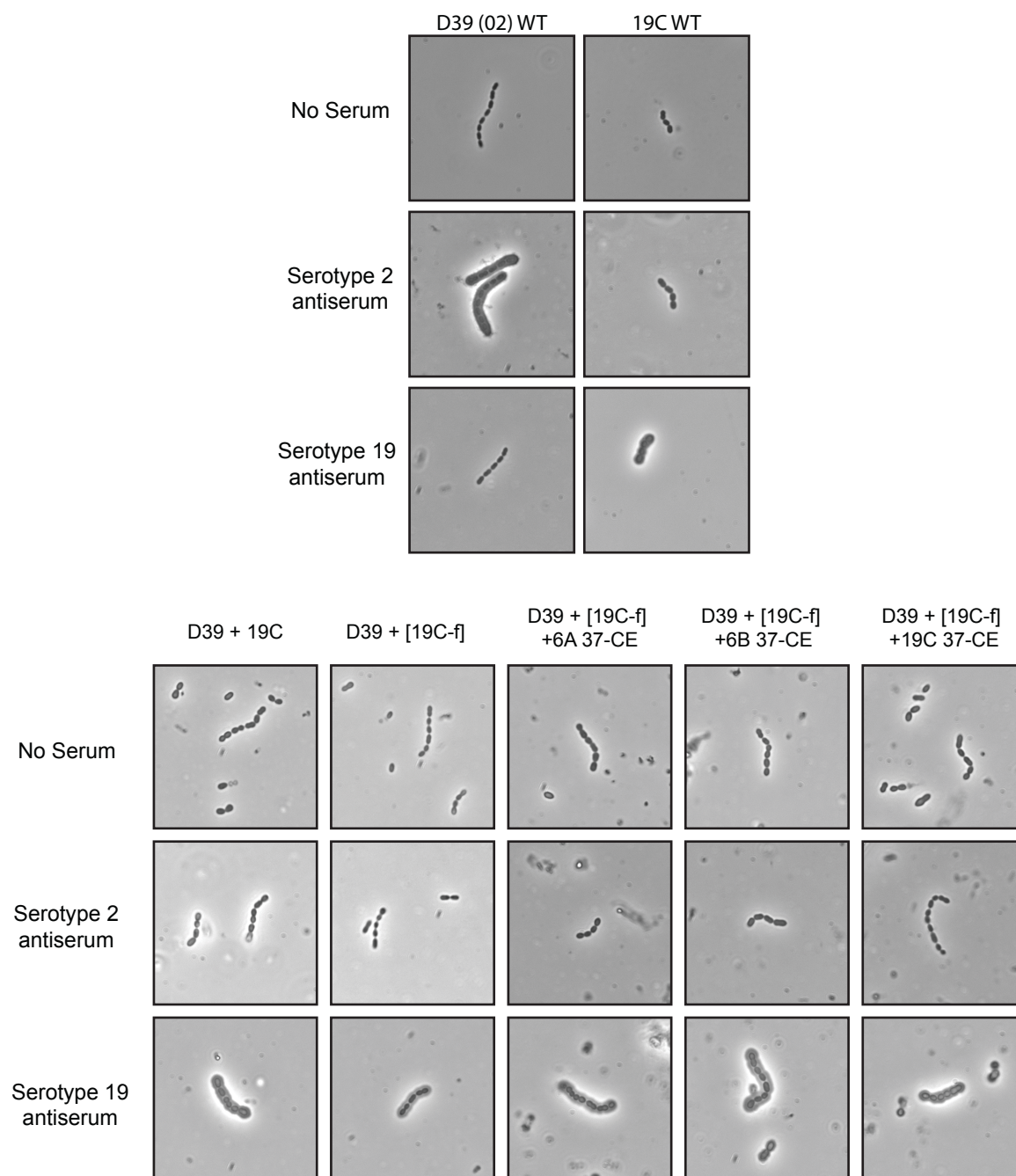
Supplementary Figure 4. Schematic diagram of the capsule chemical structures of serotype switches used and parent strain D39 (serotype 2). Note that serotype 2 and 19C do not have galactose in their capsule configuration. 6A and 6B capsules differ only by the chemical bond (1-3 versus 1-4) between rhamnose and ribitol (ref. 3).

Supplementary Figure 5



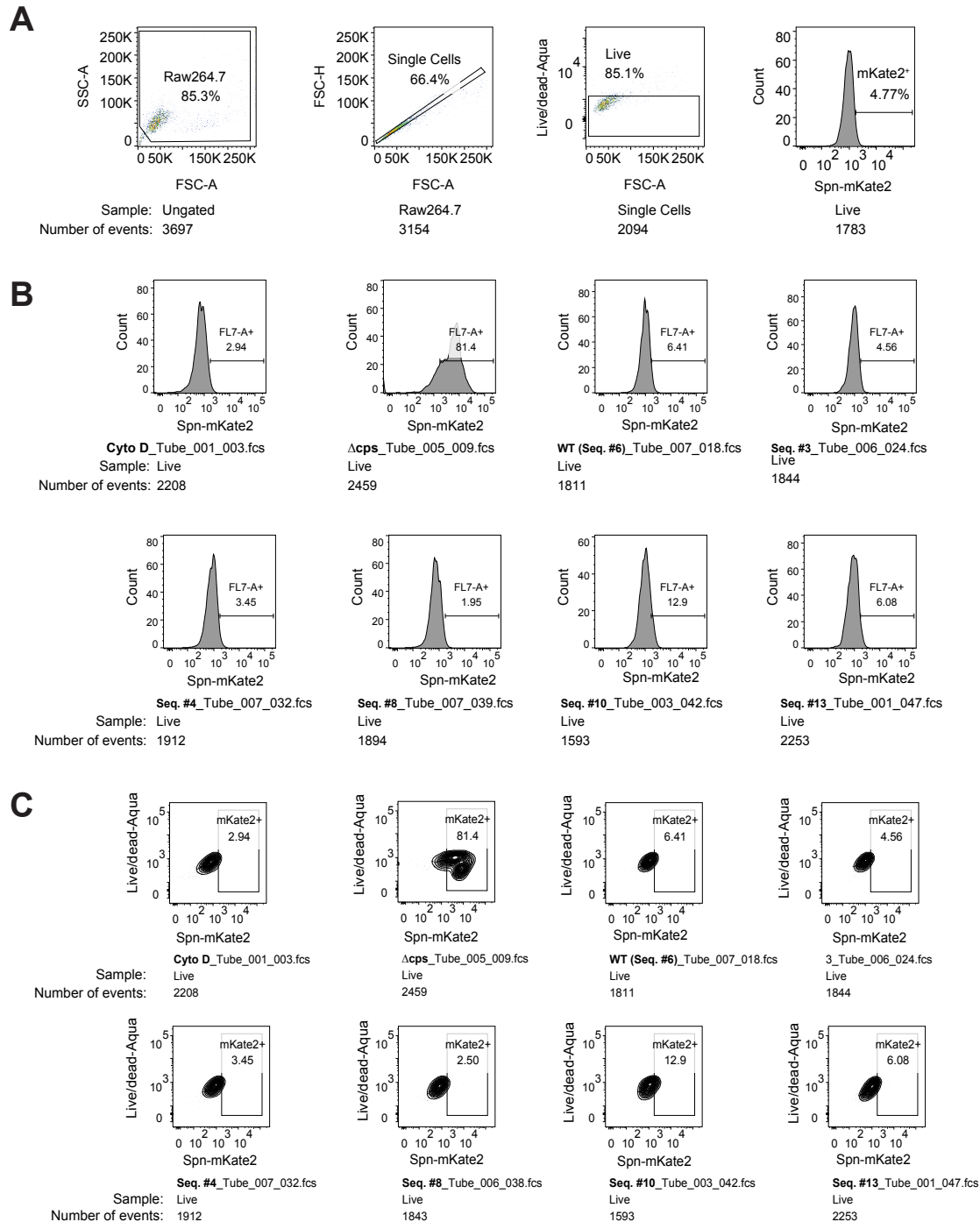
Supplementary Figure 5. Dot blot assessment of capsule switches. (A) Dot blot analysis of two 6A and two 6B wild-type strains and their respective switch strains made in a D39 serotype 2 parent strain background. Monoclonal antibodies raised against serotype 6A and 6B capsule types were provided by Dr. Moon Nahm. These dot blots were conducted two times on two separate days. (B) Dot blots using either serotype 2 or serotype 19C serum demonstrate a successful capsule switch to the 19C capsule in the D39 WT (serotype 2) genetic background. D39 WT, parent D39 serotype 2 wild-type strain; cps-J, D39 WT strain with deleted capsule locus (*cps*), where the integrated Janus construct is still in place of the *cps* locus. 19C capsule switches were then confirmed by a Quellung assay shown in Fig. S6. Source data are provided as a Source Data file.

Supplementary Figure 6



Supplementary Figure 6. Determination of the 19C capsule switch using the Quellung assay. 19C typing serum was used to identify by light microscopy the presence or absence of the 19C capsule as per ref. 4. Source data are provided as a Source Data file.

Supplementary Figure 7



Supplementary Figure 7. Gating strategy for phagocytosis assay (Fig. 3A).

(A) Gating strategy used to identify mKate2-positive (mKate2+) phagocytosing RAW264.7 macrophages. Forward scatter area (FSC-A) versus side scatter area (SSC-A) was used to gate on RAW264.7 cells and exclude debris, followed by FSC-A versus FSC-H to select single cells. Live macrophages were identified by exclusion of Live/Dead Aqua-positive events, and mKate2+ cells were defined based on D39 WT-mKate2 (Spn mKate2) fluorescence within the live, single-cell gate. (B) Representative histograms and (C) contour plots showing Spn-mKate2 fluorescence of live, single RAW264.7 macrophages. mKate2+ cells were defined based on Spn-mKate2 fluorescence within the live, single-cell gate after macrophages were incubated with unencapsulated WT D39 strain plus cytochalasin D (Cyto D), unencapsulated WT D39 strain (Δ cps), D39 WT strain (WT), or D39 WT sequence with WT 37-CE (Seq. #6) or Seq. #3, 4, 8, 10 or 13. Single representative samples are shown. Spn, *Streptococcus pneumoniae*. Flow cytometry data were analyzed using FlowJo v10.10.0. Source data are provided as a Source Data file.

Supplementary Table 1. SNPs from genomic sequencing

Strain*	Fig	Contig	Genome position	SNP	Syn vs non-syn	Coding	Gene**	Function
D39 37-CE strains (all)	3	1	317,698	G→T	S	S375Y (TCT→TAT)	<i>folC</i> ←	Dihydrofolate:folylp olyglutamate synthetase
		1	386,314	T→G	NA	intergenic (-72/+98)	<i>SpF03_</i> <i>sRNA</i> ← / ← <i>F</i> <i>CHBHC</i> <i>_02070</i>	Streptococcus sRNA SpF03/SP_0115 family bacteriocin-like peptide
		2	35,760	G→T	NS	S393S (TCC→TCA)	<i>polC</i> ←	bifunctional DnaQ family exonuclease/ATP-d ependent helicase
		5	38,368	T→G	S	F87C (TIT→TGT)	<i>pabA</i> →	Anthranilate/para-a minobenzoate synthase component II (glutamine amidotransferase)
		10	60,609	(A) _{6→7}	+NT	Coding 410 out of 540 nt (truncation)	<i>rsmD</i> ←	16S rRNA (guanine(966)-N(2))-methyltransferase RsmD
Strain*	Fig	Contig	Position	SNP	Syn vs non-syn	Coding	Gene	Function
D39 + [6B 37-CE] (Seq. #13)	3	2	228,326	C→T	NS	V276I (GTC→ATC)	<i>rimI</i> ←	Ribosomal protein S18 acetylase RimI and related acetyltransferases
		6	95,026	G→T	NS	P44Q (CCA→CAA)	<i>tarI</i> ←	Ribitol-5-phosphate cytidyltransferase
D39 + [17A 37-CE] (Seq. #1)		3	177,008	T→C	S	P421P (CCA→CCG)	<i>ugpB</i> ←	ABC-type glycerol-3-phosphat e transport system, periplasmic component
D39 + [33B 37-CE] (Seq. #8)	6	15	10,178	G→T	NS	V329F (GTT→ITT)	<i>rfaL</i> →	O-antigen ligase

* Under green bar = shared by all 37-CE strains. Under purple bar unique to designated 37-CE strains.

**Arrow indicates direction of gene transcription relative to SNP.

Supplementary Figure Legend References

1. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539. 10.1038/msb.2011.75.
2. Bentley, S.D., Aanensen, D.M., Mavroidi, A., Saunders, D., Rabinowitsch, E., Collins, M., Donohoe, K., Harris, D., Murphy, L., Quail, M.A., et al. (2006). Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet* 2, e31. 10.1371/journal.pgen.0020031.
3. Oliver, M.B., van der Linden, M.P.G., Kuntzel, S.A., Saad, J.S., and Nahm, M.H. (2013). Discovery of *Streptococcus pneumoniae* serotype 6 variants with glycosyltransferases synthesizing two differing repeating units. *J Biol Chem* 288, 25976-25985. 10.1074/jbc.M113.480152.
4. Habib, M., Porter, B.D., and Satzke, C. (2014). Capsular serotyping of *Streptococcus pneumoniae* using the Quellung reaction. *J Vis Exp*, e51208. 10.3791/51208.