Investigating the *Streptococcus sinensis* competence regulon through a combination of transcriptome analysis and phenotypic evaluation

Alec A. Brennan[1,#], Anthony Harrington[1,#], Mingzhe Guo[1], Clay P. Renshaw[1], Richard L. Tillett[2], Pedro Miura[3], and Yftah Tal-Gan[1,*]

[1] *Department of Chemistry, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada 89557, United States.*

[2] *Nevada Center for Bioinformatics, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada 89557, United States.*

[3] *Department of Biology, University of Nevada, Reno, 1664 North Virginia Street, Reno, Nevada 89557, United States.*

[#] *These authors contributed equally to this work.*

[*] *To whom correspondence should be addressed.* ytalgan@unr.edu

Additional experimental details

Solid-phase peptide synthesis. The following conditions were used for coupling and deprotection cycles in the Liberty1 automated peptide synthesizer (CEM Corp.). Deprotection of the Fmoc group was performed first using 5 mL of 2% piperidine with 2% 1, 8-Diazabicyclo [5.4.0] undec-7-ene (DBU) in DMF (90 sec, 75 °C) followed by another 5 mL of 2% piperidine with 2% DBU in DMF (90 sec, 75 °C). The resin was washed with DMF (3 x 5 mL) after each deprotection cycle. Coupling reactions were performed using 2.5 mL solution containing Fmoc-protected amino acid (5 equiv.) *N*, *N*′-Diisopropylcarbodiimide (5 equiv.) and Ethyl cyano (hydroxyimino) acetate (5 equiv.). All amino acids were coupled for 20 min (30 W, 90 °C), except Histidine. Histidine was coupled for 10 min (0 W, 25 °C) then for 40 min (20 W, 50 °C). After the synthesis was completed, the resin was washed with DMF (3 x 5 mL).

Cleavage of synthetic peptide from resin. Following synthesis, the resin was washed with diethyl ether (2 mL) and then kept in open air for 15 min before it was transferred into a 15 mL falcon tube. The peptide was cleaved with 3 mL cleavage cocktail consisting of 95% TFA, 2.5% triisopropylsilane (TIPS), and 2.5% dH_2O for 3 h with agitation. The cleaved peptide was separated from the resin by filtration, transferred to a 50 mL falcon tube and precipitated with a cold solution of diethyl ether:hexane (1:1, 45 mL, 0 °C). The precipitated crude peptide was centrifuged for 5 min at 4,600 x g, redissolved in 10 mL ACN:ddH_2O (1:1) and lyophilized before HPLC purification.

Peptide Purification. Crude peptides were purified with RP-HPLC. Standard RP-HPLC conditions were as follows: flow rates = 5 mL min^{-1} for semipreparative separations and 1 mL min^{-1} for analytical separations; mobile phase A = ddH2O + 0.1% TFA; mobile phase B = ACN + 0.1% TFA. Purities were determined by integration of peaks with UV detection at 220 nm.

Preparative HPLC methods were used to separate the crude peptide mixture to different chemical components using the following linear gradient: 5% B → 15% B over 10 min, followed by 15% B → 45% B over 23 min, and 45% B → 95% B over 13 minutes. MALDI-TOF MS was used to identify the fractions containing the desired peptides. The purity of the peptides was quantified using an analytical HPLC method with the following linear gradient: 5% B → 95% B over 27 min. Only peptide fractions that were purified to homogeneity (>95%) were used for further characterization. The identity of the isolated peptides (naturally isolated or synthesized) was confirmed by ESI+ HRMS comparing the observed mass-to-charge (m/z) ratio to the expected m/z ratio.

Temporal Determination of Endogenous CSP Production. An overnight culture of either the wild-type *S. sinensis* or *S. sinensis* transformed with the p*comX*-luciferase reporter were grown for 18 h in Todd-Hewitt broth supplemented with 0.5% yeast extract (THY) media at 37 °C and 5% CO_2. The reporter strain was incubated with 300 ng / mL of spectinomycin. Following this incubation, cultures were diluted 1:20 in fresh THY and grown to an OD_{600} of approximately 0.4 (early- to mid-logarithmic phase). Cultures were then diluted 1:100 in fresh THY media containing D-luciferin at a final concentration of 0.300 mg / mL. In a 96-well microplate, 200 µL aliquots of the 1:100 diluted culture were plated in triplicate and incubated in a BioTek BioSpa 8 automated incubator set to 37 °C and 5% CO_2. OD_{600} and luminescence readings were measured every 10 min for a total of 13 hrs, using an attached BioTek Synergy H1 multimode microplate reader. The data were processed in Excel and GraphPad Prism.

Peptide Mapping. Naturally isolated CSP was provided to the Mick Hitchcock, Ph.D. Nevada Proteomics Center for peptide mapping. Raw data files were analyzed in PEAKS Studio for confirmation of peptide sequence.

Biofilm Formation Assay. A single colony of *S. sinensis* was grown for 18 h in 10 mL THY media (pH 7.3) at 37 °C with 5% CO_2. Following incubation, *S. sinensis* was diluted 1:100 in fresh THY media containing 1% D-glucose, and 198 µL was added in triplicate to a 96-well microtiter plate. Each well contained either 2 µL of the desired concentration of peptide in DMSO, or just DMSO as a negative control. A set of wells containing only THY media containing 1% D-glucose and DMSO (no bacteria) was included for background subtraction. The plate was then statically incubated at 37 °C with 5% CO_2 for 24 h, and following incubation the absorbance at 600 nm (A600) was recorded. The contents of all wells were carefully decanted by shaking the plate gently over a glass basin. Experimental wells were then gently washed one time with 200 µL 1x PBS. To heat fix bacterial biofilms to the bottom of the well, the 96 well microtiter plate was incubated at 55 °C for 3 h. Following heat fixing, 200 µL of a 0.1% crystal violet solution was added to each well, and the solution was allowed to rest at room temperature for 5 min. The wells were then carefully decanted and washed once with 200 µL water. Following this, a total of 200 µL of a 30% (v/v) acetic acid in water was added to the wells. The plate was shaken for 15 min at 37 °C, and experimental wells were then diluted 1:5 in water. The absorbance at 595 nm (A595) was then measured for each well. Each A595 value was divided by its corresponding A600 values. Experiments were performed in triplicate. Data is presented as the percent biofilm formation relative to wild type untreated with exogenous CSP. Results are expressed as the mean +/- the standard deviation of three independent experiments.

Hemolysis Assay. A single colony of *S. sinensis* was grown for 18 h at 37 °C with 5% CO_2 in 10 mL THY media (pH 7.3). Following incubation cultures were diluted 1:25 in fresh THY media, and 198 µL was added in triplicate to a 96-well microtiter plate. Each experimental well contained 2 µL of the desired concentration of peptide in DMSO. A positive control was prepared by adding

2 µL of a 1% Triton X solution in 198 µL THY media, and negative controls were prepared by adding 2 µL DMSO in either 198 µL 1:25 diluted culture or 198 µL THY. The 96-well microtiter plate was then incubated for 5 h at 37 °C with 5% CO_2, after which the A600 was recorded and hemolysis was assessed. A 1 mL portion of defibrinated rabbit red blood cells (VWR) was aliquoted into a sterile 1.5 mL microcentrifuge tube and centrifuged at 2,000 rpm for 2 min. Following centrifugation, the top layer was aliquoted off, and red blood cells were gently washed with 1 mL 1x PBS. This process was repeated for a total of three times, until following centrifugation the top layer was mostly clear. Washed red blood cells were resuspended in 1 mL PBS, and a 15 µL aliquot was added to each well of the 96-well microtiter plate. The plate was then incubated for 30 min at 37 °C. Following incubation, the plate was centrifuged for 4 min at 4 °C at 1,600 rpm, and a 40 µL portion of the resulting supernatant was carefully removed and placed in a fresh 96-well microtiter plate. Experimental wells were then diluted 1:5 in water to prevent saturation of the detector by the positive control. The A420 for each experimental well was then recorded. Experiments were performed in triplicate. Data is presented as the percent hemolysis relative to the .01% Triton X positive control. Results are expressed as the mean +/- the standard deviation of three independent experiments.

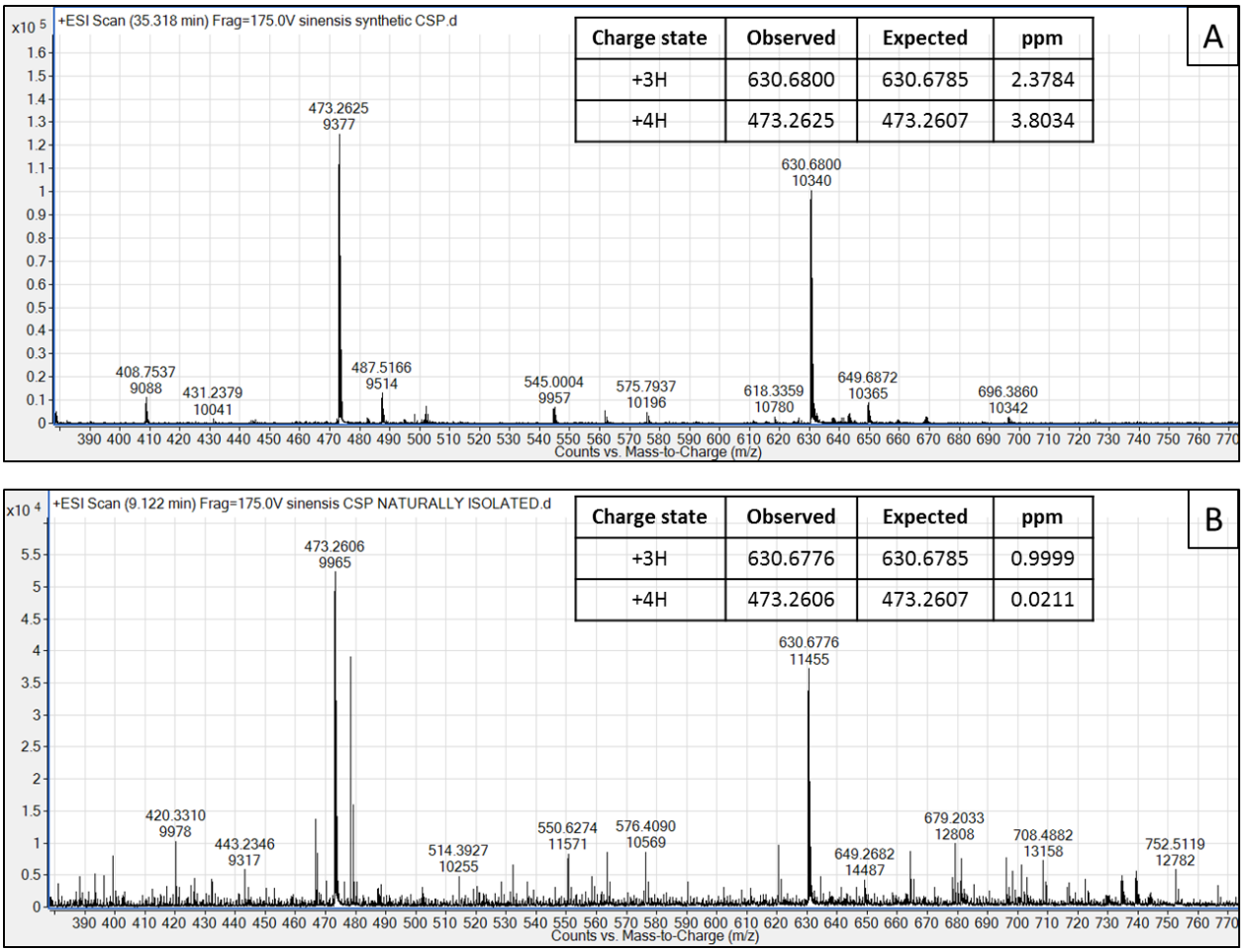
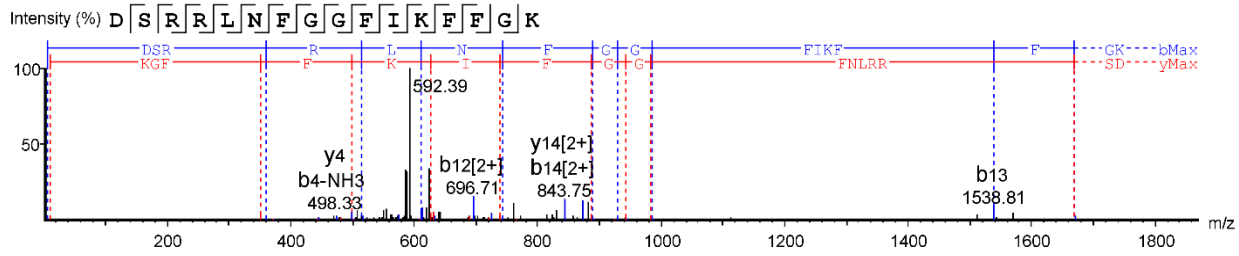


Figure S1. High Resolution Mass Spectrometry (HRMS) results: A) *S. sinensis* synthetic CSP, B) *S. sinensis* naturally isolated CSP.



#	b	b-H_2O	b-NH_3	b(2+)	Sequence	y	y-H_2O	y-NH_3	y(2+)	#
1	116.03	98.02	99.01	58.52	D	-	-	-	-	16
2	203.07	185.06	186.04	102.03	S	1773.99	1755.98	1756.96	886.73	15
3	359.28	341.16	342.14	180.08	R	1686.96	1669.05	1669.05	843.73	14
4	515.31	498.25	498.25	258.13	R	1530.86	1512.85	1513.83	766.14	13
5	628.41	610.39	610.39	314.68	L	1374.42	1357.67	1357.67	688.23	12
6	742.84	725.10	725.10	371.15	N	1261.67	1243.66	1244.65	630.53	11
7	889.54	872.18	812.18	445.07	F	1148.26	1129.62	1130.60	574.73	10
8	946.56	929.45	929.45	473.33	G	1000.09	982.76	982.76	500.22	9
9	1003.51	986.45	986.45	502.06	G	943.41	925.52	925.52	472.05	8
10	1150.48	1132.57	1133.55	576.05	F	886.73	868.50	870.39	442.78	7
11	1263.45	1245.65	1246.63	632.50	I	739.63	721.44	722.49	371.15	6
12	1391.63	1374.42	1374.42	696.76	K	625.90	608.35	608.35	313.68	5
13	1538.81	1520.81	1521.80	770.13	F	498.25	480.34	480.34	249.64	4
14	1685.87	1668.07	1668.07	843.69	F	351.23	333.19	334.18	176.10	3
15	1742.91	1724.90	1725.89	872.18	G	204.13	186.12	187.11	102.57	2
16	-	-	-	-	K	147.11	129.10	130.09	74.06	1

Figure S2. MS/MS analysis of the extracted *S. sinensis* CSP. Top: *S. sinensis* CSP sequence showing the various MS/MS fragments detected in the raw spectrum. Bottom: Complete table of expected masses for the MS/MS analysis of the predicted *S. sinensis* CSP with highlighted masses correspond to ion peaks detected.

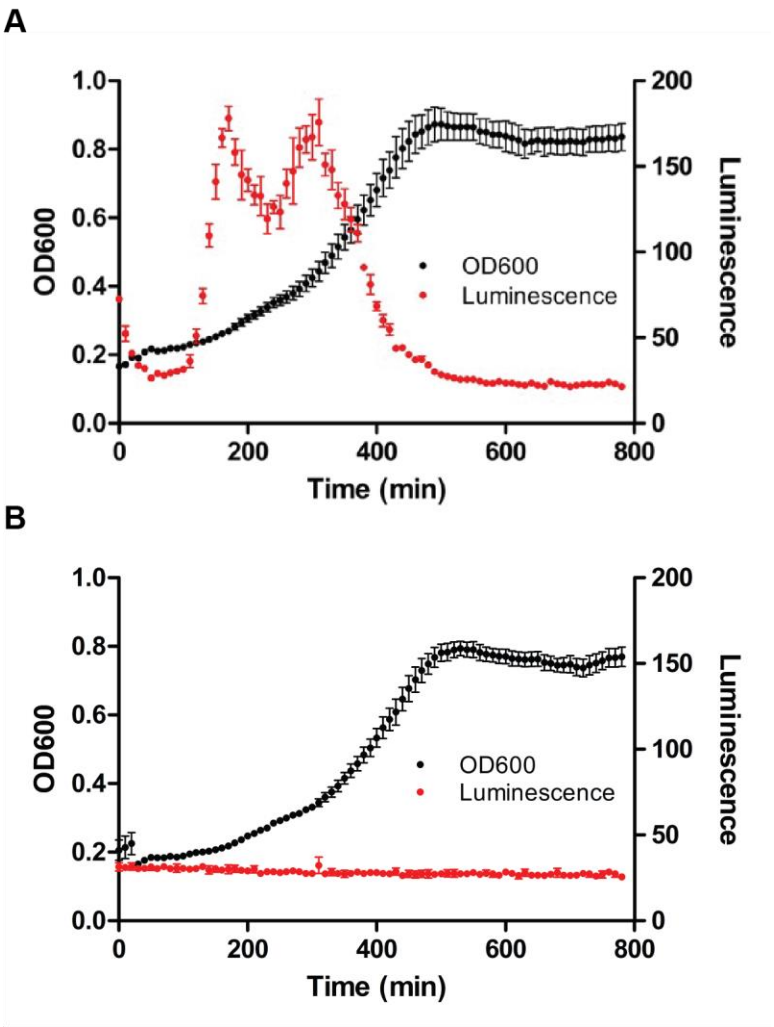


Figure S3. Temporal determination of endogenous CSP production via luminescence detection. A) *S. sinensis* transformed with the developed luciferase-based reporter system demonstrated increasing luminescence during the early exponential growth phase as a result of endogenous CSP production. B) The wild-type *S. sinensis* strain did not exhibit luminescence.

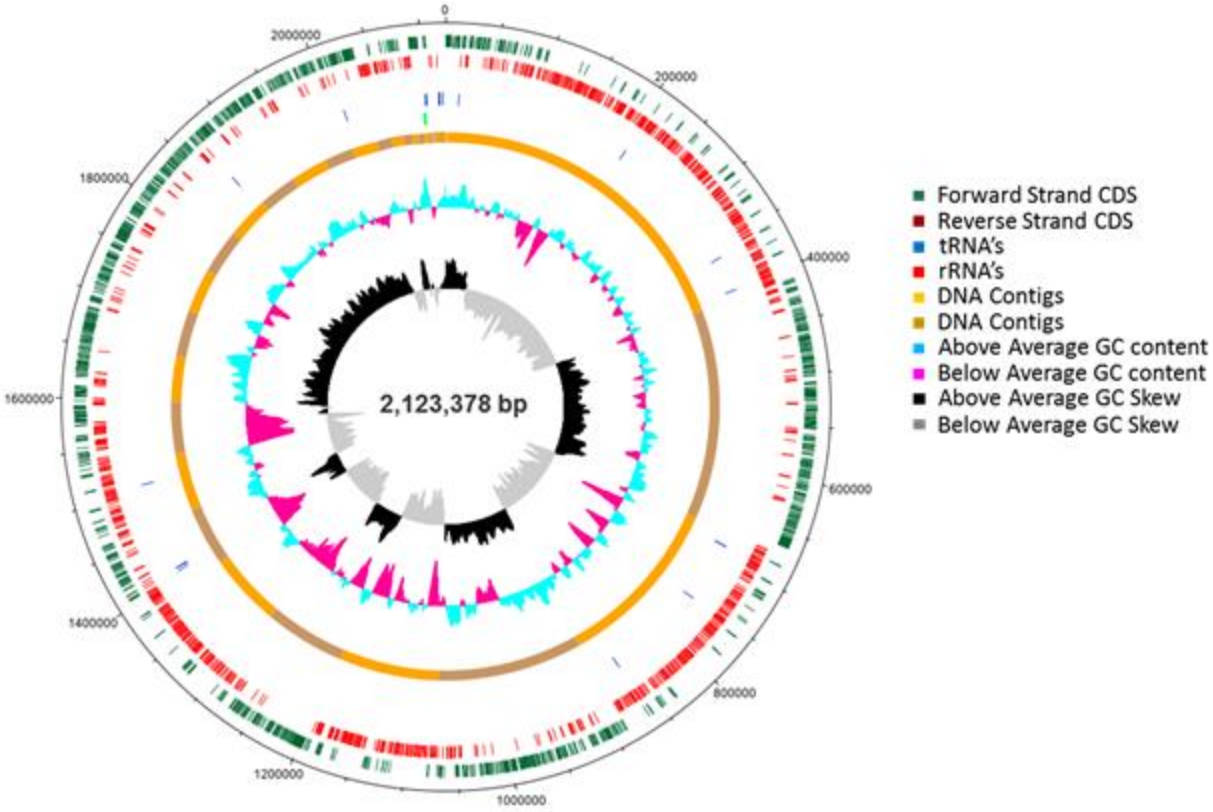


Figure S4. ***S. sinensis*** **Forsyth1A draft genome map visualized with DNAplotter.** The assembled genome is 2.12 Mbp made of 61 contigs. This was established via de novo genome assembly and is not a closed genome.

Table S1. Predicted virulence genes comparison between strains Forsyth1A, DD04, and HKU4.

Virulence Factor Class	Genes	Annotation	Locus	DD04	HKU4
Adherence	*pavA*	Fibronectin-binding proteins	FOA32_000330	☑	☑
	lmb	Laminin-binding protein	FOA32_001121	☑	☑
	srtA	Sortase A	FOA32_001697	☑	☑
	slrA	Streptococcal lipoprotein rotamase A	FOA32_000620	☑	☑
	plr/gapA	Streptococcal plasmin receptor	FOA32_001298	☑	☑
	srtB	Sortase B	FOA32_001612		☑
Biofilm & Capsule Production	*eno*	Streptococcal enolase	FOA32_000730	☑	☑
	cap3	UDP-glucose-6-dehydrogenase	FOA32_000241	☑	☑
	galE	UDP-glucose-4-epimerase	FOA32_07480	☑	☑
Manganese uptake	*psaA*	Pneumococcal surface antigen A	FOA32_000346	☑	☑
Protease	*htrA*	serine protease	FOA32_001484	☑	☑
	cppA	carbon catabolite protein A	FOA32_000227	☑	☑
	tig/ropA	Trigger factor	FOA32_001645	☑	☑
Immune Evasion	Capsule	LPS biosynthesis protein	FOA32_000253	☑	☒
		transport permease protein	FOA32_000251	☑	☑
		alpha-L-Rha alpha-1,3-L-rhamnosyltransferase	FOA32_000250	☑	☑
		rhamnosyltransferase	FOA32_000249	☑	☑
		glycosyl transferase	FOA32_000248	☑	☑
		beta-carotene 15,15'-monooxygenase	FOA32_000247	☑	☑
		glycosyl transferase	FOA32_000246	☑	☑
		glycosyl transferase	FOA32_000245	☑	☑
		membrane protein	FOA32_000244	☑	☑
		glycosyl transferase family 2	FOA32_000243	☑	☑
		glycosyl transferase	FOA32_000242	☑	☑
		glycosyl transferase	FOA32_000241	☑	☑
		sugar transporter	FOA32_000240	☑	☑
		undecaprenyl-phosphate alpha-N-acetylglucosaminyl 1-phosphate transferase	FOA32_000036	☑	☑
		UDP-glucose dehydrogenase	FOA32_000971	☑	☑
		short-chain dehydrogenase	FOA32_000983	☑	☑
		tyrosine protein kinase	FOA32_000984	☑	☑
		capsular polysaccharide biosynthesis protein CpsC	FOA32_000985	☑	☑
		tyrosine protein phosphatase	FOA32_000986	☑	☑
		LytR family transcriptional regulator	FOA32_000987	☑	☑
		UTP-glucose-1-phosphate uridylyltransferase	FOA32_001995	☑	☒

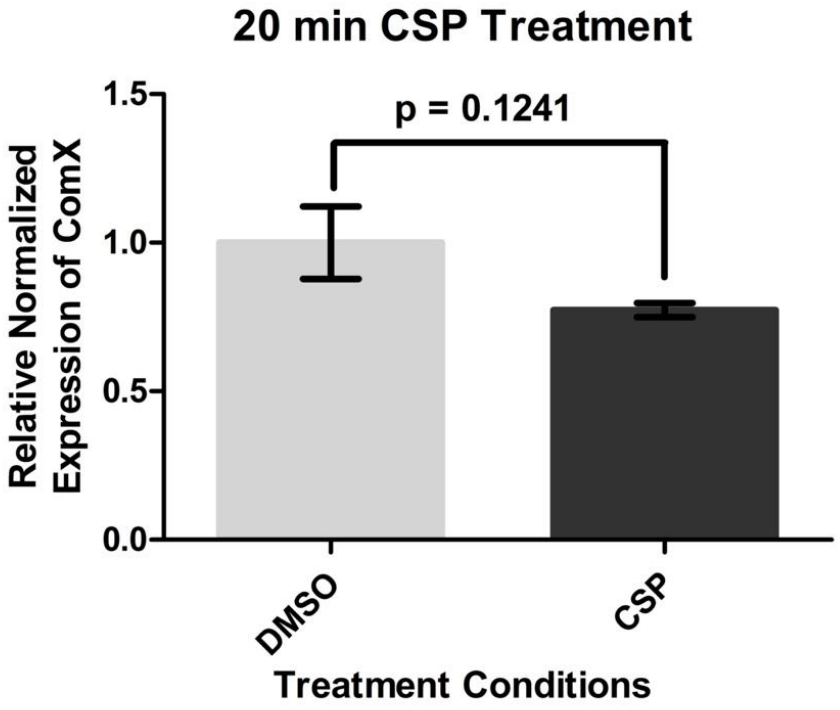


Figure S5. RT-qPCR results after 20-minute synthetic CSP treatment. Treatment with exogenous CSP for 20 minutes does not lead to continued activation of the *comX* gene. After a 20-minute incubation, stop solution was added to each sample to stop transcription. Data are normalized to *gyrB*. The RT-qPCR experiment was conducted using two biological replicates for both the CSP (positive) and DMSO (negative) conditions.

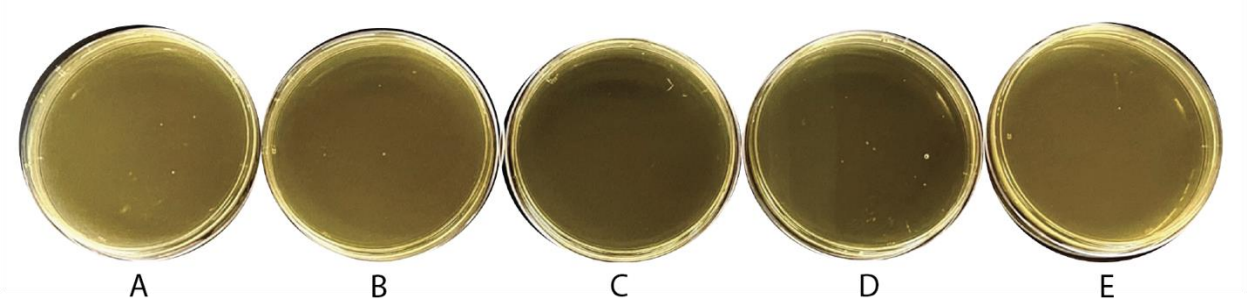


Figure S6. *S. sinensis* Forsyth1A transformation assay with HKU4 CSP. *S. sinensis* Forsyth1A is not capable of competence induction (and successful uptake of added spectinomycin-resistant pDL278) with the addition of HKU4 CSP. The synthetic CSP was added at final concentrations of 10,000 nM (A), 1,000 nM (B), 100 nM (C), and 10 nM (D). The control includes no exogenous CSP (E). See Materials and Methods for full experimental details.

Table S2. Quantification of *S. sinensis* biofilm formation. A) OD_{595} values following treatment of biofilm growth with crystal violet solution. B) Quantification of total biofilm growth in cultures incubated with exogenous CSP compared to cultures with no addition of exogenous CSP; expressed as a percentage. Media (THY)-only samples provided a negative control.

A)

Plate 1: OD_{595} *

[CSP]	10,000 nM	1,000 nM	100 nM	10 nM	DMSO	THY
Replicate 1	4.000	3.590	3.620	3.625	3.625	0.845
Replicate 2	3.645	3.955	3.650	3.495	3.725	1.245
Replicate 3	3.765	3.580	3.545	3.490	3.470	1.055
Average	3.803 ±	3.708 ±	3.605 ±	3.537 ±	3.607 ±	1.048 ±
Std. Dev.	0.1806	0.2137	0.0541	0.0765	0.1285	0.2001

B)

Plate 1: OD_{595}/OD_{600} (%) **

[CSP]	10,000 nM	1,000 nM	100 nM	10 nM	DMSO
Replicate 1	113.585	99.622	108.383	110.804	104.337
Replicate 2	89.746	127.954	115.182	95.979	102.241
Replicate 3	96.329	105.644	96.959	95.461	93.422

* Values were back-calculated to account for the dilution factor (1:5)

** Values are normalized to DMSO (no addition of exogenous CSP) as 100% biofilm formation. See calculation below.

$$\%\ Biofilm\ formation = \frac{(Replicate\ OD_{595} - Ave.\ THY\ OD_{595}) / (Replicate\ OD_{600})}{(Ave.\ DMSO\ OD_{595} - Ave.\ THY\ OD_{595}) / (Ave.\ DSMO\ OD_{600})} * 100\%$$

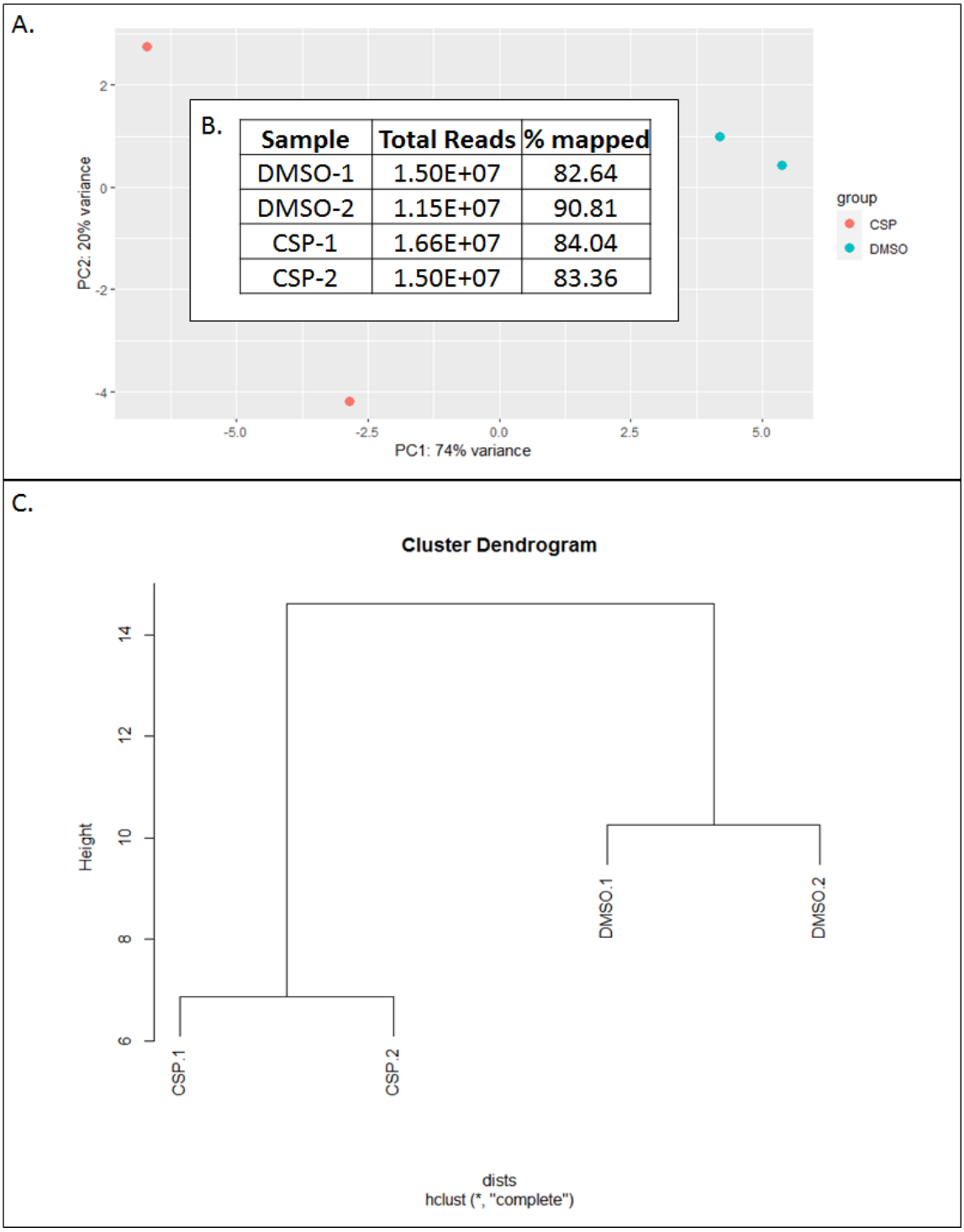


Figure S7. A) Principal component analysis using the regularized log (rlog) output of each sample count data. B) Table summarizing the total reads and percent mapped to Forsyth1A genome C) Cluster dendrogram using Euclidean distance of each sample's rlog transformation to determine grouping of replicates.