#### **Supplemental Information for**

# Saturation mutagenesis genome engineering of infective $\Phi X174$ bacteriophage via unamplified oligo pools and golden gate assembly

Matthew S. Faber<sup>1</sup>, James T. Van Leuven<sup>2,3</sup>, Martina M. Ederer<sup>3</sup>, Yesol Sapozhnikov<sup>3</sup>, Zoë L. Wilson<sup>3</sup>, Holly A. Wichman<sup>2,3</sup>, Timothy A. Whitehead<sup>4,5</sup>, Craig R. Miller<sup>2,3\*</sup>

Affiliations: <sup>1</sup>Dept. Biochemistry and Molecular Biology, Michigan State University, East Lansing MI, 48824; <sup>2</sup>Institute for Modeling Collaboration and Innovation, University of Idaho, Moscow ID, 83844; <sup>3</sup>Dept. of Biological Sciences, University of Idaho, Moscow ID, 83844; <sup>4</sup>Dept. of Chemical & Biological Engineering; University of Colorado, Boulder, CO 80303; <sup>5</sup>Dept. of Chemical Engineering & Materials Science; Michigan State University, East Lansing MI, 48824

\*Correspondence to:

jvanleuven@uidaho.edu Department of Biological Sciences 875 Perimeter MS 3051

Moscow, Idaho 83844-3051



Supplemental Figure S1. Introduction of nicking sites into shuttle vectors containing viral genes. Verification of the introduction of the BbvCI nicking site into the shuttle vector containing the viral genes is shown by the template prep step in the nicking mutagenesis protocol (generation of ssDNA). Samples were run on a 1% agarose gel with SYBR<sup>™</sup> Safe DNA gel stain (Invitrogen) added before casting, the ladder used is the 1 kb DNA ladder from GoldBio.



Supplementary Figure S2: F1 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S3: F2 tile 1 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S4: F2 tile 2 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S5: F3 tile 1 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S6: F3 tile 2 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S7: G1 tile 1 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S8: G1 tile 2 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S9: G2 heatmap of mutational frequency represented by number of sequencing read counts "counts".



Supplementary Figure S10: Heatmap of relative frequencies for mutations in the F1 fragment.



Supplementary Figure S11: Heatmap of relative frequencies for mutations in the F2T1 fragment.



Supplementary Figure S12: Heatmap of relative frequencies for mutations in the F2T2 fragment.



Supplementary Figure S13: Heatmap of relative frequencies for mutations in the F3T1 fragment.



Supplementary Figure S14: Heatmap of relative frequencies for mutations in the F3T2 fragment.



0 0.25 0.5

≥2

Relative Frequency (x1000)

Supplementary Figure S15: Heatmap of relative frequencies for mutations in the G1T1 fragment.



Supplementary Figure S16: Heatmap of relative frequencies for mutations in the G1T2 fragment.



 Relative Frequency (x1000)

 0
 0.25
 0.5
 1
 ≥2

Supplementary Figure S17: Heatmap of relative frequencies for mutations in the G2 fragment.

**Supplementary Figure S18: Agarose gel image showing effectiveness of DNase1 in ΦLB.** 25 cycles of PCR were performed after a 1-hour 37°C incubation of WT ligation mix with DNase1. Primer sequences for phix\_2375F and phix\_2949R were 5'-ACACTGACGACATGGTTCTACAtctgcttaggagtttaatc-3' and 5'-TACGGTAGCAGAGACTTGGTCTgcaccaaacataaatcacc-3'. These primers include annealing sequences for Illumina 2<sup>nd</sup> round barcoding primers.



## Supplemental Figure S19. Consistency between transformation replicates.

Virus titers for four replicate transformations were compared by Anova (global p.adj = 0.006) and pairwise tests with adjusted p-values (Wilcoxon) shown on the plot. These experiments were done using 1 uL of transformation mix, as arcing occasionally occurs with 4 uL. The G1/G2 construct is the result of ligating together the G1 and G2 libraries.



# Supplemental table S1. Primers for incorporating BbvCI nicking sites into the pCR2.1-topo shuttle vector. Blue text is the overlap region with the shuttle vector, red is

the KpnI site, green is the BbvCI site.

Nick_incorporation_fwd
GCTCTACGGGTACCGCTGAGGGAGCTCGGATCCACTAGTAACG
Nick_incorporation_rvs
GCTCTAACGGGTACCAAGCTTGGCGTAATCATGGTC

# Supplemental table S2. Inner and outer primers for PCR reactions for Illumina

**sequencing.** Red indicates overhang regions for attaching Illumina adapter primers (inner PCR primers) or overhangs for attaching to inner PCR product (outer PCR primers), black is the overlap region in the gene or the barcode, blue is the Illumina adapter.

Inner PCR Primers	Sequence (5' to 3')
Fragment F1 Fwd	gttcagagttctacagtccgacgacccttacttgaggataaatt
Fragment F2 Tile 1 Fwd	gttcagagttctacagtccgacgatcgactcactatagggcgaa
Fragment F2 Tile 2 Fwd	gttcagagttctacagtccgacgatcagcttaatcaagatgatgct
Fragment F3 Tile 1 Fwd	gttcagagttctacagtccgacgatccgtctctctgggca
Fragment F3 Tile 2 Fwd	gttcagagttctacagtccgacgatcgaaggatgttttccgt
Fragment G1 Tile 1 Fwd	gttcagagttctacagtccgacgatcgaattgggccctctag
Fragment G1 Tile 2 Fwd	gttcagagttctacagtccgacgatcggttaatgctggtaatgg
Fragment G2 Fwd	gttcagagttctacagtccgacgatcggccctctagatgca
Fragment F1 Rvs	ccttggcacccgagaattccattcgtctcacagtcgg
Fragment F2 Tile 1 Rvs	ccttggcacccgagaattccacaacggaaaccataacg
sdFragment F2 Tile 2 Rvs	ccttggcacccgagaattccattcgtctccccagag
Fragment F3 Tile 1 Rvs	ccttggcacccgagaattccatcttagacgaatcaccaga
Fragment F3 Tile 2 Rvs	ccttggcacccgagaattccatcacactcaatcttttatca
Fragment G1 Tile 1 Rvs	ccttggcacccgagaattccatgcaatgaagaaaacca
Fragment G1 Tile 2 Rvs	ccttggcacccgagaattccacttcgtctccgtacg
Fragment G2 Rvs	ccttggcacccgagaattccattgaccgcctcca
Illumina outer primer adapter	aatgatacggcgaccaccgagatctacacgttcagagttctacagtccga
Illumina outer PCR adapters and barcodes	
RPI31 (Fragment F1)	caagcagaagacggcatacgagatATCGTGgtgactggagttccttggcacccgagaattcca
RPI15 (Fragment F2 Tile 1)	caagcagaagacggcatacgagatTGACATgtgactggagttccttggcacccgagaattcca
RPI16 (Fragment F2 Tile 2)	caagcagaagacggcatacgagatGGACGGgtgactggagttccttggcacccgagaattcca
RPI17 (Fragment F3 Tile 1)	caagcagaagacggcatacgagatCTCTACgtgactggagttccttggcacccgagaattcca
RPI18 (Fragment F3 Tile 2)	caag cag a a g a c g g c a t a c g a g a t G C G G A C g t g a c t g g a g t c c t t g g c a c c g a g a a t t c c a g a g a t c c a g a g a t c c a g a g a t c c a g a g a t c c a g a g a t c c a g a t c a g a t c a g a t c a g a t c a t c a g a t c
RPI19 (Fragment G1 Tile 1)	caagcagaagacggcatacgagatTTTCACgtgactggagttccttggcacccgagaattcca
RPI20 (Fragment G1 Tile 2)	caagcagaagacggcatacgagatGGCCACgtgactggagttccttggcacccgagaattcca
RPI21 (Fragment G2)	caagcagaagacggcatacgagatCGAAACgtgactggagttccttggcacccgagaattcca

**Supplemental Table S3. Mutant library preparation summary.** Summary table of the transformants required for sufficient library coverage, transformants obtained during comprehensive mutant library preparation by nicking scanning mutagenesis, and the fold excess of the number of transformants required for coverage.

Gene mutated	F1	F2	F3	G1	G2
Number of Residues	95	144	182	132	40
Transformants obtained following NSM	620,000	890,000	370,000	620,000	640,000
Required transformants for 99.9% coverage of possible library	13781	20889	26401	19148	5803
Fold excess over amount required for coverage	45	43	14	32	110

**Supplementary Table S4. Mutant plasmid library NGS statistics.** Summary table of the libraries prior to viral genome assembly. Long fragments required separate amplicon sequencing reactions as shown.

Fragment	F1	F2		F3		G1		G2
Tile Number	1	1	2	1	2	1	2	1
Residues	1-95	99-157	158-242	246-336	337-427	2-53	55-133	136-175
Sequencing reads post								
quality filter	1163959	681935	1714460	800377	757177	747800	969275	813221
Fold oversampling of								
codon combinations	441.1	499.8	851	364.8	345.1	632.8	517.6	821.9
Percent of reads with:								
No nonsynonymous								
mutations	23.0%	63.3%	52.3%	54.0%	55.0%	68.3%	48.8%	25.2%
One nonsynonymous								
mutation	52.6%	27.5%	36.3%	33.1%	32.1%	24.1%	39.8%	59.7%
Multiple								
nonsynonymous								
mutation	24.4%	9.2%	11.3%	12.9%	12.8%	7.6%	11.4%	15.1%
Coverage of possible								
single nonsynonymous								
mutations	100%	100%	99.8%	100%	99.9%	99.6%	100%	100%

### Supplemental Table S5. Viral growth rate during transformation cell recovery.

Viral titers immediately after transformation (0 min) and then for every 30 min after up to 2 hours. As the XL1-Blue cells are not susceptible, the increase in titer is the result of cell bursts, which releases viable phage. While the virus is replicating for the first 30 minutes, they are all held within a cell and will only make one plaque.



Time	Dilution	<b>#Plaques</b>	Titer
(min)			(viruses/mL)
0	1E-3	101	101,000
30	1E-3	90	90,000
60	1E-4	177	1,770,000
90	1E-4	245	2,450,000
120	1E-5	56	5,600,000

**Supplemental Table S6. List of observed mutations in plaque sequencing.** Residues where more than one substitution was observed are bolded. Residues where mutations have been observed in previous studies (but not the exact substitution) are indicated<sup>\*</sup>.

Fragment	Mutations
G1	L20T, S28I, Q33N, T34N, T34I, A37S, S46M, S46C, A51L, T64D, A69P,

	V75C, A83S, F87Q, C90Q, S97F, T102M, L103V, <b>A106C*</b> , <b>A106Q*</b> , T122G, V127N
G2	T151G, A152F, T153L, K154A, <b>R156L</b> , <b>R156M</b>
F1	H13M, <b>D14H</b> , <b>D14Q</b> , L15C, S16E, I29G*, M44Q, D45Q*, L52M, F68V,
	G79A, K84V*, D88M, <b>N91T, N91Y, T93E, T93L, T93Y</b>
F2	N134K, T145R, E146D, P149V, E177W, S181N, Q183P, T185S*, L196M,
	H204I*, T205D*, Y211H, S222E, N233R
F3	D254S, A296Y, N304H, G307S, A308Q, A308Q, G316A, L320T, K342V,
	Y354Q, E374C, S377Q, G378Y, Q393T*

**Supplemental Table S7. Testing for the optimal amount of ligation mix to transform.** Different amounts and dilutions of wildtype ligation mix were transformed into XL1-Blue cells to test for effects of DNA and salt concentration on transformation efficiency. The most transformants results from transforming 4 uL of ligated fragments.



Dilution	Amount in	Plate	<b>#Plaques</b>	Titer
	transforma	allution		(viruses
	uon			/mlj
0	4 uL	1E-3	268	268,000
0	1 uL	1E-3,1E-2	23, 220	22,500
10X	1 uL	1E-2	57	5,700
100X	1 uL	1E-1	0	0
2X	1 uL (clean)	1E-3,1E-2	23, 99	16,450